Prodt/20-182

UNIVERSIDAD DE GRANADA
Facultad de Ciencias
Fecha 19-7-96
ENTRADA NUM. 1540

UNIVERSIDAD DE GRANADA
15 JUL. 1996
COMISION DE DOCTORADO

T
15
132

UNIVERSIDAD DE GRANADA
FACULTAD DE CIENCIAS

DEPARTAMENTO DE FÍSICA APLICADA

DESARROLLO DE FILTROS DIGITALES NO LINEALES BASADOS EN MODELOS MULTIRRESOLUCIÓN

TESIS DOCTORAL

por
Pedro Luis Luque Escamilla

BIBLIOTECA UNIVERSITARIA
GRANADA
Nº Documento 6196841(3
Nº Copia 2134711

GRANADA • 1996

DESARROLLO DE FILTROS DIGITALES NO LINEALES BASADOS EN MODELOS MULTIRRESOLUCIÓN

TESIS DOCTORAL
GRANADA • 1996

UNIVERSIDAD DE GRANADA
FACULTAD DE CIENCIAS

DEPARTAMENTO DE FÍSICA APLICADA

DESARROLLO DE FILTROS DIGITALES NO LINEALES BASADOS EN MODELOS MULTIRRESOLUCIÓN

TESIS DOCTORAL

por
Pedro Luis Luque Escamilla

Memoria presentada para optar al grado de Doctor en Ciencias Físicas

Codirigida por los profesores:

Dr. D. José Martínez Aroza
Profesor Titular de Universidad
Departamento de Matemática Aplicada
Universidad de Granada

Dr. D. Ramón Román Roldán
Catedrático de Universidad
Departamento de Física Aplicada
Universidad de Granada

GRANADA, Julio de 1996

D. José MARTÍNEZ AROZA, doctor en Ciencias Matemáticas y Profesor Titular del Departamento de Matemática Aplicada de la Universidad de Granada; y D. Ramón ROMÁN ROLDÁN, doctor en Ciencias Físicas y Catedrático del Departamento de Física Aplicada de la Universidad de Granada,

MANIFIESTAN: que la presente Tesis Doctoral intitulada

"Desarrollo de Filtros Digitales No Lineales Basados en Modelos Multirresolución"

redactada y presentada por D. Pedro Luis Luque Escamilla, corresponde a la investigación realizada por el licenciado bajo su dirección, y con esta fecha autorizan su presentación.

Granada, 15 de Julio de 1996

Fdo. José MARTÍNEZ AROZA

Fdo. Ramón ROMÁN ROLDÁN

PREÁMBULO

"Desde que llegó su libro a mis ma
parado de reirme. Algún día tendré

Grouch

Aunque tradicionalmente es difícil que matemáticos y físicos trabajen conjuntamente en un proyecto común, la presente memoria ratifica lo contrario. De la colaboración del Departamento de Matemática Aplicada con el Departamento de Física Aplicada, ambos de la Universidad de Granada, surgió ya hace tiempo el Grupo de Investigación *Procesamiento de Imágenes*, en el seno del cual se inició la tarea que hoy concluye. Sin duda, la tarea investigadora se ve enriquecida por la potencia del esfuerzo multidisciplinar.

Este trabajo se enmarca dentro del ámbito del *procesamiento digital de imágenes*, en el cual se desarrollan las técnicas encaminadas a mejorar la calidad de las imágenes digitales. Entre ellas destacan las no lineales, por su capacidad para eliminar ruido de tipo impulsivo. El filtro de mediana [CITA] ha probado ser un método excelente para la supresión de dicho ruido, además de resultar especialmente simple, y tener propiedades recomendables, como la relativa preservación de detalles en la imagen. Sin embargo, este filtro, y otros surgidos a su alrededor presentan algunos inconvenientes, los cuales han intentado subsanarse [CITAS]. En la presente Tesis Doctoral se proponen y defienden nuevos algoritmos de filtrado no lineal, destinados a la eliminación de ruido impulsivo presente en la imagen, mejorando siempre las capacidades del filtro de mediana y similares. Estos métodos de procesamiento se fundamentan en lo que hemos denominado *máscara de estimación de ruido*. Se trata de una imagen binaria que señala localiza los píxeles afectados por el ruido, de manera que se filtran sólamente aquellas partes de la imagen que realmente lo necesitan. Tanto la máscara como el filtro en sí operan utilizando la información proporcionada por las *matrices de composición*—o modelos— , las cuales indican la probabilidad de que un píxel de una cierta región de la imagen tenga un determinado nivel de gris. En este estudio, las matrices de composición definen unívocamente la clase de imágenes de la cual proceden. Así, existirá el modelo de las ecografías, radiografías, etc.

Para comprobar la eficacia de los métodos propuestos se han realizado experiencias, comparando a continuación los resultados entre sí, y con otros algoritmos de especial relevancia, ampliamente aceptados en la literatura. Con ese objetivo se han construído máscaras de composición de tres clases de imágenes: textos mecanografiados, radiografías y ecografías —estas dos últimas a partir de imágenes graciosamente cedidas por el

Hospital Clínico de Granada. A continuación se han implementado los algoritmos de filtrado, incluyendo la construcción de la máscara de estimación de ruido. Por último se han diseñado y realizado los experimentos, en los cuales se han variado los parámetros que definen el comportamiento de los filtros. Las mismas experiencias se llevaron entonces a cabo con el filtro de mediana, y otros procedimientos que surgieron como mejora de aquél, como son el filtro de mediana multietapa y el adaptativo de Sun&Venetsanopoulos [CITA].

Debo hacer hincapié en la enorme tarea de realizar, diseñar e implementar procedimientos y algoritmos en la computadora. Si bien puede suponerse que no se trata de un verdadero aporte científico, puedo asegurar que la labor llevada a cabo con el ordenador ha proporcionado, aparte de innumerables cefaleas, los mejores indicios que han originado una posterior y fructífera investigación.

La estructura del presente trabajo es la siguiente:

- El capítulo I contiene la introducción a la Tesis Doctoral y una breve descripción del marco en que ésta se encuadra. Se incluyen los conceptos y notaciones básicas que serán empleados a lo largo de la memoria. Aparece una descripción de las matrices de composición, o modelos. Se muestran además los últimos avances en el campo del procesamiento de imágenes; y los objetivos del trabajo.

- En el capítulo II se introducen los dos primeros algoritmos diseñados para mejorar el comportamiento de los filtros tipo mediana. Uno de ellos es muy simple; mientras el otro, algo más elaborado, se basa en un sistema de ventana que se adapta a la presencia de bordes en la imagen, detectados mediante un novedoso método. Se presenta un estudio profundo del primero, incluyendo un arduo, pero abordable, análisis estadístico. Del segundo se realiza una visión más determinista de sus propiedades, pues la cuestión estadística se hace impracticable.

- En el capítulo III se propone el empleo de la máscara de estimación de ruido, y se explica el proceso necesario para obtenerla. Se realiza un estudio paramétrico de la influencia de diversos factores que intervienen en su construcción. Se introduce además una medida de calidad de más-

cara, y numerosos experimentos destinados a mostrar la capacidad operativa de las máscaras, tanto en cuanto a su función de localización de pulsos, como en la de corrección del nivel de gris de una región en la imagen, alterada por el ruido.

- En el capítulo IV se explica el diseño y funcionamiento de los filtros no lineales propuestos. Se realizan numerosas experiencias encaminadas a demostrar la eficiencia de los métodos. Se muestran asímismo las propiedades más relevantes de cada uno de los algoritmos.

- En el capítulo V aparecen las conclusiones generales de la Tesis Doctoral, enumerando las ventajas e inconvenientes de las metodologías propuestas y analizando los resultados más importantes. Además, se incluyen una visión de los problemas que aún permanecen abiertos, así como posibles vías de continuación del trabajo de investigación.

Las referencias se darán entre corchetes, y constarán de varias letras en mayúsculas —las iniciales del nombre del autor o autores del libro o artículo citado—, y de dos dígitos —correspondientes al año de publicación—. En caso de hacer referencia a dos trabajos del mismo autor y año, se diferenciarán mediente el uso de una letra minñuscula añadida tras el año de publicación (a, b, ...).

Este trabajo ha sido financiado en parte por los proyectos de investigación TIC92-603, TIC93-0535, TIC94-0506 y TIC95-0607, todos ellos concedidos por la Comisión Interministerial de Ciencia y Tecnología (CICYT).

ÍNDICE GENERAL

TABULA GRATULATORIA.....I

PREÁMBULO.....III

ÍNDICE GENERAL.....VII

CAPÍTULO I.
MARCO TEÓRICO

INTRODUCCIÓN.....3

CONCEPTOS BÁSICOS..... 4

Imagen.....4

Información en imágenes.....8

Información en la visión humana.....8

Información en imágenes.....12

Información e incertidumbre de un experimento probabilístico... 13

La divergencia de Jensen- Shannon.....17

Vecindad e histograma..... 20

Resolución, interresolución y superresolución.....24

Ruido.....30

Formación de la imagen digital.....30

Sensores y almacenadores.....31

Ruido en la formación de la imagen.....35

Ruido térmico.....37

Ruido fotoelectrónico.....38

Ruido de grano de película.....39

Ruido en *display*.....43

Ruido de pérdida de transferencia.....44

Procesamiento de imágenes digitales.....45

Desarrollo histórico.....45

Áreas de procesamiento.....47

Preprocesamiento: Filtrado.....48

Formulación del problema.....49

Filtrado lineal..... .52

Dominio espacial..... 52

Dominio frecuencial.....56

Filtrado no lineal.....57

Filtro de mediana.....61

Filtro mediana separable.....69

Filtro mediana recursivo.....70

Filtro mediana ponderada..... 70

Filtro estadísticos de orden..... 74

Filtros mediana híbridos, max/min y multietapa..... 75

Filtros tipo R.....76

Filtros tipo L.....77

Filtros media recortada.....78

Filtros MTM y DW-MTM.....78

Filtros tipo M.....79

Filtros homomórficos y media no lineal..79

El problema de la medida de calidad del procesamiento.....83

Comparación entre imágenes.....83

NECESIDAD DE FILTROS ADAPTATIVOS.....92

Esquemas adaptativos.....94

Filtros adaptativos no lineales.....96

Filtros adaptativos de evolución dirigida.....101

Filtros de optimización.....104

Filtros adaptativos de media recortada.....105

Filtros racionales.....105
Filtros de media modificada, MMF.....106
Filtros DBOSF.....107
Filtros de ventana adaptativa.....108
FUNDAMENTO DE LOS FILTROS PROPUESTOS.....113
Interdependencia de histogramas a distinta resolución.....114
Matrices de composición.....114
Modelos lineales elementales.....116
Modelos empíricos.....118
El problema de las columnas indeterminadas.....119
Clases de imágenes..... 122
Representaciones gráficas.....124
Medidas de calidad de modelos.....125
Características de las matrices de composición.....128
CONCLUSIONES.....130

CAPÍTULO II.
PRIMERAS MEJORAS AL FILTRO DE MEDIANA

INTRODUCCIÓN 135
EL FILTRO DE MEDIANA RELAJADA 138
Descripción 138
Propiedades estadísticas del filtro MR 139
Preliminares 139
Cálculo de la función de distribución de
probabilidad de salida del filtro MR 143
Atenuación de ruido 149
Preservación de bordes 154
Preservación de líneas 160
Otras propiedades 163

Apilamiento o Stacking 163
Manchas o Blotches 165
Raíces 166
Características de las raíces 167
Respuesta a imágenes prueba 169
Respuesta a imágenes ruidosas 176

EL FILTRO DE MEDIANA DE VENTANA ADAPTATIVA....185
Descripición.....185
Mecanismo adaptativo de supresión de ruido 186
Segmentación basada en la divergencia de
Jensen-Shannon 187
Esquema adaptativo empleado.....189
Tipo de filtro empleado 190
Propiedades del filtro restdiver 193
Supresión de ruido y preservación de detalles..... 194
Aplicaciones a imágenes ruidosas 196

CONCLUSIONES.....211

CAPÍTULO III.
PRESELECCIÓN DE PÍXELES. LA MÁSCARA DE ESTIMACIÓN DE RUIDO.

INTRODUCCIÓN.....217

PRESELECCIÓN BASADA EN MATRICES DE COMPOSICIÓN.....222

CORRECCIÓN DE LA MÁSCARA.....235

El proceso iterativo.....235

El algoritmo del proceso iterativo.....241

Medidas de calidad de la máscara.....243

Ejemplos de obtención de la máscara de estimación de ruido.....244

El problema del acceso a la LUT. La máscara gris.....249

Influencia del umbral de selección sobre la máscara.....252

Influencia del umbral de selección en la corrección del nivel de gris de ventana de acceso a la LUT.....255

Influencia del umbral de selección en la localización del ruido.....272

Influencia de la calidad de la matriz de composición en la máscara de estimación de ruido.....279

CONCLUSIONES.....280

CAPÍTULO IV.
FILTROS BASADOS EN MATRICES DE COMPOSICIÓN.

INTRODUCCIÓN.....287

DESCRIPCIÓN DE LOS FILTROS.....290

Prefiltrado 290

Filtro Máscra-mediana.....292

Filtro Máscara-Modelo.....294

Filtro Máscara-Modelo Ponderado.....297

Filtro de Mediana Combinado.....300

Filtro Restaura-Modelo.....302

RESULTADOS Y DISCUSIÓN..... 306

Discusión..... 329

Comentarios y propiedades..... 333

Influencia de los parámetros..... 335

CONCLUSIÓN.....337

CONCLUSIONES GENERALES.....343

BIBLIOGRAFÍA.....353

CAPITULO I

MARCO TEÓRICO

I.1 INTRODUCCIÓN....3
I.2 CONCEPTOS BÁSICOS....4
I.2.1 Imagen....4
Información en imágenes....8
Información en la visión humana....8
Información en imágenes....12
Información e incertidumbre de un experimento probabilístico..13
La divergencia de Jensen- Shannon....17
Vecindad e histograma....20
Resolución, interresolución y superresolución....24
I.2.2 Ruido....30
Formación de la imagen digital....30
Sensores y almacenadores....31
Ruido en la formación de la imagen....35
Ruido térmico....37
Ruido fotoelectrónico....38
Ruido de grano de película....39
Ruido en *display*....43
Ruido de pérdida de transferencia....44
I.2.3 Procesamiento de imágenes digitales....45
Desarrollo histórico....45
Áreas de procesamiento....47
Preprocesamiento: Filtrado....48
Formulación del problema....49
Filtrado lineal....52
Dominio espacial....52
Dominio frecuencial....56
Filtrado no lineal....57
Filtro de mediana....61
Filtro mediana separable....69
Filtro mediana recursivo....70
Filtro mediana ponderada....70
Filtro estadísticos de orden....74
Filtros mediana híbridos, max/min y multietapa....75
Filtros tipo R....76
Filtros tipo L....77
Filtros media recortada....78
Filtros MTM y DW-MTM....78
Filtros tipo M....79
Filtros homomórficos y media no lineal....79
El problema de la medida de calidad del procesamiento....83
Comparación entre imágenes....83
I.3 NECESIDAD DE FILTROS ADAPTATIVOS....92
I.3.1 Esquemas adaptativos....94
I.3.2 Filtros adaptativos no lineales....96
Filtros adaptativos de evolución dirigida....101
Filtros de optimización....104
Filtros adaptativos de media recortada....105
Filtros racionales....105
Filtros de media modificada, MMF....106
Filtros DBOSF....107
Filtros de ventana adaptativa....108
I.4 FUNDAMENTO DE LOS FILTROS PROPUESTOS....113
I.4.1 Interdependencia de histogramas a distinta resolución....114
Matrices de composición....114
I.4.2 Modelos lineales elementales....116
I.4.3 Modelos empíricos....118
El problema de las columnas indeterminadas....119
Clases de imágenes....122
Representaciones gráficas....124
Medidas de calidad de modelos....125
Características de las matrices de composición....128
I.5 CONCLUSIONES....130

INTRODUCCIÓN

Este capítulo contiene los fundamentos teóricos y los resultados básicos necesarios para el desarrollo de los siguientes capítulos de este trabajo. Evidentemente, la mayor parte del material aquí presentado no es original, aunque a veces será adaptado levemente con objeto de conseguir una unificación de notación y contenido que provea a la presente tesis de cierta autosuficiencia.

Se trata, pues, de mostrar de manera homogénea tanto los conceptos básicos para comprender los resultados que se proporcionarán en lo sucesivo, como los avances más recientes que conciernen al ámbito del procesamiento de imágenes, y que por su trascendencia y relación con el tema abordado en la presente tesis, se ha considerado oportuno incluir.

El capítulo consta de tres secciones principales. La primera de ellas consiste en introducir las definiciones de las ideas fundamentales para una futura comprensión de la exposición. De entre ellas cabe destacar por su importancia la noción de imagen digital. La siguiente sección aborda el problema de la aparición de ruido en imágenes dgitales. Obsérvese que se ha comenzado tratando dos pilares básico sobre los que se fundamenta el procesamiento digital de imágenes. Es precisamente una introducción a los métodos implicados en ese marco lo que constituye la materia de la última sección en que se divide el presente capítulo.

CONCEPTOS BÁSICOS

IMAGEN

Entenderemos por *imagen* cualquier representación bidimensional que intente reconstruir un fenómeno natural sensible. Dado que la naturaleza presenta un comportamiento continuo, al menos al nivel clásico, una imagen debe ser modelizada mediante una función continua de variables reales que proporcione los valores que la caractericen. Si se trata de una imagen de intensidad, esto es, una imagen que representa una escena real mediante diferentes niveles de brillo, dichos valores serán el resultado de integrar diversas cantidades ópticas; mientras que si se trata, por ejemplo, de una imagen procedente de un mecanismo de teledetección, la función proporcionará valores de temperaturas, presiones, etc, asignándoles un color determinado previamente por convenio. Las variables de las que depende esta función imagen serán de naturaleza espacial por lo general, si bien en ciertos casos puede aparecer también dependencia temporal. En lo que respecta a nuestros objetivos, solamente emplearemos imágenes de intensidad, en las que la escena se representa mediante niveles de gris adecuados, y que no dependen del tiempo. Por lo tanto, la imagen captada por un determinado sensor de la escena natural será una función continua de dos variables espaciales.

Se desea procesar dicha imagen mediante métodos digitales, para aprovechar las enormes capacidades de cálculo que presentan los ordenadores. Desgraciadamente, los ordenadores no trabajan con funciones continuas, sino discretas. Por lo tanto, la imagen captada por el sensor habrá de ser representada mediante una estructura de datos discretos, como por ejemplo una matriz. Es decir, habremos de trabajar con un conjunto de puntos de muestra discretos obtenidos a partir de una función continua, que a su vez fue captada a partir de unos objetos reales. A este tipo de representación discreta de una escena natural se le denomina *imagen digital*, y será la que manejaremos en lo sucesivo. Obsérvese que tanto la representación continua como la discreta de la naturaleza serán igualmente válidas, cada una dentro de su ámbito, por lo que podrán ser empleadas indistintamente sin provocar confusión.

Se impone, por tanto, la conversión de la función continua proporcionada por el sensor, a la función discreta que precisa el ordenador. Ello se logra mediante la *digitalización* y la *cuantización*. El primer proceso implica que la función continua se muestrea para obtener una matriz de V filas y H columnas. El segundo consiste en asignar a cada muestra continua, que es un número real, un valor entero. Es decir, se divide el dominio continuo de la función imagen en G intervalos. El número de niveles de gris (o color) empleados en la cuantización ha de ser suficiente para la percepción humana de detalles finos en la imagen. En general, los G intervalos que se han planteado se toman iguales. Si se usan b bits para expresar los valores del brillo del píxel, entonces el número de niveles total es $G = 2^b$. Resulta evidente que, cuanto más fino sea el muestreo y la cuantización (esto es, cuanto mayores sean los valores enteros V, H y G), mejor aproximación a la función imagen continua se logra. En el presente trabajo se ha empleado, salvo excepciones advertidas, $G = 64$ (esto es, $b = 6$); y por regla general $V = H = 200$.

Un primer problema surge cuando se plantea la digitalización de la imagen. Hay que determinar el intervalo de muestreo, esto es, la distancia entre las sucesivas muestras sobre la imagen continua que proporcionarán los valores de los puntos de la imagen digital. En principio puede emplearse el teorema de muestreo de Shannon, desarrollado en el ámbito del procesamiento de señal, para este propósito. Dicho teorema viene a decir que el intervalo de muestreo debe ser menor o igual a la mitad del tamaño de los detalles significativos más pequeños que presente la imagen en cuestión. Sin embargo, aunque este criterio puede resultar muy lógico, no suele ser el más apropiado. Y es que existe un fenómeno de degradación visual aparente producido por el carácter discreto de la imagen (*aliasing*). Sólo si se emplean métodos destinados a paliar dicho problema, mediante interpolación entre niveles vecinos de brillo, por ejemplo, puede emplearse el criterio de Shannon sin más. Pero en caso contrario, conviene, por precaución, utilizar un intervalo de muestreo unas 10 veces menor de lo que indica el citado teorema.

La distribución de las muestras obtenidas a partir de la imagen continua sobre el plano de la imagen digital representa otra indeterminación que conviene resolver previamente. Dicha colocación suele llevarse a cabo sobre mallas rectangulares, si

bien en ciertos problemas se prefiere la malla hezagonal, puesto que presenta ciertas ventajas como un mayor grado de simetría circular, conectividad uniforme, mayor resolución angular y menor gasto de almacenamiento y cómputo [HER94]. Sin embargo este tipo de malla es difícil de encontrar, por lo que algunos autores han propuesto el uso de mallas pseudohexagonales, construídas a partir de mallas rectangulares [HER94].

Un punto de muestreo infinitamente pequeño en la malla corresponde a un elemento de imagen (un *píxel*) digital. Sin embargo, en la realidad el tamaño del píxel es finito, debido a que la función de muestreo no es una colección de impulsos ideales (deltas de Dirac), sino una colección de impulsos finitos. Esta unidad elemental de imagen es la que posee un determinado nivel de gris correspondiente de alguna manera con el brillo que presenta la escena real. Estos niveles de gris son de naturaleza discreta, esto es, representados por números enteros, y, a la vista de anteriores definiciones, pertenecerá al conjunto de niveles $K = \{0, 1, ..., G\}$. Si bien el píxel representa la mínima unidad, indivisible, para los fines del procesamiento de imágenes digitales, nosotros entenderemos que cada píxel está constituido por 2^b partículas ideales denominadas *dots*, las cuales sólo pueden ser blancas o negras (recordemos que b es el número de bits empleados para almacenar cada muestra). De este modo, el *nivel de gris* de un píxel es el número de *dots* blancos que contiene. Por lo tanto, un nivel de gris 0 equivale a un píxel negro, puesto que no contiene ningún elemento blanco, mientras que un nivel de gris 2^b (= 64 en nuestro caso) indica que el píxel es blanco. Es decir, si denotamos por $d_{ij}(I)$ al conjunto de *dots* que conforman el píxel (i, j), el *nivel de gris del píxel* es:

$$k = I_{ij} = \sum [d_{ij}(I)]_{rs}$$

En resumen, llamamos *imagen digital bidimensional*, o simplemente *imagen*, a una aplicación I de un conjunto producto $\mathcal{V}\times\mathcal{H}$ de índices en un conjunto K de niveles de gris:

$$I : \mathcal{V}\times\mathcal{H} \rightarrow K$$
$$(i, j) \rightarrow I_{ij} \in K$$

donde $\mathcal{V} = \{1, 2, ..., V\}$
$\mathcal{H} = \{1, 2, ..., H\}$
$K = \{1, 2, ..., G\}$
siendo V y H el número de filas y columnas, respectivamente;

esto es, la anchura y la altura, en píxeles, de la imagen. El conjunto K representa lo que llamamos *escala de resolución* de niveles de gris, dado que los píxeles son el elemento básico de la imagen, a la mayor resolución posible. Sus elementos se denominan *niveles de resolución.*

A lo largo del presente trabajo se considerarán diferentes tipos de imágenes, a saber:

- *Imagen original, O:* Es la imagen que representa más fielmente la escena real. No está alterada por ninguna clase de distorsión. Por tanto, es aquella imagen (desconocida *a priori*) a la que el proceso de filtrado trata de aproximarse.

- *Imagen de ruido, N:* Es una imagen aleatoria cuyos píxeles representan el ruido que se ha introducido a la imagen.

- *Imagen fuente, S:* Aquella proveniente de la original, pero contaminada con ruido de alguna clase.

- *Imagen resultado, R:* Es la que aparece como salida del filtro, y debe ser lo más parecida posible a la original.

En la literatura se denomina imagen original al objeto real, y fuente a la imagen digital original. Sin embargo, en la representación digital de imágenes se producen distorsiones sistemáticas que impiden en la práctica que se intente reproducir la imagen original. Por ejemplo, no se consigue igualar el grado de contraste y distribución espectral de los objetos reales, y el proceso de escaneado introduce una pérdida de resolución que no puede ser recuperada. No tenemos más remedio que aceptar estas aproximaciones, incluso cuando la calidad de la escena original y la imagen que de ella se obtiene es claramente diferente. Es por este motivo por lo que se ha decidido, en el marco de la presente tesis, definir la imagen original como la imagen *digital* sin ruido; mientras que la fuente es la que pasaria por el proceso de transmision que le introduce ruido.

INFORMACIÓN EN IMÁGENES

Si una imagen ha de ser analizada por un humano la información debe ser expresada usando variables que sean sencillas de percibir. Una buena elección serían los parámetros psicofísicos como el contraste, textura, color, intensidad, brillo, contornos, forma, etc. El ojo humano sólo distingue objetos en las imágenes si puede distinguirlos sin esfuerzo del fondo. Por eso conviene entender en primer lugar la información que recoge la visión humana para poder realizar un procesamiento de imágenes orientado al fin que se precise.

INFORMACIÓN EN LA VISIÓN HUMANA.

La capacidad de ver es una de las características más remarcables de los seres vivos. Ello les permite percibir y asimilar en un tiempo brevísimo una increíble cantidad de conocimientos acerca de su entorno a partir de radiación electromagnética en el espectro visible. Nuestro objetivo en este momento es conocer qué aspectos de la información visual humana son más importantes, y por tanto deben ser tenidos en cuenta a la hora de realizar un procesamiento de imágenes orientado a proporcionar resultados lo más parecidos posible a la realidad. En ningún caso se intenta conocer en detalle el sistema perceptual humano para emularlo empleando tecnología. Lo que se pretende es mostrar información que sea comprensible para el observador humano, por lo que interesa conocer cómo representar esas imágenes, y qué mejoras son más útiles para que nuestro sistema perceptual se sienta lo más satisfecho posible.

Incluso en el nivel más sencillo, podemos darnos cuenta de la complejidad del problema. Una imagen cualquiera muestra un objeto reconocible fácilmente por un observador. Sin embargo, la función imagen, esto es, la matriz de los valores de gris que aparecen en la imagen, tiene un aspecto que poco recuerda al objeto original. Es decir, lo que supone una fuente de información para el ordenador, lo que el ordenador es capaz de percibir, carece de información visual para el ojo humano. Si trazamos un mapa de regiones de igual intensidad dentro de la imagen, mediante un preprocesamiento, el ordenador adquiere información complementaria. El observador humano es capaz de vislumbrar algo más que en el caso anterior, pero aún está lejos de percibir el objeto con claridad. Sin embargo, mediante técnicas adecuadas de preprocesamiento podemos conseguir una imagen resultante que mantenga casi toda la información visual que tenía el objeto original.

Este ejemplo viene a significar que un sistema sensor-ordenador no puede percibir *directamente* entidades ni objetos reales tal como lo hace el sistema visual humano, sino que toma tan sólo como entrada una matriz de números espacialmente distribuída que representa las intensidades o niveles de gris correspondientes con el objeto original. El observador humano enfrentado a los datos de la imagen tomados por el ordenador, representados en forma numérica, no puede inferir nada acerca de la imagen original. Por tanto el mecanismo de visión humana no encuentra un paralelismo definido con el método de captación de información seguido por el ordenador, el cual es la base para realizar cualquier tarea dentro del ámbito del procesamiento de imágenes.

Sin embargo, sería muy interesante conocer el mecanismo de obtención de información visual humana, para así diseñar métodos de procesamiento óptimos, pues el resultado de tales artificios es la obtención de imágenes que habrán de ser reconocidas por observadores humanas como de *buena calidad*. El ojo mismo no parece muy diferente de las cámaras de video y fotográficas. Una pequeña abertura en la parte frontal permite la entrada de luz que penetra a través de una pequeña ventana transparente llamada córnea para ser focalizada a continuación por una lente flexible. En el punto focal la luz incide sobre una estructura especial, denominada retina, dónde suceden una serie de reacciones fotoquímicas que permiten reconocer los colores. Ciertas células dentro de la retina, los conos y bastones, se encargan de traducir la energía química en impulsos eléctricos para un posterior procesamiento neuronal. Esta información es llevada entonces a la parte posterior del cerebro, al llamado cortex occipital. El ojo posee características de resolución espacial ***no lineal***, puesto que la distribución de conos y bastones no es uniforme, sino de tal forma que la resolución es mejor en el centro de la retina. Si el ojo tuviese un comportamiento lineal la situación sería mucho más sencilla pues el defecto de algún estímulo podría compensarse con el aumento de otro (por ejemplo, el área de un objeto se compensa variando su contraste y brillo)..

El nivel más elevado dentro del procesamiento neuronal de la información parece tener lugar en una zona del cerebro llamada cortex asociativo. Las operaciones que tienen lugar en este área son gobernadas por las leyes de la percepción visual humana. Dichas leyes fueron descritas por la psicología expe-

rimental, y en particular por la denominada teoría de la Gestalt, que fue primero formulada como principio general de percepción por el psicólogo Charles von Ehrenfels en 1890, y posteriormente refinada por Koffka en 1935. Esta teoría pretende conocer cómo se efectúa la organización de los datos sensibles. Propone que ciertas características intrínsecas son inherentes a patrones procesados por la retina que permiten una asociación a estructuras de más alto nivel. En la fase de reconocimiento esto implica que el sistema perceptual humano no necesita contener un gran conjunto de «plantillas» con las que comparar lo que observa, sino que puede clasificar diferentes objetos y entidades en «catálogos» que pueden ser ordenados por características. Es decir, los componentes de la imagen se agrupan basándose en su proximidad, similitud de forma, continuidad y simplicidad. Existe un conjunto de principios de organización, denominados primitivas perceptuales en recientes estudios. Este es el aporte fundamental de la teoría de la Gestalt: toda percepción es *global*, no se captan las cosas aisladamente, fuera de contexto o atemporalmente, sino que las imágenes nos llegan como conjuntos unificados y son percibidas como tales. Somos sensibles a las relaciones entre los elementos de un conjunto, y con ello, elaboramos estructuras significativas de la misma manera que entendemos el conjunto de letras que componen las frases de este trabajo en lugar de percibir cada una por separado. No se debe confundir sensación con percepción, esto es, el proceso físico de la visión con la actividad mental comprensiva de organización sensorial. Así, percibir una imagen es una actividad dinámica basada en la organiuzación de las cualidades sensoriales del campo visual.

Por tanto, según estas teorías, la totalidad de la información visual viene definida por la imagen, y no por los detalles de la escena. Las regiones coherentes son identificadas como homogéneas sobre algún conjunto de atributos. Estos atributos gráficos no existen en términos absolutos, se perciben en relación a otros. La intensidad y el color proporcionan la información básica para construir esta descripción. Algunas regiones se detectan como tales basándose en los patrones de textura que presentan. En especial, el sistema visual humano es muy sensible a las estructuras regulares. Los contornos entre los objetos son discontinuidades de los atributos que definen cada región, y que las delimitan. Estos contornos y bordes tienen una importancia primordial, pues contienen información sobre la *forma*, que es fundamental para el sistema perceptual humano.

Hay otro factor muy importante en la visión humana, y que por tanto ha de tenerse muy presente a la hora de diseñar filtros de procesamiento de imágenes. Se trata de la capacidad del ojo de distinguir distorsiones. Puede ser que éstas sean tan poco definidas que no sean perceptibles por observadores humanos, por lo que no merece la pena intentar eliminarlas. Así se plantea la cuestión de saber en qué momento comienza a ser visible, subjetivamente, una distorsión. Es lo que se conoce en psicofísica como el problema de la determinación del *umbral de visibilidad*. Este parámetro puede determinarse tanto en el dominio frecuencial como en el espacial, y ninguna de las estrategias parece mejor que otra. Dado que las aplicaciones que se desarrollan en el presente trabajo se sitúan en el ámbito espacial, nos ceñiremos a él. En estas circunstancias, el umbral de visibilidad depende de la intensidad (nivel de gris) promedio del fondo local sobre el que se presenta el estímulo; y de las variaciones espaciales o temporales de forma o intensidad del estímulo. De este modo se ha comprobado que el umbral de visibilidad es mayor en regiones muy oscuras o muy claras que en regiones grises. En cualquier caso, la mayoría de las imágenes reales tienen un fondo no homogéneo, con una distribución espacial de niveles de gris compleja, presentando contornos más o menos contrastados entre diferentes regiones. Por tanto, cobra una gran importancia el conocer cómo varía el umbral de visibilidad ante estímulos producidos en la vecindad de zonas con grandes cambios de intensidad (nivel de gris). Diferentes experimentos muestran que se presenta una reducción en la visibilidad del estímulo cuando el fondo no es uniforme, sobre todo a ambos lados de un gran cambio de intensidad del fondo. Para contornos de bajo contraste el efecto es menor que cuando el borde es contrastado. En todo caso, puede asegurarse que estas variaciones del umbral en las cercanías de contornos, denominado *efecto de enmascaramiento* es muy localizado, y provoca que las distorsiones se detecten con más facilidad en zonas de nivel de gris uniforme que en la cercanía de un salto abrupto de intensidad (contornos o detalles espaciales bien definidos). Además, se ha comprobado que cuanto más largo es el contorno, mayor es la elevación del umbral de visibilidad del efecto de enmascaramiento.

De todo lo expuesto hasta el momento se pueden extraer algunas ideas que habrán de ser tenidas en cuenta a la hora de diseñar un algoritmo de procesamiento de imagen. En primer

lugar, parece necesario establecer un criterio para cuantificar la calidad de la imagen, a fin de poder asegurar que un método es mejor que otro. En segundo lugar, se deduce que el objetivo de tales procesamientos es doble. Por un lado se han de suprimir en lo posible las distorsiones que aparezcan, sobre todo en las regiones de nivel de gris uniforme, y con menos profundidad en las zonas de borde. Por otro lado, la información acerca de la forma, contenida en los contornos de la imagen, ha de ser preservada tanto como sea posible. Por lo tanto, estos serán los pilares fundamentales de los métodos de procesamiento que se pretenden diseñar en el presente trabajo.

INFORMACIÓN EN IMÁGENES.

Desde un punto de vista semántico, podemos interpretar una imagen de dos maneras radicalmente diferentes. Por una parte, existe la *aproximación de información estructural*, en la cual las imágenes se consideran como un fondo aleatorio carente de información, sobre el que aparecen una serie de *objetos* (detalles, etc) cuyos parámetros, tales como el perfil, el color o la textura, son los que aportan la información que proporciona la imagen. De otro lado, la *aproximación de información local* defiende que todas las componentes de la imagen son portadoras de información, y no sólo los objetos, si bien no en igual proporción. Por ello no se efectúa la separación entre objetos y fondo, sino que se recurre a la observación de la imagen por vecindades. Es en este último contexto dónde se puede plantear de forma natural una interpretación estadística de la imagen, que será de enorme utilidad a la hora de formalizar los métodos del procesamiento de imágenes digitales [ROS82]. Obsérvese que este tratamiento es difícil de encajar en un sistema de información estructural.

En efecto, bajo el prisma de la información local, las imágenes son el resultado de un proceso estocástico. Un tal proceso es una generalización del concepto de variable aleatoria. Podemos denotar por $f(x,y;\ \omega_i)$ a un cierto proceso estocástico; siendo ω_i un elemento del conjunto de sucesos $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$. En nuestro caso el suceso puede representar el escoger una determinada imagen de un conjunto de imágenes. Cada ω_i tiene una cierta distribución de probabilidad de suceder, en la cual la función $f(x,y;\ \omega_i)$ es la variable aleatoria, para coordenadas fijas de la imagen (x,y). Si el suceso es fijo, entonces la función $f(x,y;$

ω_i) es una función determinista denominada *realización* del proceso estocástico.

Un proceso estocástico queda enteramente descrito mediante una colección de funciones de distribución de diferentes órdenes. Así, la función de segundo orden se emplea para relacionar pares de sucesos. Si se conoce el mecanismo de formación de la imagen, es posible modelizar la densidad de probabilidad de primer orden, pero generalmente estas distribuciones suelen ser desconocidas en la práctica, al expresar relaciones complejas entre múltiples sucesos. Por lo tanto se suelen emplear otras características más simples para describir procesos estocásticos, como por ejemplo la media, la autocorrelación, la correlación cruzada, la autovarianza ...

Es por esta razón de manejabilidad matemática por lo que se suele preferir una interpretación no estructural, probabilística de la imagen. Sobre todo en análisis de imagen, puesto que permite determinar los parámetros estadísticos fundamentales de la imagen para clasificarla como un todo. Hay que reseñar que, desgraciadamente, no se dispone de una definición matemática ampliamente aceptada de las características más informativas de la imagen: los bordes y contornoes. En nuestro estudio, se considerará borde a la frontera entre dos regiones homogéneas en la imagen que posean diferentes intensidades de iluminación y/o textura.

En cualquier caso, en el presente estudio se preferirá la estrategia no estructural a la hora de considerar la información contenida en la imagen, tanto por su filosofía como por su manejabilidad matemática. Por lo tanto, conviene estudiar cómo cuantificar la información de un experimento probabilístico.

Información e incertidumbre de un experimento probabilístico.

En su Teoría de la Información de 1948 Claude E. Shannon propuso un nuevo modelo matemático para estudiar los problemas teóricos relacionados con la transmisión y codificación de información por medio de canales de comunicación. La principal innovación introducida consistía en considerar los componentes de los sistemas de comunicación como entidades probabilísticas. De esta manera, la transmisión de in-

formación no era más que una sucesión de experimentos probabilísticos, cada uno de los cuales proporcionaba una cierta cantidad de información igual a la incertidumbre eliminada al realizarlo.

El propósito de la Teoría de la Información es mostrar cómo se selecciona un mensaje en el extremo receptor de un sistema de comunicación en medio de las interferencias aleatorias y ruido a que está sometido el canal de transmisión. Para Shannon, la información podía estar constituida por una serie de palabras completamente carentes de significado: su única importancia es que podía ser cifrada, transmitida, seleccionada y descifrada. Shannon propuso una definición puramente matemática de información para cualquier distribución de probabilidad dentro del sistema. Se podría emplear esta definición para calcular la probabilidad de que se obtuviera información en medio de la confusión de la interferencia.

La entropía discreta de Shannon.

Cualquier mensaje puede representarse por una secuencia de dígitos binarios que supondremos de longitud finita. Si la longitud de una palabra es n entonces necesitamos n dígitos, es decir, la información necesaria para caracterizar esa palabra es n. El conjunto E_n de todas las palabras de longitud n contiene entonces $N = 2^n$ elementos, de modo que la cantidad de información necesaria para caracterizar un elemento es el logaritmo en base dos del número de elementos de E_n, esto es, $\log_2 N = n$. Sea ahora $E = \cup E_i$ con todos los E_i disjuntos; N_i el número de elementos de cada E_i; $p_i = N_i/N$, y $N = \Sigma N_i$. De esta manera, la información precisa para determinar una palabra será en primer lugar conocer el E_i a la que pertenece, y después calcular el número de elementos de ese conjunto. Si no podemos conocer a qué E_i pertenece la palabra, lo único que podemos saber es que se precisarán $\log_2 N$ dígitos para caracterizar la palabra. Pero si sí somos capaces, se necesitarán $\log_2 N_i$ bits para caracterizarla completamente. La cantidad promediada de información necesaria para determinar un elemento, conocido el E_i al que pertenece, será:

$$I \equiv \Sigma(N_i/N) \log_2 N_i = \Sigma p_i \log_2 N\, p_i = \Sigma\, p_i \log_2 p_i + \log_2 N$$

dónde $\log_2 N$ es la información necesaria si no conociésemos a

[1] Quisiera llamar la atención sobre la significación de la base del logaritmo empleado. En principio, y dentro del marco de la teoría de la información, la base es dos debido al carácter binario del empleo de los *bits*. Sin embargo, no hay mayor importancia en la elección de la base del logaritmo. De hecho, en la mayor parte de las aplicaciones (incluida la mecánica estadística) se suele utilizar el logaritmo neperiano.

cuál E_i pertenece un determinado elemento. En tal caso, la correspondiente pérdida de información ocasionada por desconocer E_i sería

$$H = \log_2 N - I = -\Sigma p_i \log_2 p_i$$

dónde se define $-p_i \log p_i = 0$ si $p_i = 0$.

Esta es la fórmula de Shannon[1]. Físicamente, E es un conjunto de N medidas, o un experimento probabilístico con N posibles resultados. Los E_i son los posibles sucesos o resultados de ese experimento; mientras que las p_i son las probabilidades de cada suceso. Obsérvese cómo el hecho de obtener nueva información acerca del resultado del experiemto impone determinadas restricciones a algunas de las p_i que harán decrecer de manera irremediable el valor de la entropía de Shannon, H, tal como muestra la siguiente expresión:

$$\Delta I = -\Delta H$$

Como puede verse, la fórmula de Shannon se parece a la de entropía en mecánica estadística, introducida por Boltzmann. Shannon no quiso en principio llamar entropía a su hallazgo, hasta que fue persuadido por el matemático John von Neumann[2]. Esta similitud entre *entropías* ha hecho que muchos científicos lleguen a la conclusión de que existe una relación directa entre ambos conceptos. Como la entropía de Boltzmann está íntimamente ligada con el número de maneras diferentes en que un sistema cerrado puede encontrarse en un macroestado particular, la entropía de Shannon de un sistema en un macroestado dado estaría relacionada con la cantidad de información necesaria para especificar uno de los microestados correspondientes al macroestado, considerando todos los microestados equiprobables. Así aparece de modo natural el *principio de máxima entropía*, que pretende obtener en cualquier caso la distribución de probabilidad más neutral posible incluyendo toda la información disponible, y que se ha mostrado muy potente para encontrar información útil en un mar de interferencias. Tal principio no es más que un proceso de optimización, donde se busca una distribución de probabilidad que cumpla ciertas restricciones conocidas, y a la que se le exige tener la máxima entropía. Observemos que se trata de una generalización del principio de razón insuficiente, precognizado por Laplace, que se aplica cuando en un espacio de sucesos no se pueden evaluar

[2] Se dice que le dijo: «¡Le dará una gran ventaja en los debates porque nadie realmente sabe lo que es la entropía!»

directamente las probablidades: el desconocimiento impide asignar pesos diferentes a los sucesos, por lo que se considera, a falta de otra información, que todos tienen la misma probabilidad de ocurrencia. Este caso de máxima incertidumbre *a priori* se corresponde con una distribución uniforme, cuya entropía es la máxima posible, según la definición de Shannon.

Por tanto, si la entropía es una medida de desorden, lo es de falta de información, de incertidumbre. De este modo la entropía determina la cantidad de información adicional necesaria para especificar un determinado suceso cuyo desconocimiento previo era total, por lo que se supuso equiprobable con el resto de sucesos posibles. Este punto de vista ha sido defendido a ultranza por Ed Jaynes, llevándolo a extremos subjetivistas [JAY89]. Según Jaynes, la entropía de Shannon es la medida de nuestra ignorancia acerca de procesos detallados por debajo de la escala de observación, de modo que los procesos físicos son irreversibles porque no podemos conocer todos los movimientos de todas las moléculas. Si pudiésemos hacerlo, en cada instante la situación estaría tan ordenada como en cualquier otro instante anterior o posterior, con lo que el desorden, la evolución hacia el estado de máxima entropía carecería entonces de sentido, en clara discordancia con la experiencia.

Esta breve digresión viene a colación para señalar que la entropía de una determinada distribución de probabilidad, como otras medidas de información y complejidad, dependen de la resolución, del nivel de observación, del detalle con que se estudia el sistema. Matemáticamente es totalmente correcto decir que la entropía de un sistema descrito a la máxima resolución (mínimo nivel de observación) es siempre constante, no aumenta en absoluto. Sin embargo, un sistema pluricomponente se describe en términos de unas pocas variables, y el orden de estas variables tiende a disiparse con el tiempo, es decir, la información acerca de la configuración interna del sistema se pierde en su evolución temporal.

Aunque se han propuesto otras medidas para la información o incertidumbre que genera un experimento probabilístico, la entropía de Shannon cumple ciertas propiedades que la hacen especialmente adecuada:

- La función H es continua, positiva y simétrica respecto de sus argumentos.

· Al añadir un suceso de probabilidad nula, la medida queda inalterada.

· El máximo valor de H es $\log_2 N$, correspondiente a $p_i = 1/N \ \forall i$, esto es, cuando todos los sucesos se consideran equiprobables.

· La entropía de un experimento compuesto se obtiene sumando la entropía del primer experimento con la entropía del segundo condicionado por el primero.

Estas propiedades se conocen como *axiomas de Khinchin* [KHI57]. Se puede añadir el hecho de que el valor mínimo de la entropía se obtiene para la distribución degenerada ($p_i = 0 \ \forall i \neq j$; $p_j = 1$). También que la máxima entropía crece con el número de sucesos posibles. Algunas de estas propiedades matemáticas coinciden con la interpretación física de H como medida de incertidumbre: la máxima información y mínima incertidumbre se producen cuando uno de los sucesos es seguro (distribución degenerada), mientras que una distribución uniforme apenas si lleva información, y sí la máxima incertidumbre acerca del experimento probabilístico. En algunos textos de Teoría de la Información se introducen axiomáticamente las propiedades, y se demuestra que la medida de Shannon es la única que, salvo constante multiplicativa, las verifica.

La divergencia de Jensen-Shannon.

A lo largo del presente estudio surgirá a veces la necesidad de establecer una medida de distancia entre dos distribuciones de probabilidad. En la literatura aparecen muchas definiciones de este tipo de medidas [KUL68][JOH79] con diversas aplicaciones, de entre las que cabe señalar aquí las relacionadas con el procesamiento de señal y reconocimiento de formas [KIT86][DAR93].

Entre las medidas propuestas hemos seleccionado por su especial interés en este contexto la *divergencia de Jensen-Shannon* [LIN91], así llamada por emplear la entropía de Shannon de una distribución de probabilidad y la desigualdad de Jensen, ambas usuales en el ámbito de la teoría de la infor-

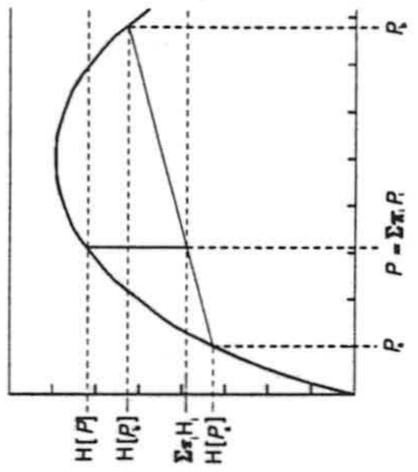


Figura 1: Representación gráfica de la divergencia de Jensen-Shannon en relación con las entropías para el caso binario.

mación. En este momento nos limitaremos a introducir dicha medida y a justificar su interés, para ahondar en sus usos cuando sea menester.

Medida de divergencia de Jensen-Shannon generalizada.

Si $\pi = \{\pi_1, \pi_2, \ldots, \pi_n \mid \pi_i > 0, \Sigma\pi_i = 1\}$ son los pesos relativos de las distribuciones de probabilidad $P_i = \{p_{ij}\}$, y la función $H(P_i) = \Sigma_j p_{ij}.\log p_{ij}$ es la entropía de Shannon, con i=1…n, la medida de divergencia de Jenssen-Shannon generalizada se define como:

$$JS(P_1, P_2, \cdots, P_n) = H\left(\sum_{i=1}^{n} \pi_i P_i\right) - \sum_{i=1}^{n} H(P_i)$$

Los dos términos del segundo término de la expresión son por un lado la entropía de P (la combinación lineal convexa de peso π_i de las distribuciones de probabilidad P_i), y por otro la misma combinación lineal convexa de las entropías respectivas.

Esta medida está especialmente indicada para ser empleada en problemas de decisión gracias a tres características fundamentales:

- Proporciona los límites superior e inferior para la probabilidad de error de Bayes.

- A cada distribución de probabilidad se le puede asignar distinto peso.

- Es generalizada para cualquier número finito de distribuciones.

Es esta última propiedad la que convierte a la medida introducida por Lin en una herramienta excepcional en nuestro trabajo. Por construírse en función de las entropías de las distribuciones, proporciona en sí misma una interpretación física relacionada con la información aportada. Y al ser aplicable a un número cualquiera de distribuciones, lo que en principio se trataba de una medida de *distancia* pasa a convertirse en una medida de *divergencia generalizada*. Esto será de vital importancia cuando estudiemos la cohesión de los modelos, pues la divergencia está relacionada con la inhomogeneidad de las distribuciones de probabilidad.

La relación entre las entropías que aparecen en la definición de la divergencia de Jensen-Shannon se ilustran, para el caso binario, en la figura 1. El eje horizontal está etiquetado con las distribuciones P_a, P_b y $P = \Sigma p_i P_i = \pi_a P_a + \pi_b P_b$, en lugar de con probabilidades. Las entropías aparecen en el eje vertical, a la izquierda, mientras que la divergencia de Jensen-Shannon aparece a la derecha. Para darle énfasis se ha marcado ésta con un segmento vertical en negrita para P. Debido a la concavidad de la entropía de Shannon, la entropía de Jensen-Shannon tiene muchas propiedades convenientes, que han sido expuestas por Lin en su trabajo ya citado. Podemos citar como las más importantes la no negatividad, la finitud, la simetría y el acotación. Además, el requisito de continuidad absoluta no tiene que ser satisfecho por las distribuciones de probabilidad involucradas, lo cual confiere una enorme versatilidad a esta medida. Obsérvese que JS = 0 si y sólo si los histogramas son iguales.

Acotación.

La búsqueda de cotas satisfactorias para la medida propuesta pretende un doble objetivo. En primer lugar, efectuar un estudio riguroso de la divergencia y los rangos de variación de sus valores. En segundo lugar, la acotación puede proporcionar una posible normalización de los resultados cuantitativos. Este último objetivo sería la base para efectuar comparaciones. En el seno del grupo investigación al que pertenezco se dedujo una cota superior alcanzable para la divergencia de Jensen-Shannon. En este sentido, esta medida mejora el comportamiento de otra distancia entre distribuciones de probabilidad, muy difundida, como es la entropía relativa. Dicha cota se deduce previa demostración de que la divergencia de Jensen-Shannon es una función $\cup$-convexa, que alcanza su máximo valor cuando las distribuciones de probabilidad son degeneradas, y cuyo dominio es un subconjunto de $C = [0,1]^{nk}$, siendo n el número de distribuciones de probabilidad que aparecen en la definición de la medida; y k el número de probabilidades que dichas distribuciones contienen. Entonces se deduce que si se tienen n distribuciones de probabilidad $P_1, P_2, \ldots, P_n$, con

$$P_i = (p_{i1}, p_{i2}, \ldots, p_{in});\ \Sigma p_{ij} = 1;\ p_{ij} \geq 0$$

y además $n = qk + r$, $0 \leq r < k$; $q, r \in \mathbb{N} \cup \{0\}$, entonces

$$JS(P_1, P_2, \ldots, P_n) \leq \log(k) - \{r(q+1)\log[(q+1)k/n] + (k-r)q\log(qk/n)\}/n$$

Como casos particulares de interés se tiene que:

- Si $n < k$, entonces $JS(P_1, P_2, \ldots, P_n) \leq \log(n)$
- Si $n = qk$ entonces $JS(P_1, P_2, \ldots, P_n) \leq \log(k)$.

Vecindad e histograma.

Si decidimos considerar una imagen digital como un suceso probabilístico, y por tanto la estudiamos basándonos en la información no estructural que nos proporciona, habrá de definirse previamente el concepto de *vecindad* que en esos casos se emplea. En efecto, bajo este punto de vista, serán los píxeles vecinos (según determinado criterio de cercanía) los que acarreen la información de la imagen. Si nos limitamos al dominio espacial, la vecindad se entenderá como un conjunto de píxeles que rodean geométricamente a uno dado. La vecindad queda unívocamente determinada por un solo parámetro: su área, o lo que es lo mismo, el número de píxeles que la constituyen. El hecho notable es que una vecindad refleja una cierta información apriorístico: en primer lugar los detalles y el fondo suelen estar agrupados; y en segundo lugar, los píxeles que pertenecen a un mismo detalle o fondo tienen valores de nivel de gris similares. Este conocimiento puede añadirse de forma natural al concepto de vecindad espacial, de modo que una vecindad puede detectar si se encuentra en un detalle o fondo, o si está ante un contorno. Además es fácil detectar píxeles cuyo nivel de gris sea claramente extraño en el entorno de vecindad, por lo que el ruido impulsivo puede ser marcado con facilidad.

En lo sucesivo emplearemos esta definición de la vecindad espacial. Surge entonces de manera natural la definición de *ventana deslizante* centrada en un píxel, como la región de la imagen I comprendida dentro de una ventana de tamaño(en píxeles) horizontal H_w y vertical V_w, centrada en el píxel ubicado en (i,j); esto es, la ventana deslizante no es más que la vecindad local que será empleada en el procesamiento. Se tomará $H_w = 2h_w + 1$, $V_w = 2v_w + 1$, y se numerará la ventana de forma que

el píxel central contenga las coordenadas locales (*0,0*). Entonces:

$$W_{ij}(I) : [-v_w, v_w] \times [-h_w, h_w] \rightarrow K$$
$$(r,s) \rightarrow [W_{ij}(I)]_{rs}$$

donde $[W_{ij}(I)]_{rs} = I_{i+r,\, j+s}$. Se trata, por tanto, de una subimagen de tamaño $H_w \times V_w$ de la imagen I comprendida entre $(i - v_w, j - h_w)$ hasta $(i + v_w, j + h_w)$. Denotaremos por w al tamaño de la ventana, esto es:

$$w = H_w \times V_w .$$

Existe un problema en la definición de vecindad espacial (y por tanto en el de ventana deslizante) cuando el píxel en cuestión cuya vecindad se analiza se halla cerca de las fronteras físicas de la imagen. En tal caso, parte de la ventana quedaría fuera de la imagen, presentando una evidente indeterminación. Esta dificultad se presenta de una manera irremediable,por lo que se precisa de una solución arbitraria para soslayarla de alguna forma. Suele ser habitual en estos casos emplear una extensión que permita determinar qué es lo que hay fuera de la imagen, para no perder información. En este sentido, se pueden escoger opciones naturales como estas:

- La imagen está inmersa en un plano blanco.
- La imagen está inmersa en un plano negro.
- La imagen prolonga su borde hacia el infinito.
- La imagen es toroidal: bordes opuestos están unidos.

En principio se escogería una opción para cada borde, de forma independiente, según las necesidades. Por ejemplo, para una imagen astronómica de cielo nocturno, la segunda elección parece correcta, mientras que para una imagen que presente objetos con periodicidad horizontal se preferirá emplear la tercera opción para los bordes horizontales, mientras que para los verticales interesa que la imagen sea cilíndrica. Sin embargo, se ha preferido emplear un criterio único de extensión de bordes, a fin de lograr una mayor homogeneidad en los resultados, que permitan una comparación directa. Supondremos que el exterior de la imagen es una extensión por repetición del píxel que ocupa justo el borde hacia el exterior del mismo.

Dado que la ventana deslizante será la que determine la vecindad del píxel en estudio, esto es, la que defina la resolución mínima de trabajo (siendo el píxel la resolución máxima), conviene definir el valor del *nivel de gris de ventana*. Para ello emplearemos la definición de nivel de gris de píxel, que será extendida de forma natural:

$$k' \equiv \sum [W_{rs}(I)] = \sum_{r=i-v_w}^{i+v_w} \sum_{s=j-h_w}^{j+h_w} I_{rs} = \sum_{r=-v_w}^{v_w} \sum_{s=-h_w}^{h_w} \left[W_{ij}(I) \right]_{rs}$$

Es fácil ver que $\Sigma[W_{ij}(I)] \in [0, G \cdot H_w \cdot V_w] \equiv [0, G']$. Este último intervalo (discreto) es lo que llamamos *escala de observación*, frente a K, que recordemos era la *escala de resolución*. Cada uno de los elementos de la escala de observación se denominan *nivel de observación*.

Este concepto de vecindad viene intrínsecamente unido a otro parámetro muy importante cuando se habla de información no estructural. Estamos hablando del *histograma local de la imagen*, que no es otra cosa que la distribución de los valores de la señal. Es decir, cada elemento del histograma proporciona la frecuencia de aparición del nivel de gris k en la imagen. Si el histograma se normaliza, representa entonces una distribución de probabilidad. Considerando una imagen digital de niveles de gris I de tamaño $V \times H$ como realización de un suceso estocástico, podremos asignar a un píxel o vecindad (según el nivel de resolución) s de la imagen una variable estocástica I_s cuyo espacio muestral es el espacio de niveles de gris (escala de grises) y cuya distribución de probabilidad es el histograma normalizado. Si k es un nivel de gris cualquiera de la escala de grises (de resolución o de observación) entonces la probabilidad $\Pr[I_s = k] \equiv P_k$, definido como:

$$P_k = \frac{\sum_{i=1}^{V} \sum_{j=1}^{H} \delta\left(I_{ij} - k\right)}{V \bullet H}$$

siendo δ la función delta de Dirac. En definitiva, denominaremos *histograma normalizado de una imagen*, o simplemente histograma, a la aplicación:

$$\begin{aligned} \mathcal{P}[I] : \mathrm{K} &\rightarrow [0,1] \\ k &\rightarrow \mathcal{P}[I](k) \equiv P_k \end{aligned}$$

Si se realiza una ordenación ascendente de los valores del histograma, se tiene el llamado *histograma ordenado* de la imagen. Esto es, una aplicación:

$$\begin{aligned} \mathcal{P}_{(\cdot)}[I] : K &\rightarrow [0,1] \\ \kappa &\rightarrow \mathcal{P}_{(\cdot)}[I](k) \equiv P_{(\kappa)} \end{aligned}$$

donde $P_{(\kappa)}$ está definida como en el histograma normalizado, pero verificando:

$$\kappa < \kappa' \Rightarrow P_{(\kappa)} \leq P_{(\kappa')}$$

siendo $(\cdot)$ una permutación de orden definida como una aplicación biyectiva:

$$\begin{aligned} (\cdot): K &\rightarrow K \\ \kappa &\rightarrow (\kappa) \end{aligned}$$

Nótese la diferencia de interpretación entre k y κ. El primer símbolo es un nivel de gris de píxel, mientras que el segundo es un ordinal. Obsérvese también que la desigualdad anterior no es estricta puesto que pueden aparecer en el histograma ordenado "mesetas" de niveles de gris con igual frecuencia de aparición

La información local, no estructurada, que va a ser empleada en nuestros procesamientos, se fundamenta en vecindades locales, esto es, en la información proporcionada por los píxeles contenidos en la ventana deslizante. Por esta razón conviene extender de manera natural el concepto de histograma aplicado a una vecindad en lugar de a una imagen completa. De este modo, podemos hablar de un *histograma de ventana* (y también de *histograma ordenado de ventana*) simplemente como:

$$\mathcal{P}[W_{ij}(I)]$$
$$\mathcal{P}_{(\cdot)}[W_{ij}(I)]$$

Usualmente, el histograma es la única información global que está disponible acerca de la imagen. Es importante reseñar que píxeles que pertenecen a un mismo detalle suelen aparecer en forma de cúmulos en el histograma, por lo que el contenido informativo de los histogramas es notable. No hay que olvidar, por contra, que in mismo histograma puede co-

rresponder a diferentes imágenes: por ejemplo, una traslación de un objeto sobre un fondo constante no altera el histograma. Sin embargo, su uso está muy extendido en el ámbito del procesamiento digital de imágenes, y el histograma local será la base de los algoritmos que se propondrán a lo largo del presente trabajo.

Podemos ahora particularizar lo expuesto al hablar de las medidas de información. Si suponemos que I_s es una fuente de información, la entropía de Shannon constituye una medida de la incertidumbre acerca del nivel de gris de un píxel o vecindad *s*. Como es la misma para todos los píxeles (vecindades) podemos emplear la entropía como definición de la información de la imagen. Por tanto, la entropía de la imagen es la entropía de su histograma. Como puede verse, dicha medida dependerá de la resolución empleada. Empleando la base 2 en el logaritmo, esta cantidad representará el número óptimo promedio de bits necesarios para codificar el nivel de gris de un píxel conociendo el histograma de la imagen.

Este tipo de medida de información en imágenes nos será de utilidad más adelante. Otros autores la han empleado profusamente. Recientemente [VOL95] se ha utilizado para clasificar imágenes, introduciendo una medida de redundancia entre píxeles vecinos a partir de la definición de información mutua, directamente deducida de la entropía de Shannon.

Resolución, interresolución y superresolución.

El concepto de *resolución* es difícil de concretar, aunque intuitivamente se relaciona con el nivel de detalle con que se observa un objeto o una imagen, o con la proximidad con que se mira. Fisiológicamente, la resolución del ojo humano es el inverso de un ángulo máximo de visión entre el observador y dos puntos próximos que están los más cerca posible para que sigan distinguiéndose como dos puntos separados. El ojo tiene su mejor resolución en 0.16 mm para objetos situados a unos 250 mm del ojo, e iluminados con 500 lux (una bombilla de 60W a 400 mm del objeto). Evidentemente, el nivel de resolución máxima en una imagen está firmemente ligado a la capacidad de resolución del ojo: no tiene sentido mostrar información con resolución menor que la que distingue el observador.

La noción de *interresolución* quiere expresar la capacidad para observar una imagen a diferentes niveles de resolución y relacionar dichas observaciones. Este concepto nos será de gran utilidad. En efecto, podemos señalar que en la visión humana se distinguen mejor las diferencias, irregularidades —ruido—, cuanto menor es la resolucion (si vemos desde mas lejos). Pero por otro lado, para eliminar esas degradaciones con más exactitud es preferible hacerlo empleando la mayor resolución posible, pues así estaremos seguros de cambiar tan sólo el valor erróneo de la imagen. Esta pareja de resoluciones diferentes proporcionan una manera inteligente de proceder a la limpieza de ruido. El establecer las relaciones entre ambas mediante la interresolución cierra definitivamente el problema. Sin embargo, obsérvese que se presenta una clara dicotomía, a mayor resolución, conseguimos mayor detalle, pero con un aumento de la complejidad. Es necesario un compromiso al fijar ambos niveles de resolución. Por ejemplo, el tamaño de la vecindad de la que obtendremos información para la detección de ruido vendrá determinada por la probabilidad de aparición del mismo —cuanto mayor sea la posibilidad de que exista ruido en la imagen, mayor habrá de ser la ventana de observción, esto es, menor la resolución, al igual que en la visión humana, donde se parecia mejor la irregularidad cuanto mayor es la región observada—. Pero no podemos hacer la resolución tan baja como queramos, pues entonces perdemos los detalles de la imagen: será el tamaño de las características de la misma que deseamos conservar la que limite el valor mínimo de la resolución.

Estos conceptos ya han sido empleados de forma implícita al introducir las definiciones de niveles de gris de píxel y ventana. Así, el nivel de observación teóricamente más bajo, que denotaremos por $m = 0$, y por tanto el mayor nivel de resolución, sería el correspondiente a los *dots*. La imagen a ese nivel de resolución es binaria (*dots* blancos y negros). Sin embargo, la imagen digital se compone de elementos discretos, los píxeles, que son en realidad el nivel más alto de resolución que en la práctica puede conseguirse. Este será para nosotros el nivel de observación más bajo, no podemos ver nada más pequeño que un píxel. Lo notaremos por $m = 1$. A partir de ahí podemos observar vecindades de píxeles cada vez más grandes, decrementando la resolución. Por la naturaleza de las imágenes digitales, cualquier vecindad observada a un nivel de resolución determinado, ($m > 1$) estará compuesto de un número

entero de píxeles ($m' = 1$). Esta restricción implica la aditividad de niveles de gris:

$$k' = \sum_{i=1}^{H_w \times V_w} k_i$$

Todas las escalas de grises están así definidas partiendo de la escala a nivel 0, teniendo para un nivel de observación cualquiera m un rango de variación $\{0, 1, 2, ..., G'\}$. Esta última restricción nos permite referirnos al nivel de observación mayor como ventana de observación. Esta ventana incluirá $H_w \cdot V_w$ píxeles y su escala de gris variará por tanto entre 0 y $H_w \cdot V_w \cdot G = G'$, donde G es el máximo nivel de gris del píxel. En definitiva, puesto que el trabajar con dos resoluciones diferentes es imprescindible para el desarrollo de los filtros que son objeto de la presente tesis, los términos con que se designarán éstas podrán ser, indistintamente, niveles de píxel y de ventana. Cada nivel de resolución tendrá asociado un histograma normalizado, tal como fue introducido en el anterior epígrafe.

Recordemos que la información contenida en el histograma era de vital importancia. Es fácil apreciar que el grado de información variará con la resolución. En efecto, al máximo nivel posible de resolución, sólo vemos píxeles individuales, los cuales constituyen por separado un valor de señal perfectamente definido. Pero nos es imposible determinar si dicha señal constituye o no una degradación. Para apreciar una irregularidad es necesario contemplar los valores de valor vecinoa para determinar si efectivamente existe una valor extraño. En otras palabra, la información es mayor a menores resoluciones en cuanto al establecimiento de regularidades, pero es mayor a menores resoluciones en cuanto a detalle.

Los dispositivos ópticos de observación poseen un límite físico (determinado por la difracción) en su resolución, de modo que existe una resolución máxima para la cual la lectura sigue siendo fiable. El hecho de que exista dicha cota puede ser un obstáculo en la práctica. Los métodos de *superresolución* están orientados a superar de manera analítica esa barrera. Usualmente se basan en métodos lineales —como el espectro de Fourier [PRA78], pirámides laplacianas, ondeletas (*wavelets*)—. Aunque son útiles en su ámbito, las características no lineales tales como los contornos y bordes en la señal no pueden ser captados como entidades singulares. La no linealidad se desga-

ja en componentes lineales a resoluciones más bajas, procesándose linealmente. Sin embargo, tales procedimientos provocan deformación y desplazamiento en los bordes, causando efectos indeseados. Y por otro lado, perceptualmente las imágenes a menor resolución son más difusas. Tales distorsiones en las características no lineales de la imagen pueden tener serias consecuencias en posteriores etapas del procesamiento.

En nuestro ámbito emplearemos el término *superresolución* para designar técnicas de estimación de imagen a un nivel de resolución m' cuando sólo podemos observarla al nivel $m < m'$. Es decir, se desea estimar los niveles de gris de los píxeles de una imagen particular empleando el conocimiento previo acerca de los niveles de gris de las ventanas que contiene. El método que se sigue es independiente de la naturaleza física del fenómeno que limita la resolución natural [MAR90].

Evidentemente, para realizar dicha estimación habrá que obtener información adicional de alguna parte. A partir de una observación única de la imagen al nivel m no obtenemos nada nuevo, pero si el dispositivo permite realizar varias observaciones de la imagen, siempre al nivel de observación; y que podemos efectuar ligeros desplazamientos—controlada o no—, entre cada dos observaciones, podremos obtener la información adicional que suple la carencia de datos proveniente de no observar la imagen al nivel de resolución. La forma más sencilla de tal proceder consiste en realizar observaciones de la imagen a nivel m desplazadas entre sí la distancia correspondiente a las dimensiones de un píxel. Como cada ventana contiene G/G' píxeles, lo anterior equivale a observar los niveles de gris de las G/G' regiones —ventanas— que lo contienen, una vez fijado un píxel.

Hay que señalar que el método indicado anteriormente posee un problema de ambigüedad. En efecto, el nivel de gris del píxel que es la intersección de todas las ventanas observadas no está unívocamente determinado por tales observaciones [MAR90]. Esta indeterminación en el valor del píxel central de un conjunto de observaciones al nivel de resolución proviene de la correlación espacial de los valores de los niveles de gris de los píxeles. Si fuésemos capaces de reunir todas las observaciones realizadas sobre la totalidad de la imagen, tampoco podríamos establecer la totalidad de las ligaduras para que la de-

terminación del nivel del píxel central fuese posible. Ello es debido en parte al elevado orden del sistema lineal de ligaduras; pero sobre todo a que la existencia de solución de tal sistema sólo se garantiza en el caso de ausencia completa de error en la medida, y aun en tal caso, la unicidad de la misma depende de algo tan remoto como la forma en que se toman las observaciones en los bordes de la imagen.

Sin embargo, es posible realizar estimaciones sobre los niveles de gris de los píxeles, empleando modelos lineales basados en sistemas de probabilidades condicionadas [MAR90][MAR95], los cuales serán empleados en la presente tesis para estimar histogramas, como se verá. Recientemente [WON95] se ha propuesto una solución basada en asignar un peso a los píxeles vecinos que contribuyen a la estimación del nivel de gris en cuestión, de acuerdo a su vecindad al mismo. Dichos pesos constituyen una distribución de probabilidad. Para elegir ésta adecuadamente se sigue el principio de máxima entropía, escogiendo aquella distribución que maximiza la entropía de Shannon bajo las ligaduras de los pesos. Entonces, aplicando consideraciones termodinámicas, se puede determinar unívocamente el nivel de gris del píxel.

RUIDO

El procesamiento de imágenes digitales surgió de la necesidad de corregir ciertos errores que se producían al adquirir dichas imágenes. Uno de los principales problemas que se presenta es la aparición de señales aleatorias que contaminan la señal original, provocando alteraciones de la misma. Es lo que conocemos por *ruido* [GUP88]. Por lo tanto, antes de estudiar detalladamente las características de estas distorsiones, conviene profundizar en el proceso durante el cual aparecen, que no es otro que el de formación de la imagen digital.

FORMACIÓN DE LA IMAGEN DIGITAL.

Una imagen es la reproducción de un objeto, siendo ese mecanismo de reproducción lo que denominamos *proceso de formación de la imagen*. Para ello se requiere básicamente:

1.- Fuente radiante: necesaria para iluminar aquello cuya imagen deseamos obtener. Las hay de diversas clases, puesto que existen diferentes tipos de radiación. En principio, esto no afectará a nuestro desarrollo.

2.- Objeto: a partir de él se obtiene la imagen.

3.- Sistemas de formación: son los encargados de registrar la energía radiante procedente del objeto tras su iluminación, para convertirla en imagen accesible. Los sistemas de formación pueden ser *biológicos* (como el ojo de algunos seres vivos), *fotoquímicos* (placas fotográficas, etc) o *fotoelectrónicos* (cámaras de televisión, scanners...). Dentro de los sistemas de formación empleados en el procesamiento digital de imágenes podemos distinguir tres fases fundamentales:

• *Sistema óptico:* es el encargado de traducir la información transportada por la distribución de energía radiante procedente del objeto (*f* en la figura 2) en una distribución de energía radiante intrínseca a la imagen del objeto (*b* en la misma figura).

• *Sensor:* es un dispositivo al que se le suministra como entrada la distribución de energía radiante de la imagen del objeto, y devuelve una corriente

eléctrica (*i* en la figura), relacionada con aquélla, mediante la utilización de un detector fotoelectrónico o similar.

• *Digitalizador:* transforma la corriente eléctrica suministrada (*i* en la figura 2) por el sensor en una señal digitalizada —imagen digital, compuesta de pixels—, susceptible de ser manipulada, visualizada y almacenada por computadoras (*I* en la figura 2).

• *Display:* se encarga de mostrar la señal digital, *I*, para que sea observada por el ojo humano. Suele ser una pantalla de ordenador.

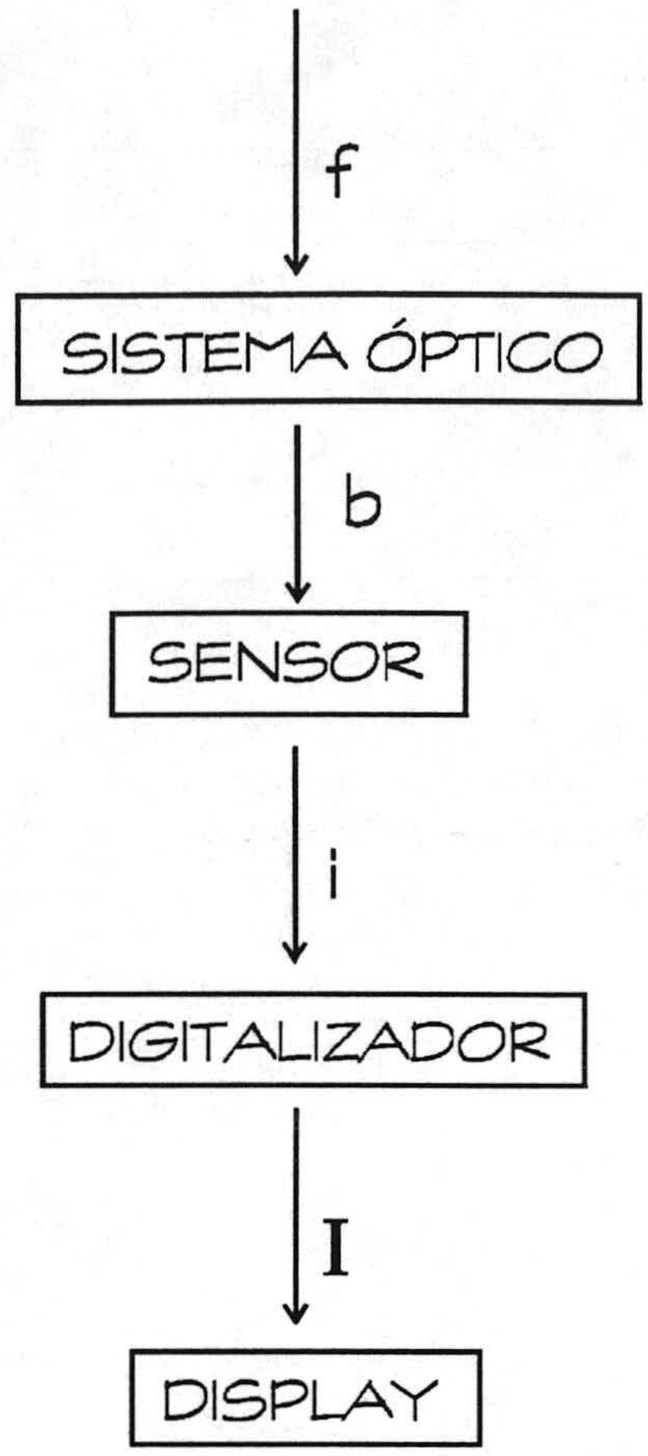


Figura 2: Organigrama del proceso de formación de la imagen.

Pese a la simplificación impuesta al intentar modelizar un proceso de formación de imagen, queda patente la enorme complejidad que un tal mecanismo encierra. Basta darse cuenta que cada elemento del mismo posee su propia función de transferencia, que a menudo no es lineal. En cualquier caso, repasemos algunas características de aquellos dispositivos fundamentales en la formación de la imagen.

Sensores y almacenadores.

Ya hemos advertido la existencia de dos familias fundamentales de sensores: los fotoquímicos y los fotoelectrónicos. Los primeros tienen la ventaja de detectar y almacenar al mismo tiempo, mientras que en contrapartida no son fácilmente discretizables. Mientras tanto los segundos son susceptibles de digitalizarse con poco esfuerzo, pero precisan de un mecanismo ulterior de almacenamiento. Pasemos a estudiar en cierto detalle cada uno de ellos.

Fotoquímicos: la fotografía.

La película fotográfica se compone básicamente de un soporte, recubierto con una emulsión de grosor comprendido entre 5 y 25 m. El soporte será opaco, de papel, para los positivos, y de acetato o cristal transparentes para los negativos. La emulsión se compone de una mezcla de cristales de haluros de plata (AgCl, AgI, AgBr) mezclados con gelatina. Los granos de sal de plata tienen forma de láminas delgadas triangulares o hexagonales de poco espesor. Sus dimensiones van desde las centésimas de micra en las emulsiones de grano fino y

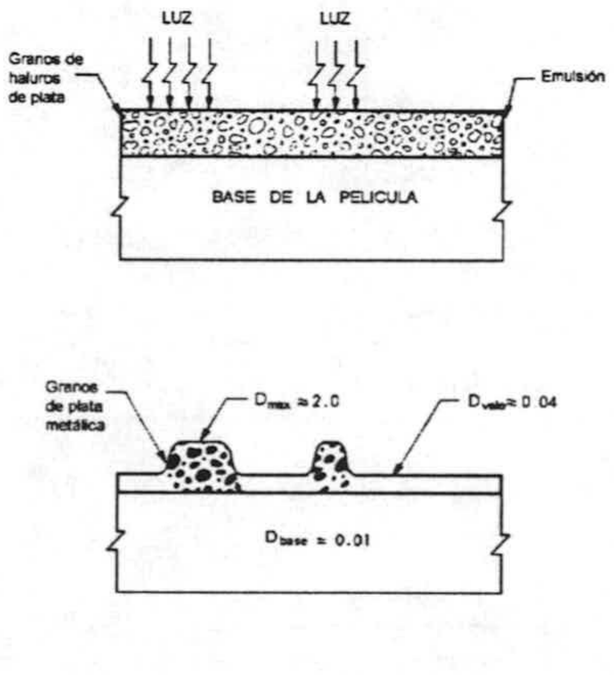


Figura 3: Croquis del comportamiento y fisonomía de la película durante la exposición y después del revelado.

gran contraste, hasta las decenas de micras en emulsiones de grano muy grueso, muy rápidas y de poco contraste. Durante la exposición, las distintas partes de la emulsión reciben diferentes cantidades (e intensidades) de radiación. Cuando un grano de sal de plata absorbe un fotón sufiecientemente energético, algunas de sus moléculas se transforman en plata, quedando el cristal expuesto. Durante el proceso químico de revelado, todos los cristales presentes en la emulsión precipitan a plata metálica, tanto los expuestos como los no expuestos, aunque siempre los primeros realizan tal precipitación de manera más rápida, pues ya contenían algunas moléculas de plata pura que actúan como centros de nucleación. De esta manera al cabo de un cierto tiempo la totalidad de los granos expuestos, y algunos (muy pocos si el tiempo es adecuado) no expuestos, son plata metálica. Con un lavado se eliminan aquellos cristales que no precipitaron, de manera que la película revelada muestra una cobertura de emulsión (ahora de plata y gelatina) de grosor variable. En el negativo, la densidad de plata es inversamente proporcional a la luz de la exposición, por lo que zonas muy expuestas mantienen su grosor inicial. Mientras tanto, en las partes no expuestas se han eliminado casi por completo los granos de sales de plata, dejando sólo una delgadísima capa llamada velo. Todo ello puede apreciarse en la figura 3.

El ennegrecimiento de la placa (negativo) se mide por su densidad óptica $D = \log_{10} 1/\tau$, siendo τ la transmitancia de la placa después del proceso de revelado y fijado, esto es:

$$D = \log_{10} \frac{1}{\tau} = \log_{10} \frac{J_i}{J_t}$$

dónde J_i es la intensidad de la luz incidente y J_t es la intensidad de la luz transmitida por el negativo. La cantidad de energía que se reciba por unidad de superficie será proporcional a la densidad de flujo de energía radiante incidente, E, y al tiempo de exposición t. El producto $H = E \cdot t$ recibe el nombre de exposición.

La respuestade la luz de una emulsión aparece plasmada en la *curva característica* dónde se representa D frente a $\log_{10} H$, y tiene el aspecto que muestra la figura 4. Se distinguen cinco tramos: la zona de *velo*; el *talón* o zona de subexposición; una región *lineal*, de exposición normal; el *hombro* o zona de sobreexposición; y una región dónde se invierte el comportamiento anterior, llamada zona de *solarización*. Obviamente esta curva varía mucho dependiendo de los materiales de que se com-

pone la placa, y del proceso de revelado empleado.

La densidad óptica del soporte de la placa, sumada a la correspondiente al velo, es lo que constituye el mínimo de la densidad óptica, D_{velo}, que suele tener el valor típico de 0.04. La densidad óptica máxima, D_m es la correspondiente al valor de saturación que se encuentraen el hombro, y que típicamente alcanza el valor de 2.0.

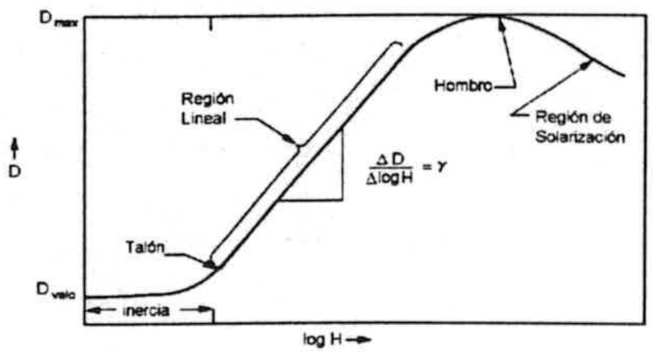


Figura 4: Curva característica de una emulsión, mostrando todas sus características.

La región lineal es en la que suele trabajar la película. La recta tiene la ecuación:

$$D(x,y) = (log_{10} H(x,y) - in)$$

dónde es la pendiente de la recta, y se le llama *contraste* o simplemente *gamma* de la emulsión, e indica la relación entre el contraste del negativo y el del objeto fotografiado. Si el negativo tiene el mismo contraste, o superior, al del objeto fotografiado. Para positivos, el es negativo y viceversa. Para un mismo material, el contraste depende del tipo de revelado. En las películas de uso común varía entre 0.8 y 2, llegando hasta 4 y 5 para radiografías.

La abscisa en el origen, *in*, se llama *inercia* de la película, cuya inversa se empleó en los comeienzos de la sensiometría como media de la rapidez de la película (hoy se emplea DIN y ASA). Por su parte, el espesor de la emulsión y el tamaño del grano determinan características muy importantes de la película, contradictorias entre sí:

- Un D_m elevado exige grosor elevado de emulsión.
- Alta definición precisa granos pequeños sobre emulsiones delgadas.
- Alta sensibilidad (velocidad), para trabajar en condiciones de poca luz, necesita pocos granos en emulsión delgada.

Así que cada cual busca un compromiso entre las tres características, de acuerdo con sus necesidades. En general, a menor sensibilidad, mayor contraste y resolución, mientras que la granularidad[3] disminuye. La máxima resolución de una imagen poco contrastada se obtiene cuando se expone y revela en-

[3] La granularidad es el efecto subjetivo que aparece sobre la fotografía, dando la sensación de que la imagen se compone de pequeños granos. Se debe a la distribución aleatoria (no uniforme) de los granos de la emulsión. El efecto es más abundante cuanto menor es la densidad óptica.

tre densidades de 0.8 a 1.2. Debajo de ese nivel, el talón reduce el contraste y la granularidad es más evidente.

Por último, señalar que todos los parámetros implicados en estos procesos dependen fuertemente del poroceso de revelado, por lo que resulta muy complicado obtener dos fotografías iguales, ni siquiera a partir de las mismas condiciones físicas.

Fotoelectrónicos.

Este tipo de sensores aparecen de muy diferentes formas: los hay que son tubos vidicon —estándar o de silicio—, CCD (*Charge Coupled Device*), CID (*Charge Injection Device*), *scanners* ... Expongamos someramente las cualidades de los más empleados:

• *Vidicon:* Un haz de electrones recorre la imagen completa. La corriente de salida, i, se relaciona con la energía radiante incidente en forma no lineal. La curva de respuesta a la intensidad incidente, b, es similar a la fotográfica. Presenta saturación cuando se alcanza el máximo de intensidad de corriente que soporta el dispositivo; y un mínimo, denominado aquí *corriente oscura*, que es la corriente residual que existe incluso en ausencia de lux incidente. También aparece una región lineal dónde se verifica que:

$$\log_{10}[i(x,y)] = \gamma \log_{10}[b(x,y)] + c_1$$
$$\Rightarrow \quad i(x,y) = c_2\,[b(x,y)]\,\gamma$$

El valor de $\gamma = 0.65$ para tubos vidicon estándar, mientras que para los de sílice no tiene un valor fijo, sino comprendido en el intervalo $]0.95,1[$, por lo que en esos casos la linealidad se pierde levemente.

• *CCD:* Emplea tecnología de estado sólido. Consiste en una matriz bidimensional de celdas sensibles a la luz. Al incidir ésta induce una carga en cada celda por efecto fotoeléctrico, la cual se desplaza ahacia la derecha celda a celda empleando un reloj bifásico, hasta alcanzar el registro de lectura. Las filas de celdas se recorren secuencialmente durante un barrido vertical. Estos dispositivos presentan la ventaja de almacenar y discretizar al mismo tiempo

la imagen, además de ser muy pequeños y cómodos. Sin embargo, pueden presentar inconvenientes graves, puesto que si una celda deja de funcionar adecuadamente, anula a las demás de su misma fila, pues impide el paso de la carga hacia el registro de lectura, produciéndose lineas completamente negras en la imagen de salida.

RUIDO EN LA FORMACIÓN DE LA IMAGEN.

El proceso de formación de la imagen digital está sujeto a errores de medida o a incertidumbres provenientes de alguno de los elementos de la cadena,—como por ejemplo el sensor—, a irregularidades en el grano de la película fotográfica, a fluctuaciones en la intensidad de la luz, etc. Por conveniencia semántica, reuniremos todos estos efectos, caracterizados por la introducción de una señal contaminante en la imagen, bajo el nombre de *ruido*.

Con frecuencia, las características intrínsecas de esta señal contaminante nos son desconocidas. Se habla entonces de *ruido aleatorio*, donde el calificativo no deja de ser un eufemismo para ocultar nuestra carencia de información acerca de un proceso físico poco entendido y/o demasiado complejo como para poder ser analizado en detalle.

De hecho, el concepto de ruido aleatorio es muy general. Las imágenes se forman a partir de energía radiante, la cual está discretizada según nos dice la mecánica cuántica. Por lo tanto existe una componente aleatoria intrínseca en toda la imagen, pues el número de fotones que impresionan un área determinada de un cierto detector fluctúa a cada instante. Esta aleatoriedad se une a la que aparece en los procesos de formación y obtención de la imagen, para conformar, sin pérdida de generalidad, el concepto de ruido aleatorio. Nótese que es posible conocer con cierto detalle el origen del ruido, pero no formalizarlo matemáticamente, como es nuestro deseo, dada la naturaleza caótica del hecho físico. Sin embargo, una observación del comportamiento del ruido a través del tiempo puede llevarnos a obtener suficiente información como para caracterizarlo de alguna manera estadística, aunque siempre sin predecirlo en detalle.

Clasificación del ruido aleatorio.

El ruido aleatorio puede clasificarse atendiendo a muy diversos criterios. Quizás el más simple sea el basado en la relación entre la entrada y la salida de la señal. De este modo, tenemos los siguientes tipos de ruido:

- *Ruido aditivo:* Se verifica que $S_{ij} = O_{ij} + N_{ij.}$

- *Ruido multiplicativo:* Entonces $S_{ij} = O_{ij} N_{ij.}$

- *Ruido impulsivo:* Es una transformación de $O \rightarrow S$ consistente en la elección aleatoria de algunos píxeles, y sustituir sus niveles de gris por otros seleccionados aleatoriamente de un conjunto de niveles, generalmente $\{0, G\}$ (ruido s*al y pimienta*), o bien el conjunto K (ruido *gris uniformemente distribuído*). Aparece, pues, como una serie de pulsos sobre la imagen. El caso *sal y pimienta* es el más evidente, pues los pulsos son blancos y negros.

Definición de la imagen máscara

En este contexto cabe definir la *imagen máscara*, M, una imagen binaria que indica qué píxeles en la imagen están afectados de ruido impulsivo. Así, $M_{ij} = 0$ indica que O_{ij} habrá de sustituírse por N_{ij}; mientras que $M_{ij} = 1$ significa que O_{ij} permanecerá intacto. En función de la imagen máscara puede definirse el ruido impulsivo como:

$$S_{ij} = O_{ij} M_{ij} + N_{ij} (1 - M_{ij})$$

Todas las operaciones entre imágenes se realizan píxel a píxel entre imágenes de igual tamaño. No se trata de álgebra de matrices.

dónde hay que tener en cuenta que todas las operaciones entre imágenes se llevan a cabo píxel a píxel, y que todas las imágenes involucradas en cualquier cálculo tienen el mismo tamaño. Si el ruido es *sal y pimienta*, $N_{ij} \in \{0, G\}$; mientras que si es unifoememente distribuído, entonces $N_{ij} \in K$. Obsérvese que en este tipo de ruido, N es aleatoria y no correlacionada con O.

Debemos hacer notar que el ruido aleatorio suele venir caracterizado por ciertos parámetros estadísticos. Así, por ejemplo, el ruido impulsivo vendrá determinado por la probabilidad de aparición de los pulsos. En el ruido aditivo hay que tener en cuenta que el ruido N_{ij} se presenta de acuerdo con una determi-

nada distribución de probabilidad $\Phi(i,j)$ que puede ser uniforme, gaussiana, laplaciana, ... Si $\Phi(i,j)$ = constante, estamos ante el *ruido aditivo uniforme*, equivalente a sumar un nivel de gris k dentro de la escala de resolución K a un determinado píxel, donde todos los k tienen la misma probabilidad de aparición. Si $\Phi(i,j)$ es una distribución normal, entonces se trata del *ruido aditivo gaussiano*. Suele ser el tipo de ruido más usual en los procesos de formación de la imagen. A veces no depende del valor de la señal (homogéneo), cosa que no ocurre nunca con el ruido multiplicativo, que siempre es dependiente de la señal. Es frecuente hablar de ruido aditivo *blanco*, cuando el espectro de potencias del ruido es constante (su intensidad no decrece cuando aumentamos la frecuencia), dado que es el más simple de manejar, y aproxima de una forma bastante aproximada al ruido que aparece realmente en la imagen.

El ruido aditivo suele aparecer en la transmisión por algún canal de la señal (por ejemplo, en una vidicom). Cuando el ruido depende de la magnitud de la señal, se obtiene un ruido multiplicativo, como el que aparece en la televisión, por el uso de líneas discretas; o el que tiene lugar en la película fotográfica debido al tamaño finito de los granos de plata. El ruido impulsivo tiene origen humano, si bien aparece también como ruido atmosférico en señales por satélite, o en determinados canales de transmisión. Es por ello por lo que conviene establecer aquí una clasificación del ruido que se produce en la formación de la imagen basándose en la fuente que lo origina. De esta manera tendremos los que a continuación se detallan.

Ruido Térmico

Los amplificadores electrónicos que modifican la señal introducen ruido térmico producido por el movimiento térmico aleatorio de los electrones en los elementos resistivos del circuito de los fotomultiplicadores. Es el ruido más simple. Es independiente de la señal, por lo que suele venir modelizado matemáticamente por un ruido aditivo blanco gaussiano, con media nula. De esta manera, el ruido queda completamente determinado por su varianza (espectro de potencias plano). Si bien es posible que aparezca alguna componente del ruido con intensidad proporcional al inverso de la frecuencia, suele ser poco importante, por lo que se desprecia a la hora de realizar el modelo del ruido que culmine con su descripción en términos matemáticos.

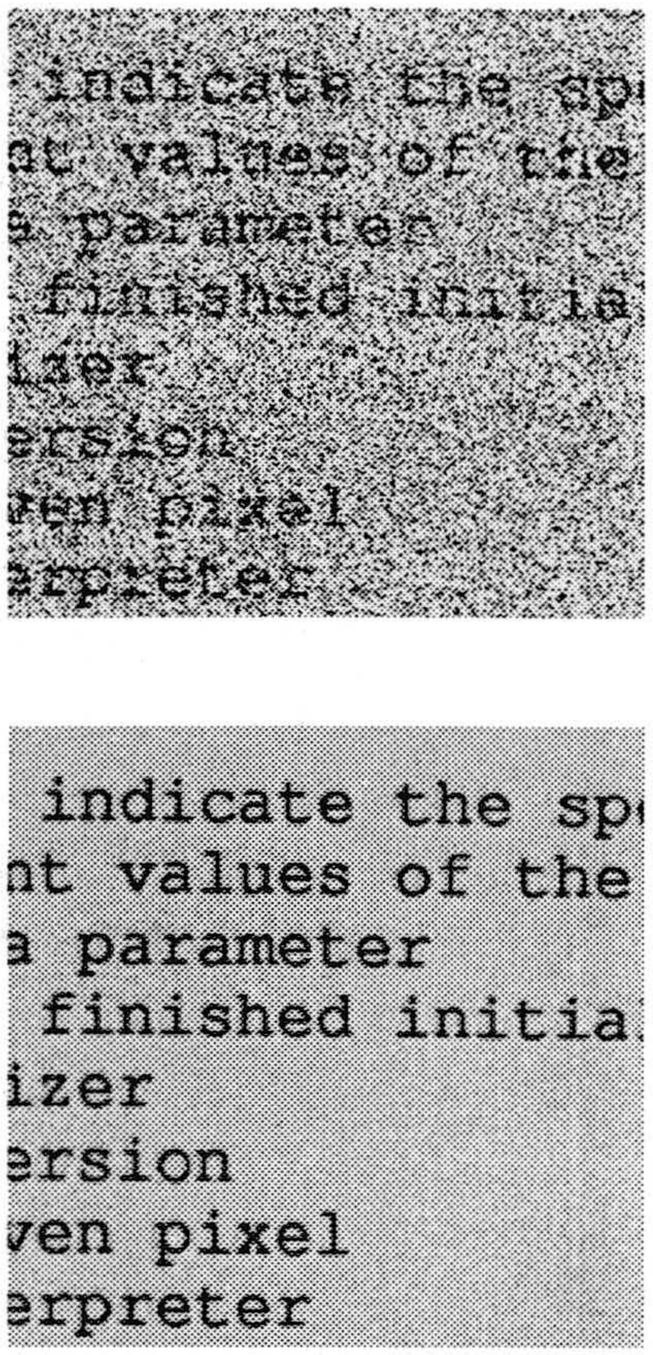


Figura 5: En la parte superior se muestra una imagen de texto afectada de ruido térmico (aditivo gaussiano de media nula y varianza 5). En el centro aparece la imagen de texto original. Abajo se ilustra la diferencia ecualizada entre las dos imágenes anteriores.

Ruido fotoelectrónico

Un fotodetector ideal actúa como un conversor de intensidad óptica a corriente, esto es, convierte fotones incidentes en electrones. Es en ese proceso dónde se genera el ruido fotoelectrónico. Si suponemos que el fotodetector es un generador de corriente proporcional a la intensidad (y número) de fotones incidentes, nos damos cuenta que el proceso de conversión fotoelectrónico es aleatorio, dada la naturaleza estadística de la luz. De ahí la presencia de ruido en esa transformación.

Debemos tener en cuenta que la intensidad de salida de un fotodetector no es la producida directamente por el efecto fotoeléctrico, sino que sufre ciertas modificaciones, para mantenerla estable, mediante un filtro eléctrico y una resistencia en paralelo (en el caso más simplificado). Si la cantidad de fotones es mucha, la corriente inducida por los electrones es casi constante (continua), y si el filtro es lineal, la densidad de probabilidad de la corriente de salida del detector es gaussiana. Sin embargo, la corriente del detector no es realmente constante, ni siquiera si la intensidad luminosa es constante, debido a la incertidumbre introducida por la naturaleza cuántica de la luz. A bajos niveles de intensidad de luz, el número de electrones producidos por el detector seguirá una distribución de Poisson. Habrá que tener en cuenta que si el número de fotones incidentes es bajo, entonces los electrones generados serán pocos, de manera que la corriente total se verá afectada por la corriente de fondo espúrea presente en todos los detectores. Así, si en un periodo τ se observa el comportamiento del dispositivo, la distribución de probabilidad de los pulsos de corriente del detector están dados por:

Se trabajará en lo sucesivo con una escala de resolución de 64 colores (6 bits).

$$\Pr\left[i_f = \frac{nq}{\tau}\right] = \frac{\left(n_d + n_f\right)^n \exp\left\{-\left(n_d + n_f\right)\right\}}{n!}$$

donde
i_f = intensidad de salida del proceso fotovoltaico.
n = número de electrones ($\in \mathbb{Z}^+$)
n_d = media del número de electrones emitidos en el proceso fotovoltaico.
n_f = media del número de electrones existentes en la corriente de fondo.

q = carga del electrón.

La incertidumbre en el conocimiento de la corriente real del detector provoca ruido impulsivo en detectores fotoemisivos; mientras que en dispositivos fotovoltaicos o fotoconductivos da lugar a ruido de excitación-recombinación. Estos tipos de ruido son superiores al ruido térmico y fotoelectrónico cuando existe amplificación interna del número de electrones (fotomultiplicadores) y el nivel de luz es extremadamente bajo. Pero en caso contrario, el ruido térmico es dominante, y la corriente del fotodetector es aproximadamente constante. En cualquier caso, el número promedio de electrones emitidos tiende al infinito a medida que aumenta la cantidad de luz. Entonces la distribución de Poisson tiende a la de Gauss, con varianza igual a la media (y se difumina el carácter discreto del fenómeno):

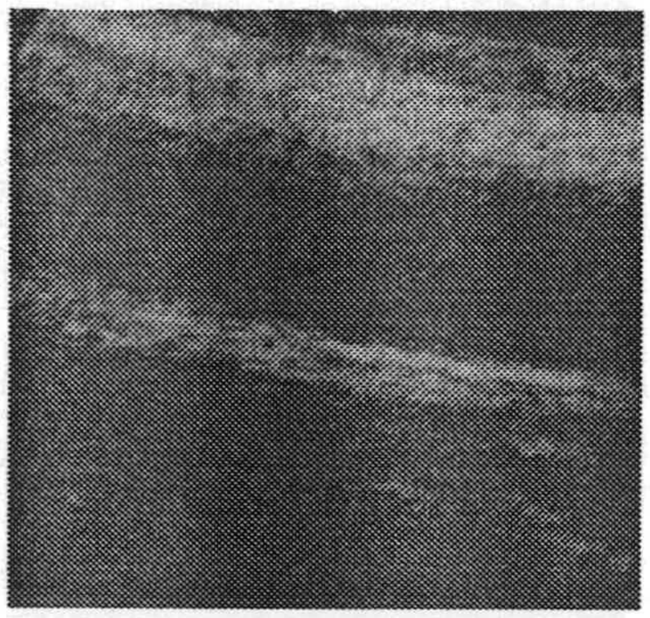
$$\Pr\left[i_f\right] = \frac{1}{\sqrt{\frac{2\pi q n_d}{\tau}}} \exp\left\{ -\frac{\left(i_f - \frac{q n_d}{\tau}\right)^2}{2.\frac{q n_d}{\tau}} \right\}$$

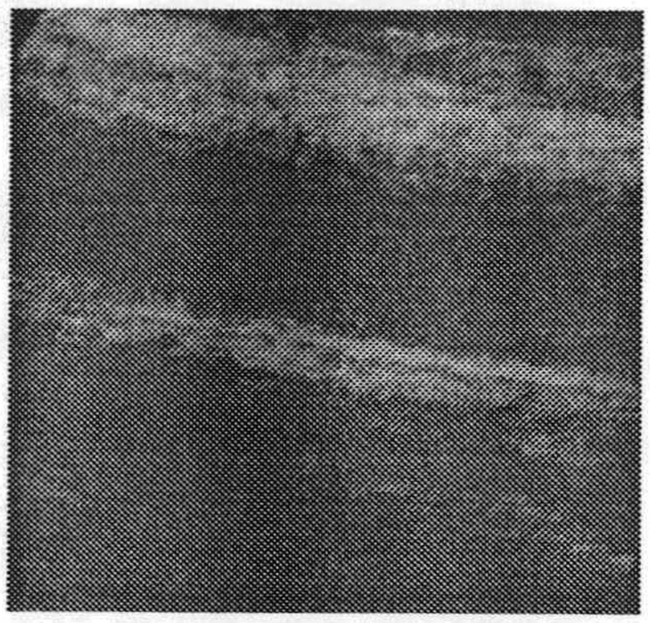
En ambos casos el ruido es dependiente de la señal. A este tipo de ruido habría que sumarle el efecto del ruido térmico ya explicado, que es independiente de la señal.

Ruido de grano de película

Se produce en el proceso de plasmar el objeto real en un soporte fotográfico. El proceso de obtención de la placa, es decir, la aparición de los granos de plata definitivos, lleva implícito cierta aleatoriedad, tanto en la exposición como en el revelado. Baste decir que el número de fotones necesarios para sensibilizar un grano de plata es variable; y que el tamaño de los propios granos es impredecible. Además, el proceso de revelado nunca es exactamente igual, puesto que el proceso de precipitación de la plata no es completamente determinista: en efecto, los granos de sal de plata que experimentan exposición similar no se convierten necesariamente en granos de plata de igual tamaño y forma. Es más, la distribución de los granos de plata es aleatoria, por lo que no esposible obtener dos fotografías idénticas. Podemos ver entonces cómo esta aleatoriedad inherente produce el ruido grano de película: conduce a una incertidumbre acerca de

Figura 6: Arriba aparece una imagen de radiografías afectada por ruido fotoelectrónico dependiente de la señal. En el centro se muestra la imagen original, sin ruido. Abajo se plasma la imagen de diferecias ecualizada entre las dos imágenes anteriores.





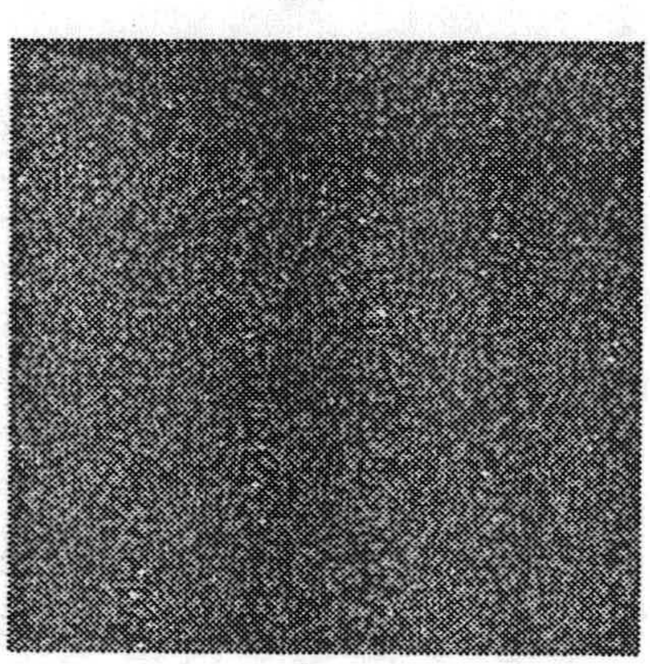


Figura 7: En la parte superior se aprecia una imagen original de ecografía. En la parte central se muestra la misma, afectada por ruido de grano de película, con tamaño de grano a = 100 micras. Abajo se muestra la imagen ecualizada de diferencias entre las dos imágenes anteriores.

la cantidad de luz que refleja (o transmite) la película. En la mayor parte de los casos, este tipo de ruido puede modelizarse por un proceso de Poisson, que en el límite continuo es gaussiano, en función de la densidad óptica D correspondiente a la *imagen fuente*:

$$\Pr[D] = \frac{1}{\sqrt{2\pi\sigma_D^2}} \exp\left\{-\frac{1}{2}\left(\frac{D-\mu_D}{\sigma_D}\right)^2\right\}$$

donde μ_D es la densidad óptica media, determinada mediante promedio espacial sobre una región A que cubra varios granos individuales; y σ_D es la desviación típica de los valores que alcanza la densidad óptica. En el caso de que se considere que el tamaño del grano es constante, podemos calcular el valor de esta desviación típica siguiendo la tesis de Nitting (1913) que supuso que la emulsión se componía de capas superpuestas de 1 grano de espesor. Entonces la densidad óptica podía expresarse como:

$$D = 0.43\frac{a}{A}n$$

donde a es la sección eficaz de un grano; A el área de examen de la película; n es el número total de granos dentro de A, y se ha supuesto que todas las variables son espaciales —esto es, dependientes del punto genérico (i,j)—. Entonces por linealidad se tiene que:

$$E(D) = 0.43\ \cancel{c}\ E(n)$$
$$\sigma_D^2 = (0.43\ \cancel{c})^2\ \sigma_n^2$$

Obsérvese que si se fija el valor de a y A, el valor de n sigue una distribución binomial, la cual, si el factor $\cancel{c}$ es pequeño, tiende a una distribución de Poisson, dónde la varianza coindide con la media. Por lo tanto se llega a:

$$\sigma_D^2 = (0.43\ \cancel{c})^{1/2}\,[E(D)]^{1/2} = 0.66\ \cancel{c}^{1/2}\,\mu_D^{1/2} = \alpha\,\mu_D^{1/2}$$

Si se escoge una hipótesis más razonable, consistente en suponer una distribución de tamaños de grano, entonces:

$$\sigma_D^2 = \alpha\,\mu_D^{\beta} \;,,\; \beta < 0.5$$

A partir de datos empíricos sensiométricos, con A pequeña, se

limita aún más el intervalo de variación del exponente, $\beta \in [0.3, 0.4]$, siendo el valor de 1/3, propuesto por Huang, el más realista [PRA78]. Sin embargo, Falconer propone el valor $\beta = 0.5$ [PRA78].

El ruido de grano de película no muestra experimentalmente correlación espacial estadística para distancias superiores al tamaño del grano, por lo que puede ser considerado como un proceso de ruido blanco bidimensional, si bien es claramente dependiente de la señal, pues la varianza es función de la media. La densidad D, descrita según la distribución gaussiana antedicha, puede expresarse como suma de la media y un proceso aleatorio normal (media nula y varianza unidad) de esta forma:

$$D = \mu_D + \sigma_D N = \mu_D + \alpha \mu^{\beta}_D N$$

dónde de nuevo todas las variables son espaciales. Vemos por tanto que, aunque el ruido es dependiente de la señal, es aditivo en el ámbito de la densidad óptica. En el dominio de la intensidad es, sin embargo, multiplicativo. En efecto, dado que la película es transparente, la densidad óptica se define como:

$$D = -log\ (\tau)$$

siendo τ la transmitancia. Entonces:

$$\tau = 10^{-D} = 10^{-\left(\mu_D + \alpha \mu^{\beta}_D N\right)} = 10^{-\mu_D} \cdot 10^{-\alpha \mu^{\beta}_D N} = \tau_0 \cdot 10^{-\alpha \mu^{\beta}_D N}$$

Por tanto:

$$\tau = \tau_0 N_1$$

donde τ_0 es la transmitancia promedio en ausencia de ruido de grano de película. La distribución N_1 es log-normal.

Puede simularse el ruido grano de película para realizar experimentos encaminados a su eliminación. Bajo la hipótesis de granos de igual tamaño, podemos suponer que cada píxel (i,j) tiene un determinado nivel de gris k, el cual provendría inicialmente de promediar densidades ópticas en la imagen original (fotográfica) sobre un área de tamaño A. Esa tarea habría sido realizada por el dispositivo digitalizador (*escanner*). De esta manera, el nivel de gris de cada píxel estará relacionado con la densidad local media μ_D. En efecto:

$$\mu_D \propto k/G$$

siendo G el máximo nivel de gris de la escala de resolución K. Bajo esta hipótesis, D sería la densidad óptica de cada grano de plata localizado en el interior de la región A, y que dió lugar al píxel (i,j). El tamaño de esa superficie viene determinada por el digitalizador. Si las imágenes se digitalizan empleando 300 píxeles por pulgada (que suele ser el valor por defecto de los dispositivos comerciales) entonces el área de fotografía que corresponde a un píxel es de 300^{-1} pulgadas cuadradas. Como este valor es muy pequeño, es razonable suponer que la densidad óptica es constante dentro de A. De este modo $D = \mu_D$, y por tanto $D = a\,(k/G) + b$.

Para determinar las constantes de la anterior expresión podemos hacer uso de valores conocidos de densidad óptica. Por ejemplo, sabemos que $D_{velo} = 0.04$ y que $D_{max} = 2.0$, típicamente. Entonces:

$$D_{velo} = 0.04 = a\,(G/G) + b$$
$$D_{max} = 2.0 = a\,(0/G) + b$$

Por tanto se tiene que:

$$a = D_{velo} - D_{max}$$
$$b = D_{max}$$

De este modo, podemos expresar el valor de la densidad óptica, y por tanto de μ_D, como:

$$\mu_D = D = (D_{velo} - D_{max})\cdot(k/G) + D_{max}$$

Recordemos que además $D = \mu_D + \sigma_D N = \mu_D + \alpha\,\mu^{\beta}_D N$, por lo que entonces podemos asegurar que:

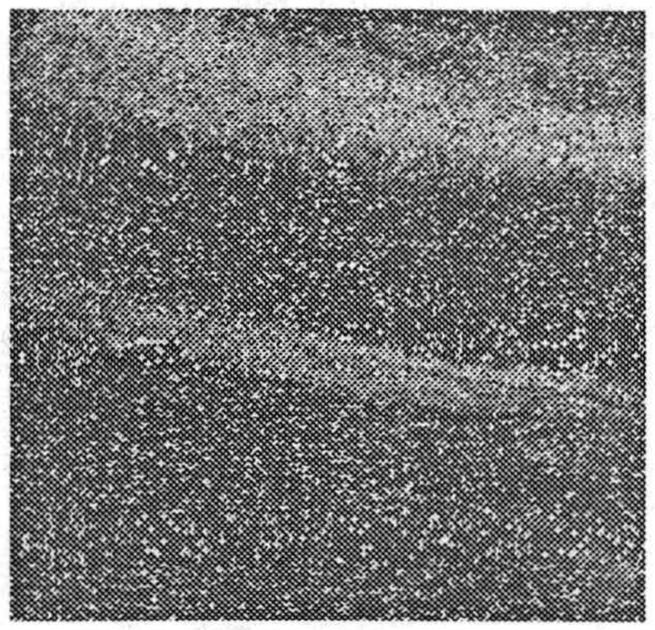
$$\mu_D = \mu_D + \alpha\,\mu^{\beta}_D N$$
$$\Rightarrow \mu_D = [(D_{velo} - D_{max})\cdot(k/G) + D_{max}] + \alpha[(D_{velo} - D_{max})\cdot(k/G)]^{\beta} N$$

De esta manera se calcula el valor del nivel de gris a partir de la densidad óptica media, siempre que se conozca la constante a, esto es, el tamaño típico de los granos de plata, en micras cuadradas. Procediendo como se ha descrito, será entonces posible realizar un algoritmo que produzca una simulación aceptable

del ruido de grano de película sobre imágenes digitales.

Ruido en el display

El número efectivo de niveles de gris en el *display* nunca es mayor que el existente en la imagen digital. En primer lugar, si el número de niveles de cuantización empleados al digitalizar la imagen es insuficiente, aparece el *ruido de cuantización*, que se presenta como aparición de falsos contornos en la imagen. Este tipo de ruido se elimina fácilmente mediante operaciones con el histograma. Existe otro tipo de ruido en el *display*. Si el ruido térmico del sistema de *display* ocupa más de un nivel de gris, el número efectivo de niveles se reduce. La raíz del error cuadrático medio (*rmse*) representa un límite inferior práctico para la resolución de la escala de grises (no se emplea la varianza, porque ésta permanece casi constante). Por ejemplo, si el *rmse* = 1% de la escala *K*, entonces la resolución fotométrica —esto es, la precisión con que el sistema reproduce la imagen—, tiene 100 niveles. Así, si trabajamos con 8 bits de resolución y *rmse* = 1% de la escala, tendremos 256 niveles de gris teóricos, pero sólo 100 efectivos. Sin embargo, si trabajamos con 6 bits, la escala no se ve alterada por este efecto. Esta es una de las razones, unida a la simplificación computacional, por la que se ha decidido emplear en los experimentos realizados en la presente memoria, una escala de resolución de 64 colores (6 bits).



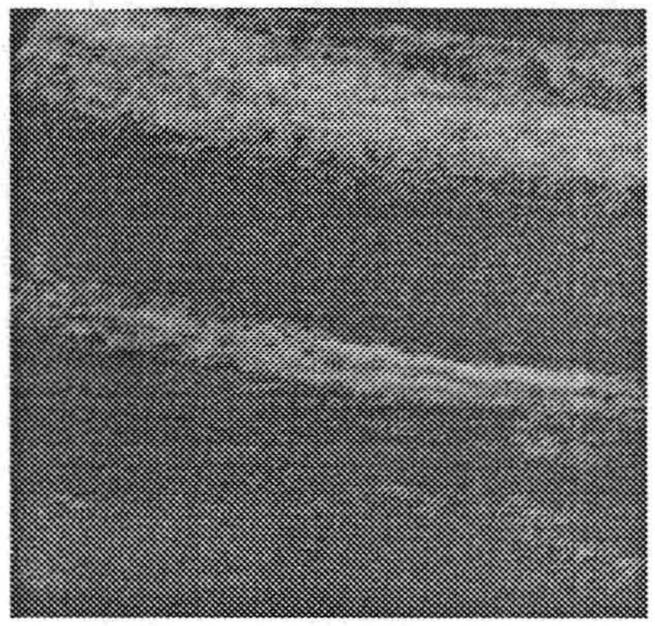


Figura 8: Ecografía afectada de un 20% de ruido impulsivo *sal y pimienta* (arriba); y la imagen de la ecografía original.

El ruido térmico en el *display* provoca tanto variaciones de intensidad como de posición. En el primer caso, se produce ruido *sal y pimienta*, visible sobre todo en regiones uniformes. Si el ruido es periódico y de intensidad razonablemente alta, puede producir molestos patrones, similares a raspas de pescado, sobre la imagen; y si está sincronizado con la señal de deflexión horizontal o vertical de barrido electrónico, puede producir un patrón de barras. El nivel óptimo de ruido —aleatorio y/o periódico) que se intenta conseguir en los *displays* comerciales para mantener buena calidad de imagen es inferior a un nivel de gris en amplitud.

Si el ruido térmico se produce en los circuitos de deflexión de electrones que determinan la mancha visual, resulta una falta de uniformidad en el espaciado de las manchas. El ruido de posición en el *display* no afecta gravemente la geometría de la imagen, a no ser que se combine con la interacción entre manchas (recordemos que la mancha de Airy determina la resolu-

ción espacial según la separación de dos de ellas que sean adyacentes). Si el ruido de posición es aleatorio, da lugar a ruido *sal y pimienta*; mientras que si no es aleatorio, como es común cuando existen fallos en el sistema de conversión analógico-digital, provocará la aparición de líneas paralelas equiespaciadas.

Figura 9: Fotografía afectada por un 20% de ruido impulsivo uniformemente distribuido, provocado por error en la conversión A/C en el digitalizador. A su lado, la imagen original.

Ruido de pérdida de transferencia.

Es uno de los ruido típicos que aparecen en las cámaras CCD, cuando la transferencia de carga entre celdas adyacentes no es completa. Ocurre sobre filas de celdas, por lo que tiene una fuerte correlación horizontal. Puede modelizarse matemáticamente, suponiendo la existencia de una región brillante que induce una carga q_0 en una celda del CCD. Ésta experimenta n transferencias antes de llegar al colector de salida del CCD. Entonces la carga final a la región brillante inicial es:

$$q_n = (1 - \varepsilon)^n q_0 \approx (1 - n\varepsilon) q_0$$

puesto que ε es del orden de 10^{-5} en la mayoría de los casos, por lo que el producto $n\varepsilon \approx 1\%$ para una cámara CCD de 512×512.

PROCESAMIENTO DE IMÁGENES DIGITALES

Existe un amplio consenso acerca de los campos que incluye el término ***procesamiento de imágenes*** entre las más notables eminencias en ese campo. Por él se entiende aquel conjunto de técnicas en las que, a partir de una imagen de entrada, se obtiene una imagen de salida (si bien en el caso de la codificación de imágenes, lo que se obtiene es una representación no pictórica). La imagen de salida presenta algún tipo de mejora frente a la de entrada, orientado a beneficiar al usuario en lo que se desee. Según sea esa mejora un realce de la imagen, una restauración de deformaciones, una compresión del número de bits necesarios para almacenar la imagen, etc, así será el tipo de procesamiento con el que se trabaja. En nigún caso se considera el contenido semántico de la imagen, es decir, se la considera como un tipo específico de señal, sin atender a su significado. Por similitud, se asocia también a este ámbito aquellas cuestiones relacionadas con la adquisición y transmisión de imágenes, y el estudio del sistema de percepción visual humano a su más bajo nivel. Algunos autores incluyen también incluyen la segmentación y extracción de características dentro del procesamiento de imágenes.

Cuando la imagen de entrada con la que se trabaja se representa en forma discreta, como conjunto de píxeles con determinado nivel de gris, se habla de *procesamiento digital de imágenes*. En él se somete una representación numérica del objeto (que es lo que constituye la imagen digital) a una serie de operaciones encaminadas a la obtención de un cierto resultado que represente una mejora de la imagen original o una extracción de información de la misma. Debido a su naturaleza discreta, el procesamiento digital de imágenes se ha concebido para ser realizado por ordenadores. Como el procesamiento de imágenes se lleva a cabo para la inspección humana, los parámetros de procesamiento suelen ser especificados por éstos.

DESARROLLO HISTÓRICO.

Puede considerarse que la óptica siempre se ha ocupado del procesamiento de imágenes. Sin embargo es a finales de los años 30 cuando se realiza un fuerte esfuerzo en acercar la óptica, hasta entonces fundamentalemte instrumental, a las teorías de comunicación e información surgidas en el seno de la matemática y electrónica. Ese esfuerzo se justificada por el hecho de que los sistemas de comunicación y de formación de imágenes

se concebían básicamente para transmitir o recibir información. Por otro lado, el formalismo matemático empleado para describir sistemas ópticos, como eran la teoría de sistemas lineales y el análisis de Fourier, se mostró muy útil para describir ciertos fenómenos eléctricos, sobre todo en el ámbito de la teoría de la señal. En el caso de los sistemas de de comunicación de señales eléctricas la información es temporal: las señales transmitidas son monodimensionales y dependen únicamente del tiempo *t*. En los sistemas ópticos la información es esencialmente espacial: las señales son distribuciones de intensidad luminosa en el espacio, función de dos variables.

A principios de los años 50 esta interrelación se concreta en el uso de técnicas ópticas de filtrado en luz coherente orientado a la mejora de la calidad de fotografías. Las ideas desarrolladas en ese ámbito fueron pronto utilizadas en el dominio de las señales sísmicas y de radar, y posteriormente en el reconocimiento de caracteres escritos.

En ese tiempo las señales eran de naturaleza continua, o analógica, de manera que el filtrado óptico era rápido, simple y relativamente asequible. Pero se planteaban muchos problemas, sobre todo de almacenamiento y de escasez de campos de aplicación. Con el advenimiento de los ordenadores se planteó la posibilidad de extender los conocimientos adquiridos durante décadas en el filtrado óptico para apicarlo a imágenes digitalizadas, fácilmente almacenables, empleando las ventajas que supone el manejo del ordenador. La máxima ventaja de este procedimiento reside, por tanto, en una mayor verstilidad. El ordenador puede programarse para realizar cualquier tipo de filtrado sobre la imagen digitalizada.

Fue la prensa la que motivó el primer empleo del procesamiento digital de imágenes. Hacia 1920 se empleaba para codificar y transmitir imágenes para los periódicos entre Londres y Nueva York, permitiendo exclusivas en unas horas cuando antes se precisaban días. Sin embargo, hasta los años 60, cuando con los ordenadores se adquiere la suficiente memoria para mejorar las prestaciones, no se puede hablar de un desarrollo del campo del procesamiento digital de imágenes. La luna, como tantas otras veces, fue ahora la fuente de inspiración. En el JPL (*Jet Propulsion Laboratory*) de Pasadena (California) se deseaba obtener la máxima información de las imágenes enviadas por los Estados Unidos para orbitar nuestro satélite. Robert Nathan

llegó a Pasadena en 1959 para establecer un sistema digital de gestión de datos. En 1961, los soviéticos conseguían imágenes espectaculares de la luna con sus misiones *Lunik*. Con el orgullo patrio dañado, Nathan se marcó el objetivo de mejorar esa gesta cuando en 1961 se lanzaron las sondas *Ranger*. Los comunistas habían empleado técnicas ópticas para limpiar áreas contaminadas por fuerte ruido de transmisión. El americano pensó que el empleo del ordenador sería una herramienta mucho más poderosa para eliminar los defectos de las imágenes de televisión provenientes de los módulos espaciales.

Del empeño de Nathan, junto con otros como Selzer y Frieden, nacieron técnicas, *software* y *hardware* que fueron el germen del desarrollo de este campo. En la primera mitad de los 60 se diseñaron técnicas de aumento de contraste, de corrección fotométricas y geométricas y eliminación de ruido periódico para las sondas *Ranger*, y luego para el proyecto *Apollo*. No contento con su victoria frente al enemigo soviético, Nathan se dedicó también a mejorar imágenes biomédicas procedentes de la reciente microscopía electrónica, y de radiografías.

Actualmente los dominios de aplicación del tratamiento numérico de imágenes digitales son muy numerosos. Podemos citar la medicina, en dónde se trabaja tanto para mejorar imágenes de todo tipo, como para desarrollar métodos de diagnóstico asistido por ordenador. En biología se emplea con profusión para análisis citológicos. Y también tenemos aplicaciones en geografía, cartografía, meteorología, astronomía, geología, recursos naturales, física, industria, estudios medioambientales... La lista no se limita al ámbito científico. Hoy día han irrumpido en este campo áreas tan dispares como los audiovisuales, la moda, el arte y la restauración.

Áreas de procesamiento.

Dentro del procesamiento digital de imágenes podemos encontrar diferentes áreas, según el tipo de mejora que demos a la imagen de salida. Por ejemplo, podemos estructurarlas como sigue:

- *Preprocesamiento*. En esencia, consisten en técnicas dirigidas a mejorar la imagen original y facilitar su análisis posterior, o bien a obtener una imagen resultante lo

más parecida posible a la original. Dicho proceso se realiza sobre las distorsiones producidas al adquirir o transmitir la imagen de un objeto real. Para poder corregir dichos defectos se emplean mecanismos que se basan en cierto conocimiento previo acerca de los métodos de captación y transmisión de señal. Algunos autores [ROS82] introducen una clasificación dentro del preprocesamiento, segregando la *restauración* del *realce*, según sea la intención última: una imagen resultado lo más fiel posible a la original, o una con ciertas mejoras adecuadas. Las técnicas de preprocesamiento pueden ser lineales o no lineales, y serán estudiadas con mayor detalle en un posterior epígrafe.

·*Compresión.* La cantidad de datos asociados a la información visual puede exceder la capacidad de almacenamiento o de transmisión de nuestro dispositivo. La compresión de imágenes intenta reducir el número de bits empleados en la representación de las mismas al mínimo valor posible. Se pueden distinguir dos fases en este tipo de procesos: en primer lugar se convierte la imagen a un espacio adecuado para su manejo, y en segundo lugar se realiza la compresión propiamente dicha.

·*Reconstrucción.* En este ámbito se persigue la reconstrucción de imágenes bidimensionales a partir de proyecciones monodimensionales. Estas técnicas son trascendentales en la generación de imágenes médicas, como tomografías y TAC, aunque también se emplean en astrofísica radar, medioambiente y estudio de la calidad de piezas industriales.

PREPROCESAMIENTO DE IMÁGENES:FILTRADO.

Bajo este epígrafe, se engloban las operaciones con imágenes que tienen lugar al más bajo nivel de abstracción. Tanto la imagen de entrada como la de salida son icónicas, del mismo tipo que el objeto original. Hay toda una pléyade de artículos, revistas, monografías y libros dedicados al estudio y diseño de mecanismos de corrección de distorsiones en imágenes [CHE92][GOR84] [GON86] [HWA95] [JAI89] [KO91] [LEE90] [NIE87] [PIT90] [PRA78] [ROC91] [ROS82] [SON93] [SON88] [SUC95] [SUN88] [WEN86] [ZEN94a]. En

gran parte de ellos, sobre todo en los que tienen cierta antiguedad, se realiza la suposición de que los sistemas, o son lineales y espacialmente invariantes, o pueden reducirse a ello. El problema de la corrección del ruido impulsivo queda relegado a un segundo plano. Además, las características de la imagen y las del ruido se suponen conocidas a priori, cosa que no sucede en las aplicaciones prácticas, dónde hay que estimar los parámetros de la imagen y ruido directamente de la imagen fuente.

Formulación del problema.

La solución al problema de la corrección del ruido se construye alrededor de la hipótesis de que es posible definir una función bidimensional que describa el comportamiento de un sistema ideal, la imagen original ya definida, O.

$$O : \mathcal{V} \times \mathcal{H} \rightarrow K$$
$$(i, j) \rightarrow O_{ij} \in K$$

donde $\mathcal{V} = \{1, 2, ..., V\}$
$\mathcal{H} = \{1, 2, ..., H\}$
$K = \{1, 2, ..., G\}$

de manera que el sistema real, la imagen fuente, puede describirse como cierta función $\mathscr{T}$ que convierta la señal ideal en la observada:

$$\mathscr{T}: K \rightarrow K$$
$$O_{ij} \rightarrow \mathscr{T}[O_{ij}] \equiv S_{ij}$$

El objetivo del filtrado es entonces determinar de alguna manera una transformación de corrección $\mathscr{F}$ de la imagen observada de modo que el resultado de su aplicación

$$\mathscr{F}: K \rightarrow K$$
$$S_{ij} \rightarrow \mathscr{F}[S_{ij}] \equiv R_{ij}$$

sea, en cierto sentido determinado por determinados criterios, tan parecido a la imagen original como sea posible. Ha de advertirse que $\mathscr{F}no$ tiene porqué ser la función inversa de $\mathscr{T}$.

Como se ha apuntado, el objetivo básico es suprimir las distorsiones y, si procede, realzar algunos detalles. De esta manera, atendiendo a la finalidad perseguida en cada caso, pode-

mos establecer nuestra propia división de las técnicas de preprocesamiento, basándonas en las de Rosenfeld [ROS82] y Sonka [SON93]. Puede esquematizarse como sigue:

· *Transformación de brillo.* Opera píxel a píxel. Podemos distinguir entre la *corrección de brillo* en la que el valor definitivo de dicho parámetro depende de la posición en la imagen; y las *transformaciones de nivel de gris* que son independientes de la localización. En este último grupo se pueden introducir los métodos de umbralización, ecualización y empleo de pseudocolor.

· *Transformaciones geométricas.* Son las que alteran la apariencia geométrica de la imagen, como las rotaciones, cambios de escala o transformaciones afines. También la interpolación de brillo entre píxeles vecinos puede incluirse bajo este epígrafe.

· *Transformaciones de vecindad local.* A partir de la información local contenida en los píxeles vecinos al procesado se realizan transformaciones adecuadas, según la finalidad. Empleando la nomenclatura del procesamiento de señal, estas transformaciones reciben el nombre de *filtrado*. También los detectores de borde se engloban dentro de este apartado.

· *Restauración.* Se intenta suprimir la degradación empleando un conocimiento previo acerca de dicho proceso *sobre la imagen completa, y no localmente.* Podemos diferenciar dentro de este ámbito métodos *deterministas*, basados en el conocimiento explícito de la función de degradación; y los *estocásticos* que intentan buscar la mejor restauración optimizando el resultado de acuerdo con dierto criterio estocástico, como por ejemplo el mínimo error cuadrático medio. En algunos casos debe estimarse previamente la función de degradación.

Nosotros centraremos nuestro interés en las transformaciones de vecindad local, si bien la frontera con la restauración de imágenes es a veces poco nítida. En estas técnicas, el preprocesamiento emplea a menudo la redundancia que aparece en las imágenes. Dicha redundancia viene originada por el hecho de que los píxeles vecinos correspondientes a un objeto real

tienen esencialmente el mismo (o similar) valor de brillo. Parece, pues, lógico emplear una pequeña vecindad del píxel en estudio para producir un nuevo valor del nivel de gris en la imagen resultante. De este modo un píxel que distorsione puede ser localizado y eliminado de la imagen, sustituyéndolo por un valor que se determine teniendo en cuenta los píxeles vecinos. Por ejemplo puede usarse el nivel de gris promedio de los vecinos, o algún estadístico de orden. La elección de la transformación local, el tamaño y la forma de la vecindad depende fuertemente del tamaño de los objetos en la imagen. Por ejemplo, el tamaño de la vecindad habrá de ser mayor que las distorsiones que desean eliminarse. De esa manera el nivel de gris del píxel que habrá de sustituírse podrá ser estimado a partir de mayor redundancia, y por tanto con mayor fiabilidad. Pero por otro lado, el tamaño de la vecindad no podrá ser mucho mayor que los detalles más finos de la imagen y que se desea preservar, puesto que si no dicho detalle podría ser considerado como distorsión.

Un filtro puede definirse matemáticamente como un operador F que transforma una señal de entrada $X(i)$ en otra de salida $Y(i)$:

$$Y(i) = F[X(i)]$$

Con objeto de hacer manejable el análisis de un filtro suelen imponerse ciertas restricciones sobre F, como por ejemplo su linealidad o invarianza espacial) pero incluso así se obtienen estupendos resultados. Si se escoge adecuadamente el operador, podemos alcanzar diferentes objetivos. Atendiendo a cuál de éstos se dirija el procesado, podemos distinguir dos técnicas contrapuestas: de un lado los métodos de *realce* y por el otro los métodos de *suavizado*. Los primeros, basándose en el empleo de derivadas locales de la función imagen, hacen más nítidos los bordes de la imagen y resaltan detalles. Equivalen a los filtros pasa-baja en el dominio de la transformada de Fourier. Presentan el inconveniente de que el ruido de alta frecuencia se ve realzado, con el consiguiente deterioro de la imagen. Las tácnicas de suavizado están encaminadas a eliminar el ruido y otras fluctuaciones, y son similares a los filtros pasa-alta en el dominio de Fourier. Como principal desventaja podemos señalar el que los bordes de la imagen y los detalles más finos se ven difuminados.

En el presente trabajo monográfico sólo se considerarán

las técnicas de suavizado. Aunque de nuevo aparece una disyuntiva, puesto que según el tipo de transformación que utilicemos, podemos tener un filtro *lineal* o *no lineal*. Esta división es tan radical que merece ser estudiada en detalle.

Filtrado lineal.

Un filtro F se llamará lineal si cumple los principios de linealidad y superposición, es decir, si la salida es una combinación lineal de la entrada:

$$F[a_1 X_1(i) + a_2 X_2(i)] = a_1 F[X_1(i)] + a_2 F[X_2(i)]$$

dónde a_1 y a_2 son dos números reales cualesquiera. La principal ventaja de los filtros lineales se la da el principio de superposición, pues basta conocer la respuesta al impulso para poder determinar el comportamiento del filtro ante cualquier señal de entrada. Los filtros lineales pueden clasificarse a su vez en dos clases, atendiendo al dominio en que operan, espacial o frcuencial. Los primeros manipulan directamente los puntos de la imagen original mediante convoluciones, mientras que los segundos se basan en la modificación de la imagen utilizando la transformada de Fourier. Ambos métodos no son equivalentes, aunque lleguen a resultados similares, por lo que se mostrarán por separado.

Métodos en el dominio espacial.

Convolución y correlación.

En el procesamiento digital de imágenes, la variación de las diferentes variables es discreta, no contínua, como ya se ha apuntado. En ese caso, la convolución de dos funciones $f(x)$ y $g(x)$ se define como:

$$f(x)^* g(x) = \sum_{y=-\infty}^{\infty} f(y)g(x-y)$$

En el caso de los filtros $f(x)$ es la señal de entrada, que se desea procesar, y $g(x)$ es la denominada respuesta finita al impulso unidad (*psf, point-spread function*), esto es, la salida que proporciona el filtro lineal a una entrada impulsiva unidad (una delta de Dirac). Si se conoce la respuesta finita al impulso para un

determinado filtro, entonces puede calcularse la respuesta a cualquier señal de entrada gracias a la linealidad del filtro. Por tanto, la función $g(x)$ caracteriza completamente al filtro. Si el rango sobre el que $g(x)$ no es nulo es $[-w,w]$, entonces:

$$f(x)^* g(x) = \sum_{y=x-w}^{x+w} f(y)g(x-y)$$

Es decir, el resultado de la convolución en un punto i se obtiene como la suma ponderada de los valores de la entrada en un entorno cercano a él , dónde los pesos vienen dados por el valor de la función $g(x)$ en el intervalo $[-w,w]$. En una imagen bidimensional, si $g(x,y)$ es no nula en el intervalo $[-w,w]\times[-v,v]$:

$$f(i,j)^* g(i,j) = \sum_{k=i-w}^{i+v} \sum_{l=j-w}^{j+v} f(k,l)g(i-k,j-l)$$

Podemos entonces comprender en qué consiste una convolución bidimensional sobre una imagen digital. Matemáticamente, se trata de invertir la función $g(x,y)$ respecto del origen, trasladar los ejes para que coincidan $g(0,0)$ con $f(i,j)$ y realizar el producto de ambas, sumando sobre los índices indicados. Si se diera el caso de que $g(x,y)$ fuese simétrica respecto del origen no sería necesario invertirla, y el proceso de convolución se reduce a lo que se denomina *correlación*:

$$f(i,j) \circ g(i,j) = \sum_{k=i-w}^{i+v} \sum_{l=j-w}^{j+v} f(k,l)g(i+k,j+l)$$

En el procesamiento de imágenes, estas operaciones equivalen a ponderar los valores de los niveles de gris de la imagen con el comportamiento del filtro. Eligiendo adecuadamente los coeficientes de $g(x,y)$ se pueden diseñar muchos filtros lineales. De hecho, extendiendo la idea básica de la convolución, los coeficientes o el dominio sobre el que $g(x,y)$ es no nula pueden variarse dependiendo de la imagen o de las coordenadas, dando lugar a filtros que se escapan de la clasificación de lineales, pero que se basan en el movimiento de una ventana $g(x,y)$ sobre la imagen.

Suavización.

Consiste principalmente en eliminar los efectos de distorsión que presenta una imagen a causa de una digitalización a baja resolución espacial o de su transmisión por una canal con ruido, como ya hemos estudiado. Ello se puede lograr muy sencillamente, por ejemplo mediante un promedio de los valores de gris vecinos a uno dado. El promedio sobre una vecindad es una

técnica directa de suavización en el dominio espacial:

$$h(i,j) = \sum_{k=i-w}^{i+v} \sum_{l=j-w}^{j+v} f(i-k, j-l)$$

Esto equivale a pasar una matriz de pesos sobre la imagen, ponderando cada nivel de gris de la imagen para obtener un valor promedio como resultado. En el caso de vecindades de 3 × 3 píxeles, la matriz de pesos sería:

$$g = \frac{1}{9}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Variando el valor de los elementos de la matriz (y el correspondiente coeficiente normalizador) se pueden obtener diferentes filtros de suavización. Obsérvese que dicha variación no es más que dar más importancia a los vecinos más cercanos frente a los más distantes al píxel en cuestión, que es el central de la ventana. Se introduce, por tanto, aunque de manera poco sofisticada, cierta información apriorística acerca de la imagen. En cualquier caso los pesos se normalizan a la unidad para no sesgar la amplitud de la imagen original.

Sin embargo, al realizar este promedio local se producen efectos colaterales indeseables, como el difuminado de las estructuras de interés, ya sean contornos, líneas o detalles pequeños. El incremento del tamaño de la ventana, esto es, de la vecindad que interviene a la hora de filtrar el píxel, conseguirá una mayor suavización a expensas de un mayor deterioro del detalle.

Nitidez.

Las técnicas de nitidez (*sharpening*) se emplean para resaltar contornos dentro de una imagen, esto es, incrementar las variaciones localizadas de intensidad. Es por tanto un proceso contrario en cierta manera al de suavización que acabamos de comentar. Como aquél se basaba en la integración de niveles de grises pesándolos adecuadamente, mediante un promedio, cabe esperar que los métodos de aumento de nitidez se fundamenten en un proceso matemático contrario: la diferenciación. Y en efecto, la aplicación de un operador de diferenciación da

lugar a un reforzamiento de los detalles. Los dos métodos de diferenciación más empleados en procesamiento de imágenes son el gradiente y el laplaciano.

El operador gradiente en un punto actúa sobre una función, dando como resultado un vector que proporciona la dirección de máxima variación de dicha función. El operador laplaciano informa acerca de la variación del gradiente. Por lo tanto ambos métodos de diferenciación permiten conocer datos precisos acerca de la situación de los contornos, al presentar éstos fuertes variaciones de contraste y brillo, por lo que pueden ser empleados adecuadamente en mecanismos de aumento de la nitidez.

Más concretamente, el modo de emplear estos operadores para realzar los detalles y contornos de la imagen puede ser alguno de los que paso a enumerar:

· Calculando la imagen de la magnitud del gradiente o el laplaciano directamente.

· Sumando el valor de alguno de esos operadores, a veces ponderado con ciertos parámetros ajustables, a la imagen original. De este modo se resaltan los contornos al tiempo que se mantiene la información subyacente. En el caso del laplaciano el filtro resultante es equivalente a una convolución, de manera que los contornos son resaltados a base de desplazar el nivel de gris de cada punto respecto al de sus vecinos. A este filtro se le denomina con frecuencia *unsharp masking* por su analogía con un método fotográfico similar del mismo nombre.

· Realzando en la dirección del gradiente. Como el laplaciano es un operador no direccional, esto es, el resultado es independiente de la orientación, conviene complementar el resultado del laplaciano con el del gradiente, que sí es direccional. Así, primero se calcula el gradiente en cada punto y después se aplica un laplaciano monodimensional en su dirección.

· Empleando mecanismos de diferencia direccional, mediante el cálculo del gradiente en una dirección predeterminada para resaltar aquellos contornos que presen-

LINEALES

- Sustento teórico.
- Fáciles de diseñar.
- Pueden eliminar interferencias en el dominio espacial y frecuencial.
- Buen comportamiento frente al ruido gaussiano.
- Mal comportamiento frente a ruido impulsivo.
- Mal comportamiento frente a contornos abruptos (espectros anchos), sobre todo si la ventana deslizante es grande.
- No se comportan bien frente a ruido dependiente de la señal.

tan la misma orientación. Un efecto de relieve muy interesante se consigue sumando una imagen de intensidad constante (nivel de gris constante) a dicho gradiente.

El modo de calcular los operadores gradiente y laplaciano en el caso discreto merecería un capítulo aparte, que no es el objetivo de este estudio. Sólo apuntar que el gradiente suele aproximarse mediante la sustitución de las derivadas parciales por diferencias en ciertas direcciones, lo que resulta en unas máscaras que se convolucionan con la imagen para obtener el valor aproximado de las componentes del gradiente. El operador laplaciano puede calcularse de una manera similar convolucionando la imagen con una máscara adecuada.

Métodos en el dominio frecuencial.

Transformada de Fourier, convolución y correlación.

La transformada directa e inversa de Fourier, en una o dos dimensiones, es una transformación lineal. Merced al teorema de la convolución, la transformada de Fourier y la convolución están relacionadas: la convolución espacial de dos funciones puede obtenerse haciendo la transformación inversa del producto de las transformadas de Fourier de ambas funciones. Sin embargo, conviene tener precauciones cuando se trabaja en la práctica con imágenes digitales, puesto que no siempre se obtiene el mismo resultado con una convolución que mediante el teorema de la convolución. Por ejemplo, un filtro acotado en el espacio de frecuencias puede tener como transformada inversa una función no acotada, es decir, con valores no nulos en todo el dominio espacial. Pero lo más importante es que tanto la convolución como el filtrado en el dominio frecuencial son operaciones lineales. El cambio de espacio nos permite elegir sobre qué componentes de la imagen actuar, si sobre sus intensidades espaciales o sobre sus componentes frecuenciales.

Suavización.

Los contornos y cambios abruptos de nivel de gris (tales como el ruido o contornos entre regiones) de una imagen contribuyen grandemente a las altas frecuencias. Por tanto se puede conseguir la suavización atenuando un rasgo específico de las componentes de alta frecuencia en la transformada de Fourier

de la imagen. Para ello se transforma la imagen al dominio de frecuencias. Allí, mediante un filtro adecuado se eliminan las componentes de alta frecuencia que se deseen, obteniéndose una nueva imagen en el dominio frecuencial. Mediante una transformación inversa de Fourier se devuelve la imagen resultante al dominio espacial. Es lo que se denomina filtro *pasa-baja*. Idealmente, se define de manera que todas las frecuencias dentro de un determinado intervalo pasan sin atenuación, y las restantes son completemente eliminadas, pero esta función de transferencia es imposible de implementar electrónicamente. Otros filtros pasa-baja, como el de *Buttetrworth* se definen de forma menos abrupta que el caso ideal. La discontinuidad de la función de transferencia ahora se suaviza. Como consecuencia, las imágenes así filtradas presentan menos limpieza de ruido de alta frecuencia, aunque muestran más respeto a los contornos.

Nitidez.

Puesto que los contornos y otros cambios bruscos de intensidad están asociados a las componentes de alta frecuencia, si deseamos resaltar dichas características podremos lograrlo mediante un filtro que permita el paso de componentes de alta frecuencia. A este tipo de filtros se le denomina *pasa-alta*. Son análogos a los definidos en la sección anterior, solo que eliminando las frecuencias bajas en este caso.

Filtrado no lineal.

En los principios del procesamiento de la señal e imágenes los filtros lineales fueron las primeras herramientas utilizadas. Su simplicidad matemática y la existencia de una elegante teoría subyacente los hacía fáciles de diseñar e implementar, además de ofrecer resultados satisfactorios en numerosas aplicaciones. Sin embargo los filtros lineales tienen un pobre comportamiento frente a ruido que no sea aditivo, o ante distorsiones que no se ajusten a un modelo lineal o gaussiano. En procesamiento de imágenes los filtros lineales tienen algunos defectos dignos de reseña, como los que se muestran en los márgenes de esta página. Además, el sistema visual humano muestra un comportamiento no lineal en sus primeros niveles de procesamiento. Muchos investigadores parecen tener el punto de vista de que es difícil obtener mejoras sustanciales mediante el uso de

NO LINEALES

- Mal sustento teórico.
- Diseños menos eficientes.
- No trabajan en el dominio frecuencial.
- Peor comportamiento frente a ruido gaussiano.
- Mejor comportamiento frente a ruido impulsivo.
- Mejor comportamiento frente a los contornos y detalles de la imagen.
- Mejor comportamiento frente a ruido dependiente de la señal.

los filtros lineales, así que el salto cualitativo hacia los filtros no lineales es completamente necesario. Por estas razones el filtrado no lineal alcanzó cierta importancia ya desde los años 60. Y desde entonces ha tenido un desarrollo muy dinámico. La enorme cantidad de artículos, libros y monografías publicados dentro de este ámbito, y la enorme cantidad de aplicaciones que han encontrado los filtro no lineales en disciplinas tan diversas como telecomunicaciones, procesamiento de imágenes, biomedicina o geofísica lo corroboran. La mayoría de los paquetes de procesamiento de imágenes que existen en el mercado incorporan filtros no lineales.

A diferencia de los filtros lineales, el incumplimiento de las propiedades de proporcionalidad y/o superposición no nos permiten conocer el comportamiento del filtro a partir de la función de respuesta al impulso. Tampoco podemos aplicar la teoría de sistemas lineales para evaluar su eficiencia. En consecuencia el estudio de los filtros no lineales es mucho más coplicado que el de los lineales, y mucho menos estructurado. Han surgido diferentes familias de filtros, cada una basada en un modelo matemático diferente. A menudo la evaluación del comportamiento de los algoritmos se realiza de una manera heurística, al no disponer de una teoría definida para el estudio de los sistemas no lineales. Por esta razón los filtros no lineales tienden a ser *ad hoc* en el sentido de que se idean algoritmos orientados a problemas muy concretos, por lo que se hace muy difícil evaluar las propiedades y optimizar los diseños. Por tanto se han tenido que desarrollar nuevas herramientas de análisis, de forma que hay que caracterizar el comportamiento de los filtros no lineales determinista y estadísticamente. DEsgraciadamente, se sigue estando lejos de la completitud teórica de los filtros lineales. Se hace imprescindible una teoría subyacente de sistemas no lineales para el investigador, con objeto de tener una fundamentación común y para lograr que los métodos sean transferibles entre diferentes aplicaciones.

Una de las clases de filtros más difundidas es la basada en al teoría de los estimadores estadísticos robustos o estadísticos de orden. Aunque no se expondrá aquí, esta base se utiliza en el estudio de la eficiencia de los diferentes filtros concretos que se han propuesto dentro de cada clase, frente a diferentes tipos de ruido, preservación de contornos y otras estructuras, etc. Existen dos herramientas dentro de la teoría de la estimación robusta, que permiten la caracterización de estos filtros no lineales, y que serán empleadas en la presente memoria:

• *Raíces:* El descriptor básico de las propiedades características de los filtros no lineales son los conjuntos de señales raíz. Estos son grupos de señales que son invariantes ante sucesivos pasos de filtrado, es decir, permanecen inalteradas aunque se filtren más veces. El hecho de que ciertas señales sean invariantes ante filtrados de mediana ofrece interesantes posibilidades. Cuando se desea eliminar ruido, suele plantearse el problema de cómo preservar ciertos detalles de la señal, manteniendo la efectividad limpiadora del filtro. Lo ideal sería poder diseñar el filtro de forma que los detalles deseados fuesen raíces de la operación de filtrado, de manera que sólo el ruido se vería afectado. Es evidente que como suele ser habitual, lo ideal no puede realizarse completamente, puesto que el filtro de mediana no es lineal, y el principio de superposición no puede emplearse. Sin embargo, cuando una señal consiste en áreas constantes y cambios bruscos entre esas áreas, puede lograrse un efecto parecido. El ruido será eliminado, pero los cambios bruscos permanecerán. En procesamiento de imágenes es común diseñar filtros de mediana que dejen inalterados ciertos patrones, como por ejemplo líneas contínuas de cierto grosor, de manera que no se vean modificados por la acción del filtro, y sólo el ruido sea eliminado.

• *Probabilidad de Colapso:* Cuando se estudian las características deterministas de un filtro no lineal, se enfrenta al algoritmo con diferentes cantidades de ruido en la imagen, y se observa su comportamiento. Sucede entonces que para una cierta cantidad de ruido la salida del filtro es un píxel aberrante, en el sentido de que es muy diferente de los de su entorno. A esa probabilidad de aparición de ruido se le denomina probabilidad de colapso (*breakdown probability*) [PIT90] [LON92]. Si el ruido es impulsivo *sal y pimienta* entonces se alcanza esa probabilidad de colapso cuando la salida del filtro es un pulso blanco o negro, es decir, cuando el proceso de supresión de ruido se ve desbordado por la aparición excesica de éste.

Dentro de los filtros de estadísticos de orden encontramos los filtros de mediana, los max/median, y los L, R o M.

Pero este tipo de filtros emplea solamente información espacial. Para señales constantes imbuidas en ruido i.i.d. (variable independiente e idénticamente distribuída), el orden temporal es irrelevante, pero ello no es cierto para señales no constantes afectadas por ruido. Con objeto de utilizar ambos tipos de información, espacial y temporal, de la señal de entrada, se han propuesto muchas clases de filtros basados en estadísticos de orden en los últimos años, como los filtros FMH (*FIR-median hybrid*) [NIE87], los filtros de mediana ponderada (*WM*) [KO91] o los filtros apilados (*stack filters*) [WEN86]. Los primeros tienen la virtud de aunar las propiedades más deseables de los filtros lineales (FIR) y no lineales (mediana). Se han escrito numerosos artículos dónde se muestra la evolución desde los filtros de mediana hasta los apilados [COY89].

Otros avances en este área han llevado al desarrollo de los filtros *Vector Median* [AST90] y *Matched Median* [AST92]. Los primeros fueron introducidos dentro del procesamiento de señal multiespectral para usar correlaciones entre distintas componentes, como las imágenes de televisión en color, o las multiespectrales enviadas por los satélites como el NOAA. Estos algoritmos se han aplicado para la eliminación de ruido impulsivo en imágenes en color y de cruzamientos de luminosidad y color. Los filtros *Matched Median*, por su parte, se diseñaron para para detección en sistemas de comunicación, y se definieron tanto para sistemas pasabanda como para basebanda. Estos procedimientos se han mostrado más eficaces que sus homólogos lineales en el caso de señales binarias contaminadas de ruido impulsivo.

Otro enfoque lo constituyen los filtros morfológicos, que abordan el filtrado utilizando propiedades geométricas de la señal, en lugar de analíticas. Se basan en la teoría de la morfología matemática [GON86] desarrollada recientemente, fundamentada en la utilización de operadores que realizan tareas de carácter geométrico.

Una de las clases con más solera dentro de los filtros no lineales son los filtros homomórficos. Se han utilizado extensivamente en procesamiento de señal e imagen, sobre todo en relación con la eliminación de ruido aditivo dependiente de la señal y ruido multiplicativo. Existen otras muchas clases de filtros no lineales, pero que no serán expuestas en este somero estudio por razones de espacio.

Filtro de Mediana.

Dentro de los filtros no lineales, los basados en la mediana se han mostrado muy útiles. Durante los últimos años su teoría se ha desarrollado muy rápido, hasta el punto que debe ser considerada como un punto de partida para la teoría general de sistemas no lineales de la que ya hemos hablado. El éxito de los filtros tipo mediana se basa en dos propiedades intrínsecas: la preservación de bordes y la eliminación eficiente de ruido, con una marcada robustez frente al ruido tipo impulsivo [PIT90] [JUS81].. Ninguna de estas propiedades puede ser mejorada por los filtros lineales tradicionales sin un oneroso gasto en tiempo de cómputo, o una manipulación *ad hoc* de los datos. El respeto a los bordes es esencial en procesamiento de imágenes debido a la naturaleza de la percepción visual humana. Y a la efectividad en limpiar ruido impulsivo hay que añadir el hecho de que la mediana es el filtrado óptimo para el ruido biexponencial, al igual que la media deslizante es el filtro óptimo para el ruido gaussiano aditivo. El filtro de mediana no es, sin embargo, una operación de procesamiento perfecta, ni es muy flexible. Puede causar manchas (*streaking, blotches*) [BOV87b][PIT90] y distorsiones de los bordes [DAV89] [PIT90] [BOV87a] sobre todo si la ventana es muy grande. Tampoco va bien con ruido uniformemente distribuido ni gaussiano. También puede eliminar detalles importantes en la imagen [RES94] [DAV88] [PIT90] [PRA78]. La razón principal de estas desventajas es que el filtro de mediana sólo utiliza información espacial contenida en los datos de entrada dentro de la vecindad del píxel procesado, y descarta la información de orden temporal original.

Desventajas del filtro de mediana

En un resumen publicado en EASCON'74, John Tukey citó el *running median* como uno de los filtros no lineales escalarmente invariantes que pueden ser útiles en el suavizado de series temporales [TUK77]. Desde entonces un cúmulo de artículos, monografías y capítulos de libros se han dedicado a explorar el mecanismo de operación de la mediana, analizando su comportamiento y proponiendo extensiones novedosas y viables. En esencia, el funcionamiento de un filtro de mediana es tan simple como esto: se ordena un número impar de muestras de la señal de entrada y se escoge el valor central, o mediana, como salida del filtro. En una dimensión, si la longitud del filtro es $2N+1$, entonces el algoritmo del filtro de mediana se reduce a:

$$Y_i = MED[X_{i-N}, ..., X_i, ..., X_{i+N}]$$

donde Y_i y X_i son las secuencias de salida y de entrada respectivamente, y $MED[\cdot]$ es el operador mediana.

En la mayoría de los casos es razonable asumir que la señal es de longitud finita, por ejemplo L, consistiendo en muestras desde X_0 a X_{L-1}. Para ser capaz de filtrar incluso las partes más externas de la señal, dónde parte de la ventana de del filtro engloba zonas fuera de la señal de entrada, hay que recurrir a soluciones de conveniencia. La más común es repetir los valores de la señal en esos bordes tantas veces como sea necesario para cubrir la ventana en su totalidad, aunque existen otras como la de suponer que la señal es equivalente topológicamente ya a un toro, ya a un cilindro, de forma que la señal de entrada no tiene límites. En la presente monografía se seguirá el camino de la extensión de la señal, método conocido en la literatura como *first and last values carry-on appending strategy*.

Una extensión inmediata del filtro de mediana son los filtros de estadísticos de orden (*ranked order filters* [PIT90] y *order statistics filters, OSF* [BOV83]). Estos procesos operan exactamente igual que el filtro de mediana excepto que se elige como salida el *i-ésimo* estadístico de orden, esto es, el *i-ésimo* valor más alto de la muestra ordenada. De este modo, el filtro de mediana es un caso particular de filtro de estadístico de orden.

$$Y_i = \textit{i-ésimo valor más alto del conjunto } \{X_{i-N}, ..., X_i, ..., X_{i+N}\}$$

Caben citarse, por su importancia, dos casos especiales de filtros de estadístico de orden como son los filtros de máximo y de mínimo, cuya salida es el máximo y mínimo valor de la lista ordenada de la entrada, respectivamente. Ambos operadores son las operaciones fundamentales en el filtrado morfológico discreto, con elementos estructurales planos, llamados erosión y dilatación [GON86] [PIT90].

Otra extensión inmediata del filtro de mediana es el filtro de mediana recursivo [NOD82] [PRA78] [PIT90], que consiste en reemplazar algunas muestras de entrada por muestras de salida previamente calculadas. Para un tamaño de $2N+1$, podemos escribirlo así:

$$Y_i = MED[Y_{i-N}, Y_{i-N+1}, ..., Y_{i-1}, X_i, ..., X_{i+N}]$$

En el filtro de mediana recursivo la operación de procesamiento de la señal se lleva a cabo *in situ*, de manera que la salida del filtro sustituye el valor antiguo de la entrada antes de que la ventana del filtro se desplace a la siguiente posición. Con la misma cantidad de operaciones, el filtro de mediana recursivo suele proporcionar una mayor capacidad de eliminación de ruido que su contrapartida no recursiva, aunque a expensas de un aumento de la distorsión.

La extensión a dos dimensiones de las expresiones utilizadas es inmediata. A continuación se describirán brevemente las propiedades deterministas y estadísticas principales de los filtros tipo mediana.

Propiedades deterministas de filtros tipo mediana: Raíces.

Definición de raíces.

Como se ha apuntado, a la hora de estudiar teóricamente el filtro de mediana no podemos emplear los métodos seguido para los filtros lineales. No tiene sentido un análisis de frecuencias ni de respuesta al impulso en un filtro de estadístico de orden. La respuesta al impulso de un filtro de mediana es nulo. La herramienta básica que define de las propiedades deterministas de los filtros tipo mediana son los conjuntos de señales raíz. El descriptor estadístico básico de los filtros tipo mediana es el conjunto de distribuciones de salida, dada la distribución de la señal de entrada, que se emplean para estudiar las propiedades de atenuación de ruido y preservación de detalles de estos filtros. Dado que la salida del filtro de mediana siempre es uno de los valores de la entrada, es lógico concebir que ciertas señales pasen inalteradas a través del filtro de mediana. Este comportamiento se ha observado de hecho para el filtro de mediana, y también en muchos filtros basados en mediana. Para un filtro de mediana de longitud $2N+1$ esto significa que:

$$X_i = MED[X_{i-N}, ...X_i, ..., X_{i+N}]$$

Si se satisface la anterior ecuación para todas las posiciones i, x_i se denomina señal raíz de un filtro de ese filtro de mediana particular.

La caracterización de las señales raíces se basa en es-

tructuras locales de la señal, como se definen en [GAL81] [HUB81].. Las resumiremos aquí para un filtro de mediana *monodimensional* de tamaño $2N+1$.

· Una *vecindad constante* es una región de al menos $N+1$ muestras consecutivas de igual valor.

· Un *borde* es un conjunto de muestras monótonas crecientes o decrecientes, rodeadas por vecindades constantes de igual o diferente tamaño.

· Un *impulso* es un conjunto de al menos 1, pero menos de $N+1$ muestras cuyos valores son diferentes de los de las regiones adyacentes, las cuales son vecindades constantes equivaluadas.

· Una *oscilación* es una estructura de señal que no corresponde con niguna de las características de las anteriores estructuras.

Obsérvese que estas definiciones están ligadas a la longitud del filtro, esto es, una vecindad constante para un filtro de longitud $2N+1$ también lo será para filtros de tamaños menores que $2N+1$.

Mediante estos conceptos, Gallagher y Wise demuestran en el artículo ya referido que cualquier señal de longitud finita arbitraria es una raíz de la mediana si está constituída sólo de bordes y vecindades constantes. También prueban que un filtrado iterativo de cualquier señal de longitud finita dará como resultado una raíz después de un número finito de pasos. La cota más ajustada conocida para el número de pasos del filtro de mediana necesarios para alcanzar una raíz es la siguiente: si la ventana del filtro es de tamaño $2N+1$ y la señal tiene un longitud L entonces se requieren al menos

$$3\left[\frac{L-2}{2(N+2)}\right]$$

pasos de filtro para producir una señal raíz. Esta propiedad del filtro de mediana es muy significativa, y se le suele llamar *propiedad de convergencia* [FIT85] [ARC82]. Este límite es bastante conservador en la práctica. En realidad, después de 5 ó 10 pasos de filtro sólo tienen lugar pequeñísimos cambios, si es que hay alguno.

Todos los filtros basados en estadísticos de orden poseen la propiedad de convergencia del filtro de mediana. Sin embargo, todas las raíces de este tipo de filtros son, al contrario que ocurre con la mediana, señales constantes, esto es, triviales. Los filtros de mediana recursivos también poseen la propiedad de convergencia. Más aún, son idempotentes, esto es, producen una raíz después de sólo un paso de filtro. Cabe señalarse que un filtr de mediana recursivo tiene exactamente el mismo conjunto de señales raíces que el filtro de mediana no recursivo de igual tamaño de ventana, pero una señal de entrada particular puede desembocar en diferentes raíces al ser tratada con ambos algoritmos. Las señales raíces de los filtros tipo mediana han sido empleadas en codificación de imágenes y voz [ARC82] [ARC83].

Propiedades estadísticas del filtro de mediana.

Los filtros de mediana se emplean como suavizadores robustos y los resultados del filtrado suelen evaluarse de manera cualitativa. El rendimiento de estos algoritmos depende en la eficacia en la supresión de las partes no deseadas de la señal y, lo que es igualmente importante, de la cantidad de información deseable es retenida en la salida del filtro. No existe una medida numérica que combine estas dos importantes propiedades para así evaluar de manera cuantitativa direscta el comportamiento del filtro. En el caso de los filtros lineales ambos objetivos pueden ser alcanzados si, y sólo si, el ruido y la señal ocupan bandas diferentes.

Las herramientas recientemente desarrolladas hacen posible e intuitivo el análisis del comportamiento estadístico de los filtros de mediana [DAV80]. Suponiendo que las muestras dentro de la ventana de filtro son combinación de una señal constante (original) corrompida de por algún ruido aditivo blanco, como suele hacerse en el análisis de los filtros lineales, puede evaluarse la distribución de salida del filtro, y de ahí estimar la varianza de salida [JUS81]. Basándose en esto, es posible la optimización del filtro, pero se llega a resultados mucho mejores si se combina con otros tipos de ligaduras (como por ejemplo restricciones en el comportamiento del filtro [GAB90]) y se aplica a extensiones del filtro de mediana que permitan mayor flexibilidad de diseño que el sólo cambio del tamaño de la ventana.

Sea la señal de entrada de un filtro de mediana un ruido blanco, modelado como variables aleatorias independientes e idénticamente distribuidas (*i.i.d.*) caracterizadas por su media *ì* y varianza σ^2. Sus funciones de distribución y densidad se denotan respectivamente por $\Phi(t)$ y $\phi(t)$ respectivamente. Los resultados de la teoría de probabilidades [DAV80] proporcionan como función de distribución $\Phi_m(t)$, y de densidad $\varphi_{med}(t)$ de la salida del filtro de mediana como:

$$\Phi_m(t) = \sum_{k=m}^{2N+1} \binom{2N+1}{k} \Phi(t)^k \left[1-\Phi(t)\right]^{2N+1-k}$$

y

$$\varphi_m(t) = \frac{(2N+1)!}{N!N!} \phi(t)\Phi(t)^N \left[1-\Phi(t)\right]^N$$

respectivamente, siendo m la mediana —esto es, el estadístico de orden $(N+1)$—.. Ambas expresiones fundamentan el análisis cuantitativo de la atenuación de ruido en los filtros de mediana. Usualmente, deben emplearse métodos numéricos para evaluar estas ecuaciones.

Bajo algunas hipótesis generales, *cf.* [CRA16] la distribución asintótica (para tamaños $w = 2N+1$ grandes) de la salida del filtro de mediana si la entrada es una normal de media y varianza conocidas, μ_m y σ^2_m, puede expresarse como:

$$\mu_m = t_{0.5}$$

$$\sigma_m^2 = \frac{1}{4(2N+1)\left[\phi(t_{0.5})\right]^2}$$

dónde $t_{0.5}$ se define a partir de $\Phi(t_{0.5})=0.5$. Este resultado explica las propiedades robustas ante la eliminación de ruido que presenta el filtro de mediana. Sin tener en cuenta la distribución de la entrada, la operación de mediana produce una estimación no sesgada de la mediana de la distribución $t_{0.5}$. Más aún, la estimación es siempre consistente, puesto que $lim_{w\to\infty}\sigma^2_{med}=0$. La varianza de salida no depende directamente de la varianza de entrada, aunque en su lugar lo hace sobre $\phi(t_{0.5})$. La varianza de la distribución crece con la amplitud de los impulsos si el ruido tiene distribución de colas anchas, o es impulsivo. Sin embargo, $\phi(t_{0.5})$ no tiene porqué cambiar. Obsérvese que la media de la muestra no posee esta propiedad.

Frente a ruido impulsivo, el filtro de mediana presenta una probabilidad de colapso que depende del tamaño de la ven-

tana. Cuanto mayor sea éste, más cantidad de ruido puede eliminar sin dar lugar a resultados aberrantes, puesto que dispone de mayor información de la vecindad:

$$B(p) = \sum_{k=m}^{2N+1} \binom{2N+1}{k} p^k p^{2N+1-k}$$

donde p es la probabilidad de ocurrencia de impulsos; y m es la posición de la mediana. Los valores numéricos de la probabilidad de colapso (denominada probabilidad de reconstrucción errónea en [PIT90]), aparecen en la tabla 1 del capítulo II. Si la probabilidad de ocurrencia de pulso es menor o igual al 30%, las medianas de tamaño 3×3 y 5×5 reducen esa ocurrencia a menos del 10%. Pero si $p = 50\%$, entonces la probabilidad de reconstruir el píxel correctamente es del 50%, independientemente del tamaño de la ventana. En cualquier caso, resulta que una ventana de tamaño $2N+1$ es capaz de eliminar N impulsos. Este no es, sin duda, un mal resultado (la presencia de un sólo impulso en la ventana del filtro de media hace que colapse), por lo que el filtro de mediana tiene una buena capacidad de eliminación de ruido impulsivo. Si el ruido es impulsivo, pero con nivel de gris de pulso uniformemente distribuido dentro de la escala K, puede probarse que la anterior expresión de $B(p)$ sigue siendo válida como probabilidad de colapso.

En la tabla 1 se muestran las varianzas de salida asintóticas para algunas distribuciones de probabilidad comunes. Los resultados para el filtro de media y mediana para ruido uniforme son válidos para todo tamaño de ventana. Puede remarcarse el hecho de que en el caso de ruido laplaciano la varianza asintótica de la operación de mediana es la mitad de la

Densidad de Probabilidad de Ruido	***Promedio***	***Mediana***
Uniforme $f(X)=\begin{cases}\frac{1}{\sqrt{12\sigma^2}}, & -\sqrt{3\sigma^2} \le X \le \sqrt{3\sigma^2} \\ 0 & \text{en otro caso}\end{cases}$	$\frac{\sigma^2}{2N+1}$	$\frac{3\sigma^2}{2N+3}$
Gaussiana $f(X)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{X^2}{2\sigma^2}\right\}$	$\frac{\sigma^2}{2N+1}$	$\frac{\pi\sigma^2}{2(2N+1)}$
Laplaciana $f(X)=\frac{1}{\sqrt{2\sigma^2}}\exp\left\{-\frac{\sqrt{2}\lvert X\rvert}{\sigma}\right\}$	$\frac{\sigma^2}{2N+1}$	$\frac{\sigma^2}{2(2N+1)}$

Tabla 1: Varianzas de salida para el filtro de media y de mediana, según la densidad de probabilidad de la entrada.

de la operación promedio. Obsérvese también que para ruido gaussiano la mediana sólo es unos 2 dB peor que el filtro de media, que es el óptimo para ese tipo de ruido.

El comportamiento tan bueno del filtro de mediana frente a ruido laplaciano se debe a una importante propiedad de optimización de la operación de mediana. La mediana de la muestra es la estimación de máxima verosimilitud del parámetro de localización de la densidad laplaciana. Esta propiedad implica que el filtro de mediana es óptimo, respecto al error medio absoluto (*MAE*) para atenuar ruido doblemente exponencial. Esta analogía proviene de alternar definiciones de media y mediana de $w = 2N+1$ valores $X_1, ..., X_w$. Consideremos la expresión

$$\Phi(\beta) = \sum_{i=1}^{2N+1} |X_i - \beta|$$

La mediana de $X_1, ..., X_w$ puede definirse como el valor β que minimiza $\Phi(\beta)$ en la anterior expresión, cuando $\beta = 1$. La definición siempre produce como resultado una de las muestras X_i. La media minimiza la misma expresión cuando $\beta = 2$.

No hay que dejar de lado el comportamiento del filtro de mediana frente a los bordes, que son los elementos que más información proporcionan en la imagen. Los filtros lineales pasa-baja destinados a eliminar ruido los distorsionan considerablemente, pues los contornos tienen altas frecuencias. El filtro de mediana tiene un buen comportamiento: puede probarse [PIT90] que si la ventana del filtro es simétrica respecto al origen, y lo contiene, el filtro de mediana correspondiente preserva cualquier borde de escalón. En poresencia de ruido gaussiano, el comportamiento del filtro de mediana puede estudiarse teóricamente a partir de la distribución de probabilidad de la salida del filtro, Φ_m, haciendo que la probabilidad de la entrada sea la de un borde escalón de altura h, — gaussiano de media 0 y varianza σ^2 a un lado del borde; y media h e igual varianza si se está al otro lado—. La esperanza y desviación de dicha distribución de probablidad proporcionan estadísticamente el comportamiento del filtro frente al borde, demostrando que la mediana distorsiona mucho menos los bordes que los filtros lineales. En el capítulo II se pueden observar gráficamente estos comportamientos.

Si el ruido que hay cerca de los bordes es impulsivo, se demuestra que la localización del borde puede ser desplazada en tantos píxeles como pulsos aparecen en la ventana. Como la

probabilidad de que ocurran n pulsos en la ventana decrece rápidamente a medida que n crece, lo más probable es que el filtro de mediana desplace el borde en un píxel, lo cual es prácticamente inapreciable si la imagen es suficientemente grande. Por otra parte, y como resulta intuitivo, se demuestra que la forma de la ventana que menos sesgo introduce en preservación de los bordes es la cuadrada.

El filtro de mediana no sólo elimina ruido en regiones homogéneas de la imagen, sino que también tiende a producir regiones de nivel de gris constante o casi constante. La forma de tales regiones depende de la geometría de la ventana de filtrado. Este efecto colateral del algoritmo es indeseable, pues se percibe como contornos inexistentes en la imagen original. Desgraciadamente, se demuestra [PIT90] que la probabilidad de que dos salidas consecutivas del filtro tengan el mismo valor es bastante alta, tendiendo a 0.5 si la ventana aumenta su tamaño.

Aparte de las propiedades generales de los filtros no lineales, el de mediana presenta las siguientes:

·*Proporcionalidad*:

$$MED[kX_1, ..., kX_w] = k\, MED[X_1, ..., X_w]$$

·*No converge necesariamente, ni es idempotente:*

$$MED[MED[X_1, ..., X_n]] \neq MED[X_1, ..., X_n].$$

·*No se generan nuevos valores, como sucede con la media.*

·*Se reduce la varianza de la señal, por lo que la capacidad de reducir el nivel de ruido es buena. Sin embargo, las texturas de una imagen pueden resultar alteradas.*

·*Se eliminan los impulsos de tamaño menor que la mitad del tamaño de la ventana de filtrado.*

Filtro de mediana separable.

Un filtro bidimensional de mediana separable mediante

una ventana rectangular de tamaño $n \times n'$ consiste en la aplicación de la mediana a lo largo de cada fila y después emplear el resultado como entrada de un nuevo cálculo de mediana, esta vez en sentido vertical. Esto equivale a aplicar un filtro de mediana con una ventana horizontal de tamaño n y al resultado un nuevo filtro de mediana vertical de tamaño n'.

Las propiedades de este filtro son similares a las de mediana en cuanto a su capacidad de suavizado del ruido, solo que su complejidad algorítmica es mucho menor, al requerir para cada punto la ordenación de $n' + n$ valores en lugar de nn'.

Filtro de mediana recursivo.

Una modificación intuitiva del filtro de mediana es la de utilizar los valores de nivel de gris ya filtrados $Y_{i-N}, \ldots, Y_{i-1}$ en el cálculo del valor de Y_i. Sus propiedades difieren de las del filtro de mediana tradicional. Además, en el caso bidimensional tienen la desventaja de que el resultado depende de la dirección en que se realiza el filtrado. A priori, el resultado tiende a estar más correlacionado.

Filtro de mediana ponderada.

Los filtros de mediana ponderada (que denotaremos por WM) son los más simples entre los filtros basados en estadísticos de orden, tras el filtro de mediana. Un filtro WM de tamaño $w = 2N+1$ puede ser descrito por w parámetros e implementado empleando una operación de ordenamiento con el mismo orden de complejidad que un filtro de mediana tradicional de igual tamaño. Por otro lado los filtros de mediana ponderada ofrecen una flexibilidad infinitamente mayor a la hora de diseñar especificaciones que el filtro de mediana. Los pesos controlan el comportamiento del proceso de filtrado. La relación de los WM con el filtro de mediana es comparable con la existente entre los filtros FIR y el filtro de media.

Los filtros WM fueron introducidos por primera vez como generalización del filtro de mediana estándar, asignando a cada posición en la ventana del filtro un peso entero no negati-

vo. Como muestra la figura 10 la estructura de un filtro WM es muy parecida a la de un filtro lineal de respuesta finita al impulso (FIR). Para señales con valores reales, los filtros WM pueden ser definidos en dos maneras diferentes pero completamente equivalentes.

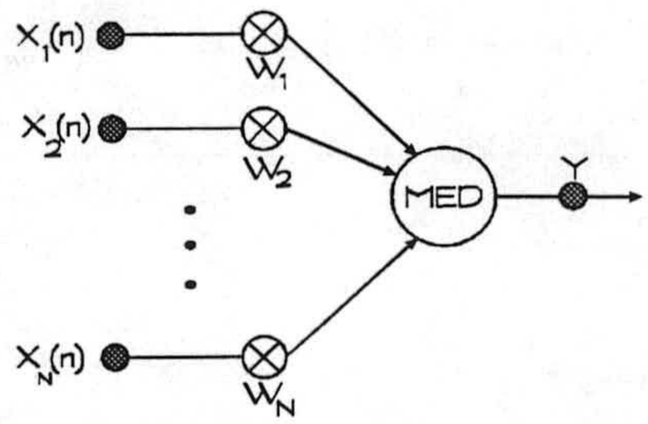


Figura 10: Estructura del funcionamiento del filtro de mediana ponderado.

Definición 1. *Para el vector de entrada de valores contínuos y tiempo discreto* $\boldsymbol{X} = (X_1, X_2, ..., X_w)$*, la salida del filtro WM de tamaño* $w = 2N + 1$ *asociado con los pesos enteros* $\boldsymbol{P} = (P_1, P_2, ..., P_w)$ *viene dado por:*

$$Y = MED[P_1 \otimes X_1, P_2 \otimes X_2, ..., P_w \otimes X_w]$$

donde $\otimes$ *es el operador duplicación, esto es:*

$$K \otimes X = \overbrace{X, \cdots, X}^{K \text{ veces}}$$

■

Este procedimiento de filtrado puede resumirse de la siguiente manera: se ordenan las muestras dentro de la ventana de filtro, se duplica cada muestra x_i un número de veces igual al del peso correspondiente P_i, y se elige como salida el valor central (mediana) de la secuencia resultante.

Definición 2. *Sea* $\boldsymbol{x}$ *el vector de entrada definido anteriormente. El resultado de aplicar el filtro WM a* $\boldsymbol{X}$ *es el valor* β *que minimiza la expresión:*

$$\Phi(\beta) = \sum_{i=1}^{2N+1} P_i |X_i - \beta|$$

■

Aquí, está asegurado que β es una de las muestras X_i debido a que $\Phi(\beta)$ es lineal y convexo si $p_i \geq 0 \ \ \forall i$.

Esta definición permite el uso de pesos no enteros positivos, cosa que no permitía la primera definición. La salida del filtro WM para pesos reales positivos puede ser calculado como sigue: se ordenan las muestras de la secuencia de entrada que aparecen dentro de la ventana de filtro, a continuación se añaden los correspondientes pesos desde la parte superior de la secuencia ordenada hasta que la suma exceda la mitad del total de la suma de los pesos. La salida del filtro es la muestra que

corresponde al último peso añadido.

Ejemplo. Consideremos un filtro WM de tamaño 5 con pesos y entrada dados por los vectores:

$$\boldsymbol{P} = (0.1, 0.2, 0.3, 0.2, 0.1), \quad X = (1, 5, 8, 11, 2)$$

Tras ordenar, obtenemos la secuencia ordenada de entrada siguiente, junto con sus correspondientes pesos:

0.2	0.3	0.2	0.1	0.1
11	8	5	2	1

Comenzando por la izquierda, añadimos los pesos hasta que la suma exceda la mitad de la suma de los pesos, esto es, 0.45. El primer peso es menor que 0.45, así que añadiremos el siguiente peso. La suma es ahora 0.5, que excede la cantidad límite de 0.45. Por tanto el resultado de aplicar el filtro WM es 8.

Como caso particular, tenemos los filtro WM simétricos, que son los más usados en la práctica. Por esta razón ceñiremos nuestro estudio a dichos filtros.

Filtros WM simétricos. Un filtro WM definido por sus pesos $\boldsymbol{P} = (P_{-N}, \ldots, P_{-1}, \boldsymbol{P_0}, P_1, \ldots, P_N)$ es simétrico si verifica que:

$$P_{-i} = P_i \quad \forall i = 1, \ldots, N$$

Debemos hacer notar que existe un número finito de filtros WM para un tamaño de ventana finito, aunque existan un número infinito de vectores peso reales. Puede verse en el anterior ejemplo que multiplicar los pesos por una cantidad constante positiva no altera el comportamiento del filtro. También podemos decir que existe una extensión natural de los filtros WM hacia la generalización que suponen los filtros WOS, o de estadístico de orden ponderado. Estos algoritmos tienen algunas propiedades diferentes a los WM, como la capacidad de eliminar sólo impulsos positivos o negativos. Incluso puede realizarse el procesamiento de un filtro WM de manera recursiva, mostrando entonces cierto paralelismo con los filtros lineales IIR, o de respuesta infinita al impulso, al contrario que los WM que presentaban semejanza con los FIR lineales. Aunque son útiles en la práctica, su manejo teórico es complejísimo. En cualquier

caso, señalemos que sólo se empleará en la presente tesis el CWM, esto es, el filtro de mediana con centro ponderado.

Existe un amplio abanico de artículos tratando el tema de los filtros WM (por ejemplo, el [KO92]) donde se puede obtener información más detallada. Aquí se recordarán las propiedades y resultados más importantes. Como el filtro de mediana, los WM suelen caracterizarse por sus propiedades deterministas y estadísticas.

Propiedades deterministas de los filtros WM.

Las señales raíces de los filtros WM han sido estudiadas por numerosos investigadores. Algunos facilitan la tarea basándose en la propiedad de descomposición de umbral (*threshold decomposition*) que permite estudiar el filtro utilizando sólo señales binarias [FIT85]. La cuestión de si un filtro WM puede hacer que una señal de longitud finita converja a una raíz en un número finito de pasos es crucial para la utilidad de ese filtro, ya que si no cumple la propiedad de convergencia su comportamiento puede ser bastante caótico. Aunque las propiedades de las raíces de los filtros de mediana son bien conocidos, para el caso de los filtros WM las cosas son mucho más complicadas. La propiedad de convergencia es conocida sólo para ciertos casos especiales. La razón principal puede ser que un filtro WM de tamaño de ventana $2N+1$ puede preservar detalles que ocupen menos de $N+1$ muestras. Wendt ha probado ciertos teoremas que caracterizan de alguna manera las señales raíces de los filtros WM. Por ejemplo, si un filtro WM simétrico con $2N+1$ variables preserva las raíces del filtro de mediana de igual tamaño, hará converger a raíz, o a un ciclo de periodo 2, en un número finito de pasos, cualquier secuencia de entrada que se añada. Si dicho filtro WM es además recursivo, siempre habrá convergencia a raíz en un número finito de pasos para cualquier secuencia de entrada añadida.

Sin embargo estos teoremas sólo son aplicables a un subgrupo de filtros de mediana ponderada. Se han encontrado algunos más generales, como el que afirma que cualquier filtro WM simétrico posee la propiedad de convergencia siempre que se verifique que:

$$P_0 \geq \sum_{i=1}^{N} P_i$$

ó

$$P_0 \leq 2\ min\{P_1, P_2, ..., P_N\}$$

o bien

$$P_0 < 2P_f - 2\sum_{i=f+2}^{N} P_i$$

donde f es el llamado *valor característico* del fitro, y verifica

$$1 \leq f \leq N + 1$$

Afortunadamente, los filtros WM que verifican lo anterior constituyen una importante subclase de filtros de mediana ponderada, pues incluye a la mayoría de los que se emplean en la práctica.

Por otro lado, los filtros de mediana ponderada en el centro, o filtros CWM, que son los que poseen mayor peso en el valor central de la ventana, siempre verifican la propiedad de convergencia [ZEN94b].

Filtros de estadísticos de orden (OSF).

Los filtros de estadísticos de orden (OSF) [BOV83] se definen como:

$$Y_i = \frac{\sum a_l X_l}{\sum a_l}$$

En lugar de tomar el valor central como salida del filtro, escogemos una combinación lineal de los valores de otras posiciones. Según los coeficientes de esa combinación lineal, los OSF pueden ser desde un filtro de mediana a un filtro de media, por lo que presentan un comportamiento mixto entre ambos filtros. Es difícil estimar los coeficientes óptimos, pues dependen fuertemente del tipo de ruido implicado.

Filtros de mediana híbridos (FMH); max/med y multietapa (multistage).

En el análisis teórico de los filtros de mediana se hacen algunas suposiciones poco realistas, como por ejemplo que la señal tiene un valor constante que es alterado por el ruido, o que los contornos son escalones ideales. Sin embargo las imágenes reales presentan unas estructuras más complejas (pequeños detalles, contornos difusos, esquinas abruptas ...) que son destruidas por el filtro de mediana que opera con ventanas deslizantes de gran tamaño. No es extraño que suceda este fenómeno, puesto que el proceso de ordenación destruye la información de relaciones de vecindad.

Este tipo de filtros están dentro de los filtros de estadísticos de orden con ventana fija, y por lo tanto mantienen sus propiedades. Se caracterizan por combinar la salida de subfiltros elementales, los cuales se ajustan a la estructura de la ventana $W_{ij}(S)$. Cada subfiltro es diseñado para preservar una característica en una dirección. Incluyendo suficientes subfiltros, el filtro respetará características básicas orientadas en cualquier dirección. La salida del filtro diferencia al max/med de los otros dos: en aquél, se computa como el máximo valor de las salidas de los subfiltros. En los filtros multietapa y FMH la salida es la mediana de la lista formada por las salidas de cada subfiltro junto con el nivel de gris del píxel central de la ventana. La salida de cada subfiltro también es distinta para cada filtro Mientras el filtro de *mediana multietapa* emplea el filtro de mediana estándar [ARC89]; el filtro FMH (*FIR Hybrid Median*) utiliza un filtro FIR (respuesta finita al impulso) [NIE87]. En este último caso el nivel de gris central no es tenido en cuenta en los subfiltros, sino sólo en la salida final del filtro. Por su parte, cada uno de los cuatro subfiltros del max/med hace la mediana de los segmentos de línea horizontal, vertical, y las dos diagonales que pasan por el píxel central de la ventana.

El tipo de características que se desea preservar determinará la capacidad del filtro: si el detalle ocupa un segmento lineal, se diseñará un filtro monodireccional; mientras que si se presenta como dos segmentos ortogonales lineales, habrán de utilizarse filtros bidireccionales. En el artículo original antedicho dónde se definen los filtros multietapa, se muestra que los filtros bidireccionales tienen mejor comportamiento como ate-

nuantes del ruido, en detrimento de sus propiedades de preservación de detalles. Pero por otro lado, los filtros multietapa monodireccionales presentan mayores probabilidades de colapso que el filtro de mediana tradicional, por lo que se comporta peor que la mediana estándar cuando la presencia de ruido es elevada. Podemos decir, en definitiva, que cualquier filtro que mejore el comportamiento del filtro de mediana tanto en preservación de bordes como en supresión de ruido, será mejor que los filtros multietapa.

Los filtros FMH muestran un comportamiento frente al respeto de detalles aún peor. Si el tamaño de ventana es $2N+1$, el filtro preserva estructuras de tamaño N; sin embargo, debido al operador promedio que emplea, los extremos de los detalles se distorsionan. En cualquier caso, los resultados en [ARCE] indican que los filtros de mediana multietapa mejoran a los FMH para bajas probabilidades de aparición de ruido impulsivo en la imagen fuente. Para mayores probabilidades, los FMH son mejores, pero con peor comportamiento frente a los bordes. Experimentalmente se muestra que cuando el ruido aparece con frecuencia (20% a 30% de ruido, como mínimo) ambos tipos de filtro dejan inalteradas muchas estructuras compuestas de ruido, en lugar de eliminarlas.

El filtro max/med, a causa del operador máximo, tiende a los valores altos, lo cual se aprecia más cuanto mayor es el tamaño de la ventana. Pero incluso en ese caso, los detalles se preservan mejor que en el filtro de mediana. Sin embargo, el filtro de max/median se ve superado tanto por el filtro de mediana multietapa, como por el FMH.

Filtros tipo R.

El resultado de filtrar con este tipo de algoritmos depende del rango relativo entre los valores de la secuencia de entrada, y no de los valores mismos. Si F es una función lineal, su forma general es:

$$Y_i = \textit{estadístico de orden}(F[X_i])$$

Tomando esa función lineal como:

$$F[x_{ij}] = 0.5\ [X_{(j)} + X_{(k)}] \qquad \forall j,k \,/\, 1 \le j \le k \le w$$

siendo $w = 2N+1$ el tamaño de la ventana de filtrado, se obtiene el llamado filtro de **Wilcoxon**. Es computacionalmente complejo y no preserva los contornos adecuadamente debido al promedio que se realiza, aunque elimina adecuadamente el ruido blanco gaussiano. Los investigadores Lee y Kao han demostrado que un filtro de Wilcoxon de tamaño $w = 5$ equivale a:

$$Y_i = MED[X_{(3)}, \tfrac{1}{2}(X_{(1)} + X_{(5)}), \tfrac{1}{2}(X_{(2)} + X_{(4)})]$$

que es generalizable a cualquier w. Si observamos detenidamente la expresión, detectamos que los valores que aparecen como entrada de la mediana son los elementos de la diagonal principal de la matriz de coeficientes de la mediana de Wilcoxon.

Filtros tipo L.

Una importante generalización de los filtros de mediana son los llamados filtros L, dónde el resultado se forma como una combinación lineal fija de estadísticos de orden:

$$Y_i = \sum_{j=1}^{2N+1} a_j X_{(j)}$$

donde $X_{(j)}, j = 1...w$ es la secuencia ordenada de las $w = 2N+1$ muestras $X_{i-N,}, ..., X_i, ...X_{i+N}$; mientras las a_j, $j=1...w$ son los coeficentes constantes. Escogiendo diferentes valores para ellos se consiguen diferentes filtros, como el de mediana, punto medio y α-TM (α-*trimmed mean*, o de media recortada). Respectivamente:

$$a_j = \begin{cases} 1 & \text{si } j = m = N+1 \\ 0 & \text{si } j \neq m \end{cases}$$

$$a_j = \begin{cases} \dfrac{1}{2} & \text{si } j = 1,\ 2N+1 \\ 0 & \text{si } j \neq 1,\ 2N+1 \end{cases}$$

$$a_j = \begin{cases} \dfrac{1}{(2N+1)(1-2\alpha)} & \text{si } j \in [2\alpha N + \alpha + 1, (2N+1)(1-\alpha)] \\ 0 & \text{si } j \notin [2\alpha N + \alpha + 1, (2N+1)(1-\alpha)] \end{cases}$$

donde debe notarse que todos los intervalos empleados son discretos.

Filtro de media recortado (α-TM).

Se ha comentado que el filtro de mediana descarta los ompulsos mientras preserva los contornos. Sin embargo, su eficiencia en cuanto a supresión de ruido blanco aditivo es inferior al de realizar la media. Por lo tanto es deseable un compromiso entre ambos tipos de filtro. El α-TM se construye con el promedio de los valores que permanecen después de eliminar los á primeros valores y los á últimos valores de la muestra una vez ordenada:

$$Y_\alpha = \frac{1}{(2N+1)(1-2\alpha)} \sum_{j=\alpha(2N+1)+1}^{(2N+1)(1-\alpha)} X_{(j)}$$

Este tipo de filtro elimina el ruido gaussiano mejor que el de mediana hasta cierto punto, y el impulsivo aproximadamente igual, pero tiende a difuminar contornos.

Filtro MTM y DW-MTM.

Una modificación del filtro α-TM consiste en que el número de valores eliminados no es fijo, sino que depende de un umbral q determinado. Se pretende con ello eliminar aquellas muestras de la ventana deslizante cuyo valor difiera considerablemente de la mediana. Es decir, promedia sólo auqellos píxeles dentro de la ventana cuyos valores están cerca de la mediana.Como se observa, constituyen un tipo especial de filtros de estadístico de orden:

$$Y_{i,j} = \frac{\sum\sum_A a_{r,s} X_{i+r,j+s}}{\sum\sum_A a_{r,s}}$$

$$a_{r,s} = \begin{cases} 1 & \text{si} \quad \left|X_{i+r,j+s} - X_{(N+1)}\right| < q \\ 0 & \text{si} \quad \left|X_{i+r,j+s} - X_{(N+1)}\right| \geq q \end{cases}$$

Una variación a su vez de este último, denominada DW-MTM (*Double-Window MTM*), se basa en emplear dos ventanas, una

de tamaño $n \times n$ y otra mayor $n' \times n'$ centradas ambas en el píxel X_{ij}. Se calcula la mediana de la primera ventana y después se realiza el promedio de aquellos valores pertenecientes al intervalo [MED[X_{ij}] - q, MED[X_{ij}] + q]. La idea es, pues, calcular la mediana de la ventana pequeña para retener los detalles, y promediar los valores de los píxeles de la ventana grande cuyos valores estén cerca de la mediana para mejorar la capacidad de eliminación de ruido [LEE85]. Preservan bordes como el filtro de mediana, y tienen mejor comportamiento frente al ruido , aunque presentan el inconveniente de no filtrar adecuadamente el ruido no estacionario, cuyo comportamiento no se ajusta al método de procesamiento. Además, el hecho de utilizar operadores de promedio, provoca que la salida pueda ser un nivel de gris que no existiese en la imagen original, lo cual es poco justificable si se quiere realizar la restauración de la imagen.

Filtros tipo M.

La forma general de los filtros tipo M es:

$$\sum_{j=i-N}^{i+N} \vartheta\left(X_j - Y_j\right)$$

dónde ϑ es una función impar, continua y que conserva el signo; Y_i es la salida del filtro y X_i son las muestras de la ventana para cada posición j. El filtro de media para X_i distribuídos según una gaussiana; y el de mediana para X_i distribuídos según una laplaciana son casos particulares de filtros M.

Filtros homomórficos y de media no lineal.

Generalizan a varios filtros de mediana preexistentes, como los filtros de mediana generalizados de Kundu o los filtros homomórficos. En una dimensión, se definen como;:

$$Y = g^{-1}\left(\frac{\sum_{i=1}^{2N+1} a_i g(X_i)}{\sum_{i=1}^{2N+1} a_i}\right)$$

donde $g(X)$ es una función analítica no lineal univaluada, y los a_i son pesos [PIT86]. Si éstos son constantes, tenemos los filtros homomórficos, que han sido aplicados tanto a procesamiento de imagen como a reconocimiento de voz y análisis de señales sísmicas.

Se basan en utilizar aplicaciones no lineales (principalmente el logaritmo) para transformar señales no lineales o convolucionadas en señales aditivas, para procesarlas a continuación con filtros lineales, y después deshacer la transformación. Las propiedades de los filtros de media no lineal deppenden de la función $g(X)$ que se utilice, y en los pesos a_i. Las siguientes elecciones de la función proporcionan filtros que son útiles en el procesamiento de imágenes:

$$g(X) = \begin{cases} X & \text{media aritmètica } \langle X \rangle \\ 1/X & \text{media armònica } Y_H \\ \log X & \text{media geomètrica } Y_G \\ X^p & \text{media } L_p \quad Y_{L_p} \\ (p \in Q - \{-1,0,1\}) & \end{cases}$$

Si los pesos a_i no son constantes, se pueden obtener otra clase de medias no lineales mediante una elección adecuada de los pesos. Un ejemplo de ello lo constituye el filtro contraarmónico:

$$Y_{CH_p} = \frac{\sum_{i=1}^{2N+1} X_i^{p+1}}{\sum_{i=1}^{2N+1} X_i^{p}}$$

donde se ha escogido:

$$g(X) = X \quad ,, \quad a_i = X_i^p \ , i = 1 \ , ..., 2N+1$$

esto es, se puede interpretar como una media aritmética que posee coeficientes dependientes de los datos. Los filtros de mediana no lineales presentan esta conocida propiedad:

$$min(X_i) \le Y_{CH_{-p}} \le Y_{L_{-p}} \le Y_H \le Y_G \le$$
$$\le \langle X \rangle \le Y_{L_p} \le Y_{CH_p} \le max(X_i)$$

Si se estudian sus propiedades estadísticas, se deduce que los

filtros de mediana no lineal son claramente superiores tanto a la media lineal como a la mediana para distribuciones de ruido gaussiano y uniforme. Para ruido impulsivo blanco o negro, pueden emplearse con relativo éxito los filtros L_p, L_{-p}, CH_p y CH_{-p}. Sin embargo, fallan al eliminar ruido impulsivo *sal y pimienta*. En tal caso habría que utilizar en caden aun filtro L_p seguido de un L_{-p}.

Aunque todos los filtros de media no lineal tienden a preservar los bordes (salvo la media geométrica), introducen sesgo: L_{-p} y CH_{-p} tienden a realzar las regiones oscuras de la imagen; mientras que los L_p y CH_p lo hacen con las regiones brillantes. Este indeseable efecto es tanto más apreciable cuanto mayor es p.

El problema de la medida de calidad del procesamiento.

Como se ha mostrado, desgraciadamente no existe una teoría general para el filtrado de imágenes. Quizás el principal inconveniente que ello acarrea sea el hecho de que no hay una medida de calidad de imágenes que sea universal y realmente eficaz. Cuando se procesa una imagen para su interpretación visual el observador es el único juez de la eficacia del método empleado. A su vez, es difícil definir cuál es la «imagen buena» con la que el observador compara mentalmente la resultante para tomar esa decisión.

Comparación entre imágenes.

El objetivo último de todo filtrado es proporcionar un resultado. Pero surge entonces un dilema: ¿es o no es bueno ese resultado? Esta cuestión sintetiza uno de los mayores problemas, tanto teórico como práctico, en el ámbito del procesamiento digital de imágenes. Démonos cuenta de que la solución no es inmediata. En primer lugar, hay que establecer el empleo que ha de tener la imagen producto de nuestro filtrado, para poder dirimir si ese resultado es o no bueno. Así por ejemplo, el filtrado con la mediana tradicional puede ser satisfactorio para un médico, pues elimina el ruido impulsivo de fondo de una radiografía, pero será inaceptable para un astrofísico, pues dicho filtro le eliminariá la mitad de las estrellas de su placa. Por otro lado, también habrá de tenerse en cuenta si la imagen resultante será luego utilizada por el ojo humano (para diagnosis médica, por ejemplo), o será tratada como un dato que se le pasa a un ordenador (para clasificar imágenes, etc).

En este último caso parece razonable escoger como medida de calidad el resultado de algún algoritmo que simule el sistema perceptual humano. Aunque nuestro conocimiento sobre la visión humana se ha incrementado significativamente en los últimos años, también se ha revelado lo complejo del sistema, por lo que estamos lejos de de desarrollar un modelo simple y potente que emule el procesamiento visual y que pueda ser empleado en sistemas de optimización de filtrado o que proporcione medidas objetivas de calidad entre imágenes. Esta es la razón por la que el examen *subjetivo* suele ser imprescindible para la evaluación última de la calidad de una imagen, a pesar de sus claros inconvenientes. Se ha establecido un campo de estudio concerniente con la búsqueda de métodos fiables de medi-

das de calidad subjetivas, denominado *psicometría*, aunque aún permanece en una fase muy primitiva. A pesar de ello, se han establecido modelos para estimar la calidad de las imágenes basándose en resultados empíricos de la psicometría y la psicofísica, si bien ninguno de ellos puede definirse unívocamente a partir de los datos experimentales.

En este sentido para evaluar la calidad de una imagen basándose en el comportamiento visual humano habría que incorporar las características de la misma que se conocen, y que fueron expuestas brevemente en un anterior parágrafo. Así, debe de existir una corrección de los efectos de enmascaramiento, promedio y no linealidad. Para tales correcciones hay que emplear parámetros obtenidos por experimentos psicofísicos [NET89] por lo que su validez es sólo cualitativa en la mayoría de los casos, aunque aún pueden realizarse esfuerzos para ajustar más los resultados a la realidad. El modelo más empleado emplea para emular el comportamiento no lineal del ojo un tipo logarítmico de no linealidad basado en los efectos globales de umbral de visión. El error se filtra para enfatizar la zona central del espectro, para seguir la respuesta frecuencial del ojo humano. Para incorporar el efecto de enmascaramiento se evalúa el cambio local de la señal en una vecindad en torno a cada píxel procesado. Esta medida se emplea después para normalizar el error de la señal. Obsérvese que a pesar de que se simplifica al máximo la respuesta visual con este modelo, el proceso parece ser demasiado costoso como para ser empleado con regularidad. Por esta razón se prefieren medidas objetivas de la calidad de la imagen, más sencillas y tratables matemáticamente, aunque no presenten relación con la sensación subjetiva.

En cualquier caso, la manera más evidente de realizar una medida de calidad consiste en comparar entre sí el resultado con la imagen original. Esto podrá llevarse a cabo cuando la imagen original sea conocida, lo cual sucede sólo en la fase de investigación. En cualquier caso, siempre podremos comparar entre sí dos imágenes resultado, para obtener una medida de calidad relativa. Parece por tanto que el problema de la medida de calidad de una imagen se reduce al de la comparación de dos imágenes de niveles de gris de igual tamaño e igual escala (de otro modo no tiene sentido).

Se han propuesto muchas medidas de similaridad entre dos imágenes binarias, entre sus partes, o entre formas poligonales

[PRA78] [HUT93]; sin embargo apenas si hay referencias a la comparación entre imágenes con niveles de gris. Uno de los primeros artículos en ese campo sugiere una medida de disimilaridad para dos imágenes de niveles de gris, pero representándolas como matrices y midiendo la norma de Frobenius de ambas matrices. Otros emplean distancias entre ambas imágenes, píxel a píxel. Pero, aparecen muchos inconvenientes. Por un lado es imposible distinguir un gran error de una suma de pequeños errores en la mayoría de las medidas. Por otra parte, curiosamente, si se comparan las mismas imágenes con cada medida, se obtienen resultados diferentes. Es decir, no hay un acuerdo unívoco entre las medidas y las imágenes. Y siempre ha de recurrirse en última instancia al observador humano. Sin embargo, estas medidas sencillas son matemáticamente tratables, e incluso permiten que se realice algún tipo de corrección basada en el conocimiento del sistema perceptual humano. Puede tenerse en cuenta el hecho de que en contraste con las medidas simples matemáticamente manejables el observador humano siempre basa su estimación acerca de la calidad de una imagen en tan sólo unas pocas áreas locales en la imagen, las que presentan mayor degradación. Estas regiones serían determinadas de manera compleja por el contexto, el fondo y otros parámetros psicofísicos. Pero es posible modelizarlos sumando el error sobre pequeñas áreas locales y usando el máximo como medida de calidad de la imagen.

Entre las diferentes propuestas podemos destacar las siguientes:

- **MSE (*Mean Square Error*).** Se trata de la medida más utilizada, y se correlaciona bastante bien con los test de calidad visual. Además es matemáticamente tratable. Se basa en la suposición de que las imágenes son el resultado de un proceso estocástico, por lo que es posible determinar sin problemas la esperanza matemática del cuadrado del valor absoluto de la diferencia, píxel a píxel, de las imágenes. Sin embargo, en la mayor parte de la literatura, se entiende por MSE a:

$$MSE = \sum_{i=1}^{V} \sum_{j=1}^{H} \left(O_{ij} - R_{ij} \right)^2$$

En su versión normalizada se tiene:

$$NMSE = \frac{\sum_{i=1}^{V}\sum_{j=1}^{H}\left(O_{ij} - R_{ij}\right)^2}{\sum_{i=1}^{V}\sum_{j=1}^{H}O_{ij}^2}$$

que son en realidad unas medidas de mínimos cuadrados (LSE).

· **PLSE (*Peak Least Square Error*).** La definición de la misma nos muestra por sí sola la pequeña variación con la anterior medida:

$$PLSE = \frac{\sum_{i=1}^{V}\sum_{j=1}^{H}\left(O_{ij} - R_{ij}\right)^2}{\left[max\{O_{ij}\}\right]^2}$$

• **PMSE (*Perceptual Mean Square Error*).** Se aproxima mejor a las características de la visión humana:

$$PMSE = \frac{\sum_{i=1}^{V}\sum_{j=1}^{H}\left(Z_{ij} - \underline{Z}_{ij}\right)^2}{\sum_{i=1}^{V}\sum_{j=1}^{H}Z_{ij}^2}$$

donde $\underline{Z}$ es la estimación de Z, la cual es la imagen percibida:

$$Z = H \cdot r(O_{ij})$$

siendo H una fundción de transferencia obtenida de un modelo de visión humana, y r(·) es una función no lineal (logaritmo o potencia). Se demuestra que PMSE tiene una fuerte correlación con la evaluación subjetiva de la imagen, si bien ese estudio se realizó en base a una sola imagen, por lo que la existencia de tal correlación para una gran cantidad de imágenes es, cuanto menos, cuestionable.

• **Haussdorf.** La medida de Haussdorf adaptada a imágenes es una distancia max-min empleada para conocer el grado de desajuste entre dos imágenes fijas, mediante la medida de la distancia del píxel de una de las imágenes que está más alejado de cualquier píxel de la otra, y vice-

versa [HUT93]. Intuitivamente, si la distancia de Haussdorf es d, entonces cualquier píxel de una de las imágenes está a una distancia d de algún píxel de la otra y viceversa. Sin embargo, esta medida se restringe a imágenes binarias, y sólo encuentra aplicación en campos como el reconocimiento de patrones y métodos de visión basados en modelos.

• **MSE subjetivo.** Otra medida subjetiva de calidad de imagen, basada en cierta medida en la forma del MSE, es la siguiente:

$$E = \left(\frac{1}{V \times H} \sum_{i=1}^{V} \sum_{j=1}^{H} \left| D_{ij} \right|^{p} \right)^{\frac{1}{p}}$$

donde

$$D_{ij} = a_{ij}\, g(O_{ij} - R_{ij})$$

Los pesos y la función g se escogen adecuadamente. El coeficiente p se elige suficientemente grande como para incrementar la importancia de los grandes errores. Como tales errores ocurren cerca de los bordes de las imágenes, o en la vecindad de un impulso no procesado, los filtros que no eliminen bien el ruido impulsivo o no preserven bordes provocarán un elevado valor de E; mientras que ocurre lo contrario si el algoritmo posee buenas características de filtrado. Se demuestra que el valor de p y de D_{ij} dependen de la imagen original. Así, se han encontrado los siguientes parámetros para diferentes imágenes:

$$\begin{aligned} D_{ij} &= O_{ij} - R_{ij} && p = 3.0 \\ D_{ij} &= m(O_{ij}) - m(R_{ij}) && p = 2.0 \text{ ó } 5.0 \end{aligned}$$

donde $m(\cdot)$ es la media local en una vecindad pequeña de la imagen que es referenciada.

• **SNR (*Signal to Noise Ratio*).** Es una medida muy empleada en teoría de señal. El resultado se mide en decibelios, y tiene muy diversas definiciones. Quizá la más general sea:

$$SNR = -10\, log_{10}\{error\}$$

• **FOM (*Figure of Merit*).** [PRAT78] Esta medida fue introducida para estudiar el comportamiento de los detectores de borde, pero puede ser adaptada para medir diferencias entre imágenes, de la siguiente manera:

$$FOM = norm\left\{\frac{1}{V \times H}\sum\frac{1}{1+\gamma\cdot\chi_{ij}^2}\right\}$$

donde $d(i,j)$ es la distancia entre los píxeles de las imágenes comparadas (puede ser la distancia euclídea, la de Manhattan, etc); mientras que χ_{ij} es una medida de distancia entre el píxel (i,j) de la imagen original y el de la imagen resultado (máximo, mínimo, media, ...). El parámetro γ es una constante de escala, que en su uso primitivo tenía un valor de 1/9 generalmente, pero que al adaptarse a la función actual suele tomar el valor 1/32768 [CITA Starovoitov].

• **δ^E(O,R) (*Medida de Baddeley*).** Tiene la siguiente representación:

$$\delta^E(O,R) = norm\left\{\left(\sum_{i,j}\chi_{ij}^E\right)^{\frac{1}{E}}\right\}$$

dónde de nuevo $d(i,j)$ es la distancia entre los píxeles de las imágenes comparadas y E un parámetro ajustable (generalmente 2).

• **M(A,B) (*Medida de Starovoitov*).** [STA95] Empleando un método de árboles de decisión, bajo ciertas restricciones que representan los requisitos que se desean para una medida de este tipo, este autor ha conseguido la siguiente, muy parecida a la definida en el anterior epígrafe:

$$M(O,R) = \frac{1}{G}\left\{\left(\frac{1}{(V\times H)^2}\sum_{i,j}\chi_{ij}^E\right)^{\frac{1}{E}}\right\}$$

• **MAE (*Mean Absolute Error*).** Esta nueva medida de

calidad entre imágenes, se fundamenta en el empleo de una cuantificación de error estadístico ampliamente aceptada. En su forma normalizada, se define como:

$$MAE = \frac{\sum_{i=1}^{V}\sum_{j=1}^{H}\left|O_{ij} - R_{ij}\right|}{\sum_{i=1}^{V}\sum_{j=1}^{H}O_{ij}}$$

Para nuestro propósito, supuesto que las aplicaciones últimas de los procesos diseñados serán para uso humano, se elegirá una distancia entre imágenes que verifique una serie de requerimientos que consideramos necesarios. Sean $A = A_{ij}$, $B = B_{ij}$ y $C = C_{ij}$ imágenes arbitrarias; entronces las propiedades deseables de la medida $M(\ ,\ )$ entre imágenes de niveles de gris son las siguientes:

·*Intuición:* La medida $M(\ ,\ )$ debe reflejar la idea intuitiva de similaridad dada por el sistema de percepción humana. En particular, debe ser insensible a pequeñas perturbaciones en los datos.

·*Monotonía:* Si $M(A,B) < M(A,C)$ entonces la imagen A se parece más a la imagen B que a la imagen C.

· *Metricidad:* Se debe verificar:

° $M(A,B) = 0 \Leftrightarrow A = B$;
° $M(A,B) = M(B,A)$;
° $M(A,B) \leq M(A,C) + M(C,B)$.

A este respecto cabe citarse un teorema [STA95]: si una medida de diversidad $M(A,B) = 0 \Leftrightarrow A = B$ y $M(A,B)$ es una función monótona normalizada, entonces $M(A,B)$ es una métrica o bien puede transformarse en una métrica por extracción de raíces.

·*Diversidad normalizada:* Entonces $M(A,B) = 1 \Leftrightarrow A_{ij} = 0, B_{ij} = G \ \forall(i,j)$.

· *Sencillez:* Su cálculo ha de ser rápido computacionalmente.

·*Difusión:* Interesa que la medida en cuestión esté lo

más ampliamente difundida entre los investigadores, y sea empleada con cierta frecuencia y consenso entre los científicos, a fin de poder realizar comparaciones entre diversos métodos de una manera directa.

Podemos entonces elegir ahora aquella medida de similaridad que más nos convenga. Las que se basan en la medida de Haussdorf, al emplear operadores de *max* o *min*, tienen algunos inconvenientes, como el considerar parecidas dos imágenes que realmente no lo son.

Las medidas de Baddeley y Starovoitov son útiles, pero más complejas de calcular, y poco empleadas en la literatura, por lo que las hace desaconsejables. En cuanto a las *MAE, MSE, PLSE* no siempre se correlacionan bien con la visión humana. Por ejemplo, podemos sumar una cantidad fija a los valores de nivel de gris de todos los píxeles (resultando una aclaración global de la imagen) que proporciona un resultado visual espléndido, pero sin embargo... ¡da lugar a un error enorme en nuestras medidas!. Sin embargo, este problema es común a todas las medidas, puesto que es realmente complicado objetivar lo subjetivo. Cabría hacerse una mejora a estas medidas, realizando una transformación previa de las imágenes que fuesen a ser comparadas, de manera que ambas se lleven a un mismo espacio perceptual en el cual sólo las diferencias notables son igualmente perceptibles. Por ejemplo, podría hacerse una transformación de acuerdo con modelos de visión humana. Pero eso siempre iría en contra de nuestra hipótesis de sencillez.

Así pues, dado que la totalidad de los métodos de medida propuestos tienen inconvenientes, nos decantamos por uno de ellos, muy simple y muy difundido, que verifica los requisitos previos que hemos enunciado. Se trata del *CMAE*, que no es más que un cociente entre los *MAE* antes y después del filtrado de todos los píxeles, de esta forma:

$$CMAE = \frac{MAE(O,S)}{MAE(O,R)} = \frac{\sum_{i,j}\left|O_{ij} - S_{ij}\right|}{\sum_{i,j}\left|O_{ij} - R_{ij}\right|}$$

que recuerda mucho al factor de realce F_e empleado por Bernstein [BER87], aunque este autor emplea el *MSE* en lugar del *MAE*. Existen otras razones que nos mueven a esta decisión. Por ejemplo, se ha demostrado que se da lugar a un estimador más ro-

busto si se busca minimizar *MAE* que si se quiere minimizar *MSE* [HUB81]. Por otro lado, el *MAE* es muy útil si trabajásemos con filtros apilados (*stack*) pues reduce el error de estimación del filtro a la suma de los errores de decisión provocados por los operadores booleanos en cada nivel de la descomposición de umbral [COY89]. Pero tales filtros están fuera del estudio de esta tesis.

También se utilizarán otros artificios para estimar la bondad de un método frente a otro. Por ejemplo, emplearemos la imagen resultante de efectuar la diferencia píxel a píxel entre dos imágenes que se pretenden comparar.

NECESIDAD DE FILTROS ADAPTATIVOS.

Ya se ha comentado que el objetivo del filtrado de imágenes digitales no es otro que el intentar conseguir una imagen resultante que sea lo más parecida posible a la original. Si la imagen original se ve afectada por la presencia de ruido de algún tipo, lo cual ocurre siempre al captar una imagen digital, será necesario conocer qué parte de la imagen es señal espúrea y qué parte es señal original. Sin embargo, no está claro qué constituye degradación, y qué no lo es si no se dispone de algún tipo de conocimiento previo. En puridad, sólo podremos llevar a cabo el filtrado conociendo en detalle las características del ruido y la señal. Es decir, se precisa de un modelo de ruido (aditivo, impulsivo...) como los que hemos estudiado en anteriores párrafos que explique la corrupción aleatoria de la señal; y un modelo de la imagen original. Pero esos modelos son desconocidos a priori en los casos prácticos, pues no se dispone de la imagen original. Entonces pueden seguirse dos caminos diferentes:

• *Filtros no adaptativos*: Se diseñan los filtros para ser usados sólo con un determinado tipo de ruido, e incluso a veces con un tipo de señal específico, tal como sucede con la mayoría de los filtros no lineales que suelen emplearse en procesamiento de imagen, como el filtro de mediana. Este proceder resulta muy rápido y sencillo. Sin embargo, lo que interesa a la hora de procesar una imagen no es eso. Recordemos que podíamos interpretar que las imágenes son procesos estocásticos bidimensionales, cuyas estadísticas no tienen por qué ser iguales en cada región de la imagen. Así por ejemplo, la probabilidad de aparición de ruido es diferente en cada región de la imagen, aunque globalemente sea constante. La máxima variación en las características del ruido dentro de una misma imagen suceden cuando ésta se obtiene en el canal de ruido de la transmisión de la imagen (ruido impulsivo) y en el ruido atmosférico de las imágenes de satélite. En tales casos un filtro no adaptativo no funcionará bien, pues tiene unas características definidas orientadas a unas propiedades muy específicas de ruido y señal, las cuales son aquí desconocidas.

• *Filtros adaptativos:* De lo dicho en el anterior parágrafo, se desprende que la elección de filtros que se acomode a la señal es lo más razonable. La mayoría de estos filtros adaptativos basan su funcionamiento en la estimación de las características del ruido y de la señal (generalmente la media y la varianza) a partir de la imagen fuente. Esa información previa acerca de la señal y el ruido ha de ser obtenida empleando métodos interresolución. En efecto, intuitivamente, a la máxima resolución cada píxel es una señal en sí mismo, y no puede distinguirse si es o no un ruido. A menores resoluciones entra en juego un efecto promediador no necesariamente lineal. Frente a ese promedio podemos ver si el píxel, a máxima resolución, constituye o no una irregularidad. El filtrado es, por lo tanto, dependiente de la escala de resolución; es decir, sólo podemos distinguir el ruido de la señal en el espacio de resoluciones. De este modo, a resoluciones bajas se clasifica cierta información como ruido, descartándola en la descripción a mayor resolución. Dicha estimación suele ser local —esto es, realizado sobre ventanas pequeñas—, y la precisión con que se calculan define la bondad del método. Una importante propiedad que han de tener estos estimadores es la robustez frente al ruido impulsivo, dada su freceunte aparición.

Otra razón para elegir filtros adaptativos en el procesamiento de imágenes es su capacidad para preservar la información contenida en los contornos de la imagen, la cual es de primordial importancia en el sistema visual humano. Muchos filtros no adaptativos, como el de media, se comportan bien en regiones homogéneas pero fallan estrepitosamente cerca de los bordes de la imagen. Otros, como el filtro de mediana, no se comportan tan mal, aunque siguen difuminando los contornos de alto contraste.

Queda, pues, de manifiesto, que el empleo de filtros adaptativos parece la elección más racional, pese a su mayor complejidad y lentitud de ejecución. Obsérvese además que cada parte de una imagen difiere generalmente lo suficiente de otras dentro de la misma imagen como para justificar el empleo de un filtro adaptativo. Sus ventajas pueden superar este inconveniente si el algoritmo logra comportarse adecuadamente.

ESQUEMAS ADAPTATIVOS

Desde hace unos años uno de los temas científicos más en candelero ha sido el que concierne al estudio de los *sistemas complejos adaptativos* (*SCA*). Se trata de sistemas que aprenden y evolucionan de forma similar a la de los seres vivos. Ejemplos sencillos podrían ser el caso de una bacteria que se hace inmune a los antibióticos, un perro lazarillo que aprende a guiar a un invidente, o un filtro de imagen que se adapta a la señal de entrada. Si se realizaran estudios parciales de estos ejemplos, claramente no lineales, no podrían dar más que resultados parciales: para tener una idea del comportamiento global hay que estudiar el sistema en su totalidad.

Realicemos, por tanto, un pequeño ejercicio de síntesis. Lo que tienen en común estos procesos es la existencia de un SCA que adquiere información tanto de su entorno como de la interacción entre el propio sistema y el entorno, identificando regularidades, condensándolas en una especie de esquema o *modelo* y actuando en el mundo real conforme a dicho esquema. En cada caso hay diversos esquemas en competencia y los resultados de la acción en el mundo real influyen de modo retroactivo en dicha competencia. Cualquiera de estos modelos nos permite hacer predicciones. Lo normal es que esas predicciones resulten erróneas al principio, pero si la estructura del problema es fácil de captar, la discrepancia entre la predicción y observación hará que los esquemas erróneos sean descartados en favor de otros mejores; o bien que aquéllos evolucionen hacia otros más precisos. De este modo, pronto se puede prever con exactitud el comportamiento futuro del sistema.

En resumen, aunque cada SCA posee sus propias características físicas, todos tienen en común los siguientes puntos:

- **Adquisición de información.** Se lleva a cabo a partir de un flujo de datos captados de algún modo.

- **Procesamiento de información.** Se realiza en dos fases:

  - **Abstracción.** De la información adquirida se extraen regularidades, separándolas de lo meramente arbitrario, condensándolas en un *modelo* sujeto a variación, que incluye reglas que gobiernan el com-

portamiento del sistema.

· **Predicción.** Cada esquema resultante se combina con información adicional (de la clase de información incidental dehechada en la fase de abstracción) para generar una estimación de un resultado aplicable al mundo real, como por ejemplo la descripción de un sistema que puede ser observado, la predicción de un suceso o una prescripción del propio SCA.

· **Retroalimentación (bucle).** La predicción tiene efectos retroactivos, de manera que se seleccionan aquellos modelos que más se ajustan a la realidad. Para ello se exige que la predicción minimice una adecuada *función de error*.

La información de la que se trata está relacionada con la selección entre diversas alternativas, y puede expresarse simplemente si éstas se reducen fácilmente a una serie de elecciones entre alternativas binarias equiprobables. De este modo cada elección proporciona un bit de información. Las regularidades encontradas en la fase de abstracción permiten *comprimir* la información global, haciendo más manejable el estudio del sistema. Parece, pues, que la captación de regularidades es un proceso trascendental en el SCA. Para identificar esas regularidades dentro de una sucesión de datos, un SCA divide dicha sucesión en numerosas partes comparables entre sí e investiga rasgos comunes, los cuales se van incorporando a un modelo que describe la sucesión de datos (o un sistema que los genere). Este modelo conviene que sea *entrenado* con un gran número de datos, con objeto de eliminar las probabilidades de que aparezcan fluctuaciones aleatorias. Se incrementa así la posibilidad de discriminar entre modelos diferentes, y se asegura que cada modelo corresponde a una clase particular de regularidades.

Sin embargo, la elaboración de modelos no está exenta de errores. Por un lado, podemos confundir componentes aleatorias dentro del flujo de datos con regularidades. Por otro, los rasgos que el SCA cree comunes son en realidad aleatorios, y no deberían incluirse en el modelo. Es razonable por tanto suponer que los SCA tenderían a evolucionar hacia una situación de relativo

equilibrio, en el cual la identificación correcta de algunas regularidades estaría acompañada de ambos tipos de confusión.

Podemos distinguir dos tipos de adaptación que no son exactamente un SCA:

· **Adaptación directa.** Da una respuesta directa a las señales del medio. Se detecta una señal de alarma que activa un mecanismo de defensa. Es lo que ocurre en un termostato: cuando la temperatura sobrepasa un cierto umbral, se desconecta el sistema de generación de calor.

·**Sistema experto.** La información suministrada por expertos humanos se introduce en un ordenador en forma de un modelo interno que sirva para interpretar datos de entrada. El modelo es fijo, a diferencia de los SCA, dónde el modelo es susceptible de variación: en los SCA existe aprendizaje, y en sistemas expertos no.

Sin embargo, si el modelo fijo que emplean los sistemas expertos, o bien la señal de alarma de los sistemas de adaptación directa, sufre modificaciones por los expertos humanos a la vista de los resultados, estamos ante un *SCA con humanos en el bucle*, o un *sistema de evolución dirigida*. El investigador que decide si un algoritmo es bueno o no, y lo va modificando, forma junto con el algoritmo un SCA. Este último puede a su vez englobarse dentro de otro SCA de orden superior, en el que el carácter de la dirección humana podría evolucionar a su vez, por ejemplo ante presiones externas como las de la necesidad de resultados para publicar una tesis.

FILTROS ADAPTATIVOS NO LINEALES

A la luz de lo dicho, podemos entender como el hecho fundamental de la adaptación la capacidad de emplear cierta información a priori para mejorar el comportamiento de un sistema. Según establecimos a la hora de hablar de la información contenida en las imágenes, se seguirá un esquema de información local, no estructural, en la que se usan vectores de muestreo, adquiridos de la imagen mediante una ventana deslizante. Por lo tanto, seguiremos aquí esa metodología, de manera que estudiaremos filtros adaptativos que empleen información no estructural. Dentro de esta categoría podemos distinguir de nuevo entre

filtros lineales y no lineales. Nuestro estudio se centrará en estos últimos, dado que el proceso físico que da lugar a la contaminación de ruido en la imagen es claramente no lineal. Por otro lado, los filtros adaptativos lineales presentan el inconveniente de precisar información previa para cada aplicación particular, perdiendo de este modo la generalidad que se les supone a los esquemas adaptativos. En definitiva, centraremos nuestra atención en los filtros adaptativos no lineales que empleen nociones de interresolución para obtener información con la que seleccionar y corregir las irregularidades de la imagen.

Podemos esquematizar el sistema general para el procesamiento de imágenes digitales adaptable a la señal de entrada como en la figura 11. En ésta aparecen dos bloques. Se aprecia cómo los niveles de gris de la imagen fuente S_{ij} actúan como señal de entrada, una vez que se ha introducido el ruido. El primer bloque puede corresponder a un filtro linal o no lineal que se va ajustando en cada momento a la señal de entrada S_{ij}. Dicho ajuste viene gobernado por un segundo bloque denominado *sistema de control*. Su cometido consiste en llevar a cabo la adaptación del filtro mediante algún criterio apropiado, basado en algunas características locales de la señal de entrada ruidosa S_{ij}, teniendo en cuenta alguna información *a priori* acerca de la señal sin ruido O_{ij}, proveniente de la imagen original O, y el ruido, basándose en el modelo de formación de ruido.

La información visual que adquiere un observador humano cuando percibe un hecho natural debe preservarse en la medida de lo posible al enfrentar al observador ante una *imagen* de dicho hecho natural. Los métodos de preprocesamiento local no adaptativos no emplean información a priori. Es difícil inferir esa información mientras se procesa pues la vecindad que se maneja suele ser pequeña, pero se puede emplear con gran provecho el conocimiento de propiedades muy generales acerca de la naturaleza de la degradación, como pueden ser los parámetros estadísticos del ruido. Sin embargo es obvio que además del conocimiento del sistema de adquisición y transmisión de imágenes (que determinan las características del ruido) hecho de poseer cierta información previa acerca de la propia imagen permitirá una corrección más eficiente de las degradaciones.

Parece que nos encontramos ante un círculo vicioso: para estimar los parámetros del ruido y la señal a través de la imagen

fuente hay que separar el ruido de la señal, pero eso sólo puede hacerse si los parámetros del ruido y señal son conocidos. Para salir de esta aparente contradicción sólo hay dos caminos:

- *Método determinista.* Se separan los píxeles afectados de ruido de los no distorsionados si se conocen a priori al menos las características del ruido o de la imagen original. Como se puede comprender, en la práctica este método está descartado.

- *Métodos estadísticos.* Podemos tratar el caso como un problema estadístico de estimación de parámetros. Entonces las características de la señal, de los píxeles de la imagen, son tratados como variables aleatorias, cuya realización es el valor observado de esa característica.

Por tanto el único modo de estimar automáticamente los parámetros de ruido y señal son los métodos estadísticos. Habría que considerar un problema añadido: la estimación estadística de parámetros precisa de un modelo estadístico acerca de las características en consideración, por lo que para cada modelo habría una estimación diferente. Afortunadamente, el ruido suele ser un objeto estadístico sencillo, esto es, descrito por muy pocos parámetros. Como vimos, en el caso de ruido gaussiano, basta dar la media y la desviación típica; si el ruido es impulsivo se describía mediante la probabilidad de aparición de un pulso. Luego el problema de la estimación de las características del ruido puede resolverse con herramientas relativamente simples, incluso si sólo disponemos de la imagen fuente. Sólo hay que elegir entre los parámetros medibles de la señal aquellos para las que las distorsiones inducidas por el ruido se maniefiesten como anomalías de comportamiento detectable de la manera más simple posible.

En el ámbito del procesamiento de imágenes, el conocimiento apriorístico más sencillo lo constituye la noción de *vecindad espacial*, entendida como el conjunto de píxeles que rodean a uno dado, y representada por su área (esto es, el número de píxeles que la constituyen). En efecto, los detalles en una imagen y las características primordiales de las regiones uniformes (textura, por ejemplo), suelen formar cúmulos, y no permanecer aislados. Por tanto, *sabemos* de antemano que los detalles de la imagen me aparecerán en grupos.

Esta información previa resulta a todas luces insuficiente. Pero es posible encontrar más fuentes de conocimiento a priori sin necesidad de profundizar mucho. Así, podemos advertir sin demasiada dificultad que los píxeles que pertenecen a un mismo detalle tienen niveles de gris muy similares. Este conocimiento puede añadirse de una manera natural al concepto de vecindad espacial, mediante la noción de *histograma*. Podemos entonces concluir que los píxeles que pertenecen a un mismo detalle forman cúmulos en el histograma. Esta información ya no es una trivialidad, sino que encierra una enorme significación, como se irá demostrando a lo largo de la presente tesis.

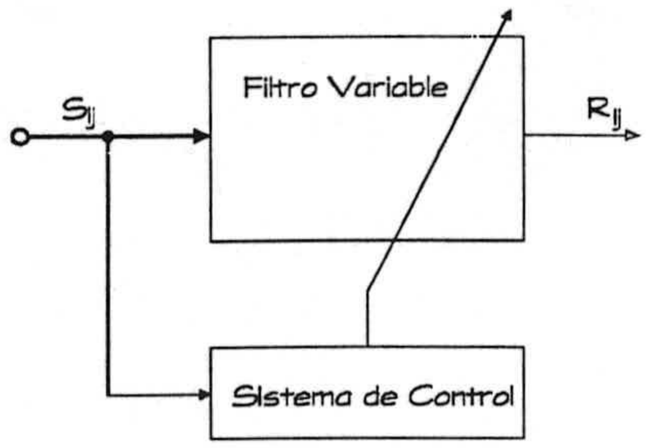


Figura 11: Esquema del funcionamiento genérico de uns filtro adaptativo.

Los filtros de rango (*rank filters*) y los filtros de estadísticos de orden (*order statistics filters*) ya presentados pueden proporcionar el primer ejemplo de filtro adaptativo, al menos localmente adaptativo, puesto que sus parámetros son función de su histograma local, esto es, de las características locales de la imagen. En primera instancia puede argumentarse que estos filtros en realidad no emplean ninguna relación espacial extrínseca entre los píxeles, porque funcionan reordenando los de una cierta vecindad en una distribución ordenada, y que es esta falta de información espacial su mayor defecto. Sin embargo, tal característica constituye tan sólo un aspecto más de su naturaleza adaptativa, y resulta ser más una ventaja que un inconveniente. En efecto, no se emplean relaciones espaciales extrínsecas, pero éstas se manifiestan intrínsecamente en la distribución ordenada a través de los parámetros de un histograma de distribución de valores de píxeles dentro de una vecindad, puesto que, como hemos señalado, los píxeles de un mismo detalle son vecinos también en el histograma. Y la ventaja viene de la mano del hecho de que la representación de relaciones espaciales en los histogramas locales es independiente de la orientación de los detalles, por lo que este tipo de filtros no precisa información previa sobre este particular, al contrario de lo que sucede en el diseño de algoritmos más sofisticados como los filtros lineales optimizados.

Sin embargo, para obtener mejores resultados es necesario otro tipo de conocimiento apriorístico. Por ejemplo, si tenemos un filtro iterativo para eliminar ruido emplearemos:

$$R_{ij} = F_n(F_{n-1}(F_{n-2}(\dots F_0(S_{ij}) \dots)))$$

dónde F es una operación de filtrado que se aplica a una vecin-

dad. Como las características del ruido van variando de una iteración a otra, la elección de F y del tamaño de la vecindad (y por tanto de la ventana) se establece mediante una información previa acerca del ruido (en nuestro caso, si sabemos de antemano que es ruido gaussiano blanco aditivo, se elegirá como filtro la media, que es el mejor estimador respecto al error cuadrático medio; si es ruido impulsivo, se usará la mediana, ...) y de la imagen (como el tamaño medio de los detalles, etc). En resumen, aunque la información acerca de la imagen queda muy bien representada por el empleo del histograma, sobre las propiedades del ruido no puede saberse nada en principio, por lo que ha de introducirse a priori, antes de la elección del tipo de filtro.

Este ejemplo constituye, por tanto, una muestra de un *SCA con evolución dirigida* pues es el usuario quien decide a la vista de los datos y de los resultados el tipo de modelo de ruido y de imagen que ha de emplearse. Sin embargo, estamos obligados a estimar las características estadísticas del ruido a partir de la imagen fuente. Sería deseable encontrar mecanismos que obrasen automáticamente de ese modo, lo que englobaría a estos filtros no lineales en un esquema claramente adaptativo, e independiente en gran medida de la intervención del usuario.

En esa dirección se encaminan los filtros adaptativos que emplean estimadores de varianza y media locales como manera de determinar las características estadísticas del ruido y la señal. Es lo que se denomina *estimación de escala* en estadística. Para ello simplemente aplican la definición de tales cantidades y la aplican a una ventana deslizante. La desviación típica insesgada constituye el estimador de escala de máxima verosimilitud para una distribución normal. Pero desgraciadamente, la mínima presencia de ruido impulsivo destruye su comportamiento —esto es, posee un probabilidad de colapso nula—. Por lo tanto se impone la necesidad de hallar estimadores de escala más robustos frente al ruido impulsivo. De este modo se han diseñado los llamados M-estimadores y L-estimadores, con similar fundamento a los M- y L-filtros que fueron expuestos anteriormente. Los M-estimadores tienen el inconveniente de que tienen una definición implícita, por lo que precisan de iteraciones para su resolución. Los L-estimadores son fáciles de implementar, pero son complejos, por lo que su cálculo es computacionalmente costoso. Afortunadamente existen estimadores más simples y rápidos, como el MAD (mediana de las desviaciones absolutas de la mediana) [PIT90]:

$$\sigma^* = 1.483\ MED\ \{\ |X_i - MED[X_j]|\ \}$$

Este estimador tiene excelentes propiedades de robustez.

Para realizar la estimación de parámetros se recurre a veces a métodos de minimización de alguna función de error. Por ejemplo, pueden escogerse los parámetros de tal manera que la diferencia entre la señal de entrada y la estimación hecha por el filtro se minimice. Si tal diferencia se estima mediante el error cuadrático medio, el método se denomina MMSE (*minimum mean square error*).

También existen los filtros que operan según el *modelo de dos componentes*. [PIT90]. Se basan en la hipótesis de que toda imagen está formada por una componente pasa-baja, relacionada con las regiones homogéneas; y otra pasa-alta vinculada a la presencia de contornos en la imagen. Cada componente puede procesarse de manera diferente, recombinando al final los resultados en una única imagen, en la cual el ruido se ha reducido considerablemente, e incluso pueden haberse realzado los detalles.

En resumen, podemos decir que los filtros adaptativos son en su mayoría de evolución dirigida, esto es, que varían su comportamiento según el valor de ciertos parámetros. No se han construido filtros que sean auténticamente adaptativos en el sentido de aprender a partir de filtrados previos —salvo, tal vez, en algoritmos que empleen redes neuronales, y que quedan fuera del propósito de este trabajo—. En cualquier caso, dentro de los filtros de evolución dirigida podemos distinguir entre los que precisan que los parámetros y umbrales se los suministre el usuario, y aquéllos que realizan una estimación automática de parámetros. Estos últimos son los que guardan cierto interés para completar el ámbito en el que se enmarca esta tesis. Pasemos a esbozar someramente las propiedades y características de algunos filtros adaptativos no lineales:

Filtros adaptativos de evolución dirigida.

Los filtros de evolución dirigida modifican su comportamiento según decisiones tomadas de acuerdo con los valores

que alcancen ciertos parámetros informativos. El propio usuario puede ser el encargado de modificar los parámetros, en cuyo caso el esquema adaptativo ha de extenderse hasta incluir a aquél; si bien los que nos interesan aquí son aquellos en los que el parámetro varía con la señal. Dentro de este ámbito los filtros suelen tener en cuenta información acerca de la presencia tanto de bordes como de ruido. Recordemos que la información que proporcionan los bordes en una imagen es muy importante para la percepción visual humana, y debe por lo tanto ser conservada al máximo. Como ya dijimos, cierto filtros, como el de media móvil, funcionan bien en regiones homogéneas pero yerra en los bordes, al contrario de lo que sucedía con otros filtros, como el de mediana. Parece apropiado el diseño de filtros que funcionen de manera diferente en los bordes que en las regiones uniformes de la imagen, para lo que se empleará un detector de bordes, que debe ser introducido previamente por el diseñador, para decidir si hay presente un borde o no. Asímismo se suele incluir un detector de ruido impulsivo. Cuando se encuentran pulsos, pueden ser eliminados mediante una estimación basada en la información proporcionada por los píxeles vecinos. La detección de la presencia de bordes puede emplearse para reducir el tamaño de la ventana de filtrado en la cercanía de los contornos para minimizar el efecto de distorsión que el filtrado provoca en los mismos.

Se han propuesto muchos detectores de borde, todos ellos con ventajas e inconvenientes, de acuerdo con el modelo matemático de borde que emplea, y de la metodología seguida. Puede encontrarse una excelente recopilación de los detectores de borde más ampliamente empleados en [PRAT78] [LEV85] [DAR93], aunque la comparación entre los diferentes métodos suele ser cualitativa, debido a la inexistencia de medidas cuantitativas globalmente aceptadas por la comunidad científica. Están basados en una estimación estadística que suele ser local, esto es, que emplean unas ventanas relativamente pequeñas para obtener las características del ruido y la señal. Una importante propiedad de estos estimadores es su robustez frente a ruido impulsivo, que aparece en muchas aplicaciones del procesamiento de imagen. Permítaseme realizar una somera descripción de los más empleados.

En primer lugar, puede detectarse un borde como la diferencia entre los estadísticos de orden máximo y mínimo (detector de rango). La idea es generalizable, pudiéndose diseñar de-

tectores de borde basados en la diferencia entre los resultados de dos filtros no lineales en lugar de los operadores min/max. Por ejemplo, se ha ensayado con ciertos OSF como los L_p y $CH_{\pm p}$, proporcionando detectores que solo encuentran contornos contrastados, pero la salida que dan como resultado posee menos ruido que la obtenida mediante detector de rango.

Otros detectores son esencialmente medidas de la varianza local de la imagen, dado que ésta adquiere valores pequeños en regiones homogéneas, y elevados en las zonas de borde. Empleando un umbral adecuado, este método conduce a resultados aceptables. Por extensión, se puede utilizar cualquier cuantificador de la dispersión de los niveles de gris en la imagen en sustitución de la varianza. Sin embargo, como dijimos, esta estrategia resulta poco robusta frente al ruido impulsivo, pues su presencia eleva el valor de la varianza, o cualquier medidor de dispersión, conduciendo a erróneas localizaciones de borde. El uso de estimadores robustos como el *MAD* pueden ser la solución.

Existe un enfoque diferente para estimar la presencia de contornos, basado en estimaciones locales de la desviación típica de la imagen, σ_x; y del ruido, σ_n. En zonas homogéneas ambas tienen valores similares, pero en la cercanía de un contorno $\sigma_x > \sigma_n$. De esta manera se mejora enormemente la robustez frente al ruido, sobre todo si se emplean estimadores de la desviación típica como el *MAD*. Pero a cambio se pierde conectividad en los bordes detectados.

También puede emplearse la estadística de otro modo con el mismo objetivo. Bovik y Munson [BOV86] detectan bordes midiendo diferencia de niveles de gris promedio entre dos ventanas adyacentes: si dicho valor excede el de cierto umbral, se considera borde. Algo parecido, aunque algo más sofisticado, realiza el detector de Prewitt; sin embargo es más susceptible al ruido.

En cualquier caso, para decidir la presencia o ausencia de borde es necesario un umbral, el cual ha de elegirse a partir del histograma de salida del detector. El umbral será grande si la supresión de ruido es más importante que la preservación de los contornos, y pequeño en caso contrario. Por ello a veces se emplean umbrales adaptativos, basados en el cociente entre las varianzas de ruido y señal.

En cuanto a la detección de ruido impulsivo en la imagen, puede también hacerse mediante el uso de funciones de decisión booleanas, trabajando con prueba de hipótesis; o mediante comparación con píxeles vecinos, como se estudiará más adelante.

Filtros de optimización.

En este tipo de filtros, la estimación del nivel de gris del píxel central se realiza a partir de sus píxeles vecinos de manera que el resultado minimice una determinada función de error (que suele ser el error cuadrático medio, MSE, en cuyo caso la minimización se denomina MMSE). Como el criterio define un valor promedio del error en una vecindad local, estos algoritmos son localmente adaptativos, y precisan ventana deslizante de forma y tamaño adecuados. Si se tiene una imagen fuente afectada de ruido aditivo:

$$S_{ij} = O_{ij} + N_{ij}$$

entonces la estimación que minimiza el MSE es:

$$R_{ij} = \underline{O}_{ij}$$

$$R_{ij} = \left(1 - \frac{\underline{\sigma}_n^2}{\underline{\sigma}_x^2}\right) \cdot S_{ij} + \frac{\underline{\sigma}_n^2}{\underline{\sigma}_x^2} \cdot \underline{\mu}_x$$

donde los símbolos subrayados representan valores *estimados*, σ_n es la desviación típica local del ruido; σ_x lo es de la señal; y μ es la media de la muestra. Si la región es homogénea, la varianza de ruido y señal es similar, por lo que se realizará un filtrado de media, que eliminará el ruido. Y si hay un borde en las cercanías, la varianza del ruido es mucho menor que la de la imagen, por lo que el filtro dejará inalterado el píxel central de la ventana, para evitar la distorsión del borde. Obsérvese cómo los filtros de optimización, como este MMSE, son de evolución dirigida.

El comportamiento frente al ruido depende de la elección de las medidas locales de la media y varianza de la imagen fuente, y de la varianza del ruido. Estos valores pueden conocerse mediante alguno de los estimadores locales ya discutidos. Tan sólo presenta problemas el conocimiento de la varianza del ruido, que sólo puede ser estimada mediante una gran ventana en una región homogénea de la imagen fuente, donde la varianza

de la señal y del ruido son similares. Es difícil evaluar la varianza del ruido si éste no se distribuye espacialmente de forma homogénea. Si la imagen fuente también está afectada por ruido impulsivo, éste ha de ser detectado y eliminado previamente.

Una variación del filtro de optimización con MMSE lo constituye el filtro adaptativo propuesto por Sun y Venetsanopoulos [SUN88], que será expuesto dentro del epígrafe dedicado a los filtros de ventana adaptativa.

Filtro de media recortada (α-TM) adaptativo.

Ya vimos que el filtro α-TM tiene un comportamiento híbrido entre la media y la mediana, por lo que está indicado para eliminar ruido gaussiano e impulsivo. Pero falla para ruidos con distribuciones con colas meedias o cortas. El filtro α-TM complementario —que emplea los valores deshechados en el filtro α-TM—, es precisamente el apropiado para ese tipo de degradaciones. Por lo tanto, podemos emplear un estimados de colas de distribuciones de ruido (como la curtosis, o $V_i(\alpha)$ [PIT90]), para decidir qué filtro emplear.

De este modo podemos construir un filtro de evolución dirigida, en el que según el tipo de distribución de ruido, podemos aplicar el filtro de media si es gaussiano; el de mediana si impulsivo; el punto medio si uniforme; y el α-TM o su complementario en los casos intermedios. El principal inconveniente de este filtro adaptativo es que es difícil determinar los umbrales entre una distribución y otra, que ha de hacerse empleando métodos de Monte Carlo.

Filtros racionales.

Este tipo de filtros [RAM95a] [RAM95b] intentan aprovechar las capacidades de los filtros lineales para eliminar ruido, y la de los no lineales para preservar bordes. Deben su nombre a que la salida del filtro depende de la entrada mediante una función racional. En una dimensión, si la entrada es $\boldsymbol{X} = (X_{n-1}, X_1, X_{n+1})$:

$$Y_n = \frac{X_{n-1} + X_{n+1}}{S(X) + 1/w} + X_n\left[1 - \frac{2}{S(X) + 1/w}\right]$$

donde $S(X)$ es un estimador de borde y w es el valor del umbral de dicho detector, decidido por el usuario . En el artículo original se escoge como $S(X)$ el operador laplaciano, aunque pueden emplearse otros. Alcanza valores elevados cerca de los contornos. De este modo, si $S(X)\ w^{-1}$ el valor de la salida tiende a la de un filtro lineal con coeficientes [w, 1-2w, w]; mientras que si $S(X)\ w^{-1}$ entonces $Y_n X_n$, y el filtro no actúa, preservando así los bordes. Vemos cómo el filtro opera como un sistema de evolución dirigida, según el valor de $S(X)$. Sin embargo, este algoritmo no es capaz de eliminar el ruido en una sola pasada, por lo que precisa de aplicación iterativa, aunque ello no merme en exceso a los bordes, que quedan preservados por el propio comportamiento del filtro. Además, el empleo de un parámetro de comportamiento desconocido como es el w supone una dificultad que hace que el filtro pierda generalidad de aplicación, y por tanto merma su carácter adaptativo.

Filtro de mediana modificado (MMF).

Entre las características del filtro de mediana está la de preservar la posición de los bordes. Eso no era del todo cierto cuando había ruido impulsivo presente. En efecto, si existe un pulso dentro de la ventana, el borde puede ser desplazado en un píxel. Si son dos lo spulsos presentes en la ventana, el desplazamiento puede ser hasta de dos píxeles. Sin embargo, la probabilidad de que ocurran i impulsos dentro de la ventana decrece a medida que i lo hace, por lo que la probabilidad de hallar dos pulsos en $W_{ij}(S)$ es mucho menor que la de hallar un pulso. De este modo, es más probable que ocurra un desplazamiento de un píxel por efecto del ruido impulsivo. El filtro de media modificado MMF [LEE90] se diseña para evitar este tipo de alteración de los contornos de la imagen. Para ello precisa detectar la presencia de borde, para lo cual emplea un método de prueba de hipótesis, que es un caso especial de medida de dispersión:

$$T = (\underline{X}_{m+1}) - (\underline{X}_{m-1}) \quad ,, \quad \text{Si } T > T_m \text{ (umbral)} \Rightarrow BORDE$$

donde $\underline{X}_m$ es la salida del filtro de mediana, y T_m se determina a partir de la pdf de T. En caso de que se detecte un borde, MMF reemplaza la salida de la mediana tradicional por otra más probable, para lo cual emplea otra prueba de hipótesis más sofisticada, basada en la comparación del valor del píxel central

con el de la salida de la mediana. Comparando con los umbrales adecuados es posible determinar si la salida de la mediana es un impulso, e incluso si se trata de un pulso blanco o negro.Entonces, en lugar de ese valor resultado de la mediana, se da como salida del filtro el estadístico de orden $\underline{X}_{m+1}$ ó $\underline{X}_{m-1}$ dependiendo del tipo de impulso. De este modo, sólo se modifica el valor de la mediana en la proximidad de los contornos, y sólo si la salida de dicho filtro va a resultar ser un pulso. Los resultados del MMF son levemente mejores que los del filtro de mediana para ruido impulsivo, si bien manifiesta un pobre comportamiento frente a ruido gaussiano.

Filtros de estadísticos de orden de evolución dirigida (DBOSF).

Para eliminar el pobre resultado de los MMF con ruido no impulsivo, el mismo autor [LEE90] diseña los DBOSF, fundamentados en un a hibridación de filtros tipo L con el MMF. La idea es eliminar los impulsos y luego realizar un filtrado L.

A partir de la salida del MMF, X_{ij}, podemos suponer con bastante aproximación que un determinado píxel X_{ij} de la ventana es un pulso, siempre que:

$$X_{ij} \notin [X_{MMF} - 3\sigma, X_{MMF}+3\sigma]$$

donde σ es la varianza local. Además, para evitar que la salida del filtro emplee valores pertenecientes a regiones por encima de un borde abrupto, se trabaja sólo con los píxeles de la ventana tales que pertenezcan al intervalo $[X_{MMF} - C, X_{MMF}+C]$, siendo $C = \min\{3\sigma, h/2\}$; y h la altura del borde abrupto que queremos preservar, y que es elegido por el usuario. De este modo, la salida del DBOSF es la media de los valores de los píxeles de la ventana dentro del intervalo anterior. Este filtro se ha mostrado efectivo con ruido impulsivo y gaussiano, e incluso refuerza algunos detalles. Sin embargo, la dependencia con el valor h lo hace demasiado *as hoc*, perdiendo en parte su finalidad adpatativa, que tal vez no justifique el costo computacional.

Filtros de ventana adaptativa.

Para que los píxeles dentro de una ventana de observación sean estadísticamente representativos, el tamaño de la ventana ha de determinarse correctamente. Si la ventana fuese demasiado grande, podrían aparecer píxeles pertenecientes a diferentes regiones de la imagen, provocando que el resultado del filtrado difumine los contornos y los detalles, que son los elementos que llevan la información de la imagen. Si, por el contrario, la ventana es demasiado pequeña, la muestra puede resultar demasiado escasa para generar la estadística que desemboque en el filtrado, por lo que los resultados pueden presentar una insuficiente supresión del ruido en las regiones homogéneas de la imagen. Por tanto, sería lógico emplear ventanas pequeñas si la señal presenta una fuerte variación; y grandes si se está ante regiones constantes.

Para llevar a cabo esa tarea, Song y Pearlman [SON88] emplean lo que denominan *índice de actividad de señal*, que no es otra cosa que un detector de borde basado en prueba de hipótesis, del mismo tipo que los empleados en el MMF ó DBOSF anteriormente citados. Para ello emplea el valor de la varianza local y de la varianza del ruido (estimada a partir de una ventana grande sobre una región homogénea de la imagen), pero teniendo en cuenta tan sólo los elementos que están en los bordes de la ventana, puesto que son éstos los que representan de manera efectiva las características de la vecindad si la ventana crece. Ese índice de actividad proporciona información acerca de la homogeneidad de la señal. Por lo tanto, existe un umbral a partir del cual el índice indica que la región es probablemente un borde, por lo que la ventana habrá de decrecer (y en caso contrario, aumentar). Ese umbral es adaptativo, dependiente del tamaño de la ventana actual, de modo que si ésta es grande, la siguiente será menor. De este modo se garantiza una convergencia de tamaño de ventana.

Para evitar que la región interna de la ventana contenga pulsos de ruido cuando ésta toca un borde, se emplea un condicionante más para permitir el decrecimiento de la ventana: ello sucederá si y sólo si la ventana anterior fue más pequeña, y el índice de actividad supera el umbral. Este esquema puede aplicarse a diferentes tipos de algoritmos de filtrado. De este modo, si se emplea filtro de media se tiene un filtro de media adaptativo, y lo mismo con el de mediana ... También se puede

emplear un método de optimización de filtro, como la minimización del MSE, aunque sólo diferirá del de media adaptativo en que se da cierto énfasis relativo a los datos vecinos al píxel que se filtra.

Hay que señalar que el umbral depende de un parámetro de diseño, que hay que determinar *a priori*. Los resultados para ruido aditivo gaussiano son aceptables; sin embargo, aunque en el artículo citado se hace hincapié en la importancia de eliminar el ruido impulsivo, no se realiza ningún experimento que muestre la eficiencia del método propuesto frente a este tipo de contaminación de la imagen, lo cual lleva a sospechar que el comportamiento no es demasiado bueno.

El proceso de filtrado puede plantearse desde otro punto de vista. Bernstein [BER87] propone utilizar la interpretación estructural de la imagen, considerándola compuesta de regiones homogéneas y bordes. Mientras las primeras constituyen elementos de bajas frecuencias, los últimos lo son de altas frecuencias. Se pueden procesar cada componente por separado (lo cual permite realizar un realce de detalles si se considera oportuno), con objeto de no dañar los contornos de la imagen. Para separar ambas componentes se emplean algoritmos pasa banda. En particular, Bernstein emplea la media para determinar la componente de baja frecuencia, mientra que la diferencia de la señal y esta última estimación constituye la componente de alta frecuencia. De este modo, para una imagen fuente genérica contaminada de mezcla de ruido aditivo y dependiente de señal, se empleará un filtro adaptativo basado en el modelo de dos componentes. Para ello se utilizan los algoritmos pasa banda antedichos, realzados con ciertos pesos normalizados, calculados teóricamente de modo que minimicen la diferencia (en términos del error cuadrático medio, MSE) entre la señal de salida del filtro y la imagen fuente, y que surgen en un intento de generalizar una analogía con los filtros de Wiener:

$$Y_{ij} = \langle X_{ij} \rangle + b_{ij}(X_{ij} - \langle X_{ij} \rangle)$$

Dichos pesos, $b_{ij.}$ que dependen del cociente entre las varianzas locales del ruido y la señal, son tales que se anulan si la región es homogénea (la varianza del ruido es menor que la de la señal), con lo que se filtra tan solo la componente de altas frecuencias. Si ambas varianzas son similares, los pesos tienden a la unidad, y el filtro se comporta como un pasa alta. Existen además dos

parámetros ajustables de los que depende el valor asignado a los pesos. Uno de ellos controla el umbral a partir del cual se suprime la componente de alta frecuencia; mientras que el otro gobierna la supresión de ruido cerca de los bordes.

Los pesos, que como queda de manifiesto proporcionan información acerca de la presencia de bordes en la image, pueden ser empleados con otra finalidad: la de adaptar la ventana del filtro. Podemos comenzar el filtrado empleando un tamaño inicial 5×5 o 7×7, y si b_{ij} es suficientemente grande, la ventana decrece hasta que baja el valor del peso, o se llega a la ventana mínima de 3×3. En caso contrario, la ventana crece hasta alcanzar su máximo tamaño. Aun quedaría el problema del ruido impulsivo, que falsearía los valores de $b_{ij.}$ Para evitarlo, se detecta y elimina dicho ruido, de manera que los valores de medias y varianzas locales se tomarían sobre los elementos limpios de la imagen. Bernstein propone un detector de pulsos basado en hipótesis de decisión sobre la media y varianza de la señal y ruido frente a valores de los pulsos, lo cual añade un nuevo parámetro empírico. Siguiendo el esuqema expuesto se construye el SAM (*signal adaptive median filter*) de Bernstein. Su comportemiento frente a bordes y ruido es bueno, pero el coste computacional es enorme.

En un intento de extender la potente base matemática de los filtros lineales al ámbito de la no linealidad se diseñaron los filtros L adaptativos. En éstos, los coeficientes que multiplican a los estadísticos de orden se adaptan a la señal mediante una minimización del error cuadrático medio (MMSE, *minimum mean square error*). Para resolver las ecuaciones conducentes a esa optimización se pueden emplear métodos de mínimos cuadrados o iterativos. Aunque exigen cierto tiempo de cómputo para lograr una adpatación a la señal, ofrecen resultados positivos frente a diferentes tipos de ruido.

Más recientemente Hwang y Haddad [HWA95] han propuesto dos filtros de ventana adaptable. Por una parte, el RAMF (*ranked order adaptive median filter*), que emplea dos etapas. En la primera de ellas, se observa la salida del filtro de mediana, y se decide si es o no un impulso mediante operaciones booleanas, en función de los valores de la tabla de verdad de las mismas. Es tan sofisticado que permite discriminar un impulso de una señal real con idéntico nivel de gris. Si se decide que existen impulsos residuales a la salida del filtro de mediana, se aumenta la venta-

na, y se vuelve a realizar la comprobación. Sólo cuando se está seguro de que el filtro de mediana no proporcionará un valor espúreo, se sale de esta primera etapa, y se pasa a la segunda.

En esta última parte del proceso, se verifica si el píxel central es o no un pulso. En caso afirmativo, se filtra con la mediana tradicional, con la ventana que se determina en la primera etapa, cuya salida es con toda seguridad un valor de señal no impulsivo. Si el píxel central no es un pulso, se deja inalterado, y se avanza la ventana, reiterando el procedimiento completo.

El filtro RAMF puede ser mejorado por el SAMF (*impulse-size based adaptive median filter*), que puede ser empleado para ruido mezcla de impulsivo y gaussiano. La filosofía de este filtro se basa en cuestionarse sucesivamente acerca del tamaño del pulso, mediante pruebas de hipótesis. En primer lugar examina si el pulso es de tamaño unidad, en cuyo caso filtra con mediana estándar de tamaño 3×3. Si la respuesta es negativa, el algoritmo se pregunta si el pulso es de tamaño 2. En caso afirmativo, se filtra con mediana tradicional, primero con ventana 5×5 y a continuación con 3×3; y si no, se cuestiona acerca de si el tamaño del pulso es 3. De nuevo, una respuesta positiva lleva al filtrado de mediana tradicional con tamaños sucesivos de 7×7, 5×5, y 3×3; mientras que una negativa llevaría a una nueva pregunta acerca del tamaño del pulso. Procediendo de esta manera, y fijando un tamaño máximo de pulso posible, se llega a la eliminación completa del ruido impulsivo, mediante el empleo de filtrado de mediana anidado. Una vez suprimido ese tipo de distorsión, se procede al procesamiento con un filtro de media para eliminar ruido gaussiano.

Obsérvese que estos dos últimos métodos permiten eliminar ruido con casi total seguridad, evitando además filtrar los píxeles no afectados por ruido, eliminando las distorsiones introducidas por el algoritmo de mediana tradicional. Sin embargo, su comportamiento en la vecindad de bordes y contornos puede ser defectuoso, pues no tiene en cuenta su presencia a la hora de aumentar la ventana, que puede ser demasiado grande como para mantener las cualidades óptimas de preservación de borde de los filtros de mediana.

Por último, exponemos los filtros adaptativos propuestos por Sun y Venetsanopoulos [SUN88], que abundan en el empleo de optimización mediante MMSE pero teniendo en cuenta

el tipo de ruido e introduciendo un elemento novedoso: una ventana deslizante, diferente de la ventana de observación, para evaluar la desviación típica del ruido. Se emplean estimadores locales, calculados sobre la ventana de observación, para aproximar la esperanza ($\underline{E}_x$) y y varianza ($\underline{V}_x$) de la señal necesarios para resolver la ecuación de optimización del MSE. Entonces se realimenta la entrada con la salida del filtro, y se emplea la diferencia entre entrada y salida como estimación del ruido. La ventana de ruido se desplaza sobre la secuencia de estimaciones de ruido, y a través de ella se calculan localmente la media y varianza del ruido ($\underline{E}_n$ y $\underline{V}_n$ respectivamente). Existe un estimador de media ($\underline{E}_x$, $\underline{E}_n$) para cada tipo de ruido.

El esquema que proporciona los mejores resultados en supresión de ruido y preservación de bordes es el siguiente:

$$\begin{aligned} Y_{ij} &= (1 - a)\, X_{ij} + a\, \underline{E}_x && si\ a \geq c \\ Y_{ij} &= MED(X_{ij}) && si\ a < c \end{aligned}$$

donde los X_{ij} pertenecen a la ventana de observación, y los parámetros se determinan como:

$$a = \begin{cases} b + \dfrac{\underline{V}_n}{\underline{V}_x} & \text{si } b + \dfrac{\underline{V}_n}{\underline{V}_x} \leq 1 \\ 1 & \text{si } b + \dfrac{\underline{V}_n}{\underline{V}_x} > 1 \end{cases}$$

$$b \in [0.3, 0.5]$$

$$c \in [0.5, 0.6]$$

siendo los valores aconsejados $b = 0.32$ y $c = 0.58$. Los resultados tanto en supresión de ruidos aditivos y multiplicativos son satisfactorios, con una buena preservación de bordes.

FUNDAMENTO DE LOS FILTROS PROPUESTOS.

De lo dicho hasta el momento parece deducirse que los algoritmos de procesamiento más eficientes en cuanto a aplicabilidad son los adaptativos. Además, el filtro ha de preservar los contornos y detalles de la imagen, que son los elementos de la misma que más información proporcionan. Y, por supuesto, el procesamiento ha de dar lugar a una imagen resultado carente del ruido que pudiera afectar a la imagen fuente. El tipo de ruido más molesto y evidente es el impulsivo, y es además el más difícil de erradicar. Los filtros tipo mediana lo hacen aceptablemente bien, pero a costa de difuminar la imagen.

El propósito que se pretende con la presente tesis es el diseñar algoritmos de filtrado no lineales que se basen en algún tipo de información ariorística que los dote de cierto carácter adaptativo, y cuya finalidad básica sea la eliminación de ruido impulsivo y el respeto de los contornos y detalles de la imagen. Para ello se desarrollará una metodología tendente a la detección de los impulsos que habrán de ser procesados, de modo que el resto de los píxeles de la imagen, no contaminados por el ruido, permanecerán inalterados. Este modo de operar no sólo provoca un filtrado más rápido, dino que permite mantener los elementos de la imagen casi intactos, pues sólo se procesa lo que realmente precisa arreglo. A continuación, se desarrollará un sistema de filtrado adecuado al tipo de ruido en cuestión.

Evidentemente, el mecanismo de detección de impulsos precisa de algún conocimiento *a priori*, al igual que la mayor parte de los sistemas adaptativos. Esa información previa será obtenida en nuestro caso empleando la noción ya introducida de interresolución. La idea consiste en estimar los niveles de gris de los píxeles (máxima resolución) de la imagen a partir del conocimiento de los niveles de gris de las ventanas contenidas en la imagen (mínima resolución). Mediante un entrenamiento adecuado, es posible construir modelos que representen de forma fiable la probabilidad de aparición de un determinado nivel de gris k en una ventana de nivel de gris k'. Del modo de construcción de tales modelos, también denominados *matrices de*

composición, y su empleo en el procesamiento de imágenes digitales se hablará en un posterior capítulo, pero a continuación se expondrán las características básicas que poseen estos modelos, y la filosofía subyacente en ellos, que permitirá su posterior aprovechamiento en tareas de filtrado.

INTERDEPENDENCIA DE HISTOGRAMAS A DISTINTAS RESOLUCIONES.

Como ha sido indicado, los histogramas locales de una imagen contienen información relevante que puede ser utilizada por un algoritmo inteligente para procesar adecuadamente dicha imagen. Pero la información disponible era diferente según el grado de resolución empleado al observarla, por lo que en principio puede resultar problemático elegir acertadamente la resolución precisa para lograr el mejor procesamiento. Sin embargo, esta dependencia entre información y resolución puede ser manejada ventajosamente.

Consideremos una imagen, o una porción de imagen, observada a un nivel m' de la cual calculamos su histograma a dicho nivel, $\mathcal{P}[I]_{m'}$. Si se quiere conocer su histograma $\mathcal{P}[I]_m$ a una resolución más fina $m < m'$ hay que tener en cuenta una restricción de compatibilidad entre ambos histogramas, puesto que ha de conservarse el nivel de gris global de la imagen (o porción): sean cuales sean los tamaños escogidos para las regiones de observación, el número total de *dots* elementales negros es siempre el mismo. Esta restricción es necesaria pero no suficiente. No existe una relación unívoca entre los histogramas, tan sólo ciertas ligaduras. Este marco de posibilidades origina el estudio de los modelos lineales de dependencia, o *matrices de composición* [GAR94] [ROM91] [MAR90] [MAR95] [ROM??].

MATRICES DE COMPOSICIÓN.

El objetivo que ahora perseguimos es conocer la relación entre histogramas a diferentes resoluciones. Para una imagen I de tamaño $V \times H$ dicha interdependencia se expresará como:

$$\mathcal{P}[I]_m = Q_{mm'}(I) \cdot \mathcal{P}[I]_{m'}$$

donde los histogramas son vectores, y $Q_{mm'}(I)$ es una matriz. Para determinar de manera más explícita esta relación entre

histogramas a diferente resolución, esto es, para calcular los valores de las columnas de esa matriz, se contabilizan el número de vecindades a mayor resolución m, con nivel de gris k que aparecen dentro de la vecindad a menor resolución m' con nivel de gris k', y se computa la suma de los histogramas de las vecindades a nivel m' que tengan nivel de gris k'. A continuación se divide por el número total de vecindades a nivel m que hay dentro de las regiones a nivel m' con nivel de gris k', independientemente del nivel de gris de la vecindad menor.

En el trascurso de la investigación, y como se ha apuntado, se fijó el nivel de resolución más fino en la escala del píxel, mientras el más burdo (menor resolución) se determinaba por el tamaño de la ventana deslizante. Por tanto, en nuestro caso, la proporción que acabamos de introducir se reduce a calcular el cociente entre:

- la suma de los histogramas de las ventanas con nivel de gris k'.

- el número total de ventanas con nivel de gris k'.

De este modo, para una imagen I de tamaño $V \times H$ dicha relación entre los histogramas a diferentes resoluciones, siendo siempre la de mayor resolución la de píxel ($m = 1$, por lo que no aparecerá en el subíndice), es:

$$\mathcal{P}[I] = Q(I)\,\mathcal{P}[I]_{m'}$$

Por lo tanto, en el caso que nos ocupa tenemos que:

$$Q(I) = \{Q_{k'}\}_{k'=0}^{G\cdot w}$$

donde la columna de la matriz es

$$Q_{k'} = \frac{\sum_{i=1}^{V}\sum_{j=1}^{H}\mathcal{P}\left[W_{ij}(I)\right]\delta\left(\sum\left[W_{ij}(I)\right]-k'\right)}{\sum_{i=1}^{V}\sum_{j=1}^{H}\delta\left(\sum\left[W_{ij}(I)\right]-k'\right)}$$

y $G\cdot w = G'$. A la matriz $Q(I)$ se le denomina *matriz de composición*, o modelo.

Consistencia de la matriz de composición.

Una matriz de composición es *consistente* si existe, o se puede sintetizar, al menos una imagen real cuyas $Q_{k'}$ sean iguales a las de la matriz. En este sentido se considera que todas las imágenes que verifican el modelo forman una *clase*, que por tanto está caracterizada por una misma matriz de composición que relacione sus histogramas.

Para un modelo consistente se deben cumplir ciertas propiedades. Por ejemplo, se ha mencionado ya que se debe preservar el nivel de gris, pero también habrá ciertas restricciones respecto a la existencia de vecindades o píxeles oscuros en regiones claras y viceversa. La condición necesaria y suficiente para la consistencia [MAR90] es que la matriz de composición cada columna $Q_{k'}$ sea una combinación lineal convexa de las distribuciones elementales asociadas a una composición fija para todas las vecindades a una resolución $m'>1$ dada.

Por lo tanto es posible distinguir una serie de modelos que serán elementales, y a partir de los cuales se formarán, por combinaciones lineales convexas, los demás. El estudio de tales modelos elementales es por tanto primordial.

MODELOS LINEALES ELEMENTALES

Los modelos o matrices de composición elementales son una solución matemáticamente asequible al problema de encontrar matrices de composición consistentes: consideramos que la composición de cada vecindad en píxeles es fija, y presuponemos una cierta forma en la composición. De esta manera podemos conseguir los siguientes modelos [MAR90] [GAR94] [MAR95], por ejemplo:

- *Modelo Alto Contraste*. Una hipótesis de alto contraste se puede plantear aceptando que los píxeles que componen una vecindad son de niveles de gris tan diferentes entre sí como sea posible.

- *Modelo de Bajo Contraste*. Planteamos ahora una hipótesis opuesta al caso anterior. Las ventanas de nivel de gris k' se componen de de píxeles cuyos niveles de gris son tan cercanos al valor central de la escala de gris de

píxel, $G/2$, como sea posible.

• *Modelo Conservativo*. Es una matriz de composición próxiam a la matriz identidad, por lo que supone una conservación de histogramas frente a cambios de resolución. Responde a la idea de ausencia de contraste en el interior de las regiones, es decir, supone la homogeneidad, en la medida de lo posible, de éstas.

Modelo Hipergeométrico. Los modelos elementales anteriores se basan en la exigencia de ciertas hipótesis, que implica una composición fija de ventanas en píxeles. Sin embargo, otro enfoque del problema puede conducir a buscar una solución más neutral que no imponga ninguna restricción a la composición de las ventanas. Si no se dispone de información sobre qué composiciones son las más probables, se acepta que en cualquier ventana de nivel de gris k' todas las distribuciones posibles de los valores de niveles de gris k de los píxeles son equiprobables (de nuevo el principio de razón insuficiente). La solución ahora para la matriz de composición se deduce de la llamada distribución hipergeométrica: cada columna de la matriz será una distribución geométrica

$$(Q_{k'})_k = \frac{\binom{k'}{k}\binom{G'-k'}{G-k}}{\binom{G'}{G}} = \frac{\binom{G}{k}\binom{G'-G}{k'-k}}{\binom{G'}{k'}}$$

Aplicando el modelo hipergeométrico a un histograma conocido a una determinada resolución m' resulta otro histograma a nivel más fino $m<m'$ que representa el límite de las frecuencias observadas en imágenes con el mismo histograma a resolución menor cuando sus tamaños tienden al infinito.

Se han elaborado representaciones gráficas de las matrices de composición [MAR90] [GAR94] [MAR95] que resultan esclarecedoras. En el eje de abscisas se representan los niveles de gris posibles al nivel de resolución de píxel. En el eje de ordenadas aparecen los niveles de gris de la ventana. En el eje OZ se muestran los valores estimados para las distribuciones.

Evidentemente, $Q(I)$ depende de los niveles de observación. Como a lo largo del resto de la presente memoria, se ha elegido como escala de grises a nivel de píxel $K = \{0,...,G=63\}$. En las representaciones que acompañan estas páginas, en las figuras 12 y 13, la ventana de observación se ha tomado con un tamaño 3×3, lo que implica un nivel de gris de ventana $k' \in K' = \{0,...,G'=567\}$. En estas gráficas se aprecia claramente una fuerte tendencia hacia un tipo determinado de composición de ventanas a resolución inferior. Para el modelo de alto contraste, las distribuciones tienen máximo peso en los valores extremos de la variable k. En el modelo de alto contraste el mayor peso se asigna al nivel central de la escala de píxel, mientras el modelo conservativo implica regiones homogéneas en su interior, asemejándose a la matriz identidad. El modelo hipergeométrico presenta, en cambio, distribuciones más suaves, menos abruptas en la asignación de probabilidades a los valores de k.

MODELOS EMPÍRICOS

Hasta este punto se han aportado varias soluciones, todas ellas teóricas, al problema de determinar las proporciones que constituyen la matriz de composición, de acuerdo con ciertos criterios referentes a la posible distribución de un nivel de gris k' en niveles de gris de píxel k. En este parágrafo se planteará la determinación empírica de los elememntos de la matriz de composición, partiendo sólo de las restricciones de consistencia. Para ello se aplicará estrictamente la definición de matriz de composición, obteniéndose cada columna como promedios **reales** calculadas a partir de una serie de ventanas de nivel k' tomadas de ciertas imágenes similares que constituirán la muestra de entrenamiento. El procedimiento es, en esencia, el siguiente:

· Se dispone de una o varias imágenes de muestra, similares entre sí, de las cuales se conoce su histograma a la resolución deseada, mayor que el nivel de píxel.

· Para cada ventana de tamaño consecuente con el nivel de resolución elegido, y de nivel de gris k' que aparezca en dicha muestra se determina su histograma local.

· Se promedian todos los histogramas locales de todas las ventanas de nivel de gris k' fijo obtenidos como se indicó en el anterior punto.

· El resultado de dicho promedio es la columna de la matriz de composición $Q_{k'}$.

· Se repite el proceso para todos los niveles de gris de ventana posibles, hasta completar la matriz de composición.

Más adelante se planteará el estudio de los factores que intervienen al realizar este promedio y su influencia en el resultado. Como se observa, el proceso de entrenamiento que da lugar a un modelo definitivo es una caracterísitca propia de los esquemas adaptativos. La matriz de composición así construida se empleará como información previa para detectar y corregir ruido impulsivo.

La matriz que interrelaciona los histogramas a diferentes resoluciones no queda prefijada inicialmente, como en el caso de los modelos lineales elementales, sino que es construida a partir de un conjunto de imágenes de muestra. De este modo, la elección de dicha muestra de entrenamiento determinará la forma de la matriz de composición.

El proceso de elaboración descrito es en sí conceptualmente simple, y asequible en su realización a un sólo ordenador. Un primer paso, por tanto, para la realización de modelos lineales empíricos y posterior procesado de imágenes es la elaboración de los programas y algoritmos de ordenador capaces de analizar la imagen a distintas resoluciones: localizar las vecindades a nivel m', calcular su nivel de gris k' y computar cuántos píxeles, elementos de mayor resolución, de cada nivel k, hay dentro de esas ventanas.

La consistencia del modelo está garantizada por la propia naturaleza del mecanismo de construcción: un modelo es consistente cuando existe, o se puede construir sistemáticamente, sl menos una imagen cuyos histogramas verifican ese modelo. Nuestras matrices de composición han sido obtenidas a partir de las imágenes.

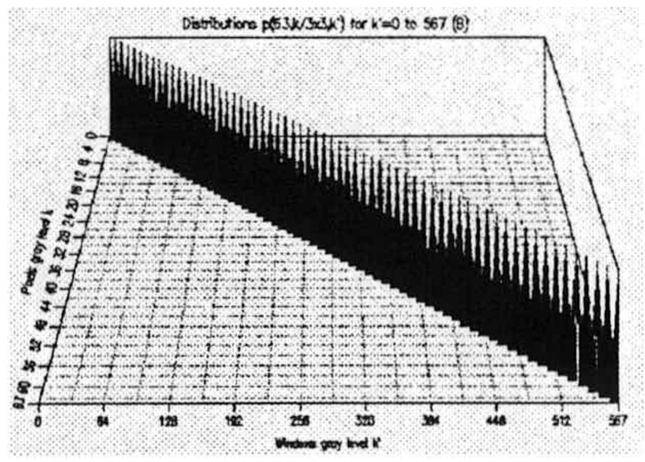
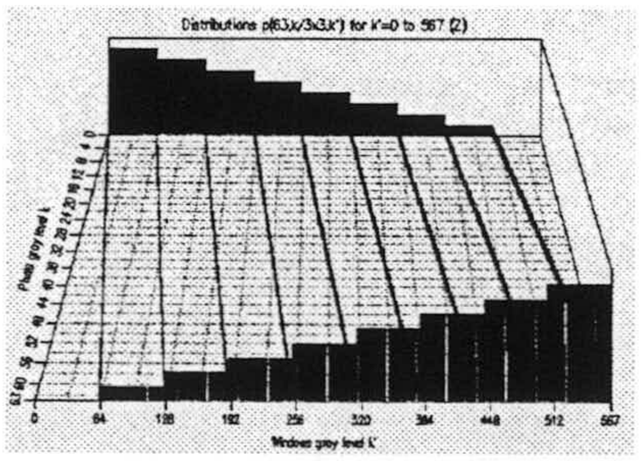
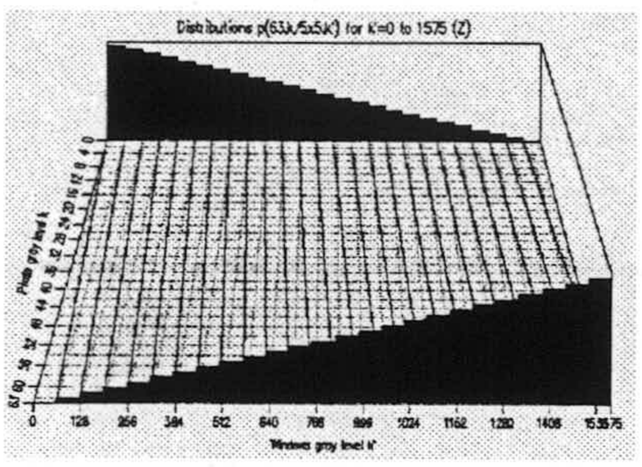


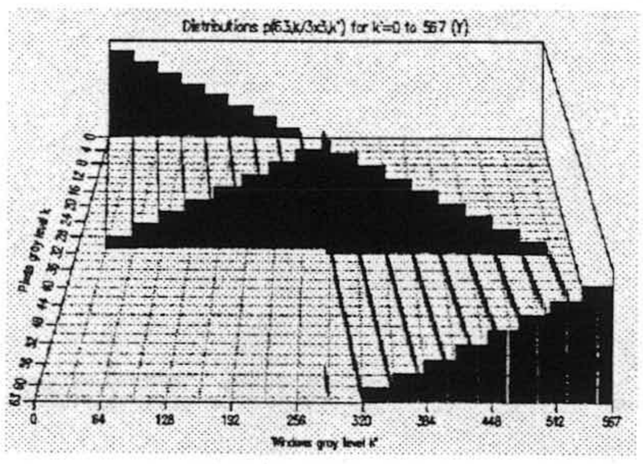
Figura 12: Las dos gráficas superiores corresponden al modelo conservativo (la superior con tamaño de ventana 3×3, y la otra 5×5). Por otra parte, las dos imágenes inferiores representan el modelo hipergeométrico con tamaño de ventana 3×3 y 5×5 (la que aparece sobre estas líneas).

EL PROBLEMA DE LAS COLUMNAS INDETERMINADAS.

En el ámbito de la construcción de modelos se presentan







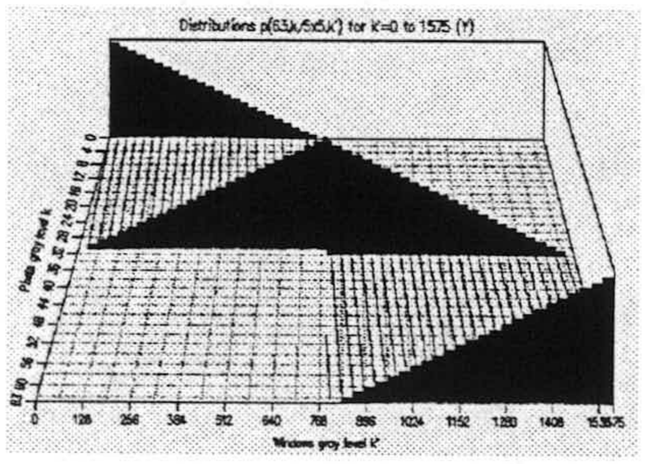


Figura 13: En la parte superior aparece el modelo de alto contraste para una ventana de tamaño 3×3. Inmediatamente bajo él se muestra el mismo modelo para tamaño de ventana 5×5. Las imágenes inferiores representan el modelo de bajo contraste para ventanas de tamaño 3×3 y 5×5 (la que aparece justo sobre estas líneas).

algunos problemas de naturaleza meramente técnica, pero que merece la pena detallar. Supongamos que se estudia un conjunto de imágenes reticulares o texturadas, es decir, formadas por repetición espacial de una celda elemental, como por ejemplo un tablero de ajedrez. Si la ventana deslizante tiene un tamaño múltiplo del de dicha celda, entonces se comprobará que todas las vecindades observadas tendrán el mismo nivel de gris, por lo que el histograma presentará un único pico correspondiente a dicho nivel. La matriz de composición se reduce a una sola columna. Todas las ventanas tienen el mismo nivel de gris k' y presentan la misma distribución a resolución de píxel, por lo que se promediarán regiones idénticas. La desviación de cada distribución respecto a la media es nula, y el modelo será «perfecto», en el sentido de que se especificará más adelante.

En el caso de las imágenes reales dicho efecto es prácticamente despreciable. Puesto que se efectúa un promedio estadístico a partir de una muestra determinada, en principio, y para garantizar la fiabilidad, es deseable que dicha muestra sea lo más extensa posible. Por tanto, no es fácil que se presenten probleams derivados de la periodicidad, multiplicidad o cambio de fase. Sin embargo no deja de ser ilustrativo de un hecho que sí suele suceder en las matrices de composición elaboradas con imágenes reales: la existencia de columnas vacías en el modelo.

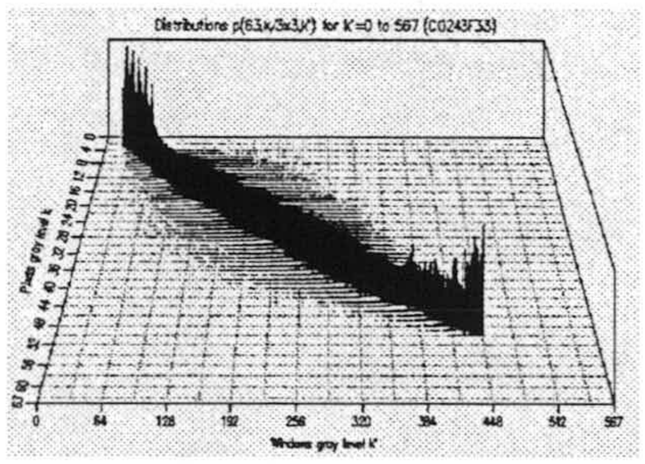
Hay ocasiones en que la muestra no es tan amplia como cabría desear, o simplemente se da el caso de que un determinado nivel de gris k' no aparece representado en la muestra. De esta manera sucede que para uno o varios niveles de gris de ventana no existe muestra, por lo que no aparecerá ninguna distribución en la matriz de composición. Esta situación se corresponde con el caso de indeterminación que existe en la definición de modelo dada, cuando el denominador se anula.

Como hemos señalado, la aparición de columnas vacías se puede producir por dos causas diferentes. Si la responsable de tales indeterminaciones es la escasez de muestra, sin duda puede resolverse el problema. La manera más sencilla es aumentar la muestra mediante la adición de nuevas imágenes al proceso de recuento ya citado que da origen a la matriz de composición. Si las columnas indeterminadas siguen apareciendo, aun cuando la muestra sea significativa, existen otras alternativas [GAR94], más o menos artificiales, encaminadas a dotar a la matriz de composición de toda la información admisible. Así, Maria José

García Salinas rechaza soluciones como la interpolación entre columnas vecinas debido a la carencia de significado físico del proceso, que implica suponer una continuidad que realmente no existe, aparte de las ambigüedades técnicas que aparecen. La autora supone que la primera muestra que se toma para realizar el modelo da lugar a un modelo base, que se irá enriqueciendo y completando a medida que se añaden sucesivas muestras. Si al concluir el entrenamiento con suficientes muestras siguen existiendo indeterminaciones en ciertos niveles *k'* entonces se puede añadir al modelo empírico un modelo teórico cuyas distribuciones sean más parecidas a las existentes. El grado de parecido se realizaría mediante una medida basada en la divergencia de Jensen-Shannon.

Sin embargo esta solución no deja de ser un intento de arreglo un tanto artificioso. Si bien puede argumentarse que se trata de una adición de información basada en la ya existente, no cabe duda que el exceso de información no proviene de la muestra de entrenamiento. Es del juicio del autor de esta tesis que, siempre que la muestra esté lo suficientemente entrenada, con una muestra estadísticamente significativa, la presencia de columnas indeterminadas implica la existencia de regiones de nivel de gris *k'* que *realmente* no aparecen en la muestra. por tanto, esa ausencia de regiones con ese nivel de gris constituye una fuente de información, y no una carencia de ella. Es por esta razón por la que se ha decidido mantener la presencia de columnas indeterminadas, sin intentar arreglarlas posteriormente de ninguna manera, siempre que el entrenamiento sea adecuado.

En principio no se puede intentar cambiar la imagen de entrenamiento modificando los parámetros de escaneado tales como el brillo y el contraste. Un aumento de brillo significa comprimir el histograma contra los niveles de gris superiores, mientras una variación del contraste supone una contracción o expansión del histograma en toda la escala de grises. Por tanto, un modelo empírico lleva implícito un ajuste estándar de brillo y contraste (también de resolución) tanto para las imágenes de entrenamiento como las imágenes que se han de procesar. Tales restricciones implican una pérdida de generalidad en el ámbito de la construcción y validez del modelo, el cual queda especificado para una clase concreta de imágenes, con unos determinados parámetros característicos. En realidad se está empleando un método usual de análisis, descomponiendo un problema en otros más sencillos o con menos variables. Agrupando previa-



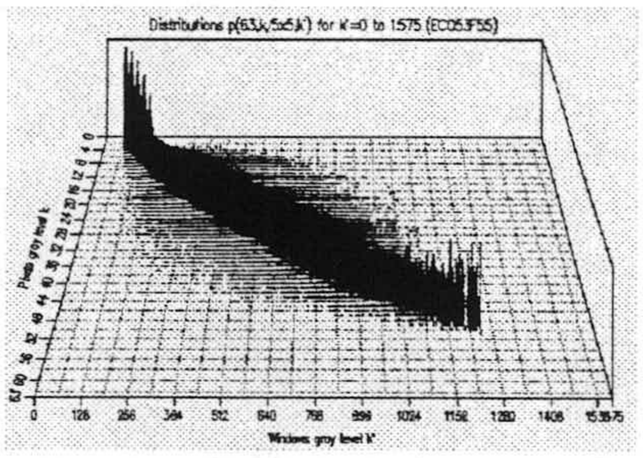


Figura 14: Modelo empírico de ecografías, para una ventana 3×3 (arriba) y 5×5. Obsérvese la presencia de columnas indeterminadas en las regiones extremas de la escala de observación, debido a la inexistencia de regiones demasiado oscuras y demasiado claras en las imágenes originales que sirvieron de entrenamiento.

mente las imágenes buscamos una especialización de los modelos. Además, este procedimiento queda avalado por la naturaleza de las aplicaciones del filtrado, en la que se trabaja con clases de imágenes muy particulares, y con un equipo informático concreto. En cualquier caso, la variación de brillo o contraste de las imágenes de entrenamiento da lugar a modelos ligeramente diferentes.

CLASES DE IMÁGENES.

La elaboración de una matriz de composición a partir de imágenes originales implica que dicho modelo representa a un determinado tipo de imágenes. Es decir, si el modelo se realiza a partir de un conjunto representativo de imágenes similares (por ejemplo, radiografías, textos, ecografías...), esa matriz de composición proporcionará información solamente para ese tipo de imágenes. Surge inmediatamente la necesidad de definir qué tipo de imágenes.

En principio, la definición matemática es evidente: una ***clase de imágenes*** la constituyen aquellas que tienen la misma matriz de composición, para las mismas resoluciones. Por tanto, solo en el caso excepcional de que para cada nivel de gris todas las regiones tengan la misma composición, podríamos referirnos a los elementos de nuestra matriz como los coeficientes del modelo lineal empírico. En este supuesto, las imágenes que originaron la matriz de composición y todas las que tengan el mismo juego de proporciones formarían claramente una clase. Pero en general esto no sucederá, pues cada ventana de nivel k' estará compuesta por un histograma particular, mientras el definitivo es un promedio de todos ellos.

Por tanto parece evidente que debemos trabajar con una definición menos restrictiva del concepto de clase. Lo más sencillo y lógico es no exigir que los coeficientes de la matriz de composición sean exactamente iguales, pero sí similares. Para valorar esta aproximación se siguen estos tres pasos:

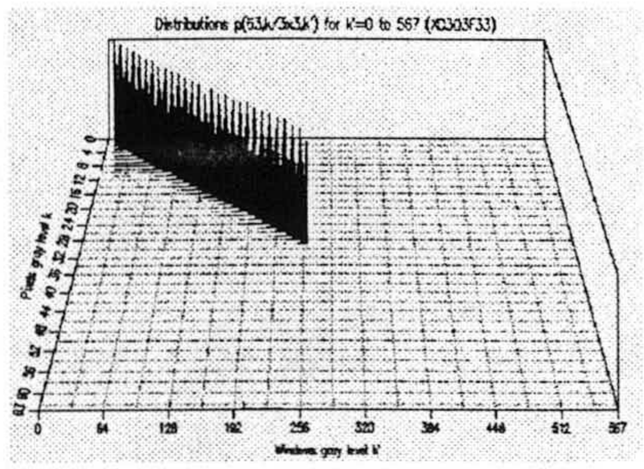
1) Antes de elaborar la matriz de composición se selecciona el conjunto de imágenes originales que formarán la muestra de entrenamiento de forma subjetiva, asegurando que todas tienen un cierto parecido, en el sentido de presentar el mismo tipo de objetos, niveles de gris,

etc, o bien proceder de un mismo mecanismo de obtención, o cualquier otra similitud formal. Por ejemplo, se pueden seleccionar radiografías, fotografías del planeta Marte, ecografías de vasos sanguíneos, etc.

Conviene mencionar aquí que la definición matemática no es fácil de visualizar. Matrices de composición con elementos muy similares pueden haberse obtenido con imágenes de aspectos muy diferentes, independientemente de cómo sean sus histogramas. Y también sucede el caso contrario, si tenemos imágenes semejantes en las que cambia apreciablemente el brillo o el contraste.

2) Una vez seleccionada la muestra, se procede al cómputo de la matriz de composición.

3) Posteriormente, se estudian y comparan diferentes modelos empíricos teniendo en consideración parámetros tales como resolución, tamaño de ventana, cantidad de muestra y tipo de imágenes con que se obtuvieron. El objetivo es asociar indefectiblemente un determinado modelo con una clase de imágenes concreta, para lo cual es preciso encontrar las posibles diferencias entre matrices de composición creadas a partir de diferentes tipos de imágenes.



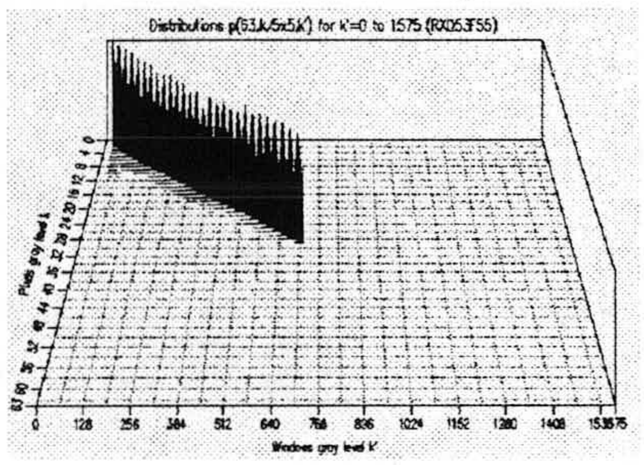
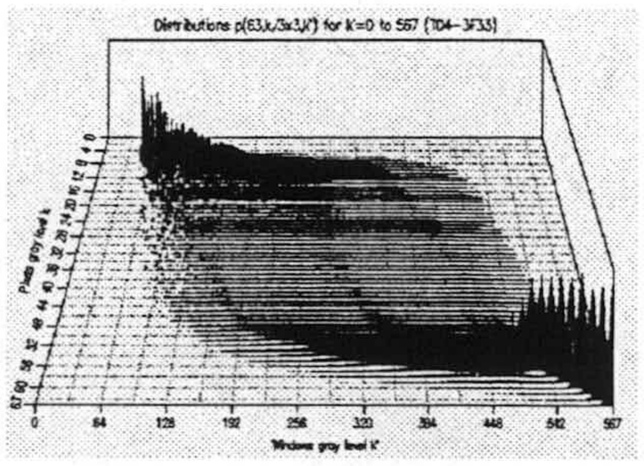


Figura 15: Se muestra el modelo empírico para imágenes radiográficas, tanto para ventana de tamaño 3×3 (arriba) como para tamaño 5×5. Obsérvese el gran parecido con el modelo conservativo. Se aprecian además el gran número de columnas indeterminadas, a pesar del numeroso entrenamiento, que presenta la matriz de composición, evidenciando la ausencia de ventanas de niveles de gris muy claros, como era de esperar en imágenes de esta clase.

Teniendo en cuenta los criterios esgrimidos en este parágrafo, si un conjunto de imágenes se ajusta razonablemente bien a una cierta matriz de composición, con indicadores de aceptabilidad elevados, dicho modelo representará a esa clase de imágenes. Para realizar los experimentos que fundamentan la presente tesis, se han escogido tres clases de imágenes representativas, ecografías, radiografías y textos mecanografiados. Las razones que han llevado a elegir estas clases son diversas. Por un lado, son fáciles de conseguir, y además cada una de ellas tiene características muy definidas, por lo que determinan la clase nítidamente. Por otro lado, el procesamiento de esas clases de imágenes es el objetivo de gran parte de la investigación que actualmente se lleva a cabo a nivel mundial en nuestro ámbito.

REPRESENTACIONES GRÁFICAS.



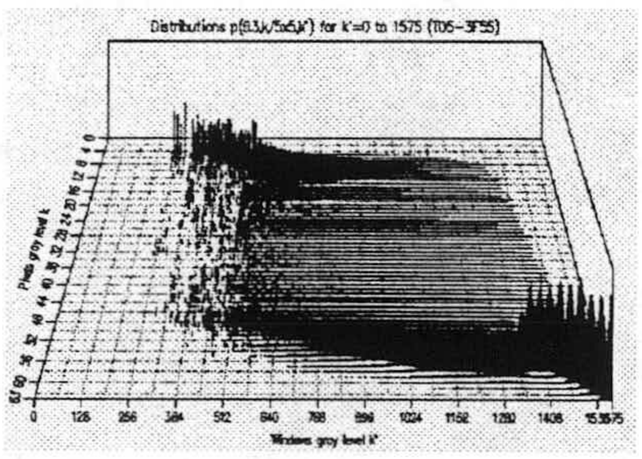


Figura 16: Modelo empírico de texto , realizado a partir de ventanas de tamaño 3×3 (arriba) y 5×5. La dispersión del modelo indica que una ventana de nivel de gris *k'* está compuesta de píxeles que poseen diversos niveles de gris *k*. Existen columnas indeterminadas en la zona menos oscura de la escala de observación.

Como ejemplo de lo dicho hasta el momento, podemos mostrar unas gráficas, análogas a las de los modelos teóricos antes expuestas, de ciertos modelos empíricos elaborados para el procesamiento de imágenes. Se han elegido tres clases de imágenes, que serán las que se utilicen a lo largo de todo el estudio que refleja la presente memoria. Se trata de imágenes de radiografías (figura 15), textos mecanografiados (figura 16) y ecografías (figura 14). Todas las imágenes de muestra pertenecientes a esas clases han sido escaneadas a una escala de 256 niveles de gris, aunque después fue desplazada a 64 niveles de gris con objeto de facilitar la tarea computacional. El tamaño de la ventana empleada, y que determina el nivel de resolución mínimo empleado, fue fijado en 3×3 y 5×5, por limitaciones en la memoria del ordenador.

Podemos observar que las matrices de composición representadas muestran una cierta regularidad indicativa del tipo de imágenes, pues se presentan distintas estructuras según cómo y con qué muestra se han elaborado, a pesar de que la consistencia de los modelos impone muchas restricciones. Existen otras características que pueden deducirse de esta representación gráfica. Así, por ejemplo, la matriz de composición de radiografías muestra columnas indeterminadas a partir de un cierto nivel de gris, demostrando que las imagenes de esa clase no poseen regiones claras, sino grises oscuros sobre fondo casi negro. Además, las distintas columnas de la matriz muestran ser distribuciones casi degeneradas, con valores nulos en casi todos sus elementos exceto unos pocos. Es un comportamiento muy parecido al del modelo teórico conservativo, que demuestra una cierta homogeneidad interresolución, puesto que cada ventana es homogénea en su interior, en el sentido de que están compuestas por píxeles que tienen casi todos el mismo nivel de gris. de esta interpretación se deduce que para los parámetros elegidos no se van a detectar formas o características especiales, y se podría adoptar el nivel de ventana como nivel de píxel.

Para el modelo de imágenes de texto las distribuciones son mucho menos degeneradas, más uniformes, indicando que las ventanas de nivel de gris *k'* pueden estar compuestas de píxeles de diferentes *k* de muchas maneras diferentes. No se aprecia similitud con ningún modelo teórico, y además las columnas indeterminadas son poco numerosas. Esto implica que la muestra contiene imágenes con zonas de grises muy variados, es decir, que las imágenes de texto no son todas oscuras, como las radio-

grafías, ni todas claras, sino que hay zonas claras y oscuras.

En cuanto a la matriz de composición de las ecografías, se observa un comportamiento intermedio entre los casos anteriores, que tampoco parece mostrar un claro parecido con ningún modelo teórico. Es de destacar la diferencia existente entre los modelos de una y otra clase de imágenes.

MEDIDAS DE CALIDAD DE LOS MODELOS.

Hasta el momento hemos visto cómo se construyen los modelos empíricos. Básicamente son el resultado de un *entrenamiento* a partir de una muestra determinada, siendo su principal objetivo el *procesamiento* de imágenes de la clase de las que constituyen la muestra. Por lo tanto, cabe preguntarse por la calidad de un modelo, en cuanto al resultado del procesamiento (calidad extrínseca) y en cuanto a calidad intrínseca, esto es, en cuanto a la calidad de la muestra. El caso de la calidad en relación al procesamiento se abordará en el momento en que se introduzcan los métodos de filtrado, por lo que por ahora no será abordado ese problema, aunque bien es cierto que cabe esperarse una cierta interrelación entre la calidad extrínseca y la intrínseca. Es decir, un modelo intrínsecamente bueno debe dar resultados de procesamiento buenos. Se pretende por tanto en este punto abordar la obtención de una medida cuantitativa de la calidad intrínseca de un modelo, para lo cual empleamos la única información disponible, que consiste en:

- histograma de la imagen o imágenes que conforman la muestra.

- histograma de ventana, para todos los niveles *k'* presentes en la muestra.

Procedamos a ello. Como se ha apuntado, la calidad intrínseca está relacionada con la calidad de la muestra de entrenamiento. Para que la matriz de composición construída a partir de ella sea de calidad, es deseable que las muestras sean a la vez *representativas* y *aleatorias*. Una muestra se dice *aleatoria* si cualquier grupo del mismo número de individuos que reúna los requisitos exigidos para pertenecer a la muestra, y que procedan de la misma población, tiene la misma probabilidad de ser seleccionado. En nuestro caso el sistema de recuento de regiones

está completamente determinado a nivel de algoritmo, y el recuento para un determinado nivel viene impuesto por la propia imagen en estudio. El único grado de subjetividad yace en la elección de las imágenes de entrenamiento. En cuanto a la *representatividad* sería deseable que en ningún caso se extrapolasen conclusiones extraídas del estudio de una muestra determinada a otras poblaciones cuya composición no esté representada por dicha muestra.

En cualquier caso, la representatividad de una muestra nunca llega a ser total, pues no se emplea toda la población. Los errores derivados de esta incompletitud se denominan errores de muestreo, que no han sido considerados en los estudios llevados a cabo sobre este tema. Sin embargo, sí se ha considerado la *confianza estadística* de los resultados. Este factor será el que determine si la muestra, y por extensión el modelo, tiene o no calidad.

Dicha confianza estadística dependerá del *tamaño* y de la *homogeneidad* de la muestra utilizada en la construcción de cada columna de la matriz de composición. En efecto, puesto que para cada nivel de gris de ventana *k'* se promedian los histogramas a nivel de píxel de todas las ventanas con ese nivel de gris que aparecen en la totalidad de la muestra, habrá de considerarse tanto el número de distribuciones promediadas como la similitud entre ellas.

Podemos intentar un tratamiento separado de la homogeneidad y del tamaño de la muestra, si bien ambos influirán de manera conjunta en la calidad del modelo, como quedará reflejado al proponer la medida de calidad final.

Homogeneidad de la muestra.

Si una columna es promedio de distribuciones muy similares entre sí, esto es, muy *cohesionadas*, la confianza en los resultados de un procesamiento será mayor si dicha columna era promedio de otras muy distintas. Siguiendo el criterio establecido en análisis inferencial, se podría estudiar la distribución muestral del estadístico media de los histogramas, que es una distribución normal con media igual a la media de la muestra y desviación típica igual a la de la muestra dividida por la raíz del número de elementos. Desgraciadamente, aparece un problema

técnico a la hora de calcular estas desviaciones típicas, debido a que los elementos de la población no son números sino distribuciones de probabilidad. Hay que establecer por tanto una medida alternativa de distancia entre distribuciones. Existen en la literatura un buen número de ellas [KUL68] [JOH79] [DAR93] [KIT86], aunque en nuestra investigación nos hemos decantado [GAR94] por su especial ínterés por la *divergencia de Jensen-Shannon* [LIN91] y que ya ha sido expuesta en una sección anterior. Por ser aplicable a cualquier número finito de distribuciones, establece por sí misma una medida de la inhomogeneidad de la muestra, y al construírse a partir de entropías de las distribuciones, proporciona en sí misma una interpretación física relacionada con la cantidad de información aportada.

Tamaño de la muestra.

Queda claro que no podemos emplear todas las imágenes pertenecientes a la misma clase para determinar su matriz de composición. Estos modelos se elaboran por tanto a partir de un subconjunto (muestra) de la población total. Dicha muestra es pequeña comparada con la totalidad de la población, por lo que si queremos extrapolar el significado de y utilización de una matriz de composición para todos los elementos de la clase tendríamos que hacer uso de la estadística inferencial y comprobar los parámetros y sus incertidumbres del conjunto a partir de los datos obtenidos de las muestras empleadas.

No se va a abordar el estudio estadístico de las técnicas de muestreo, pues no es el objetivo. Si no existe homogeneidad, una muestra con más elementos no siempre será representativa que otra con menos elementos, pero no detallaremos cómo. Por el contrario, nos interesa simplificar y homogeneizar la muestra al máximo, como hemos apuntado en el anterior epígrafe. Para no perder generalidad, sería interesante conocer qué cantidad de muestra puede ser suficiente para asegurar la representatividad de la misma. Esto equivale a reducir el error de muestreo. Se impone relacionar el tamaño de la muestra con el error de muestreo:

· Para muestras de tamaños iguales el error de muestreo es tanto menos probable cuanto mayor sea su valor.

· La probabilidad de que un error sea grande es

inversamente proporcional al tamaño de la muestra.

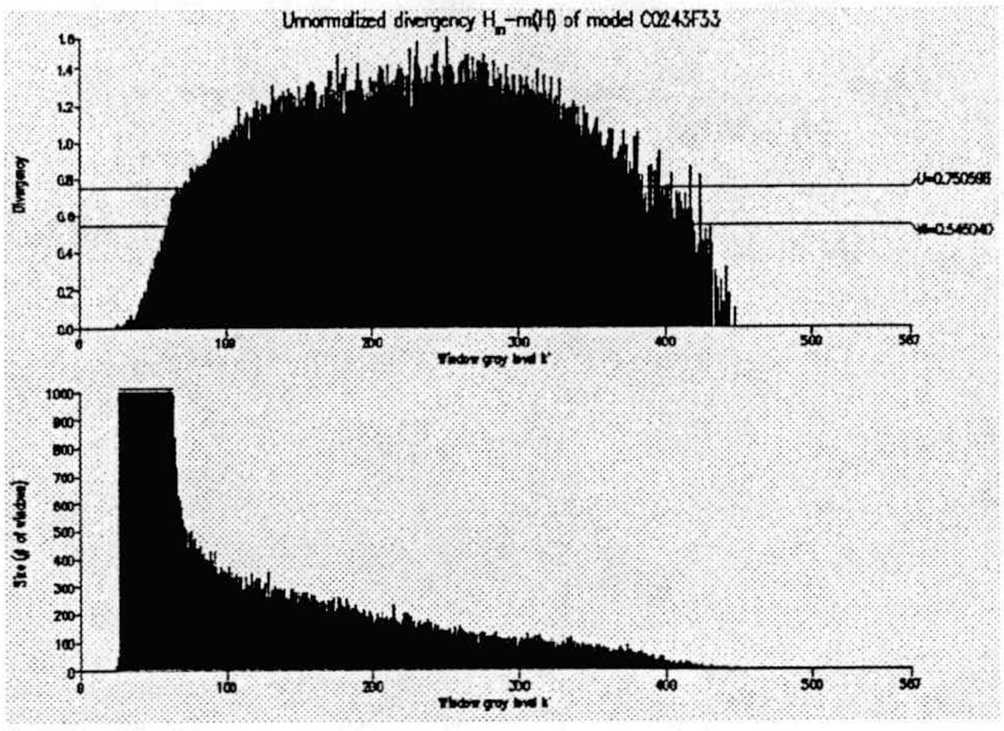
Los datos directos relativos al tamaño forman un vector de dimensión $G' + 1$ cuyos elementos son los tamaños de cada muestra de nivel de gris k' de la muestra de entrenamiento. Se trata del histograma a nivel de observación mayor m' de la muestra. Se puede derivar alguna medida de dicho histograma-muestra que deberá ser función creciente del tamaño. En general, para cada k' el número de muestras puede ser nulo, grande o pequeño. La distribución asociada puede ser uniforme, degenerada, etc. Todo ello debe reflejarse en la medida que se adopte. Cualquier opción acorde con lo anterior es válida. Sólo habría un problema al realizar la normalización, puesto que para cada matriz de composición, depende de la muestra escogida.

Se ha elegido de nuevo la divergencia de Jensen-Shannon para este objetivo. Se demuestra [GAR94] que la divergencia comienza fluctuando al crecer la muestra, pero alcanza un valor estacionario a partir de un tamaño determinado, que se denomina *tamaño límite* y que puede ser interpretado como el tamaño mínimo para que la muestra sea representativa. Este tamaño podría ser empleado para normalizar la medida, como se dijo anteriormente. Las fluctuaciones iniciales pueden explicarse debido al hecho de que al principio hay poca muestra, y se van adicionando más elementos. Si la muestra que se va añadiendo no es homogénea con la anterior, la divergencia crece, y si no, puede disminuir. Llega un momento en que la adición de muestra no implica mayor o menor homogeneidad, la muestra ha alcanzado su representatividad y la divergencia tiene un valor estable.

Empleando estos conceptos de cohesión y representatividad, Maria José García Salinas ha llegado a determinar un criterio de óptima divergencia para elegir los modelos base que, añadidos a uno empírico, permitían eliminar las columnas indeterminadas [GAR94]. Como ese punto de vista no es compartido en esta tesis, no profundizaremos en el tema, remitiendo al lector interesado a la tesina citada.

Características de las matrices de composición.

Se ha expuesto en este parágrafo la influencia de la muestra utilizada en la construcción de una matriz de composición



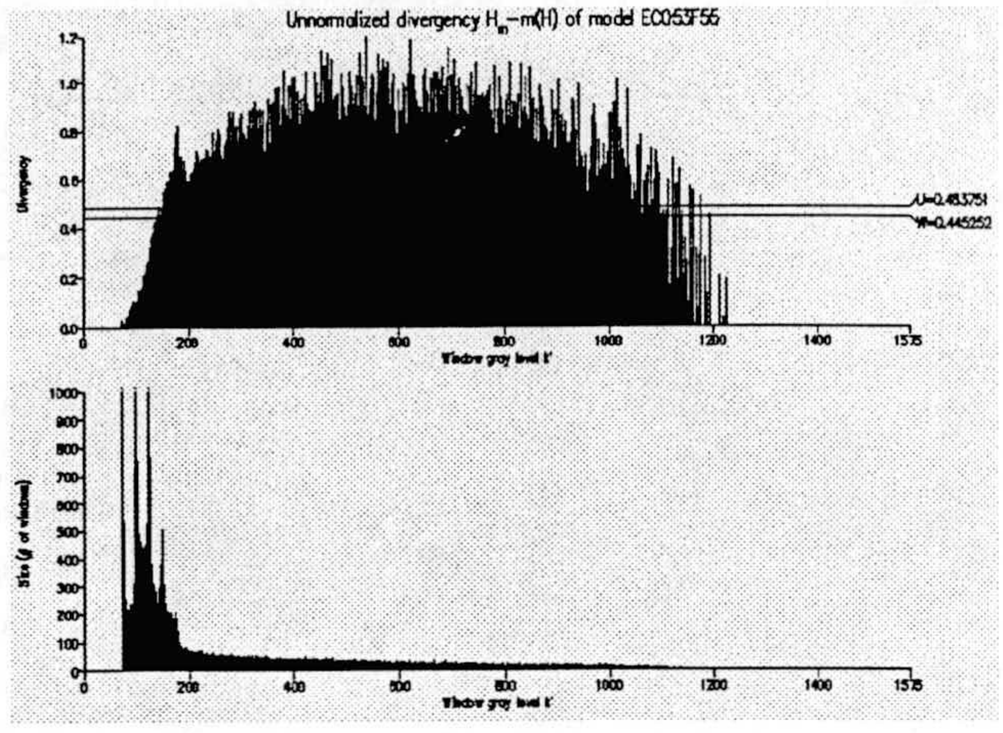
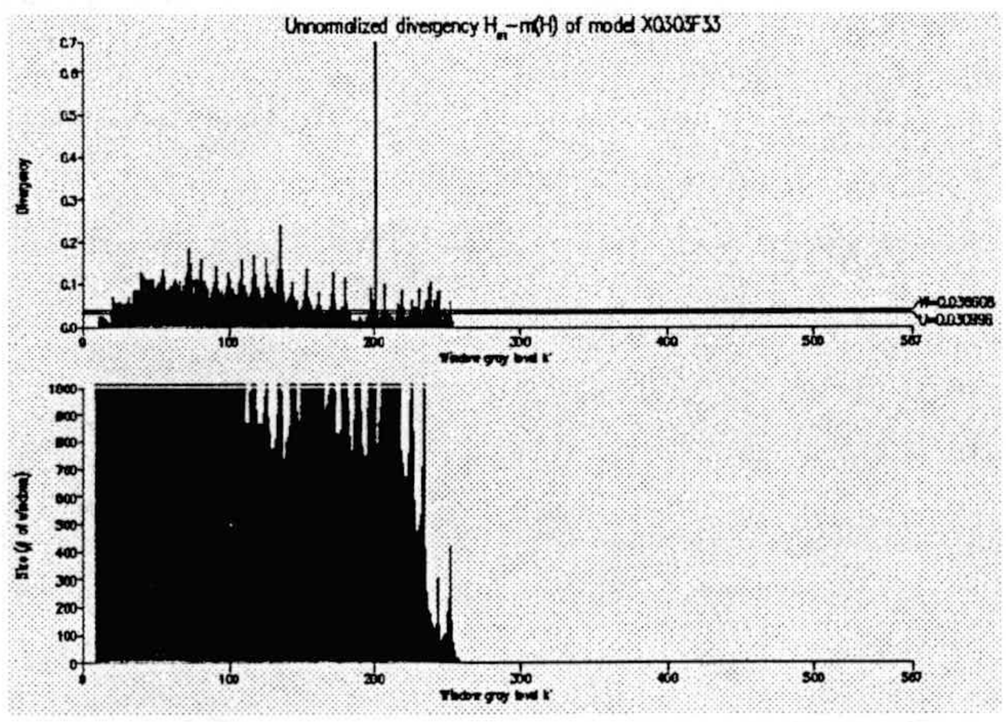


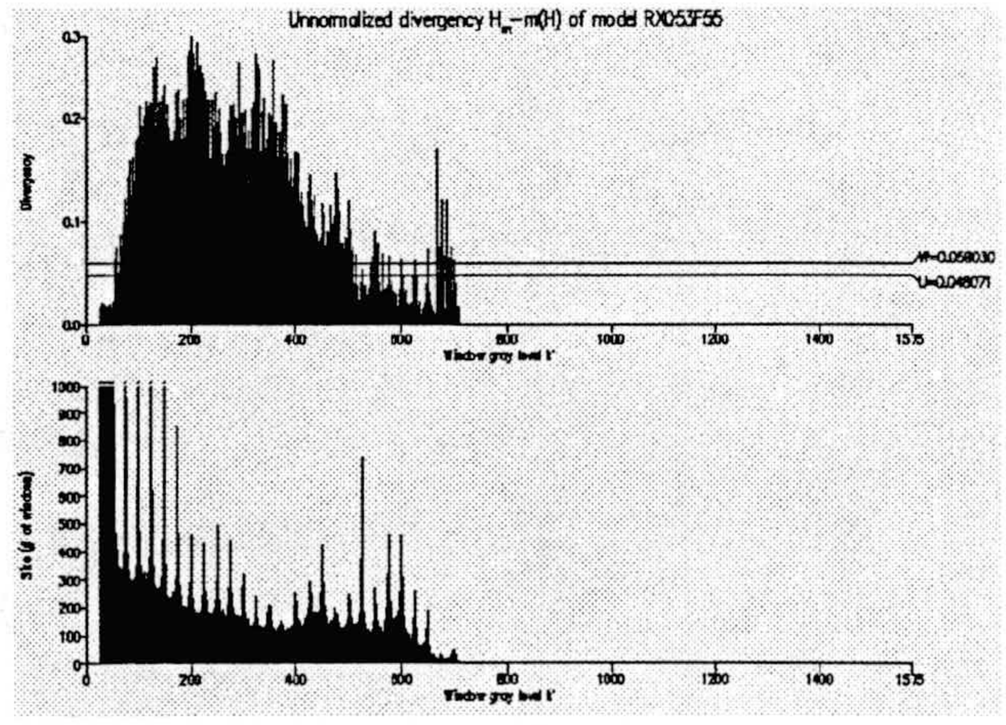
Figura 17: Divergencia y tamaño para cada k', en el caso de una matriz de composición de ecografías. A la derecha, con ventana de tamaño 5×5, y a la izquierda 3×3.

empírica. La principal exigencia para dicha muestra viene impuesta por el hecho de que el modelo se construye para un conjunto concreto de imágenes. Por tanto, las ventanas de gris k' han de tener histogramas semejantes a nivel de píxel. Para valorar esa similitud se ha introducido una importante herramienta, la divergencia de Jensen-Shannon. Esta medida podría ser también eficaz en la resolución de otros problemas como la clasificación de imágenes [GAR94]. Por otra parte, también se ha apuntado la influencia del tamaño de la muestra en la calidad intrínseca del modelo, y se ha determinado la existencia de un determinado tamaño mínimo representativo.

En la figura 17 hay una representación de la divergencia de la muestra para cada k' para el caso del modelo de ecografías, en la que se indica además la divergencia de Jensen-Shannon media (U) y ponderada (W) según el tamaño de cada muestra de nivel k'. El histograma a nivel menor de resolución m' de toda la muestra también se incluye, para muestra de nivel $k' < 1000$ por facilitar la comprensión de la gráfica.

Figura 18: Modelo de radiografías. Representación gráfica de la divergencis JS y el tamaño de la muestra para cada k'. A la izquierda se ha empleado un tamaño de ventana de 3×3. A la derecha, 5×5.





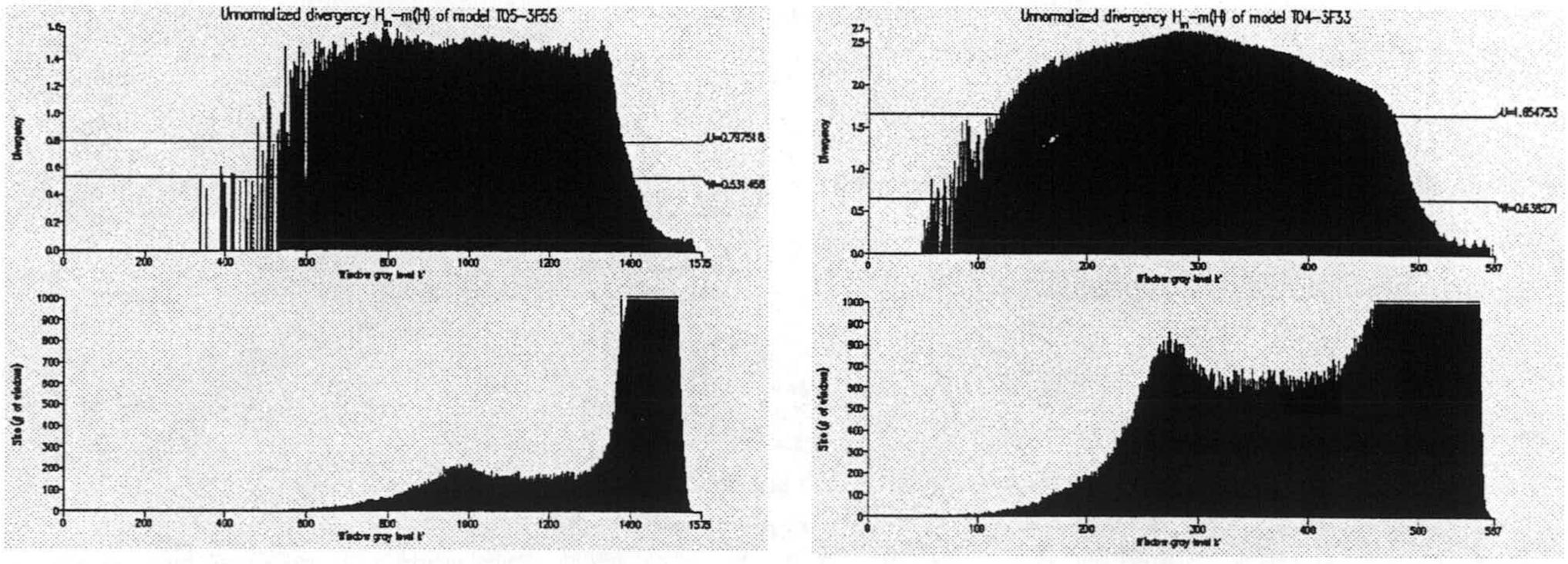


Figura 19: Se representan la divergencia JS y el tamaño de la muestra para un modelo de texto mecanografiado. A la derecha, se ha utilizado en la construcción de la matriz de composición una ventana de tamaño 5×5. A la derecha, 3×3.

La figura 19 contiene la misma información que la anterior, pero para una matriz de composisión de textos mecanografiados. Se aprecia cómo la divergencia adquiere valores mayores que en el caso anterior, y que la muestra es escasa o nula para valores de nivel de gris $k' < 100$. Predominan las regiones de tonos claros, aunque no de manera muy llamativa. En cualquier caso, no presentan una divergencia muy elevada.

En cuanto a la figura 18, análoga a las anteriores, pero representando un modelo de radiografías, podemos apreciar cómo predominan los colores oscuros, con divergencias muy bajas en todos los casos, indicando la gran cohesión interna del modelo.

Para apreciar la influencia de la resolución de observación, se han representado también en las figuras 14, 15, 16, 17, 18 y 19 los modelos de cada clase para ventana de 5×5.

CONCLUSIONES

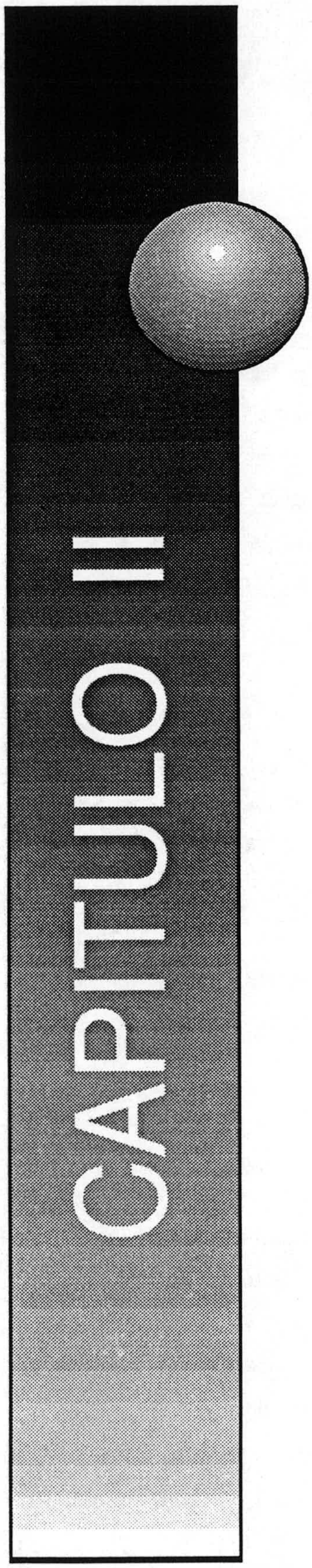
En el presente capítulo se han presentado los resultados básicos y las investigaciones ya realizadas que enmarcan adecuadamente el trabajo que se va a exponer a continuación. Aparte

de definiciones fundamentales, se han introducido de forma somera las bases del procesamiento digital de imágenes. Además, se ha realizado una descripción a vuelapluma de las características de los diferentes ruidos que pueden contaminar una imagen a lo largo del proceso de formación de la misma. También se ha hecho hincapié en el tipo y cantidad de información que una imagen puede proveer; y en la manera en que podemos emplear conocimientos previos acerca de la imagen para realizar un filtrado adecuado.

Es posible extraer ciertas conclusiones de todo lo dicho hasta el momento. En primer lugar, reconocer que el objetivo del procesamiento, y por ende del que aquí se propone, es la eliminación de las distorsiones de la imagen que conocemos por ruido. Sin embargo, no se debe olvidar lo que es casi tan importante como lo anterior: el respeto que hay que mantener por aquellos elementos de la imagen que proporcionan la información que lleva la imagen, como son bordes, contornos y detalles significativos. Es, pues, doble, la meta que debe alcanzar todo algoritmo de filtrado: por un lado, supresión de ruido; por el otro, preservación de bordes.

También se ha demostrado que el ruido más molesto y evidente que surge en la formación de una imagen digital es el impulsivo. Este tipo de ruido es filtrado de forma eficiente por algoritmos no lineales, y de entre ellos por los tipo mediana. Es por esta razón por lo que en el presente trabajo se centrarán los esfuerzos en eliminar el ruido impulsivo mediante métodos no lineales, de manera que se mejore en todo momento las capacidades de los filtros de mediana. Al elegir este esquema, habremos de describir los filtros determinista o estadísticamente..

Para ello se diseñarán filtros que detecten y filtren sólamente los píxeles de la imagen afectados realmente por ruido. Esa tarea habrá de ser llevada a la práctica empleando conocimientos apriorísticos acerca de la imagen. En ese sentido se han creado las matrices de composición, o modelos, representativas de una determinada clase de imágenes (radiografáis, textos ...), y que informan acerca de los niveles de gris más probables en las regiones de una imagen de la clase. Será en los siguientes capítulos donde se desarrollen los métodos fundamentados en estos modelos.



PRIMERAS MEJORAS AL FILTRO DE MEDIANA

III.1 INTRODUCCIÓN 135

III.2 EL FILTRO DE MEDIANA RELAJADA 138
III.2.1 Descripción 138
III.2.2 Propiedades estadísticas del filtro MR 139
Preliminares 139
Cálculo de la función de distribución de probabilidad de salida del filtro MR 143
Atenuación de ruido 149
Preservación de bordes 154
Preservación de líneas 160
III.2.3 Otras propiedades 163
Apilamiento o Stacking 163
Manchas o Blotches 165
Raíces 166
Características de las raíces 167
Respuesta a imágenes prueba 169
III.2.4 Respuesta a imágenes ruidosas 176

III.3 EL FILTRO DE MEDIANA DE VENTANA ADAPTATIVA.....185
III.3.1 Descripición.....185
III.3.2 Mecanismo adaptativo de supresión de ruido 186
III.3.3 Segmentación basada en la divergencia de Jensen-Shannon 187
III.3.4 Esquema adaptativo empleado.....189
III.3.5 Tipo de filtro empleado 190
III.3.6 Propiedades del filtro restdiver 193
III.3.7 Supresión de ruido y preservación de detalles.....194
III.3.8 Aplicaciones a imágenes ruidosas 196
III.4 CONCLUSIONES.....211

INTRODUCCIÓN

En el capítulo precedente, se han introducido de forma somera los métodos actuales de filtrado de imágenes digitales afectadas por ruido. Mediante un ajuste de sus parámetros, dependiendo de las características locales de la imagen de entrada, esos métodos de procesamiento eliminan el ruido aceptablemente, intentando siempre mantener inalterados los bordes que aparecen en la imagen. La mayoría de los filtros son muy útiles para la supresión del ruido gaussiano aditivo, tanto los lineales como los no lineales, y sin excesivo costo computacional. Pero, en general, son poco efectivos a la hora de eliminar el especialmente molesto ruido impulsivo, consistente en la aparición de pulsos de corta duración, y que se genera con demasiada frecuencia en el proceso de obtención de la imagen digital. Los filtros adaptativos, que se comportan tan bien frente a ruidos dependientes de la señal y gaussianos, no lo hacen tanto frente al ruido impulsivo.

Bien es cierto que hemos mostrado algunos algoritmos especialmente robustos frente a este tipo de ruido. De entre ellos destaca por su sencillez el filtro de mediana. Este método de procesamiento es particularmente efectivo en la supresión del ruido impulsivo, y preserva la posición de los bordes muy bien. Pero desgraciadamente tales mecanismos poseen molestos efectos colaterales. Aparecen manchasde gris uniforme (*blotches*), se satura el filtrado cuando el ruido supera cierto umbral (probabilidad de colapso, o *breakdown point*), y lo que es más importante, se difuminan, e incluso se eliminan, los bordes, líneas, esquinas y otros detalles de la imagen. A pesar de los esfuerzos y las múltiples variantes que de este tipo de filtros se han diseñado, y que en buena parte se han enumerado en el anterior capítulo, se está lejos de la solución definitiva.

Si tenemos en cuenta la enorme importancia de la información visual que contienen los bordes y detalles de una imagen, podemos comprender que sea un objetivo fundamental en el ámbito del procesamiento digital el encontrar algoritmos que eliminen el ruido impulsivo, pero respetando en todo momento la presencia de bordes y detalles. Este problema surge por ejemplo en visión artificial, teledetección, tomografía y ecografía médica, y en otros sistemas de imágenes de rayos X, donde la dosis de radiación debe ser minimizada. En estos casos la presencia de bordes agudos y estrechas líneas en diferentes orientaciones es una característica de especial importancia por la información que proporcionan. Tal objetivo es el fin último de la presente tesis doctoral.

A pesar de sus defectos, los filtros tipo mediana poseen propiedades que los hacen especialmente indicados para procesamiento de imágenes. Es por ello por lo que los primeros intentos de diseño de filtros que realizamos fueran encaminados a mejorar las características del filtro de mediana. En este capítulo se introducen dos nuevos algoritmos de filtrado. El primero de ellos se denomina *filtro de mediana relajado* (filtros RM), que se ha basado en una idea muy simple y que generaliza, de alguna manera, los filtros de mediana (y por ende los filtros de estadísticos de orden). El otro de los métodos propuestos se llama *restdiver*, y es un algoritmo daptativo de mediana que ajusta el tamaño de la ventana en función de la presencia de bordes en las cercanías. Ambos filtros se muestran especialmente adecuados para el tratamiento de ruido impulsivo; y su comportamiento frente al filtro de mediana es muy bueno. La preservación de bordes siempre supera al del filtro tradicional de mediana, mientras que la capacidad de eliminación de ruido se mantiene en niveles más que aceptables. Como se ve, en el presente capítulo se ha empleado el filtro de mediana como referencia.

En la primera parte se expondrá en detalle la metodología que emplea el filtro RM. A pesar de su simplicidad, su estudio teórico resulta bastante complicado. Aún así, se muestra un análisis estadístico del filtro, de acuerdo con los métodos ampliamente utilizados en la literatura, fundamentados en la estadística robusta. Se derivará una expresión para la función de distribución de probabilidad, de la cual se deducirá el comportamiento del algoritmo frente a la presencia de ruido aditivo e impulsivo. También se calcularán expresiones que pondrán de manifiesto

las propiedades de preservación de líneas y bordes del método. Empleando simulaciones por ordenador, se obtendrán las respuestas estadísticas del filtro frente a diferentes señales de entrada, y se compararán los resultados con el filtro de mediana.

A continuación se muestran otras interesantes propiedades del método, analizando la respuesta del mismo frente a señales deterministas. De particular importancia son las señales raíces del filtro (aquellas que permanecen invariantes ante un procesamiento iterativo). Los resultados se comparan de nuevo con el filtro de mediana tradicional.

Por último, se expondrá en este capítulo cómo a partir de un esquema adaptativo simple se puede diseñar un filtro muy capaz. El *restdiver* adapta el tamaño de su ventana, creciendo hasta encontrar un borde en la imagen. Con ello aprovechamos las cualidades del filtro de mediana, al eliminar ruido tanto mejor cuanto mayor la ventana, y difuminar menos los bordes cuanto menor es la misma. La principal ventaja de este método frente a otros mecanismos de mediana con ventana adaptativa [BER87] [SUN88] estriba en que en los que aparecen en la literatura varían su tamaño basándose en el resultado de un prefiltrado de mediana, por lo que estos esquemas siguen eliminando detalles como el filtro de mediana tradicional. Sin embargo, *restdiver* lo hace en función de un preprocesamiento que le informa acerca de las posiciones de los bordes en la imagen. Para realizarlo ha sido necesario diseñar un método de detección de bordes que funcione adecuadamente en presencia de ruido impulsivo, el cual aparece detallado. De nuevo los resultados se comparan con los del filtro de mediana tradicional.

EL FILTRO DE MEDIANA RELAJADA

DESCRIPCIÓN

El objetivo es mejorar el comportamiento del filtro de mediana, intentando aumentar la capacidad de eliminación de ruido, pero procurando no difuminar la imagen. Para ello se diseñó un algoritmo que, aunque extremadamente sencillo, permitía paliar alguno de los problemas que planteaba el filtro de mediana tradicional. Fue bautizado con el nombre de *Filtro de Mediana Relajada*, (MR), y opera de la siguiente manera: una ventana deslizante $W_{ij}(S)$de tamaño $w = 2N + 1$ centrada en la posición (i, j) recorre la imagen fuente. De este modo se tiene una región ventaneada en cada instante. Ordenando los niveles de gris de esa región se obtiene una lista $W_{ij}(S)_{(\cdot)}$. Dentro de esta lista se define un intervalo simétrico centrado en la mediana, limitado por dos valores previamente definidos: uno inferior, $\boldsymbol{l}$, y otro superior, $\boldsymbol{u}$. Por ejemplo, podemos elegir el intervalo de manera que contenga los tres valores centrales, esto es, $\boldsymbol{l} = 4$ y $\boldsymbol{u} = 6$ para una ventana 3×3. Este intervalo, denominado *intervalo de bondad*, contiene un cierto número de niveles de gris vecinos a la mediana de $W_{ij}(S)_{(\cdot)}$, por lo que se les supone una probabilidad de aparición suficientemente buena como para no ser considerados como ruido. Teniendo ello en cuenta, el filtro MR es:

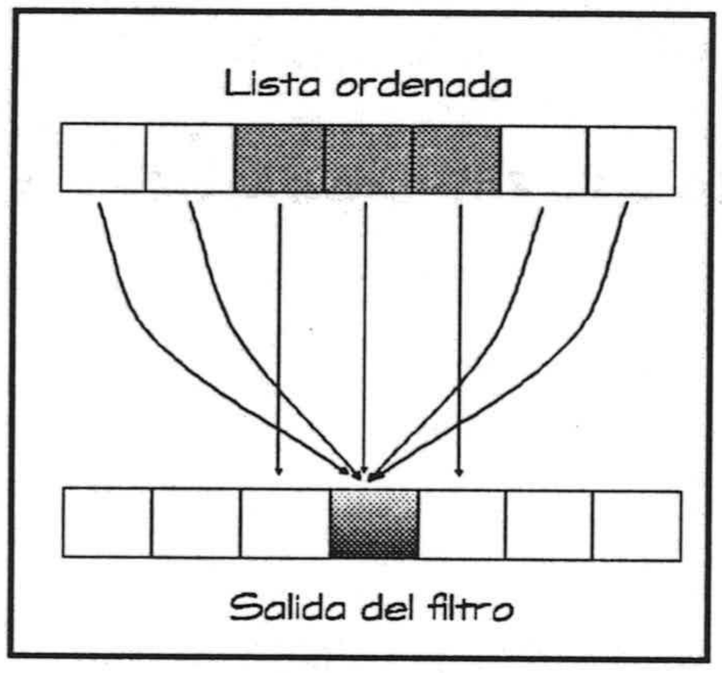


Figura 1: El filtro MR actúa como el filtro de mediana sólo en los extremos de la lista ordenada. Los valores que pertenezcan al *intervalo de bondad* permanecerán inalterados.

$$R_{ij} = MR_{lu}\left[W_{ij}(S)\right]$$

$$= \begin{cases} R_{ij} & \Leftrightarrow R_{ij} \in \left[\left[W_{ij}(S)\right]_{(l)}, \left[W_{ij}(S)\right]_{(u)}\right] \\ \left[W_{ij}(S)\right]_{(m)} & \Leftrightarrow R_{ij} \notin \left[\left[W_{ij}(S)\right]_{(l)}, \left[W_{ij}(S)\right]_{(u)}\right] \end{cases}$$

siendo $\boldsymbol{m}$ la posición central de la lista, esto es, la mediana. Por tanto, si el nivel de gris del píxel central de la ventana deslizante pertenece al intervalo de bondad se le considera suficientemente probable y no se filtrará. En caso contrario, se le procesa con el filtro de mediana tradicional. Obsérvese que si $\boldsymbol{l} = \boldsymbol{u}$ entonces el filtro MR se reduce al filtro de mediana, y si $\boldsymbol{l} = 1$, $\boldsymbol{u} = 2N + 1$ se tiene el filtro identidad. El mecanismo de funcionamiento del filtro MR se ha ilustrado en la figura 1.

PROPIEDADES ESTADISTICAS DEL FILTRO DE MEDIANA RELAJADA

Aparte de ser sencillo de implementar, es de vital importancia que el filtro sea capaz de atenuar el ruido. Debido a la no linealidad del método de procesamiento, esta propiedad es difícil de analizar en una situación general. La supresión del ruido depende fuertemente de la señal de entrada. Por ejemplo, el comportamiento del filtro MR es muy diferente en una región de nivel de gris uniforme conrespecto a las cercanías de los bordes de la imagen. Sin embargo es posible calcular las características estadísticas de la salida del filtrado para muchos tipos de señal de entrada [DAV80]. En esta sección analizaremos las propiedades de atenuación del ruido y de respeto a los bordes y líneas de la imagen para el filtro MR y las compararemos con las del filtro de mediana tradicional. Para ello es preciso obtener algunos resultados previos. Como norma para el resto del capítulo, y a no ser que se especifique lo contrario, por W_{ij} entenderemos que la ventana deslizante $W_{ij}(S)$.

PRELIMINARES.

Lema 1. *Sea A un suceso, y X una variable aleatoria con función de distribución* $\Phi(t)$ *y densidad de probabilidad* $\varphi(t)$, *entonces se tiene:*

$$\Pr[A] = \int_{-\infty}^{+\infty} \Pr[A / X = x]\varphi(x)dx = \int_{-\infty}^{+\infty} \Pr[A / X = x]d\Phi(x)$$

Además se tiene:

$$\Pr[A; X \in B] = \int_B \Pr[A / X = x]\varphi(x)dx$$
$$= \int_B \Pr[A / X = x]d\Phi(x)$$

Lema 2. $\forall\, x \in [0,1],\ \forall\, n, k \in \mathbb{N}$, *se tiene:*

$$\int_0^1 \binom{n}{k} x^k (1-x)^{n-k}\, dx = \frac{1}{n+1}$$

Demostración:

La demostración puede hacerse fácilmente por inducción, si bien es más sencillo si nos percatamos que por definición de la función beta $\beta(m,n)$ se tiene:

$$\int_0^1 x^k (1-x)^{n-k}\, dx = \beta(k+1, n-k+1)$$

De las propiedades de dicha función se sigue lo pedido.

■

Proposición 1. *Supongamos que X_{ij} son variables aleatorias independientes e idénticamente distribuidas (i.i.d.) con una función de distribución $\Phi(t)$, y densidad $\varphi(t)$. Denotemos por $\Phi_r(t)$ la función de distribución del estadístico de orden $[W_{ij}]_{(r)}$. Entonces:*

(i) $[W_{ij}]_{(2N+1)}$ *tiene función de distribución:*

$$\Phi_{2N+1}(t) = \Phi(t)^{2N+1}$$

(ii) $[W_{ij}]_{(1)}$ *tiene función de distribución:*

$$\Phi_1(t) = 1-[1-\Phi(t)]^{2N+1}$$

(iii) $[W_{ij}]_{(r)}$, *r=1,...2N+1 tiene función de distribución:*

$$\Phi_r(t) = \sum_{k=r}^{2N+1} \binom{2N+1}{k} \Phi(t)^k [1-\Phi(t)]^{2N+1-k}$$

Demostración:

(i) La función de distribución de $[W_{ij}]_{(2N+1)}$ viene dada por:

$$\begin{aligned}\Phi_{2n+1}(t) &= \Pr[[W_{ij}]_{(2N+1)} \le t] \\ &= \Pr[\text{todos los puntos de } W_{ij} \text{ son} \le t] \\ &= \Phi(t)^{2N+1}\end{aligned}$$

(ii) Análogamente se tiene:

$$\begin{aligned}\Phi_1(t) &= \Pr[[W_{ij}]_{(1)} \le t] \\ &= 1 - \Pr[[W_{ij}]_{(1)} > t] \\ &= 1 - \Pr[\text{todos los puntos de } W_{ij} \text{ son} > t] \\ &= 1 - [1 - \Phi(t)]^{2N+1}\end{aligned}$$

(iii) La función de distribución de $[W_{ij}]_{(r)}$ es:

$$\begin{aligned}\Phi_r(t) &= \Pr[[W_{ij}]_{(r)} \le t] \\ &= \Pr[\text{al menos r puntos de } W_{ij} \text{ son} \le t]\end{aligned}$$

Luego se tiene que:

$$\Phi_r(t) = \sum_{k=r}^{2N+1} \binom{2N+1}{k} \Phi(t)^k [1-\Phi(t)]^{2N+1-k}$$

■

Proposición 2. Para $t_r \leq t_s$, la función de distribución conjunta de $[W_{ij}]_{(r)}$, $[W_{ij}]_{(s)}$ viene dada por:

$$\Phi_{r,s}(t_r,t_s) = \sum_{k=s}^{2N+1} \sum_{h=r}^{k} \binom{2N+1}{k} \binom{k}{h} \Phi(t_r) \cdot \cdot [\Phi(t_s)-\Phi(t_r)]^{k-h} [1-\Phi(t_s)]^{2N+1-k}$$

dónde *r,s* son tales que $1 \quad r \quad s \quad 2N+1$.

Demostración:

$$\begin{aligned} \Phi_{r,s}(t_r, t_s) &= \Pr[[W_{ij}]_{(r)} \leq t_r, [W_{ij}]_{(s)} \leq t_s] \\ &= \Pr[\text{al menos } r \text{ puntos de } W_{ij} \text{ son} \leq t_r, \\ &\qquad \text{al menos } s \text{ puntos de } W_{ij} \text{ son} \leq t_s] \end{aligned}$$

Por tanto:

$$\Phi_{r,s}(t_r,t_s) = \sum_{k=s}^{2N+1} \sum_{h=r}^{k} \Pr[\text{exact. h puntos de } W_{ij} \text{ son } \leq t_r, \text{ exact. k puntos en } W_{ij} \text{ son } \leq t_s]$$

Ahora bien, la siguiente probabilidad corresponde a una ley multinomial, como se ve en la figura 2. Entonces:

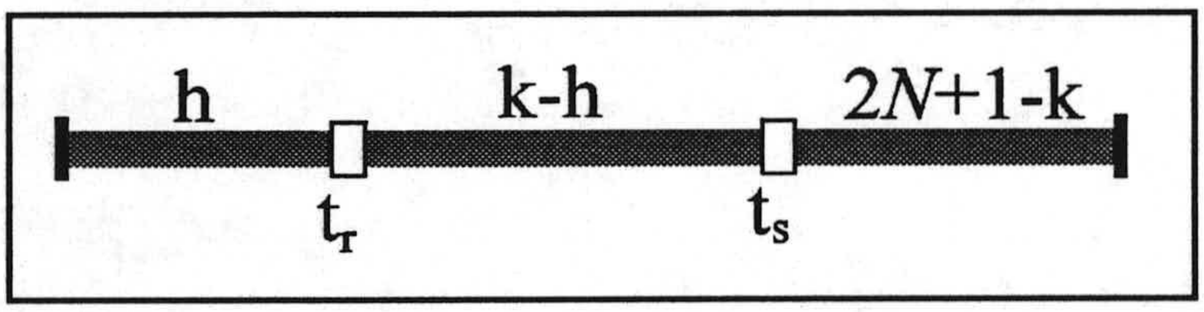


Figura 2: Ilustra el caso multinomial que se presenta a la hora de estudiar la distribución de probabilidad de la salida del filtro MR.

$$\Pr[\text{exact. } h \text{ puntos de } W_{ij} \text{ son} \leq t_r, \text{ exact. } k \text{ puntos de } W_{ij} \text{ son} \leq t_s] =$$

$$= \frac{(2N+1)!}{h!(k-h)!(2N+1-k)!} \Pr\left[X_{ij} \le t_r\right]^h$$
$$\cdot \Pr\left[t_r < X_{ij} < t_s\right]^{k-h} \cdot \Pr\left[X_{ij} > t_s\right]^{2N+1-k}$$
$$= \binom{2N+1}{k}\binom{k}{h} \Phi(t_r)^s \left[\Phi(t_s) - \Phi(t_r)\right]^{k-h}$$
$$\cdot \left[1 - \Phi(t_s)\right]^{2N+1-k}$$

Con lo que se concluye la demostración. ■

Para calcular la distribución de probabilidad de la salida de un filtro de mediana relajada nos será de gran utilidad trabajar con la ventana $\hat{W}_{ij}$, que representa los elementos de la ventana deslizante W_{ij} *excepto* el píxel central. Hagamos la siguiente definición:

Definición 1. Sea

$$\hat{\Phi}_r(t) = \Pr\left[\left[\hat{W}_{ij}\right]_{(r)} \le t\right]$$
$$= \Pr\left[\text{al menos r puntos de } \hat{W}_{ij} \text{ son } \le t\right]$$
$$= \sum_{k=r}^{2N} \binom{2N}{k} \Phi(t)^k \left[1 - \Phi(t)\right]^{2N-k}$$

Y para $t_r \le t_s$, con $1 \le r \le s \le 2N$, notemos por:

$$\hat{\Phi}_{r,s}(t_r, t_s) = \Pr\left[\left[\hat{W}_{ij}\right]_{(r)} \le t_r, \left[\hat{W}_{ij}\right]_{(s)} \le t_s\right]$$
$$= \sum_{k=s}^{2N} \sum_{h=r}^{k} \binom{2N}{k}\binom{k}{h} \Phi(t_r)^h \left[\Phi(t_s) - \Phi(t_r)\right]^{k-h}$$
$$\cdot \left[1 - \Phi(t_s)\right]^{2N-k}$$

donde $\hat{W}_{ij} = W_{ij} - S_{ij}$ es una ventana de tamaño $2N$.
Como propiedad, para $t_r \le t_s$, se tiene que:

$$\hat{\Phi}_{r,r}(t_r, t_s) = \hat{\Phi}_r(t_r)$$

■

CÁLCULO DE LA FUNCIÓN DE DISTRIBUCIÓN DE PROBABILIDAD PARA EL FILTRO DE MEDIANA RELAJADA.

Sea $\Psi(t) = \Pr[R_{ij} \leq t]$ la función de distribución de la salida del filtro de mediana relajada. Para facilitar su cálculo, estudiemos los distintos casos en que podemos obtener $R_{ij} \leq t$ dependiendo del valor de la entrada (imagen fuente) S_{ij}:

(*a*) *Caso 1:* Aquí $S_{ij} \leq t$. Entonces la salida verificará $R_{ij} \leq t$ sólamente por dos causas excluyentes:

1. $[W_{ij}]_{(m)} \leq t$
2. $[W_{ij}]_{(l)} \leq S_{ij}$ y $[W_{ij}]_{(m)} > t$

tal y como se muestra en la figura 3. Entonces:

$$\begin{aligned}\Pr[R_{ij} \leq t, S_{ij} \leq t] &= \Pr[S_{ij} \leq t, [W_{ij}]_{(m)} \leq t] \\ &+ \Pr[S_{ij} \leq t, [W_{ij}]_{(m)} > t, [W_{ij}]_{(l)} \ S_{ij}] \\ &= \Psi_1(t) + \Psi_2(t)\end{aligned}$$

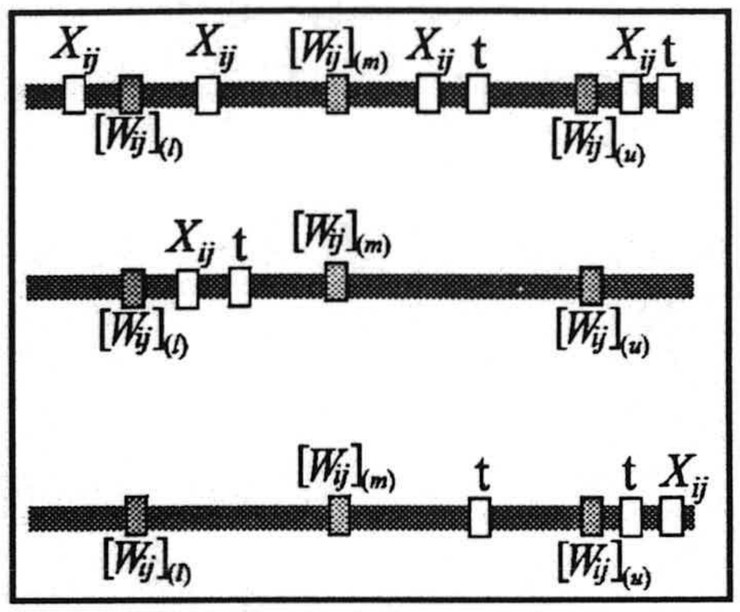


Figura 3: Los dos primeros gráficos ilustran el caso 1. Se señalan las posiciones relativas de los valores más significativos de la distribución.

(*b*) *Caso 2*: Si $S_{ij} > t$, entonces la única posibilidad de que $R_{ij} \leq t$ es que se verifique que $S_{ij} > [W_{ij}]_{(u)}$ y que $[W_{ij}]_{(m)} \leq t$, tal como se muestra en la figura 4. Entonces:

$$\begin{aligned}\Pr[R_{ij} \leq t, S_{ij} > t] &= \Pr[[W_{ij}]_{(m)} \leq t, S_{ij} > t, [W_{ij}]_{(u)} < S_{ij}] \\ &= \Psi_3(t)\end{aligned}$$

Por lo tanto, podemos asegurar que

$$\begin{aligned}\Psi(t) &= \Pr[R_{ij} \leq t] \\ &= \Pr[R_{ij} \leq t, S_{ij} \leq t] + \Pr[R_{ij} \leq t, S_{ij} > t] \\ &= \Psi_1(t) + \Psi_2(t) + \Psi_3(t)\end{aligned}$$

Llegados a este punto hagamos una breve digresión. Se darán valores especiales a las variables, de manera que se obtengan filtros de comportamiento conocido que nos permitan verificar si el anterior resultado no es incorrecto.

Filtro identidad: Tal filtro corresponde a $l = 1$ y $u = 2N + 1$, y resulta que:

$$\Psi_2(t) = \Pr[S_{ij} \le t,\ [W_{ij}]_{(m)} > t]$$

y

$$\Psi_3(t) = 0$$

Por tanto,

$$\begin{aligned}\Psi(t) &= \Pr[S_{ij} \le t,\ [W_{ij}]_{(m)} > t] + \Pr[S_{ij} \le t,\ [W_{ij}]_{(m)} \le t] \\ &= \Pr[S_{ij} \le t] \\ &= \Phi(t)\end{aligned}$$

tal y como era de esperar.

Filtro de mediana tradicional: En este caso se tiene $\boldsymbol{l} = \boldsymbol{u} = m$, y queda:

$$\Psi_2(t) = \Pr[S_{ij} \quad t,\ [W_{ij}]_{(m)} > t,\ [W_{ij}]_{(m)} \le S_{ij}] = 0$$

y

$$\Psi_3(t) = \Pr[S_{ij} > t,\ [W_{ij}]_{(m)} \le t]$$

Por tanto,

$$\begin{aligned}\Psi(t) &= \Pr[S_{ij} \le t,\ [W_{ij}]_{(m)} \le t] + \Pr[S_{ij} > t,\ [W_{ij}]_{(m)} \le t] \\ &= \Pr[[W_{ij}]_{(m)} \le t] \\ &= \Phi_m(t)\end{aligned}$$

lo cual coincide con la distribución de probabilidad bien conocida del filtro de mediana tradicional.

Una vez que podemos confiar en nuestros cálculos, establezcamos la expresión explícita de $\Psi(t)$. La función de distribución del filtro de mediana relajada puede expresarse como:

$$\Psi(t) = \Psi_1(t) + \Psi_2(t) + \Psi_3(t)$$

dónde

$$\begin{aligned}\Psi_1(t) &= \Pr[S_{ij} \le t,\ [W_{ij}]_{(m)} \le t] \\ \Psi_2(t) &= \Pr[S_{ij} \le t,\ [W_{ij}]_{(m)} > t,\ [W_{ij}]_{(l)} \le S_{ij}] \\ \Psi_3(t) &= \Pr[S_{ij} > t,\ [W_{ij}]_{(m)} \le t,\ [W_{ij}]_{(u)} < S_{ij}]\end{aligned}$$

La expresión explícita de $\Psi_1(t)$, aplicando el concepto de pro-

babilidad condicionada, es la siguiente:

$$\begin{aligned}\Psi_1(t) &= \Pr[[W_{ij}]_{(m)} \le t \,/\, S_{ij} \le t]\ \Pr[S_{ij} \le t] \\ &= \Pr[[\hat{W}_{ij}]_{(m-1)} \le t]\ \Pr[S_{ij} \le t]\end{aligned}$$

Por tanto,

$$\Psi_1(t) = \hat{\Phi}_{m-1}(t) \cdot \Phi(t)$$

Aplicando el lema 1, se obtiene la siguiente expresión para $\Psi_2(t)$:

$$\Psi_2(t) = \int_{-\infty}^{t} \left[\hat{\Phi}_{l-1}(\tau) - \hat{\Phi}_{l-1,m-1}(\tau, t)\right] \varphi(\tau) d\tau$$

En efecto,

$$\begin{aligned}\Psi_2(t) &= \int_{-\infty}^{t} \Pr\left[\left[W_{ij}\right]_{(m)} > t, \left[W_{ij}\right]_{(l)} \le S_{ij} \,/\, S_{ij} = \tau\right] \varphi(\tau) d\tau \\ &= \int_{-\infty}^{t} \Pr\left[\left[\hat{W}_{ij}\right]_{(m-1)} > t, \left[\hat{W}_{ij}\right]_{(l-1)} \le \tau\right] \varphi(\tau) d\tau \\ &= \int_{-\infty}^{t} \left(\Pr\left[\left[\hat{W}_{ij}\right]_{(l-1)} \le \tau\right] - \right. \\ &\qquad \left. - \Pr\left[\left[\hat{W}_{ij}\right]_{(m-1)} \le t, \left[\hat{W}_{ij}\right]_{(l-1)} \le \tau\right] \right) \varphi(\tau) d\tau \\ &= \int_{-\infty}^{t} \left[\hat{\Phi}_{l-1}(\tau) - \hat{\Phi}_{l-1,m-1}(\tau, t)\right] \varphi(\tau) d\tau\end{aligned}$$

De manera similar:

$$\Psi_3(t) = \int_{t}^{+\infty} \hat{\Phi}_{m,u}(t, \tau) \varphi(\tau) d\tau$$

En efecto,

$$\begin{aligned}\Psi_3(t) &= \int_{t}^{+\infty} \Pr\left[\left[W_{ij}\right]_{(m)} \le t, \left[W_{ij}\right]_{(u)} < S_{ij} \,/\, S_{ij} = \tau\right] \varphi(\tau) d\tau \\ &= \int_{t}^{+\infty} \Pr\left[\left[\hat{W}_{ij}\right]_{(m)} \le t, \left[\hat{W}_{ij}\right]_{(u)} < \tau\right] \varphi(\tau) d\tau \\ &= \int_{t}^{+\infty} \hat{\Phi}_{m,u}(t, \tau) \varphi(\tau) d\tau\end{aligned}$$

Reuniendo estos cálculos, resulta la siguiente expresión de $\Psi(t)$:

COMISION DE DOCTORADO

$$\Psi(t) = \hat{\Phi}_{m-1}(t)\Phi(t) + \int_{-\infty}^{t}\left[\hat{\Phi}_{l-1}(\tau) - \hat{\Phi}_{l-1,m-1}(\tau,t)\right]\varphi(\tau)d\tau$$

$$+ \int_{t}^{+\infty}\hat{\Phi}_{m,u}(t,\tau)\varphi(\tau)d\tau$$

Sin embargo, aún no conocemos las expresiones explícitas de $\Psi_2(t)$ y $\Psi_3(t)$, necesarias para la elaboración definitiva del cálculo de la función de distribución de probabilidad del filtro de mediana relajada. Establezcámoslas a continuación. Para ello nos serviremos de un nuevo lema, que pasamos a enunciar.

Lema 3. *Si $r \le s$, sean*

$$I_{r,s}(t) = \int_{-\infty}^{t}\hat{\Phi}_{r,s}(\tau,t)\varphi(\tau)d\tau$$

y

$$J_r(t) = \int_{-\infty}^{t}\hat{\Phi}_r(\tau)\varphi(\tau)d\tau$$

Se tiene:

$$I_{r,s}(t) = \sum_{k=s}^{2N}\left(\frac{k+1-r}{k+1}\right)\binom{2N}{k}\Phi(t)^{k+1}\left[1-\Phi(t)\right]^{2N-k}$$

Cuando $r = s$ se sabe que para $\tau \le t$:

$$\hat{\Phi}_{r,r}(\tau,t) = \hat{\Phi}_r(\tau)$$

y resulta que:

$$J_r(t) = \sum_{k=r}^{2N}\left(\frac{k+1-r}{k+1}\right)\binom{2N}{k}\Phi(t)^{k+1}\left[1-\Phi(t)\right]^{2N-k}$$

Demostración:

$$I_{r,s}(t) = \int_{-\infty}^{t}\hat{\Phi}_{r,s}(\tau,t)\varphi(\tau)d\tau$$

$$= \sum_{k=s}^{2N} \sum_{h=r}^{k} \binom{2N}{k} \binom{k}{h} \int_{-\infty}^{t} \Phi(\tau)^h \left[\Phi(t) - \Phi(\tau)\right]^{k-h} \cdot$$

$$\cdot \left[1 - \Phi(t)\right]^{2N-k} d\Phi(\tau)$$

$$= \sum_{k=s}^{2N} \sum_{h=r}^{k} \binom{2N}{k} \left[1 - \Phi(t)\right]^{2N-k} \int_{-\infty}^{t} \binom{k}{h} \Phi(\tau)^h \cdot$$

$$\cdot \left[\Phi(t) - \Phi(\tau)\right]^{k-h} d\Phi(\tau)$$

Realizando el cambio de variable $\Phi(\tau) = x\Phi(t)$, y aplicando el lema 2, se obtiene:

$$I_{r,s}(t) = \sum_{k=s}^{2N} \sum_{h=r}^{k} \binom{2N}{k} \Phi(t)^{k+1} \left[1 - \Phi(t)\right]^{2N-k} \cdot$$

$$\cdot \int_0^1 \binom{k}{h} x^h (1-x)^{k-h} dx$$

$$= \sum_{k=s}^{2N} \sum_{h=r}^{k} \binom{2N}{k} \Phi(t)^{k+1} \left[1 - \Phi(t)\right]^{2N-k} \left(\frac{1}{k+1}\right)$$

$$= \sum_{k=s}^{2N} \left(\frac{k+1-r}{k+1}\right) \binom{2N}{k} \Phi(t)^{k+1} \left[1 - \Phi(t)\right]^{2N-k}$$

■

Aplicando ahora el lema 3, se tiene:

$$\Psi_2(t) = J_{l-1}(t) - I_{l-1,m-1}(t)$$

$$= \sum_{k=l-1}^{2N} \left(\frac{k-l+2}{k+1}\right) \binom{2N}{k} \Phi(t)^{k+1} \left[1 - \Phi(t)\right]^{2N-k}$$

$$- \sum_{k=m-1}^{2N} \left(\frac{k-l+2}{k+1}\right) \binom{2N}{k} \Phi(t)^{k+1} \left[1 - \Phi(t)\right]^{2N-k}$$

$$= \sum_{k=l-1}^{m-2} \left(\frac{k-l+2}{k+1}\right) \binom{2N}{k} \Phi(t)^{k+1} \left[1 - \Phi(t)\right]^{2N-k}$$

De modo análogo, se tiene:

$$\Psi_3(t) = \int_t^{+\infty} \hat{\Phi}_{m,u}(t,\tau) \varphi(\tau) d\tau$$

$$= \sum_{k=uh=m}^{2N} \sum^{k} \binom{2N}{k} \binom{k}{h} \int_{t}^{+\infty} \Phi(t)^h \left[\Phi(\tau) - \Phi(t)\right]^{k-h} \cdot$$

$$\cdot \left[1 - \Phi(\tau)\right]^{2N-k} d\Phi(\tau)$$

$$= \sum_{k=uh=m}^{2N} \sum^{k} \binom{2N}{k} \Phi(t)^h \int_{t}^{+\infty} \binom{k}{h} \left[1 - \Phi(t)\right]^{2N-k} \cdot$$

$$\cdot \left[\Phi(\tau) - \Phi(t)\right]^{k-h} d\Phi(\tau)$$

Realizando el cambio de variable $\Phi(\tau) = x[1 - \Phi(t)] + \Phi(t)$, y aplicando el lema 2 podemos simplificar el resultado anterior, para reducir a $\Psi_3(t)$ a una forma más sencilla y manejable. De hecho, queda:

$$\Psi_3(t) = \sum_{k=uh=m}^{2N} \sum^{k} \binom{2N}{k} \Phi(t)^h \left[1 - \Phi(t)\right]^{2N+1-h} \cdot$$

$$\cdot \int_{0}^{1} \binom{k}{h} x^{k-h} (1-x)^{2N-k} dx$$

$$= \sum_{k=uh=m}^{2N} \sum^{k} \binom{2N}{k} \binom{k}{h} \Phi(t)^h \left[1 - \Phi(t)\right]^{2N+1-h}$$

$$\cdot \beta(k-h+1, 2N-k+1)$$

$$= \sum_{k=uh=m}^{2N} \sum^{k} \frac{(2N)!}{k!(2N-k)!} \frac{k!}{h!(k-h)!} \Phi(t)^h \cdot$$

$$\cdot \left[1 - \Phi(t)\right]^{2N+1-h} \frac{(k-h)!(2N-k)!}{(2N+1-h)!}$$

$$= \sum_{k=uh=m}^{2N} \sum^{k} \left(\frac{1}{2N+1-h}\right) \binom{2N}{h} \Phi(t)^h \cdot$$

$$\cdot \left[1 - \Phi(t)\right]^{2N+1-h}$$

Acabamos de obtener la expresión explícita de todas y cada una de las funciones que componían de manera aditiva la distribución de probabilidad de la salida de un filtro de mediana relajada. Lo único que resta por tanto es reunir los resultados pertinentes para conseguir la expresión final de $\Psi(t)$. Dicha expresión es:

$$\Psi(t) = \sum_{k=m-1}^{2N} \binom{2N}{k} \Phi(t)^{k+1} [1-\Phi(t)]^{2N-k}$$
$$+ \sum_{k=l-1}^{2N} \left(\frac{k-l+2}{k+1}\right) \binom{2N}{k} \Phi(t)^{k+1} [1-\Phi(t)]^{2N-k}$$
$$+ \sum_{k=u}^{2N} \sum_{h=m}^{k} \left(\frac{1}{2N+1-h}\right) \binom{2N}{h} \Phi(t)^{h} [1-\Phi(t)]^{2N+1-h}$$

Expresión final de la función de distribución de probabilidad de la salida del filtro de mediana relajada (MR)

Se acaba de establecer la expresión matemática que proporciona la función de distribución de probabilidad para la salida del filtro de mediana ralajada [BEN??]. Este resultado servirá como base para el análisis cuantitativo de la atenuación de ruido y del comportamiento frente a bordes y líneas en la imagen para el caso del filtro MR.

Atenuación de ruido

Ruido impulsivo. Para la eliminación del ruido impulsivo se puede calcular la probabilidad de ocurrencia de un pulso a la salida cuando la probabilidad de aparición p de los pulsos es conocida. Si dicha probabilidad es alta, significará que el filtro proporcionará una salida ruidosa a una entrada ruidosa, es decir, que el filtro no elimina adecuadamente el ruido impulsivo. Denotemos por $B(p)$ dicha probabilidad, que suele denominarse de *colapso* (*breakdown*), para el filtro de mediana relajada. Su valor se deduce directamente de la expresión de $\Psi(t)$, pues en este caso la probabilidad de la entrada es una binomial. Entonces:

$$B(p) = \sum_{k=m-1}^{2N} \binom{2N}{k} p^{k+1} [1-p]^{2N-k}$$
$$+ \sum_{k=l-1}^{2N} \left(\frac{k-l+2}{k+1}\right) \binom{2N}{k} p^{k+1} [1-p]^{2N-k}$$
$$+ \sum_{k=u}^{2N} \sum_{h=m}^{k} \left(\frac{1}{2N+1-h}\right) \binom{2N}{h} p^{h} [1-p]^{2N+1-h}$$

La tabla 1 muestra las probabilidades de colapso, calculadas por integración numérica, para el filtro MR con diversos intervalos de bondad *centrados*, y para ventanas de tamaño 3×3 y 5×5. También aparecen los valores de $B(p)$ para el filtro de mediana de igual tamaño de ventana. Se ha variado la probabilidad p de aparición de ruido impulsivo en la imagen S.

Tabla 1.a: Se muestran los resultados obtenidos por integración numérica de la función $B(p)$ que proporciona los valores de la probabilidad de colapso para el filtro de mediana relajada y para el filtro de mediana tradicional, ambos con ventana 3×3. En la cabecera, p indica la probabilidad de aparición de ruido impulsivo, mientras que los números entre corchetes son los valores de l y u del intervalo de bondad.

p	INTERVALO DE BONDAD			
	[2,8]	[3,7]	[4,6]	mediana
0,05	0,008924	0,001019	0,000097	0,000033
0,625	0,013569	0,001932	0,000241	0,000097
0,1	0,032134	0,007314	0,001626	0,000891
0,125	0,047847	0,013645	0,003985	0,002482
0,2	0,010815	0,049859	0,025087	0,019581
0,25	0,158104	0,091225	0,057579	0,048927
0,3	0,215328	0,147904	0,1097	0,098809
0,375	0,313627	0,26074	0,227179	0,216618
0,4	0,349217	0,304759	0,275857	0,266568
0,5	0,5	0,5	0,5	0,5

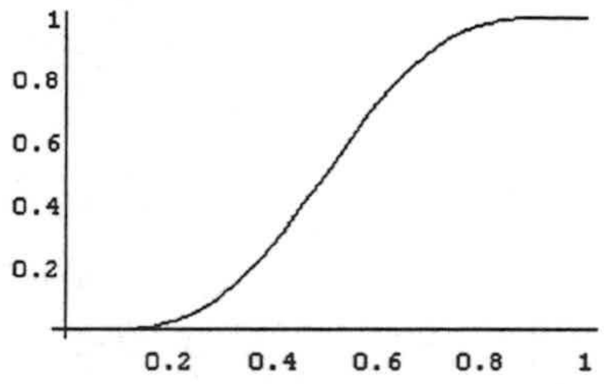
De los resultados se deduce que la probabilidad de colapso siempre es mayor en MR que en el filtro de mediana como era de esperar, puesto que el filtro de mediana equivale a un MR con $l = u$, y cuanto mayor es el intervalo de bondad (*i.e.* $|l - u|$), mayor es la probabilidad de que el ruido caiga dentro de él y se deje sin procesar. Por lo tanto el filtro de mediana relajada es siempre menos robusto que el filtro de mediana tradicional frente a ruido impulsivo.

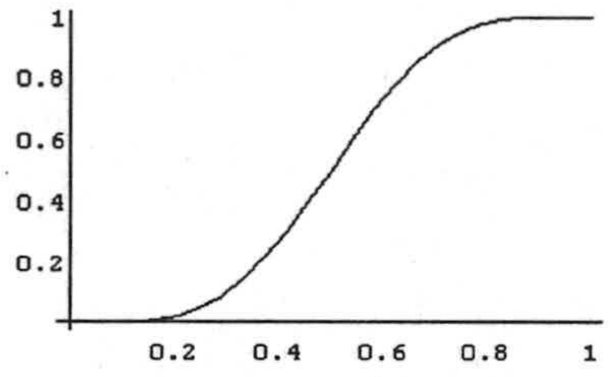
Sin embargo, cuando el intervalo de bondad es pequeño, los valores de las probabilidades de colapso son muy parecidos a los correspondientes a la mediana. Por ejemplo, en una ventana 3×3, con $l = 4$, $u = 6$, se tiene un $B(p)$ como muestra la figura 4(a). En la figura 4(b) aparece representada la misma función para un filtro de mediana 3×3. Como puede apreciarse, la diferencia entre ambas es muy pequeña. De hecho, podemos representar dicha diferencia, que aparece plasmada en la figura 4(c). De ella se deduce que la máxima diferencia entre ambas proba-

INTERVALO BONDAD	PROBABILIDAD DE APARICIÓN DE RUIDO									
	0,05	0,0625	0,1	0,125	0,2	0,25	0,3	0,375	0,4	0,5
[2,24]	0,021096	0,030468	0,062872	0,08642	0,160181	0,2103	0,261403	0,342979	0,372302	0,5
[3,23]	0,006791	0,011716	0,03372	0,052911	0,121306	0,17085	0,222863	0,310963	0,344605	0,5
[4,22]	0,001706	0,003587	0,015203	0,028958	0,085264	0,132404	0,184619	0,278981	0,316924	0,5
[5,21]	0,000343	0,000888	0,005747	0,012802	0,054654	0,096523	0,147347	0,247156	0,28932	0,5
[6,20]	0,000056	0,000181	0,001828	0,004988	0,03151	0,065342	0,112363	0,215848	0,262	0,5
[7,19]	7,617·E-	0,000031	0,000492	0,001664	0,016207	0,040743	0,815001	0,185843	0,235473	0,5
[8,18]	8,667E-7	4,363E-6	0,000113	0,000476	0,007438	0,023456	0,056524	0,158442	0,210706	0,5
[9,17]	8,349E-8	5,278E-7	0,000022	0,000118	0,003103	0,012786	0,038395	0,135272	0,189101	0,5
[10,16]	6,875E-9	5,483E-8	3,810E-6	0,000026	0,001261	0,007078	0,026865	0,117776	0,172171	0,5
[11,15]	5,098E-1	5,265E-9	6,634E-7	6,271E-6	0,000597	0,004493	0,020667	0,106595	0,16093	0,5
[12,14]	6,061E-1	8,623E-1	2,091E-7	2,623E-6	0,000404	0,003567	0,018082	0,101166	0,155287	0,5
[13,13]	3,591E-1	5,642E-1	1,621E-7	2,165E-6	0,000369	0,00337	0,01747	0,099737	0,153768	0,5

Tabla 1.b: Análoga a la tabla anterior, pero para ventana de tamaño 5x5.

bilidades de colapso se produce cuando el porcentaje de ruido es del 30% ($p = 0.3$). Pero para mayores o menores probabilidades de aparición de impulsos, la diferencia es muy poco importante.





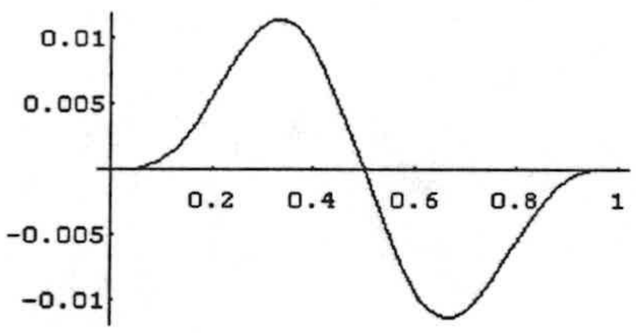


Figura 4: La gráfica superior muestra la función B(p) para el filtro $MR_{4,6}$, mientras la gráfica intermedia es una representación de la misma función para la mediana de ventana 3×3. Se observa un gran parecido entre ambas. La última gráfica ilustra la diferencia entre ellas. Se aprecia que la máxima disimilitud se produce para una probabilidad de aparición de ruido impulsivo de 30%

Si el ruido impulsivo es del tipo *sal y pimienta*, esto es, aparecen pulsos de máximo y mínimo valor (*i.e.* blancos y negros) equiprobables, entonces la probabilidad conjunta de colapso es $B(p) + B(q)$, siendo p la probabilidad de aparición de pulsos blancos, y q de aparición de pulsos negros. El resultado puede apreciarse en la figura 5, donde se ha representado gráficamente la probabilidad conjunta de colapso para un $MR_{4,6}$ de ventana 3×3. El comportamiento de esta probabilidad no difiere apreciablemente de la que se ha mostrado en la tabla 1 para $B(p)$.

Ruido gaussiano. Supongamos que la imagen de entrada S_{ij} es resultado de la imagen original O afectada por ruido blanco aditivo ν_{ij} de media nula y varianza unidad, es decir:

$$S_{ij} = O_{ij} + \nu_{ij}$$

Entonces las características estadísticas de la entrada son:

$$E(S_{ij}) = 0$$
$$E(S_{ij}^2) = \sigma_\nu^2$$

Usando la expresión calculada de $\Psi(t)$, se puede calcular la varianza de la salida del filtro de mediana relajada mediante integración numérica para una entrada S_{ij} con distribución conocida $\Phi(t)$.

Ello se ha llevado a cabo empleando ruido aditivo blanco gaussiano con distribución $N(0,1)$, tanto para el filtro MR como para el de mediana tradicional. Se ha trabajado con dos diferentes tamaños de ventana (3×3 y 5×5) ; y en el caso de la mediana relajada, se han empleado distintos valores de intervalo de bondad, con objeto de mostrar cómo evoluciona el comportamiento del algoritmo. Los resultados han quedado plasmados en la tabla 2. De ella se desprende en primer lugar que todos los filtros hacen disminuir la varianza. Ello significa que la dispersión de niveles de gris que tiene lugar en la señal por la adición del ruido queda reducida. En otras palabras, el ruido gaussiano se ve suavizado, y tanto más eficientemente cuanto mayor sea la reducción de la varianza. En cualquier caso, ninguno de los algoritmos propuestos presenta un comportamiento demasiado bueno frente a este tipo de ruido.

También observamos que el filtro de mediana relajada siempre proporciona valores de la varianza local de salida mayores que los correspondientes al filtro de mediana tradicional de igual tamaño de ventana. Ello significa que el filtro MR es peor que el de mediana, en cuanto a suavización del ruido gaussiano. Sin embargo, comprobamos que cuanto menor es el intervalo de bondad, más se asemeja la disminución de varianza a la de la mediana. Ello era de esperar, puesto que el límite superior del intervalo de bondad es el filtro identidad, que deja la varianza inalterada, y el inferior es el filtro de mediana.

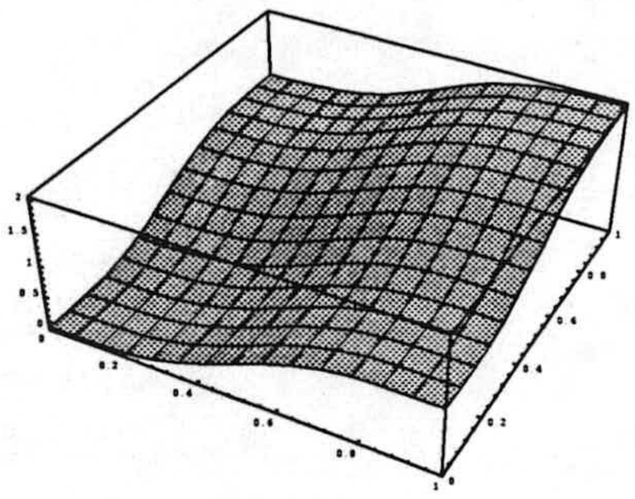


Figura 5: Aquí aparece representada la probabilidad de colapso respecto a la presencia de ruido compuesto de pulsos de dos valores diferentes, cada uno de los cuales posee su propia probabilidad de aparición.

TAMAÑO DE VENTANA	INTERVALO DE BONDAD	VARIANZA
3×3	[2,8]	0,467441
	[3,7]	0,261047
	[4,6]	0,18384
	Mediana	0,166
5×5	[2,24]	0,675262
	[3,23]	0,48258
	[4,22]	0,351042
	[5,21]	0,25743
	[6,20]	0,190181
	[7,19]	0,142308
	[8,18]	0,109072
	[9,17]	0,086988
	[10,16]	0,073325
	[11,15]	0,06583
	[12,14]	0,062556
	Mediana	0,061746

Tabla 2: Aquí aparecen los valores de la varianza que han sido calculados mediante integración numérica a partir de la función de distribución de probabilidad de la salida del filtro MR. También se incluyen los correspondientes a la mediana tradicional. Se han empleado diferentes tamaños de ventana, y en el caso del filtro de mediana relajada, también se han usado distintos intervalos de bondad.

A la vista de los resultados, tanto para ruido gaussiano como para impulsivo, podemos por tanto concluir que el filtro MR es siempre menos potente suprimiendo ruido que el filtro de mediana tradicional; pero que cuando el intervalo de bondad es moderado o pequeño, la diferencia de comportamiento entre ambos métodos es poco importante: ello parece que condiciona a nuestros filtros a ser empleados solo con reducidos intervalos de bondad. Sin embargo, debemos tener en cuenta que la supresión de ruido no es el único fin del procesamiento. La preservación de detalles de la imagen es un objetivo también muy importante, que habrá de tenerse en cuenta.

PRESERVACIÓN DE BORDES

En este apartado consideraremos el efecto del filtro de mediana relajada sobre un borde de escalón afectado por ruido blanco gaussiano:

$$S_{ij} = \begin{cases} \nu_{ij} & \Leftrightarrow j \leq 0 \\ h + \nu_{ij} & \Leftrightarrow j \geq 1 \end{cases}$$

dónde h es una constante que representa la altura del escalón del borde, y ν_j es un ruido *i.i.d.* con función de distribución $\Phi_A(t)$ que corresponde a la región del borde anterior al escalón, que denotamos por A. La función $\Phi_B(t)$ representa la distribución de la variable $h + \nu_{ij}$ en la región superior del escalón, que denominamos B. Está claro que $\Phi_B(t) = \Phi_A(t\text{-}h)$.

El comportamiento del filtro de mediana relajada cerca del borde ruidoso se examina calculando la esperanza, varianza y la raíz del error cuadrático medio, $E(R_{ij})$, $Var(R_{ij})$, $rmse(i,j) = [E(R_{ij} - O_{ij})^2]^{½}$. Aquí, como en toda esta sección, R_{ij} representa la imagen filtrada, salida del filtro, mientras que O_{ij} es la imagen original, en este caso el borde de escalón de altura h.

Notemos por m_{ij} el número de puntos pertenecientes a A en la ventana W_{ij}. Evidentemente, este número depende de la posición de la ventana sobre el borde. El número de puntos pertenecientes a la región B será de $2N + 1 - m_{ij}$.

Proposición 3. *La función de distribución $\Psi_{Rij}(t)$ del filtro de mediana relajada para el borde ruidoso defini-*

do anteriormente viene dada por:

$$\Psi_{R_{ij}}(t) = \sum_{k=m-1}^{2N} \sum_{n_0=max\{0,k-(2N-d_{ij})\}}^{min\{k,d_{ij}\}} \binom{d_{ij}}{n_0}\binom{2N-d_{ij}}{k-n_0} \cdot$$

$$\cdot \Phi_A(t)^{n_0} \cdot \left[1-\Phi_A(t)\right]^{d_{ij}-n_0} \cdot$$

$$\cdot \Phi_B(t)^{k-n_0} \cdot \left[1-\Phi_B(t)\right]^{2N-d_{ij}-k+n_0} \cdot \Phi_{ij}(t)$$

$$+ \sum_{k=l-1}^{m-2} \sum_{n_0=max\{0,k-(2N-d_{ij})\}}^{min\{k,d_{ij}\}} \binom{d_{ij}}{n_0}\binom{2N-d_{ij}}{k-n_0}\left(\frac{k-l+2}{k-n_0}\right) \cdot$$

$$\cdot \Phi_A(t)^{n_0} \cdot \left[1-\Phi_A(t)\right]^{d_{ij}-n_0} \cdot$$

$$\cdot \Phi_B(t)^{k-n_0} \cdot \left[1-\Phi_B(t)\right]^{2N-d_{ij}-k+n_0} \cdot \Phi_{ij}(t)$$

$$+ \sum_{k=uh=m}^{2N} \sum^{k} \sum_{n_0=max\{0,h-(2N-d_{ij})\}}^{min\{h,d_{ij}\}} \binom{d_{ij}}{n_0}\binom{2N-d_{ij}}{h-n_0}$$

$$\left(\frac{1}{2N+1-h}\right)\Phi_A(t)^{n_0} \cdot \left[1-\Phi_A(t)\right]^{d_{ij}-n_0} \cdot$$

$$\cdot \Phi_B(t)^{h-n_0} \cdot \left[1-\Phi_B(t)\right]^{2N-d_{ij}-h+n_0} \cdot \left[1-\Phi_{ij}(t)\right]$$

donde

$$\Phi_{ij}(t) = \begin{cases} \Phi_A(t) & \Leftrightarrow j \le 0 \\ \Phi_B(t) & \Leftrightarrow j \ge 1 \end{cases}$$

y

$$d_{ij} = \begin{cases} m_{ij}-1 & \Leftrightarrow j \le 0 \\ m_{ij} & \Leftrightarrow j \ge 1 \end{cases}$$

Demostración:

A partir de la expresión de $\Psi(t)$, se deduce el de la distribución pedida, que puede escribirse como:

$$\Psi_{R_{ij}}(t) = \left(\sum_{k=m-1}^{2N} \Pr\left[\text{exact. k puntos en } \hat{W}_{ij} \text{ son } \leq t\right] \right)$$

$$\cdot \Pr\left[S_{ij} \leq t\right]$$

$$+ \left(\sum_{k=l-1}^{m-2} \Pr\left[\text{exact. k puntos en } \hat{W}_{ij} \text{ son } \leq t\right] \cdot \right.$$

$$\left. \cdot \left(\frac{k-l+2}{k+1} \right) \right) \cdot \Pr\left[S_{ij} \leq t\right]$$

$$+ \left(\sum_{k=u}^{2N} \sum_{h=m}^{k} \Pr\left[\text{exact. h puntos en } \hat{W}_{ij} \text{ son } \leq t\right] \right.$$

$$\left. \cdot \left(\frac{1}{2N+1-h} \right) \right) \cdot \Pr\left[S_{ij} > t\right]$$

Ahora bien, a partir de la definición del borde ruidoso resulta que:

$$\Pr\left[\text{exact. k puntos en } \hat{W}_{ij} \text{ son } \leq t\right] =$$

$$= \sum_{n_0 = max\{0, k-(2N-d_{ij})\}}^{min\{k, d_{ij}\}} \Pr\left[\text{exact. } n_0 \text{ puntos } \in A \text{ son } \leq t, \right.$$

$$\left. \text{exact. } k - n_0 \text{ puntos } \in B \text{ son } \leq t\right]$$

$$= \sum_{n_0 = max\{0, k-(2N-d_{ij})\}}^{min\{k, d_{ij}\}} \binom{d_{ij}}{n_0} \binom{2N-d_{ij}}{k-n_0} \cdot$$

$$\cdot \Phi_A(t)^{n_0} \left[1 - \Phi_A(t)\right]^{d_{ij}-n_0} \cdot$$

$$\cdot \Phi_B(t)^{k-n_0} \left[1 - \Phi_B(t)\right]^{2N-d_{ij}-k+n_0}$$

donde d_{ij} es el número de puntos pertenecientes a la ventana $\hat{W}_{ij}$.

Obsérvese que n_0 varía desde $max\{0,k-(2N-d_{ij})\}$ hasta $min\{k,d_{ij}\}$ porque por un lado $n_0 \leq k$ y $n_0 \leq d_{ij}$ implica que $n_0 \leq min\{k,d_{ij}\}$, y por otro lado $k - n_0 \leq k$ y $k - n_0 \leq 2N - d_{ij}$ implica que $n_0 \geq max\{0,k-(2N-d_{ij})\}$.
Por otra parte se tiene que:

$$\Pr[S_{ij} \leq t] = \Phi_{ij}(t) \quad \text{y} \quad \Pr[S_{ij} > t] = 1 - \Phi_{ij}(t)$$

Con tales resultados se demuestra lo pedido.

■

Usando esta función de distribución recién calculada, se puede conocer el valor de $E(R_{ij})$, $Var(R_{ij})$, $rmse(i,j) = [E(R_{ij} - O_{ij})^2]^{½}$ del filtro de mediana relajada mediante integración numérica para un ruido de función de distribución conocida $\Phi_A(t)$. Como ejemplo de aplicación supongamos que $\Phi_A(t) = N(0,1)$, esto es, la distribución normal de media nula y varianza unidad. Integraremos a lo largo de un recorrido horizontal de filtro (*i* fija; *j* variable). Las figuras 6, 7 y 8 muestran los valores de $E(R_{ij})$, $Var(R_{ij})$ y $rmse(i,j)$ para el filtro MR de ventana 3×3 con diferentes valores de intervalo de bondad, suponiendo que el borde es un escalón de altura $h = 4$, degradado con un ruido gaussiano $N(0,1)$. Es interesante observar cómo el $MR_{2,8}$ tiene los valores esperados más cercanos al borde ideal, pero posee los errores $rmse(i,j)$ más elevados, debido a su pobre capacdad de suprimir ruido. A medida que el intervalo de bondad se reduce, los valores esperados se separan más del comportamiento ideal, haciendo que el escalón sea gradual, acercándose más a la media local del filtro de mediana. Obsérvese cómo $MR_{4,6}$ y la mediana tradicional son casi coincidentes, indicando que su comportamiento frente al ruido es muy similar, aunque $MR_{4,6}$ preservará el borde mejor. Ello queda patente en la figura 8, donde se aprecia que el $rmse(i,j)$ del filtro de mediana es peor que todos los MR (excepto el $MR_{2,8}$) en la región del borde. En el borde el error $rmse(i,j)$ de la mediana es el doble que en la región de gris constante. En cualquier caso puede decirse a la luz de la figura 6 que todos los filtros analizados preservan el borde de una forma similar. Las varianzas locales, que se muestran en la figura 7, se manifiestan mejores en la zona del borde, al igual que el $rmse(i,j)$, evidenciando que es en esa región dónde el comportamiento de los filtros es peor. La menor varianza es la del filtro de mediana, debido principalmente a que emplea más píxeles dentro de la ventana deslizante que el resto de los filtros en estudio. Como cabía esperar, cuanto mayor es el intervalo de bodad, peor la capacidad de supresión de ruido del filtro MR.

Resulta curioso observar que aunque la variaza de la mediana es la menor de todas las estudiadas, su $rmse(i,j)$ es muy elevado. Ello es achacable al gran error de sesgo presente al

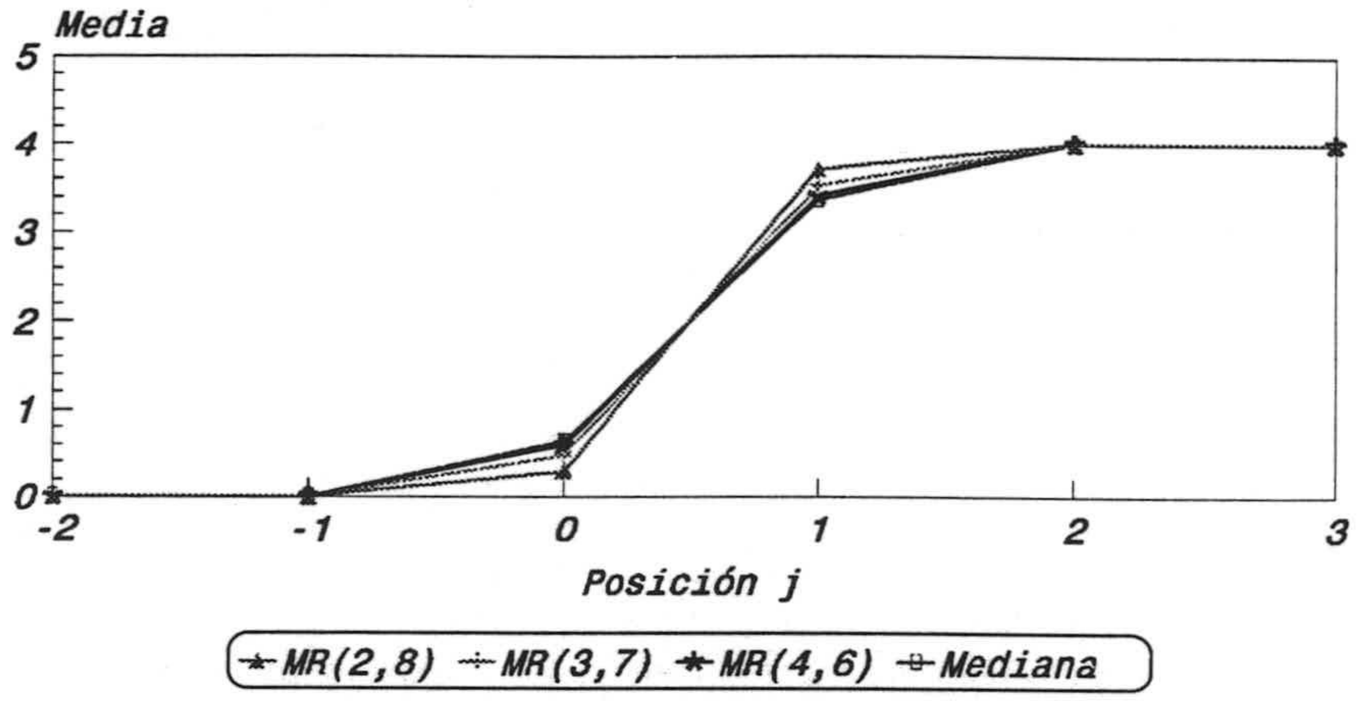


Figura 6: Resultados del cálculo numérico de la media de la salida de los filtros MR y mediana de ventana 3×3, cuando la entrada es un borde de escalón $h = 4$. Se ha variado el intervalo de bondad en el filtro de mediana relajada.

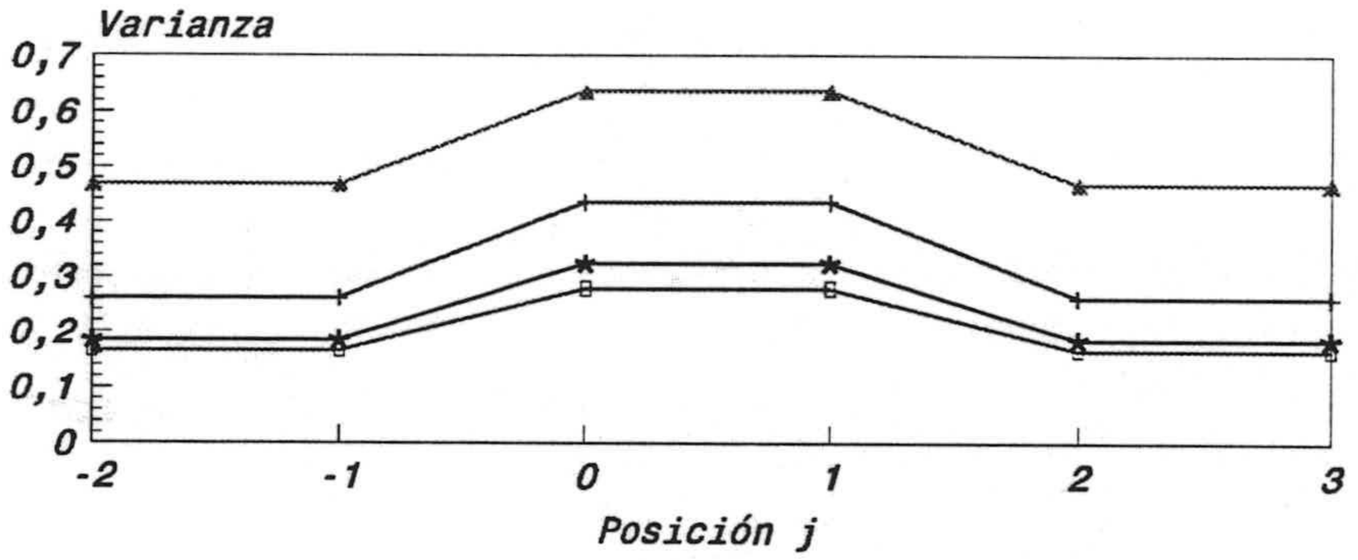


Figura 7: En este gráfico se muestra la varianza local de salida cuando la entrada es un borde abrupto de escalón $h = 4$. Se contempla tanto el caso de MR con diferentes intervalos de bondad como el del filtro de mediana tradicional. La ventana es de tamaño 3×3.

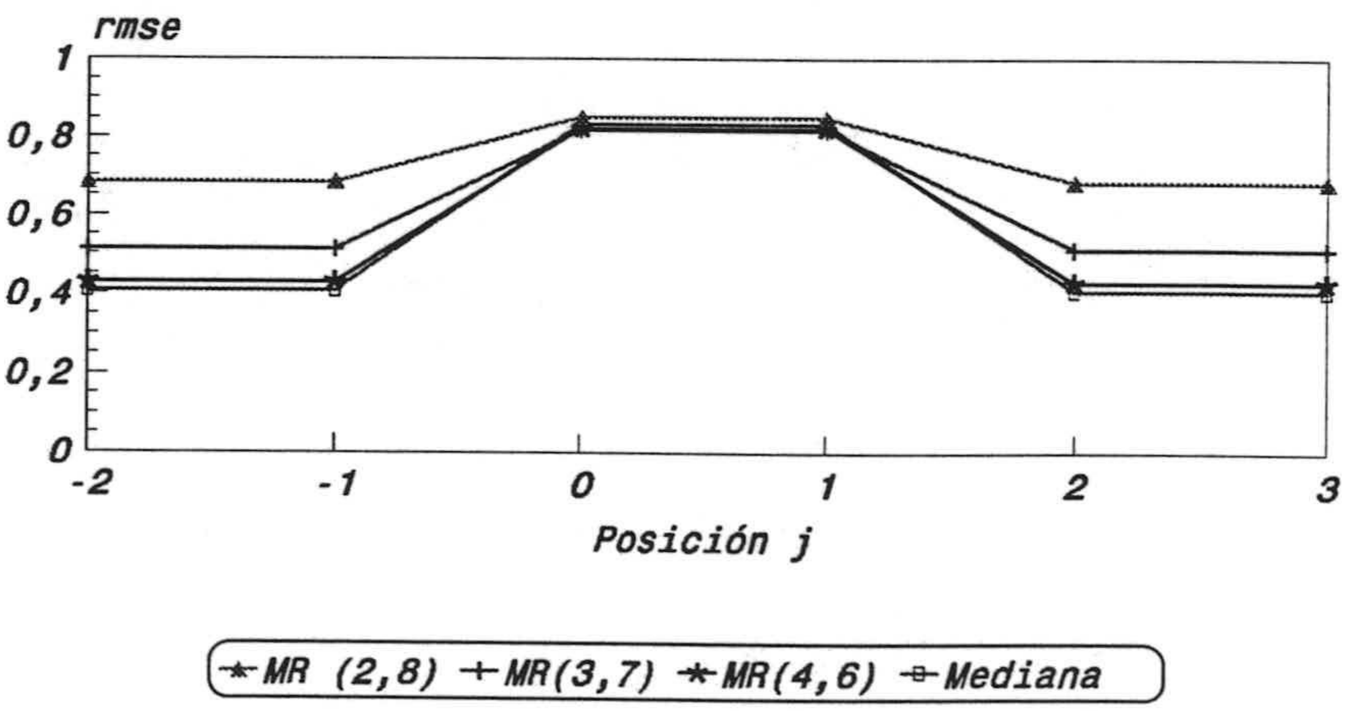


Figura 8: Aquí aparecen representados los valores obtenidos para la raíz del error cuadrático medio (*rmse*) tanto para MR con distintos intervalos de bondad como para el filtro de mediana tradicional, ambos con ventana 3×3. Dicho error da cuenta de la diferencia existente entre el borde abrupto original (h=4) y la señal resultante.

estar la ventana a caballo entre dos regiones de nivel de gris muy diferentes. Este error afecta menos a los filtro MR, al emplear menos píxeles de la ventana, y por ello no se aprecia una diferencia tan grande entre los comportamientos de su varianza y el error cuadrático medio.

PRESERVACIÓN DE LÍNEAS

Consideremos la línea ruidosa S_{ij} definida por:

$$S_{ij} = \begin{cases} \nu_{ij} & \Leftrightarrow j \neq 0 \\ h + \nu_{ij} & \Leftrightarrow j = 0 \end{cases}$$

dónde h es una constante que representa la altura de la línea, y ν_{ij} es un ruido *i.i.d.* con función de distribución $\Phi_A(t)$ que corresponde a la región que circunda la línea, que denotamos por A. La función $\Phi_B(t)$ representa la distribución de la variable $h + \nu_{ij}$ en la región comprendida por la línea, que denominamos B. Está claro que, como en el caso del borde, $\Phi_B(t) = \Phi_A(t-h)$.

Proposición 4. *La función de distribución $\Psi_{Rij}(t)$ del filtro de mediana relajada para la línea ruidosa definida anteriormente viene dada por:*

$$\Psi_{R_{ij}}(t) = \sum_{k=m-1}^{2N} \sum_{n_0=max\{0,k-(2N-d_{ij})\}}^{min\{k,d_{ij}\}} \binom{d_{ij}}{n_0}\binom{2N-d_{ij}}{k-n_0}\cdot$$

$$\cdot \Phi_A(t)^{n_0} \cdot [1-\Phi_A(t)]^{d_{ij}-n_0} \cdot$$

$$\cdot \Phi_B(t)^{k-n_0} \cdot [1-\Phi_B(t)]^{2N-d_{ij}-k+n_0} \cdot \Phi_{ij}(t)$$

$$+ \sum_{k=l-1}^{m-2} \sum_{n_0=max\{0,k-(2N-d_{ij})\}}^{min\{k,d_{ij}\}} \binom{d_{ij}}{n_0}\binom{2N-d_{ij}}{k-n_0}\binom{k-l+2}{k-n_0}\cdot$$

$$\cdot \Phi_A(t)^{n_0} \cdot [1-\Phi_A(t)]^{d_{ij}-n_0} \cdot$$

$$\cdot \Phi_B(t)^{k-n_0} \cdot [1-\Phi_B(t)]^{2N-d_{ij}-k+n_0} \cdot \Phi_{ij}(t)$$

$$+ \sum_{k=uh=m}^{2N} \sum^{k} \sum_{n_0=max\{0,h-(2N-d_{ij})\}}^{min\{h,d_{ij}\}} \binom{d_{ij}}{n_0}\binom{2N-d_{ij}}{h-n_0}$$

$$\left(\frac{1}{2N+1-h}\right)\Phi_A(t)^{n_0} \cdot [1-\Phi_A(t)]^{d_{ij}-n_0} \cdot$$

$$\cdot \Phi_B(t)^{h-n_0} \cdot [1-\Phi_B(t)]^{2N-d_{ij}-h+n_0} \cdot [1-\Phi_{ij}(t)]$$

donde:

$$\Phi_{ij}(t) = \begin{cases} \Phi_A(t) & \Leftrightarrow j \neq 0 \\ \Phi_B(t) & \Leftrightarrow j = 0 \end{cases}$$

y

$$d_{ij} = \begin{cases} m_{ij} - 1 & \Leftrightarrow j \neq 0 \\ m_{ij} & \Leftrightarrow j = 0 \end{cases}$$

Demostración:

El procedimiento es similar al de la proposición 3, excepto que ahora:

$$\Phi_{ij}(t) = \begin{cases} \Phi_A(t) & \Leftrightarrow j \neq 0 \\ \Phi_B(t) & \Leftrightarrow j = 0 \end{cases}$$

y

$$d_{ij} = \begin{cases} m_{ij} - 1 & \Leftrightarrow j \neq 0 \\ m_{ij} & \Leftrightarrow j = 0 \end{cases}$$

■

Empleando la expresión de la función de densidad que acabamos de establecer se puede calcular el valor de $E(R_{ij})$, $Var(R_{ij})$, $rmse(i,j) = [E(R_{ij} - O_{ij})^2]^{1/2}$ del filtro de mediana relajada mediante integración numérica para ruido de función de distribución $\Phi_A(t)$ conocida. Para analizar la capacidad de los filtros MR para preservar líneas estrechas afectadas de ruido gaussiano $N(0,1)$ se ha diseñado, además de la línea binaria ruidosa de un píxel de grosor ya definida, una línea binaria ruidosa con dos píxeles de grosor, cuya definición se sigue de la anterior.

Las medias locales de las distribuciones de probabilidad de la salida del filtro MR con ventana 3×3 para la línea de un píxel se muestran en la figura 9. Mientras la mediana elimina completamente la línea, los filtros MR la preservan tanto mejor cuanto más amplio es su intervalo de bondad, como es fácil de comprender.

Las varianzas locales y los $rmse(i,j)$ de las distribuciones de salida se muestran en las figuras 10,11. De nuevo la varianza de la mediana es la más pequeña aunque su $rmse(i,j)$ es insatisfactoriamente elevado debido al gran error de sesgo. En

Figura 9: En esta gráfica se muestra la media de la salida de los filtros MR y mediana tradicional de ventana 3×3, cuando la entrada es una línea vertical de 1 píxel de grosor. Se ha variado el intervalo de bondad del filtro de mediana relajada para mostrar la evolución del comportamiento.

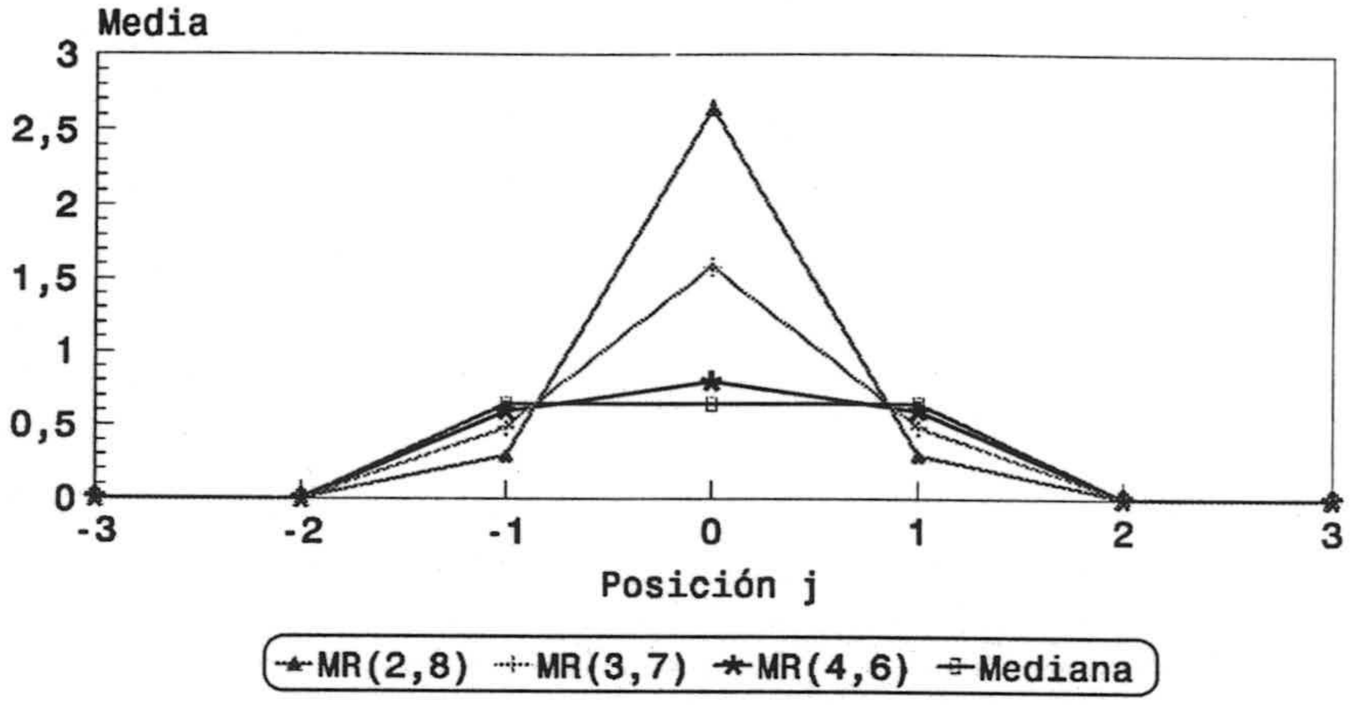


Figura 10: Se presentan los valores calculados por integración numérica de la varianza de salida de los filtros MR y mediana tradicional de ventana 3×3. De nuevo la entrada es una línea de 1 píxel de grosor. Obsérvese la influencia de la variación del intervalo de bondad en el filtro MR.

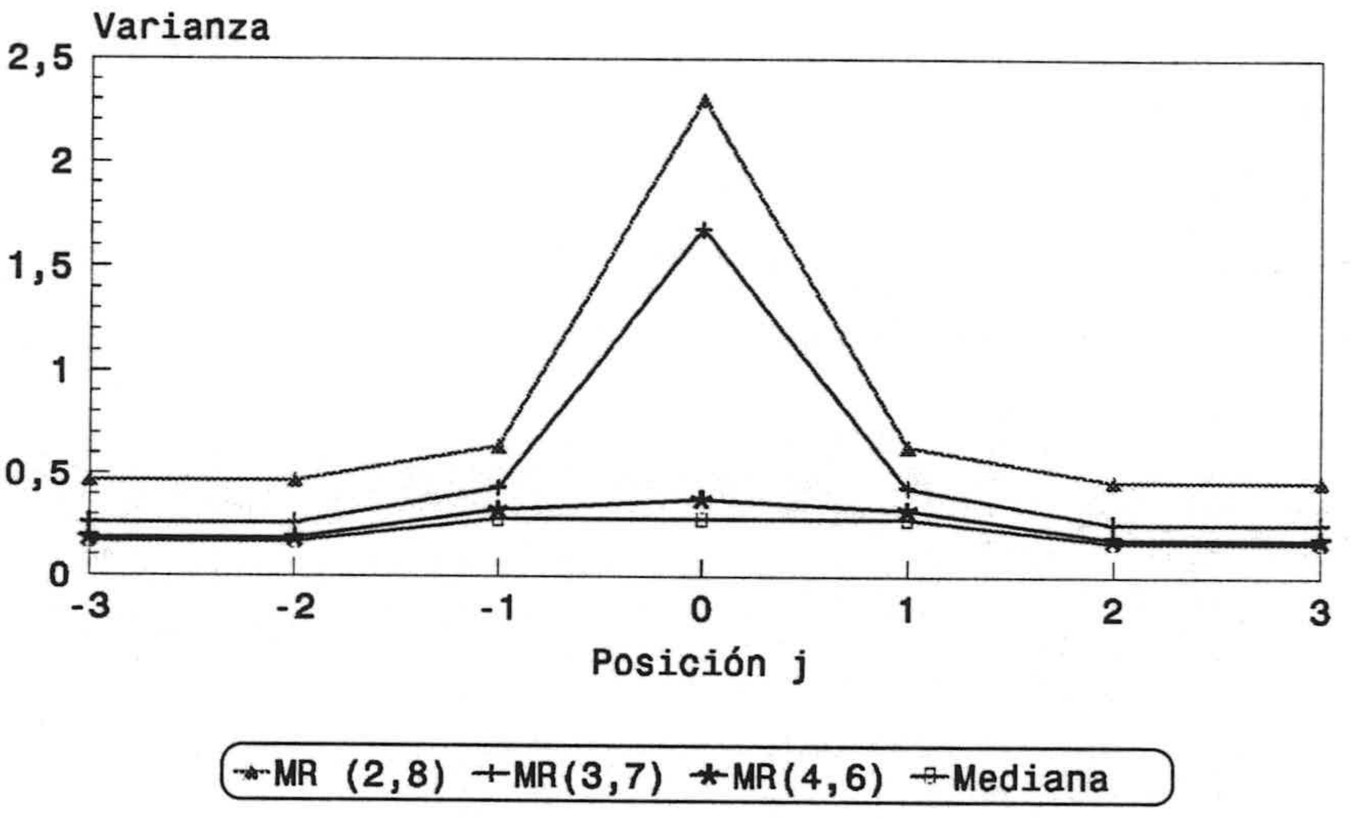
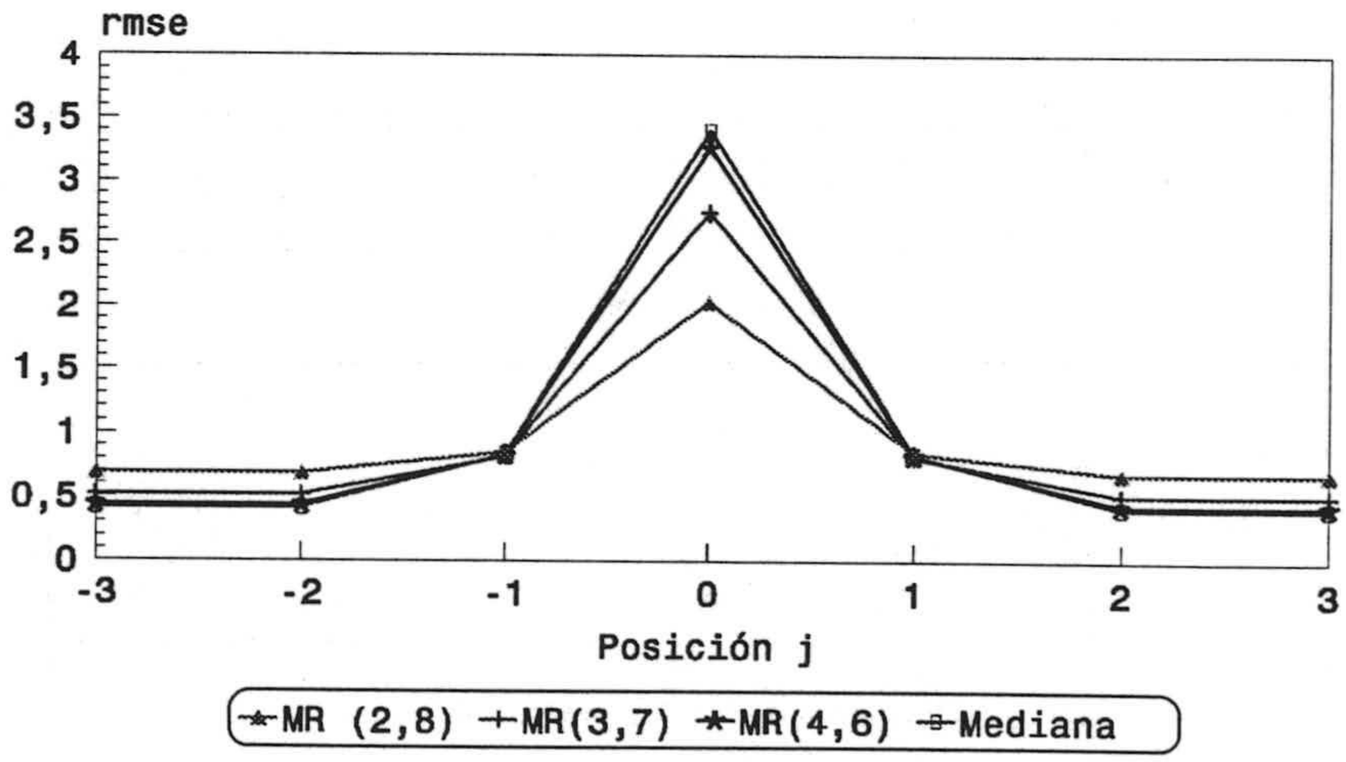


Figura 11: Siguiendo con el ejemplo de la línea de 1 píxel de grosor, aquí se muestra la raiz del error medio cuadrático para el filtro de mediana tradicional y el MR con distintos intervalos de bondad. El tamaño de la ventana es 3×3.



la línea ($j = 0$) el $rmse(i,j)$ del $MR_{2,8}$ es el menor. Existe una diferencia escasa entre la mediana tradicional y el $MR_{4,6}$.

Los valores esperados de las distribuciones de salida para la línea de dos píxeles se muestran en la figura 12. Los filtros MR son los que mejor preservan la imagen. El filtro de mediana decrece la amplitud de la línea en un 16.25%, y la ensancha levemente. Las varianzas locales y $rmse(i,j)$ se muestran en las figuras 13 y 14. En la línea ($j = 0,1$) el peor resultado es para $MR_{2,8}$, seguido de la mediana, que se muestra peor que el resto de los filtros MR, aunque las diferencias son pequeñas.

OTRAS PROPIEDADES

Pueden añadirse otras propiedades a las ya indicadas, muy importantes. Algunas de ellas muestran mejoras significativas frente al filtro de mediana. Otras no tanto.

APILAMIENTO O *STACKING*

Puede verse fácilmente que el filtro de mediana relajada no opera de igual forma que los llamados filtros *stack* o apilados [WEN86]. Basta considerar el contraejemplo que se muestra en la tabla 3, en dónde se filtra una señal monodimensional con una ventana de tamaño 5 y un intervalo de bondad comprendido entre $l = 2$ y $u = 4$.

Filtro	Señal Fuente	Resultado
MR	12O23	2
Stack	OOOO1	O
	O1O11	O
	11O11	1
		1

Tabla 3: Se muestra con este ejemplo que el filtro de mediana relajada no es apilable. Se propone una señal de entrada sencilla cuyo filtrado por MR no puede ser igualado mediante apilamiento.

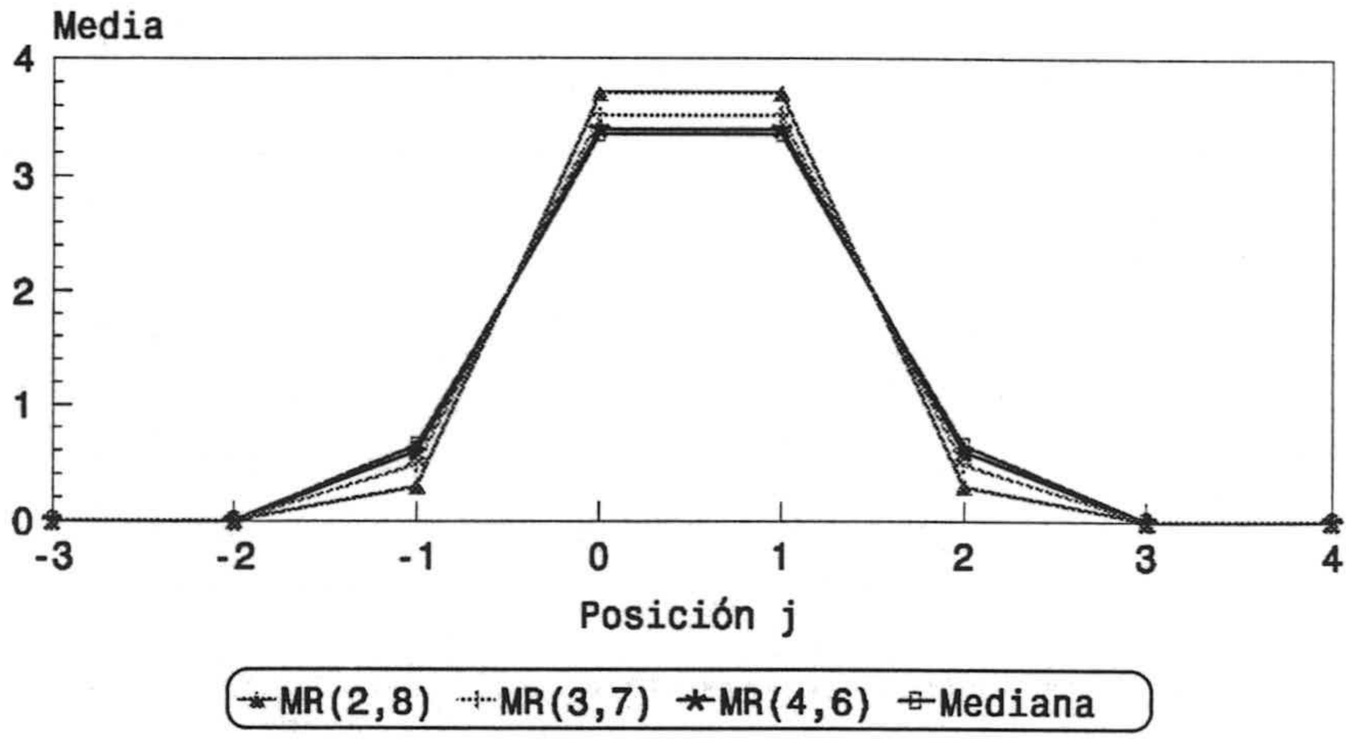


Figura 12: Resultado de calcular la media de la salida de los filtros de mediana y MR de ventana 3×3 para una señal e entrada compuesta por una línea de 2 píxeles de grosor. Se muestran resultados para diferentes intervalos de bondad.

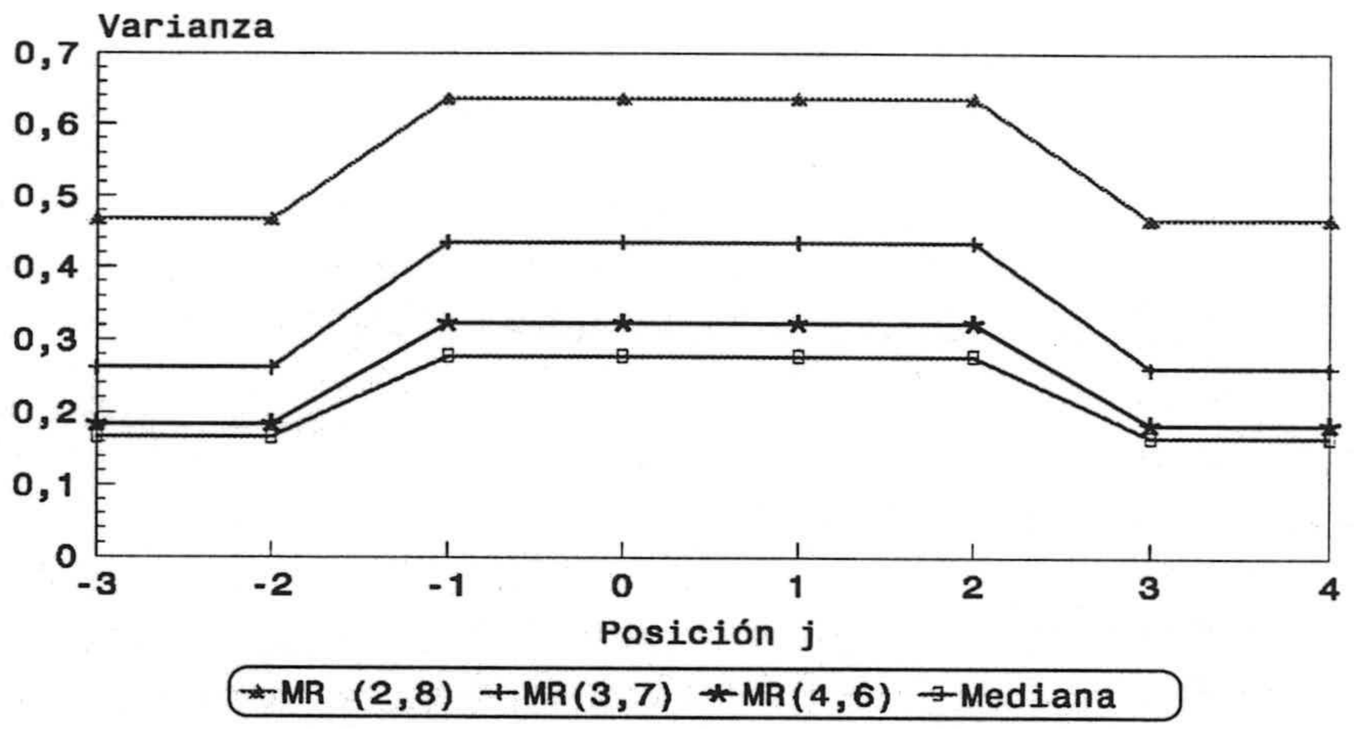


Figura 13: Varianza local de salida del filtro de mediana tradicional y relajada, de tamaño 3×3 y diferentes intervalos de bondad. La entrada es una línea de 2 píxeles de grosor.

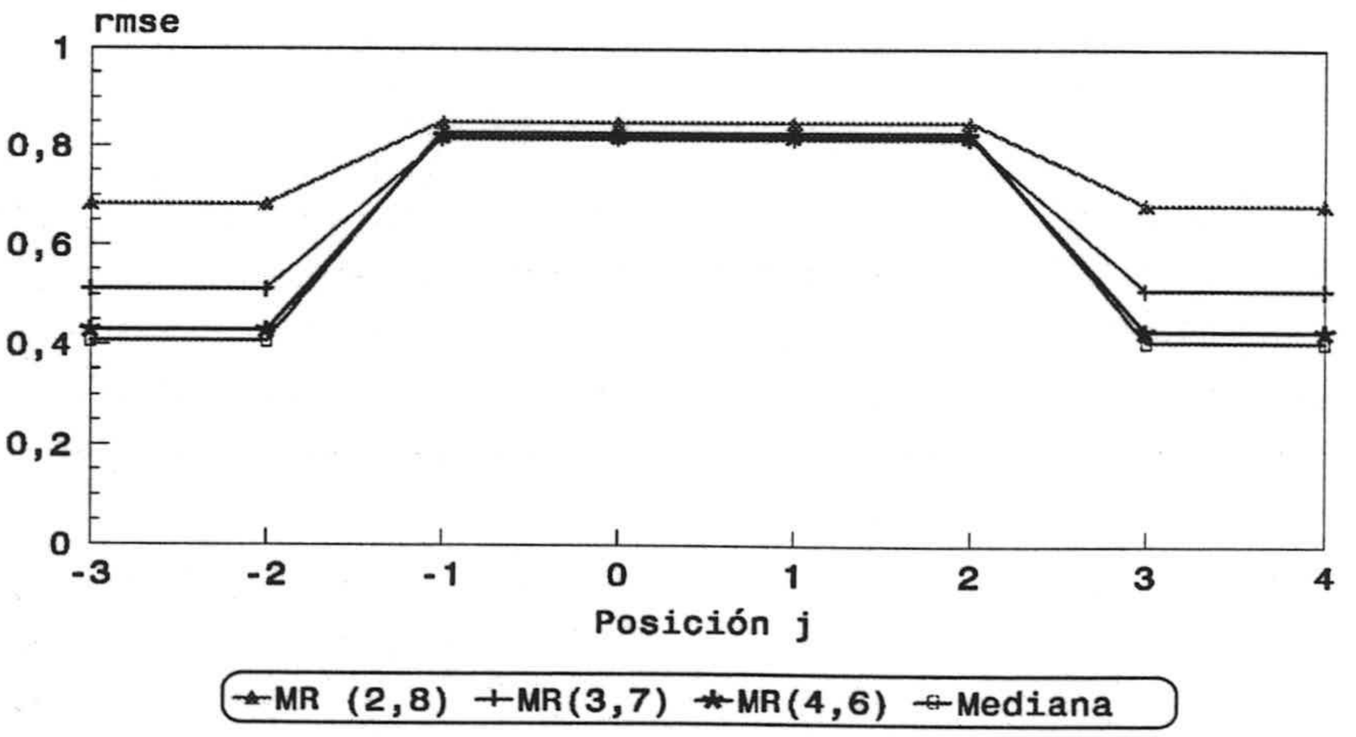


Figura 14: Representación gráfica de los valores calculados para la raiz del error medio cuadrático de la salida de los filtros de mediana tradicional y MR cuando la imagen original es una línea de 2 píxeles de grosor. La ventana de filtro tiene tamaño 3×3. Aparecen resultados para diferentes intervalos de bondad.

BLOTCHES

El filtro de mediana no sólo suaviza el ruido en regiones homogéneas de la imagen, sino que también tiende a producir regiones de intensidad constante o casi constante. La forma de esas regiones depende de la geometría de la ventana deslizante empleada por el filtro. Suelen tener la apariencia de "parches" lineales o manchas amorfas (*blotches*). Estos efectos son indeseables, pues se perciben como líneas o contornos entre *blotches* que no existen en la imagen original. Pues bien, el filtro de mediana relajada presenta este efecto muy minimizado.

Para realizar un estudio teórico del problema, Pitas y Venetsanopoulos [PIT90] han propuesto una medida de este efecto basado en la probabilidad, bajo ciertas hipótesis, de que el filtro de mediana proporcione la misma salida para dos píxeles diferentes, en función de la distancia que los separa.

Este, o cualquier otro tipo de análisis puede emplearse en el filtro de mediana relajada, pero teniendo en cuenta una diferencia primordial con el filtro de mediana habitual: ahora hay píxeles que no son filtrados. Cualquier medida del efecto de generación de *blotches* válida para dos píxeles empleando el filtro de mediana se reducirá en nuestro caso por la probabilidad de que esos dos mismos píxeles sean filtrados o no. De esta manera, si $Pr_{Med}(R_{ij} = R_{i+d,\, j+f})$ es la probabilidad de que el filtro de mediana proporcione dos salidas iguales para dos píxeles separados por una distancia horizonta d y una vertical f, entonces la misma probabilidad para el filtro MR será:

$$Pr_{MR}(R_{ij} = R_{i+d,j+f}) = Pr_{Med}(R_{ij} = R_{i+d,j+f}) \cdot Pr(R_{ij} \text{ y } R_{i+d,j+f} \text{ sean filtrados})$$

La probabilidad de que esos píxeles sean filtrados puede calcularse de manera inmediata si se asume la hipótesis simplificadora de que todos los píxeles afectados de ruido van a ser filtrados. Coo el ruido impulsivo es espacialmente independiente, si p es la probabilidad de filtrar un píxel dado, la probabilidad de filtrar dos es p^2. Así, para una imagen que esté contaminada por un 30% de ruido impulsivo, el factor de reducción del efecto de generación de *blotches* para el filtro MR en relación con el de mediana tradicional es $Pr(R_{ij} \text{ y } R_{i+d,j+f} \text{ sean filtrados}) = p^2 = (0.3)^2 = 0.09$. Por supuesto, cabría incluir un factor de corrección que tuviera en cuenta el alejamiento de la hipótesis empleada que suele tener lugar en aplicaciones reales del filtro.

RAÍCES.

Una herramienta muy importante en el análisis de filtros lineales es el estudio del comportamiento estacionario de los mismos frente a señales de entrada senoidales. Como es bien sabido, la frecuencia de una señal senoidal no cambia al pasar a través de un filtro lineal.Sólo su fase y su amplitud varian. Pero este hecho no es válido para filtros no lineales, como la mediana o el MR. Sin embargo, aunque el método no es aplicable, no así la idea subyacente: si se encuentran señales que permanezcan inalteradas al pasar po un filtro no lineal, se podrá caracterizar al mismo. Tales señales invariantes se denominan *raíces* del filtro [ARC82][FIT85]. Su significado para filtros de mediana es el mismo que las señales senoidales en el pasabanda de los filtros lineales. Estudiando las raíces puede obtenerse información cualitativa acerca del efecto del filtro sobre una secuencia de entrada. Por ello, se ha hecho un gran esfuerzo durante los últimos años en este sentido por parte de numerosos investigadores, tanto para el filtro de mediana como para otros tios de filtros no lineales (*stack*, multietapa ...).

El estudio de las raíces en el caso del filtro MR está orientado sobre todo a comprender su habilidad en preservar líneas, esquinas y otros detalles. Como ya se ha comentado, para el filtro de mediana cualquier región de nivel de gris constante dentro de otra región de nivel de gris constante diferente es una raíz, siempre que su tamaño sea mayor que un determinado umbral. Pues bien, esto también sigue siendo válido en el filtro de mediana relajada, aunque en este caso el umbral es menor. Por tanto, el algoritmo no eliminará ruido que aparezca en tal cantidad que aparezca agrupado en regiones suficientemente grandes para ser consideradas raíces para el filtro MR.

Desgraciadamente, el estudio completo y riguroso de las raíces para el filtro MR no puede llevarse a cabo, al menos en los términos convencionales, puesto que este filtro no es apilable, como ya se ha apuntado. Y es que cuando un filtro no lineal es apilable, entonces basta encontrar las raíces para señales bidimensionales para conocer todas las raíces posibles del filtro. Pero no estamos en esas condiciones, así que el abordaje del problema dista de estar al alcance de nuestra mano. En cualquier caso, para tener una idea del comportamiento general del algoritmo, podemos realizar un análisis heurístico: primero con-

sideraremos señales binarias bidimensionales para a continuación filtrar imágenes más complejas. Los resultados nos reportarán la información que buscamos.

Características básicas de las raíces.

Arce *et al.* han analizado en detalle las propiedades deterministas del filtro de mediana monodimensional [ARC83], aunque las propiedades del caso bidimensional no están completamente entendidas. Las señales bidimensionales compuestas de áreas uniformes y bordes monótonos rectos entre ellas son raíces para el filtro de mediana de ventana $(2k + 1) \times (2k + 1)$ si las regiones uniformes tienen una anchura mínima de $k + 1$ píxeles. Esto mismo también es válido para líneas o detalles de $k + 1$ píxeles de grosor. Por tanto, para un filtro de mediana de tamaño 5×5 no es posible preservar líneas o detalles de uno o dos píxeles de grosor.

Una señal compuesta de regiones monótonas separadas por bordes rectilíneos será raíz de un filtro de mediana relajada monodimensional de tamaño $(2k + 1)$ y límites de intervalo $\boldsymbol{l}$ y $\boldsymbol{u}$ si tales regiones uniformes tienen una anchura de al menos $(2k + 1) - u + 1 = l$.

Proposición 5. *El filtro de mediana relajada monodimensional preserva detalles más sutiles que el filtro de mediana monodimensional de igual tamaño.*

Demostración.
Se tiene, por definición de límites de intervalo, que

$$l + u = 2k + 2 \qquad k + 1 = \tfrac{1}{2}(l + u)$$

La media aritmética de dos valores está comprendida entre ambos, siempre que $l \quad u$ (nótese que el caso contrario implica que se está empleando la mediana tradicional), por lo tanto

$$l < \tfrac{1}{2}(l + u) < u$$

expresión que, teniendo en cuenta la anterior igualdad, adquire esta forma:

$$l < k + 1 < u$$

Por otro lado, la condición de raíz para el filtro de mediana relajada es que el ancho de la región uniforme ha de ser al menos l; para el filtro de mediana de igual tamaño esta anchura debe ser como mínimo de $k + 1$. Por tanto, en virtud de la última desigualdad, el filtro de mediana relajada monodimensional conservará regiones y detalles de menor tamaño, y por tanto más sutiles, que el filtro de mediana monodimensional de igual tamaño. ■

Una señal compuesta de regiones monótonas separadas por bordes rectilíneos será raíz de un filtro de mediana de tamaño $(2k + 1) \times (2k + 1) = (2k + 1)^2$ siempre que su anchura sea al menos de $k + 1$ [CITAAA]. Podemos demostrarlo fácilmente teniendo en cuenta que $(2k + 1)^2 = 2\,[2k\,(k + 1)] + 1 = 2N + 1$, y reduciéndonos al caso monodimensional, tomando ahora N en lugar de k y teniendo en cuenta la geometría de la ventana. En el caso de la mediana relajada del mismo tamaño, y límites de intervalo l y u, una señal compuesta de regiones monótonas separadas por bordes rectilíneos será raíz si tales regiones uniformes tienen una anchura $a \quad l/(2k + 1)$. Esto es válido también para líneas de grosor a. Así por ejemplo, un MR_{818} preserva líneas de grosor 2.

Proposición 6. *El filtro de mediana relajada MR bidimensional preserva detalles más sutiles que el filtro de mediana bidimensional de igual tamaño.*

Demostración.

En el caso bidimensional, y por definición de los límites del intervalo de bondad, se tiene que:

$$(2k + 1)^2 + 1 = l + u$$

Por lo tanto:

$$4k^2 + 4k + 2 = l + u \Rightarrow 2k^2 + 2k + 1 = \tfrac{1}{2}(l + u)$$

Empleando que la media aritmética verifica que:

$$l \leq \tfrac{1}{2}(l + u) \quad u$$

dándose la igualdad sólo en el caso de que $l = u$, que

corresponde con el filtro de mediana tradicional, tenemos que:

$$u > 2k^2 + 2k + 1 > l$$

Si dividimos por $(2k + 1)$ la desigualdad se transforma en:

$$\frac{u}{2k+1} > \frac{2k^2}{2k+1} + 1 > \frac{l}{2k+1}$$

Por otro lado, como $k > 0$, tenemos que $2k^2 < 2k^2 + k$, y por tanto:

$$\frac{2k^2}{2k+1} < k \Leftrightarrow \frac{2k^2}{2k+1} + 1 < k + 1$$

De este resultado, junto con el anterior, se deduce que:

$$k + 1 > \frac{2k^2}{2k+1} + 1 > \frac{l}{2k+1}$$

o, lo que es lo mismo:

$$k + 1 > \frac{l}{2k+1}$$

Queda, pues, demostrado que el filtro MR bidimensional deja inalterados detalles, líneas y regiones más sutiles que los que deja el filtro de mediana bidimensional de igual tamaño. ■

Podemos generalizar estos resultados diciendo que el filtro de mediana tradicional bidimensional de tamaño $(2k + 1) \times (2k + 1)$ preserva detalles siempre que éstos se extiendan en un área (entendiendo área por número de píxeles) mínima de $2k(k + 1) + 1 \equiv N + 1$. Para el caso del filtro MR el área mínima que han de ocupar los detalles que sean raíces del algoritmo ha de ser de l. Se puede demostrar fácilmente que $l < N + 1$, reduciendo el problema bidimensional a uno monodimensional.

Respuesta a imágenes prueba.

Respuesta a señales monodimensionales. Para comprender el funcionamiento real de este filtro, y conocer sus aplicaciones y

limitaciones, comenzaremos dando unos sencillos ejemplos numéricos, para concluir la sección con algunos filtrados de imágenes complejas. En todos los casos se dará una comparación con el filtro de mediana.

En la figura 15 se observa el filtrado de una señal monodimensional tanto con el filtro de mediana como con el filtro MR, para distintos tamaños de ventana e intervalo de bondad. A la vista de la misma, podemos realizar los siguientes comentarios:

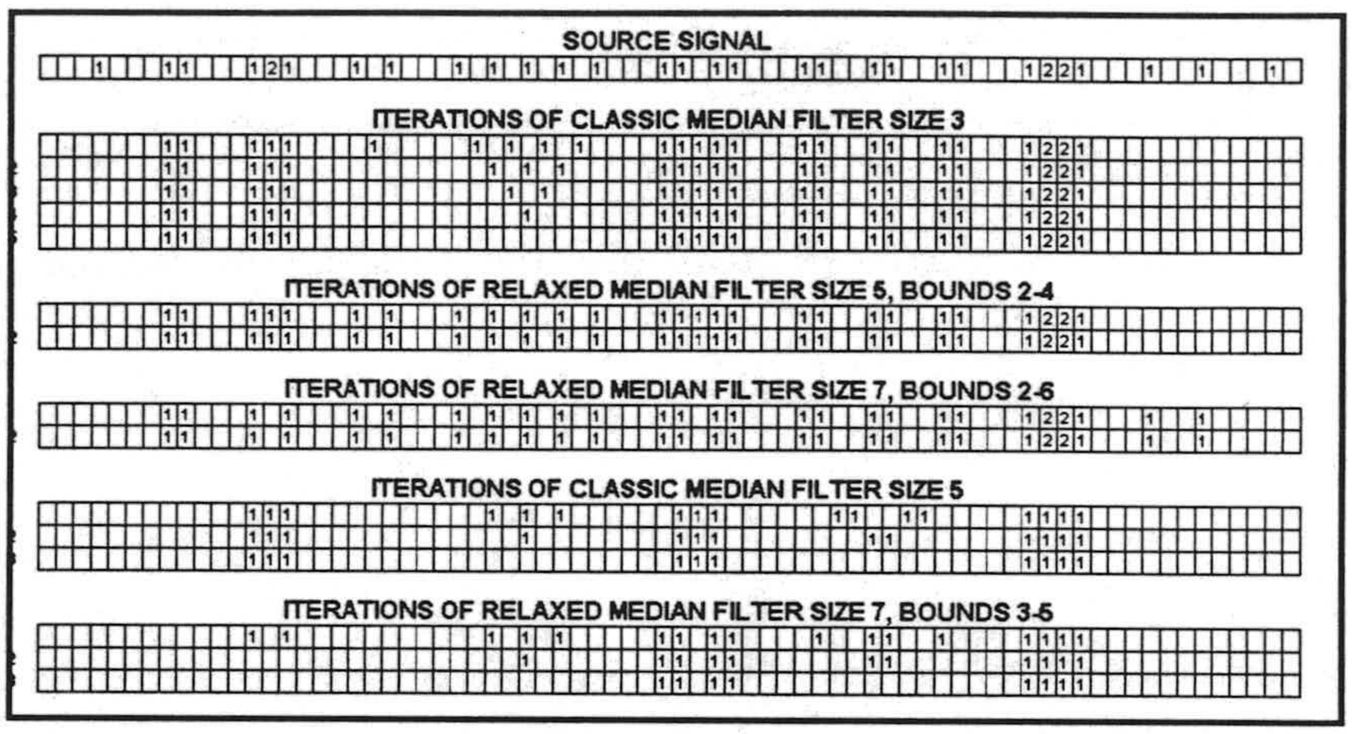


Figura 15: Se muestra en esta figura el comportamiento del filtro de mediana relajada y de mediana tradicional para diferentes señales de entrada.

- En general, se obtienen resultados similares para el filtro de mediana con ventana de tamaño 3×3 y los filtros de mediana relajada con ventana 5×5 e intervalo de bondad (2,4) y con ventana 7×7 e intervalo (2,6). Sólo difieren en detalles menores, como al decidir si un elemento está o no aislado , y por tanto siendo o no eliminado. Sin embargo, los filtros MR convergen a raíz más rápidamente que el filtro de mediana habitual, ya que aquellos respetan la alternación de los impulsos, incluso si sólo hay dos impulsos.

- El filtro de mediana relajada de tamaño de ventana 7×7 e intervalo (3,5) es comparable al filtro de mediana de tamaño de ventana 5×5. En cualquier caso, el primero respeta patrones mayores que el filtro de mediana.

Respuesta a imágenes prueba bidimensionales simples.

Estructura de las imágenes prueba: En esta sección examinaremos las propiedades deterministas del filtro MR y la mediana utilizando unas imágenes de prueba especialmente generadas. Todas tienen un tamaño de 54×54 píxeles. Para eliminar los efectos de borde analizaremos los algoritmos sólo en el área central de 50×50 píxeles de la imagen. Las imágenes prueba que se han empleado son las siguientes:

·*Imagen circunferencia*. Se trata de una circunferencia completa trazada sobre un fondo oscuro. El nivel de gris de la circunferencia es 255 y el del fondo es 0. Existen distintas imágenes de circunferencia, dónde se ha variado el grosor de la misma, de manera que se dispone de circunferencias de 1, 2, 3 y 4 píxeles de anchura. La elección de este tipo de figura viene justificada por la necesidad de estudiar el comportamiento del filtro frente a cualquier orientación del borde.

·*Imagen circunferencias concéntricas*. En este caso se ha construido una imagen en la que aparecen dos circunferencias concéntricas de grosor unidad, separadas también por un píxel. de nuevo se han trazado con un nivel de gris de 255 sobre un fondo de nivel de gris 0. La importancia de esta imagen radica en que será capaz de mostrar el comportamiento del filtro frente a bordes muy juntos.

·*Imagen de cuadrados concéntricos*. Para estudiar cómo actúa el algoritmo frente a la presencia de esquinas, se ha diseñado una imagen en la que aparecen dos cuadrados concéntricos sobre un fondo oscuro. El nivel de gris del cuadrado más interno es de 255, el del cuadrado mayor es de 119 y el del fondo es 0.

Todas estas imágenes quedan recogidas en la figura 16.

Análisis de las imágenes prueba filtradas: Centremos primeramente nuestra atención sobre las imágenes de circunferencias. Las de grosor de 3 y 4 píxeles son raíces tanto para el filtro de mediana como para el MR para cualquier tamaño de ventana empleado (3×3 y 5×5). Ello es debido a que la anchura de los bordes en la figura son siempre mayores o iguales que *n*

+ 1, siendo N tal que $(2N + 1)^2$ es el tamaño de la ventana del filtro de mediana. Sin embargo, para imágenes de circunferencias con grosor menor de 3 píxeles comienzan a aparecer diferencias entre el filtro de mediana tradicional y el aquí introducido.

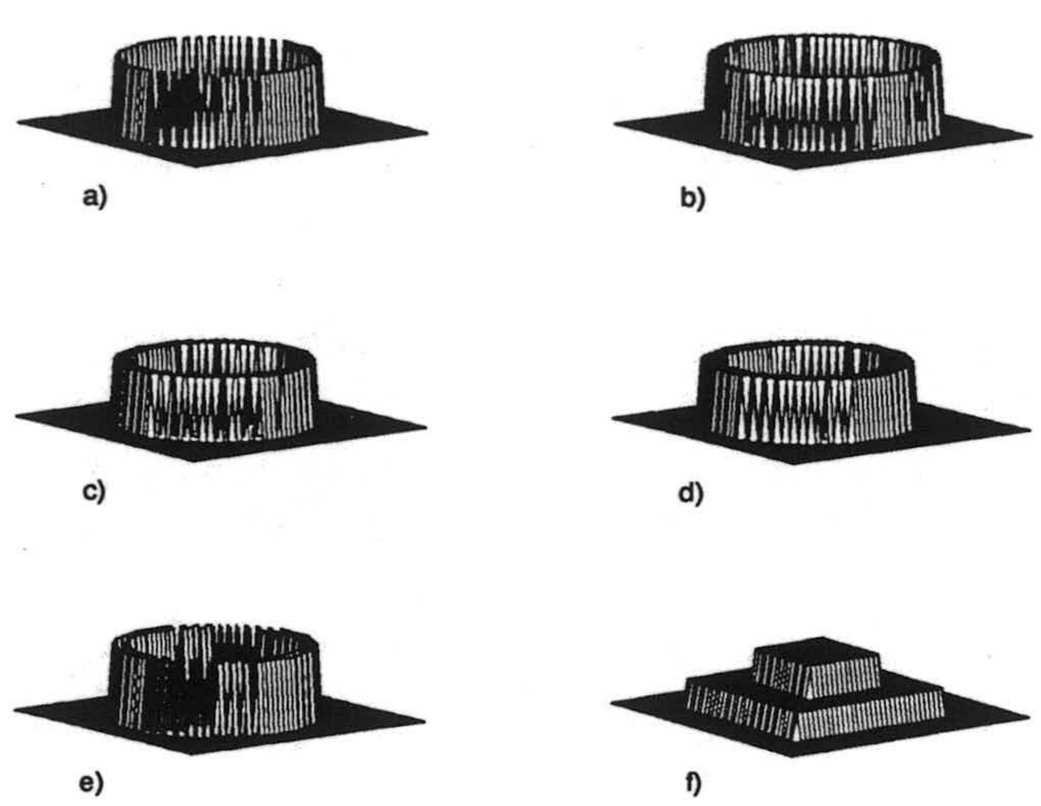


Figura 16: Imágenes originales que se han empleado en la realización de los experimentos encaminados al estudio de las raíces en el filtro MR. En *a), b), c)* y *d)* se observa la imagen circunferencia de 1,2,3 y 4 píxeles de anchura, respectivamente. En *e)*se muestra la imagen de circunferencias concéntricas y en *f)* la de cuadrados concéntricos.

Para la imagen de circunferencia de grosor de 2 píxeles, observamos un comportamiento invariante frente al filtro de mediana y MR, ambos con tamaño 3×3. De nuevo la anchura del borde es en este caso mayor o igual a $N + 1$. Pero cuando filtramos dicha imagen con ventana de tamaño 5×5 (figura 18) el filtro de mediana tradicional ya no la deja inalterada (al no verificar las condiciones para ello) sino que elimina casi completamente la circunferencia, dando como resultado una imagen prácticamente uniforme. Si repetimos el filtrado con el mismo tamaño de ventana, pero con el MR, encontramos que la imagen en cuestión es raíz para intervalos de bondad mayores o iguales a $(l, u) = (10,16)$. Para un intervalo más estrecho, como (11,15) ó (12,14) la imagen resultante queda alterada con respecto a la original, perdiéndose parte de la circunferencia, quedando sólo arcos de ella. Puede observarse que la destrucción de la imagen original se produce de manera simétrica, como cabía esperar de la geometría de los bordes. Aunque lo más interesante es darse cuenta que la zona que primero se elimina es la parte horizontal y vertical de la circunferencia, mientras

que permanecen las diagonales. Es posible justificar este comportamiento, si tenemos en cuenta las limitaciones gráficas del ordenador. Así, una circunferencia hay que trazarla sobre un entramado cuadriculado de píxeles. Este hecho condiciona la manera en que habrá de dibujarse la circunferencia, sobre todo

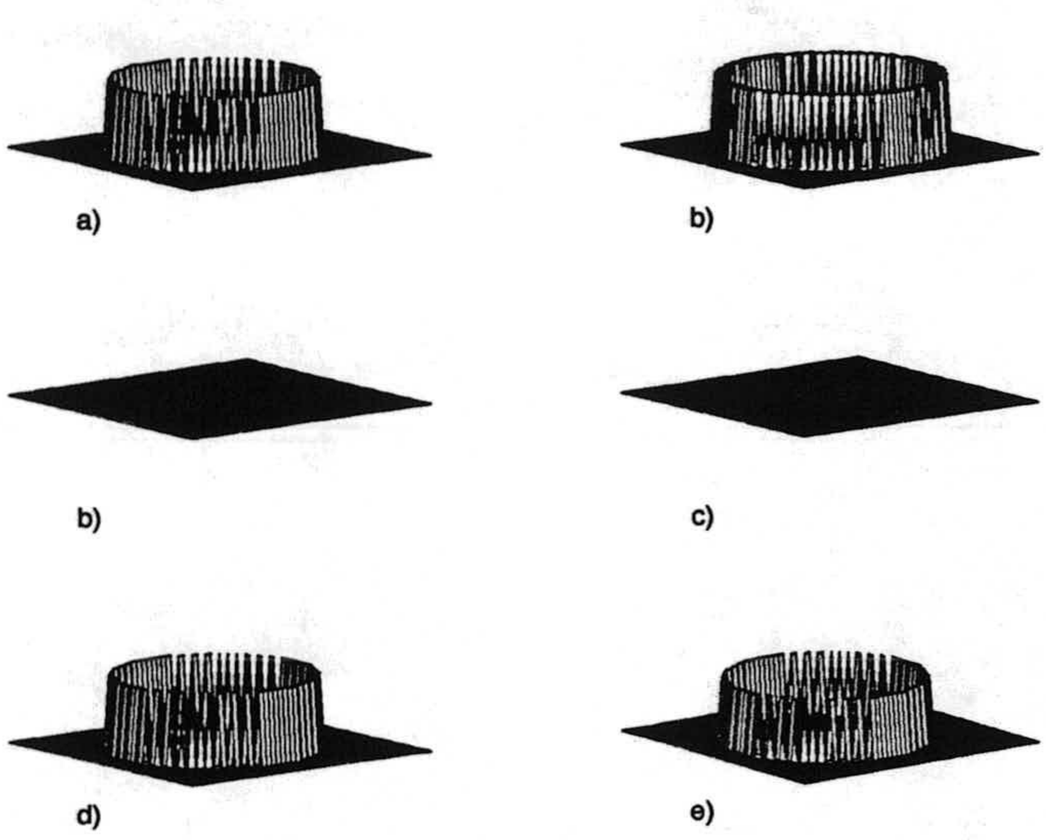


Figura 17: Resultados de filtrar la imagen circunferencia de 1 píxel de grosor. En *a)* se muestra la imagen original; *b)* ofrece el resultado de filtrar con mediana tradicional de ventana 5×5, mientras que *c)* lo hace para el mismo filtro con ventana 3×3. En *d)* se puede contemplar el filtrado con $MR_{5,21}$ y en *e)* con $MR_{3,7}$.

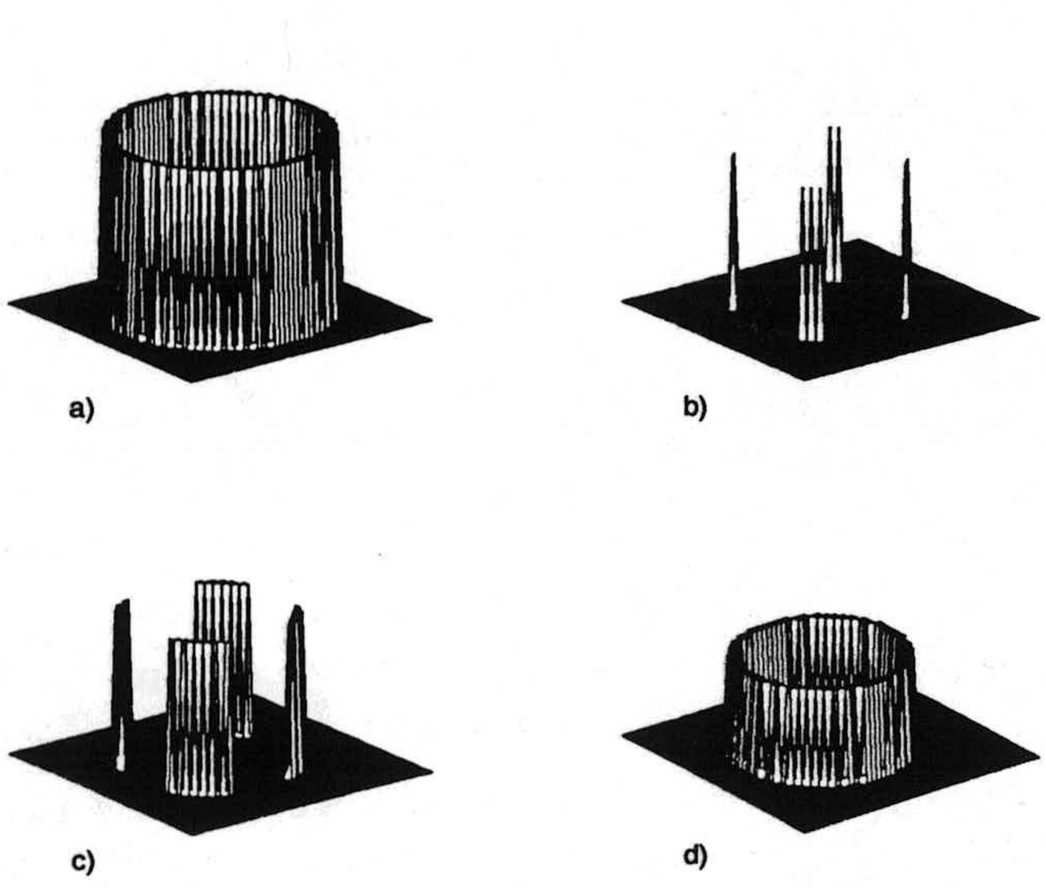


Figura 18: Filtrados de la imagen circunferencia de 2 píxeles de grosor con filtro de mediana tradicional y MR. En *a)* tenemos la imagen original; *b)* muestra el resultado de la mediana de ventana 5×5; *c)* ilustra el filtrado con $MR_{11,15}$ y *d)* con $MR_{10,16}$. Esta imagen es raiz para los filtros con ventana 3×3. Las imágenes circunferencia con grosores superiores a 2 son raíz para los dos tamaños de ventana que se utilizan.

si el tamaño de la misma es pequeño, como es nuestro caso. Así, si ampliamos las imágenes de circunferencias que hemos empleado en este estudio descubriremos que las mismas se componen de trazos horizontales y verticales muy definidos, unidos por otros más o menos diagonales que conforman la apariencia de circunferencia pese a estar trazada sobre una cuadrí-

cula. Pues bien, en esas zonas diagonales la ventana que recubre la imagen al realizar el filtrado se puede encontrar con una cantidad de píxeles pertenecientes a la circunferencia que exceden a los que tendrían si la misma fuese perfecta. De este modo, en las zonas diagonales de la circunferencia puede darse la condi-

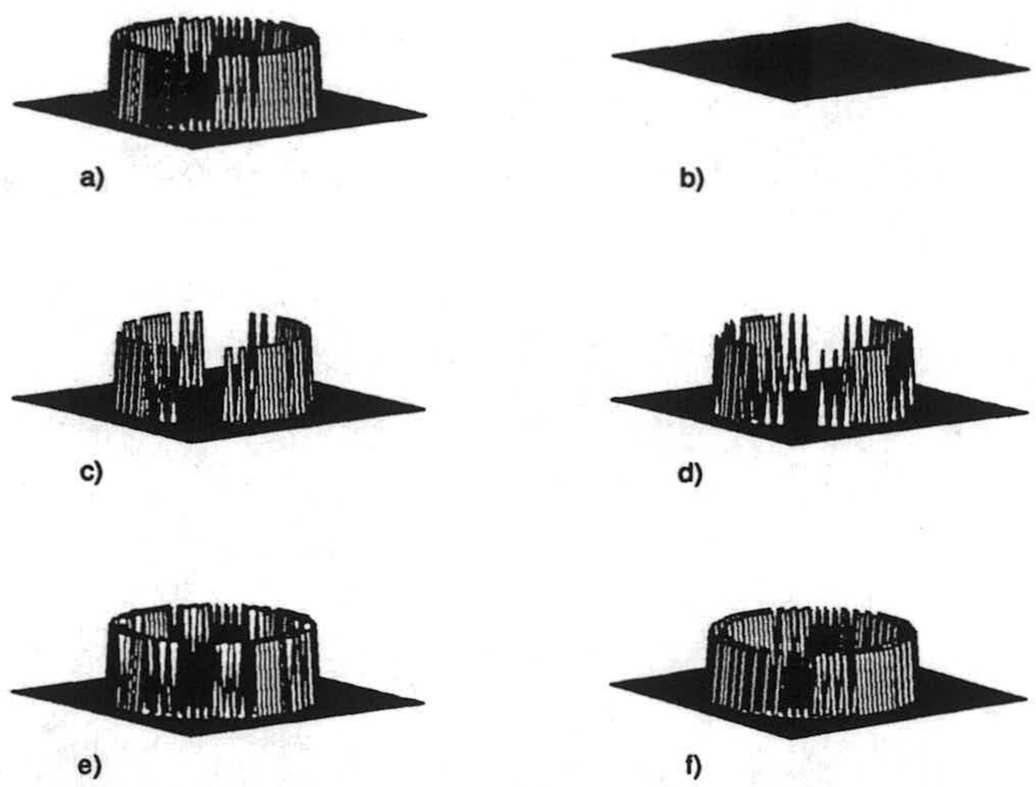


Figura 19: Procesamiento de la imagen de circunferencias concéntricas con filtro de mediana relajada y tradicional. En el apartado *a)* puede observarse la imagen original. En *b)* se expone lo obtenido empleando el filtro de mediana de ventana 5×5 mientras que en *c)* se hace lo propio para la mediana de tamaño 3×3; *d)* es el resultado de filtrar con $MR_{9,17}$; *e)* con $MR_{2,8}$ y *f)* con $MR_{7,19}$.

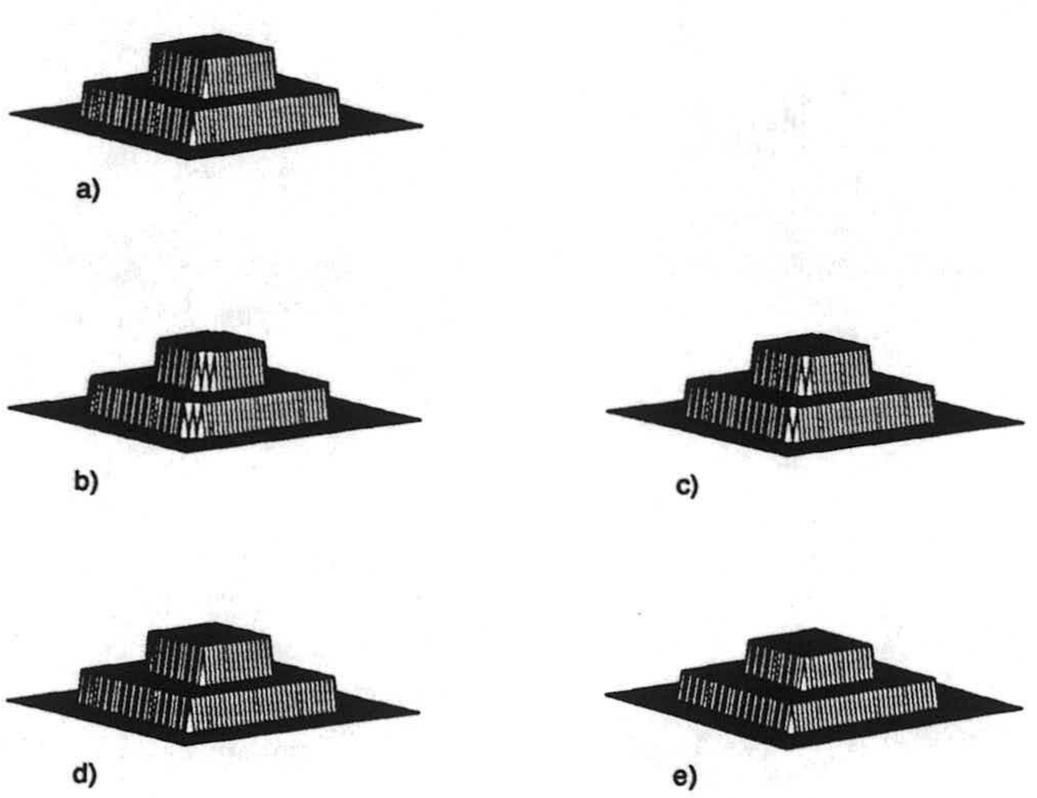


Figura 20: Se ilustra el procesado de la imagen de cuadrados concéntricos con los filtros de mediana tradicional y relajada. Así, en *a)* se observa la imagen original. En *b)* puede verse el resultado para la mediana 5×5, mientras en *c)* aparece el correspondiente a la mediana 3×3. En *d)* se muestra lo obtenido para el filtro $MR_{9,17}$, y en *e)* para $MR_{4,6}$.

ción de raíz más facilmente que en los trazos rectos. Es por esta razón por lo que el filtro MR comienza a erosionar los bordes horizontales y verticales antes que los diagonales, cuando el grosor de la circunferencia es de 2 píxeles.

Cuando los bordes de la circunferencia sólo ocupan 1 píxel de anchura, el comportamiento del filtro de mediana tradicional es eliminarlo completamente, para los tamaños de ventana que venimos manejando, pues los píxeles que componen la circunferencia no llegan a ocupar en la ventana de observación la cantidad necesaria para que sean raíces del filtro. En caso de emplear el algoritmo MR encontramos ese mismo comportamiento para intervalos de bondad estrechos. Sin embargo, para intervalos mayores que $(l, u) = (6,20)$ en el caso de tamaño de ventana 5×5 la imagen de circunferencia de 1 píxel de grosor es raíz del filtro. Si empleamos una ventana de 3×3, entonces la imagen queda inalterada incluso para intervalos estrechos, a partir de $(l, u) = (3,7)$ en adelante, mientras que para $(l, u) = (4,6)$ la circunferencia desaparece completamente. Estos comportamientos pueden observarse en la figura 17.

La imagen de circunferencias concéntricas, filtrada con MR y mediana de diferentes tamaños se muestra en la figura 19. Cuando se emplea un tamaño de ventana 3×3 el filtro de mediana tradicional modifica apreciablemente la imagen, debido a que la anchura de los anillos es menor de la mitad de la ventana. El filtro $MR_{3,7}$ muestra una alteración de las circunferencias concéntricas similar a la capilaridad, debida a que la distancia entre las circunferencias concéntricas es de 1 píxel.

Para tamaños de ventana 5×5 el filtro de mediana tradicional elimina completamente los anillos concéntricos. Esto ocurre ya para el filtro de mediana relajada $MR_{11,15}$, y para intervalos de bondad menores se observa una ruptura de la figura. Sin embargo, para intervalos mayores o iguales que $(l, u) = (7,19)$ la imagen es raíz del filtro de mediana relajada.

Por último, podemos observar el comportamiento de los filtros frente a la imagen de cuadrados concéntricos en la figura 20. En ella se aprecia que el filtro de mediana tradicional elimina las esquinas, degradando con ello la imagen, tanto para tamaño de ventana 3×3 como para 5×5. Ello no sucede para el filtro de mediana relajada con tamaño de ventana 3×3, dónde la imagen es raíz para todos los intervalos de bondad posibles. En el caso de tamaño 5×5, de nuevo la imagen queda inalterada para intervalos de bondad mayores o iguales que $(l, u) = (9,17)$.

Queda, pues, de manifiesto que los filtros MR preservan detalles más sutiles que el filtro de mediana tradicional. Si

bien es cierto que el lograr tal objetivo mediante el empleo de intervalos de bondad elevados resta utilidad en cuanto a supresión de ruido, para intervalos de bondad pequeños este aserto deja de ser válido. Puede emplearse un intervalo adecuado para lograr un compromiso entre eliminación de ruido y respeto a los detalles. en algunos casos los resultados son manifiestamente mejores que los que proporciona la mediana tradicional, como las esquinas de la imagen cuadrada o las líneas de la imagen circular.

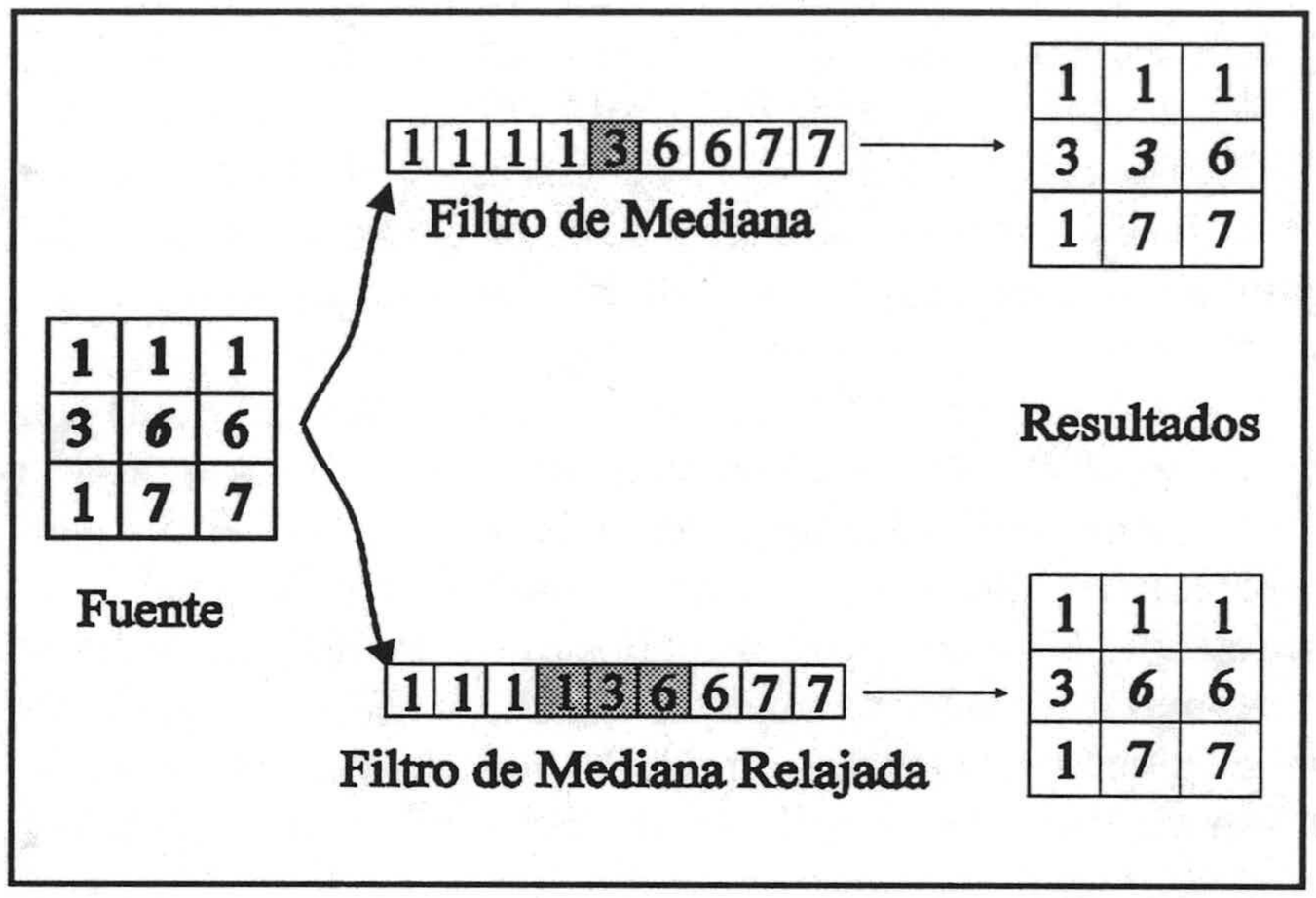


Figura 21: Se ilustra gráficamente la propiedad del filtro MR de evitar la erosión de las esquinas, mostrando cómo supera en este aspecto al filtro de mediana tradicional.

RESPUESTA A IMÁGENES RUIDOSAS

Para continuar indagando en las propiedades del filtro de mediana ralajada, esta vez someteremos a los algoritmos a señales de entrada más complejas. De este modo se pretende tanto validar los resultados obtenidos hasta el momento como hallar otros nuevos.

Respuesta a imágenes prueba bidimensionales complejas.

Consideremos a continuación el sencillo ejemplo de señal bidimensional que aparece en la figura 21. Se observa una ventana 3×3 sobre la imagen fuente, en la que aparece una esquina, y se calcula el resultado aplicando el filtro de mediana y

el de mediana relajada. Se observa que en el primer caso, la esquina se redondea, mientras que el filtrado con MR conserva ese detalle.

Evidentemente, los niveles de gris dentro de la ventana pueden determinar el comportamiento del filtro MR, haciendo que estropee la esquina. Sin embargo, esto parece ser inusual, como ilustra la figura 22. En este ejemplo se usa una ventana $W_{ij}(S)$ de tamaño 5×5 para filtrar una imagen original O de 16 niveles de gris con motivos geométricos. El resultado del filtrado con $MR_{9,17}$ exhibe claramente un mayor respeto a las formas que aparecen en la imagen, si lo comparamos con el resultado del filtro de mediana. Incluso el error medio absoluto (MAE) que vamos a emplear como medida para evaluar la diferencia entre dos imágenes es menor en el caso del MR que en el de la mediana. Los resultados numéricos de estos ejemplos aparecen en la tabla 4.

Podría argumentarse que estos datos no son relevantes, ya que a medida que el intervalo de bondad se hace mayor, el filtrado disminuye, por lo que es evidente que respetará mucho mejor que cualquier filtro la imagen original sin ruido. Sin embargo, hemos de decir que los límites del intervalo, ***l*** y ***u*** se han escogido de acuerdo con el criterio de buena limpieza de ruido impulsivo de sal y pimienta. Para hallar este valor se ha diseñado un conjunto de experimentos empleando imágenes fuente con motivos geométricos con diferentes cantidades de ruido, variando ***l*** y ***u*** pero manteniendo el intervalo de bondad centrado en la lista ordenada. Se han empleado dos tamaños de ventana, 3×3 y 5×5. Los resultados numéricos aparecen en la tabla 4.En las figuras 23 y 24 se muestran gráficamente estos valores. Se ha empleado el CMAE como medida de calidad de mejora de la imagen tras el proceso de filtrado. Como puede verse, todos los MR proporcionan mejores resultados que la mediana para ventana de tamaño 5×5, aunque no para 3×3. Incluso para ruidos elevados, donde la probabilidad de colapso es superior que la de mediana, los resultados son mejores debido al mayor respeto por bordes y detalles en la imagen. Para un 10% de ruido impulsivo *sal y pimienta* el mejor resultado corresponde a un intervalo de bondad (9,17); para el 20% de (10,16) y para 30% un (12,14). Esta evolución es comprensible, pues cuanto mayor sea la probabilidad de aparición de ruido, mayor es la probabilidad de colapso, por lo que conviene cerrar el intervalo de bondad. Para ruido impulsivo uniforme-

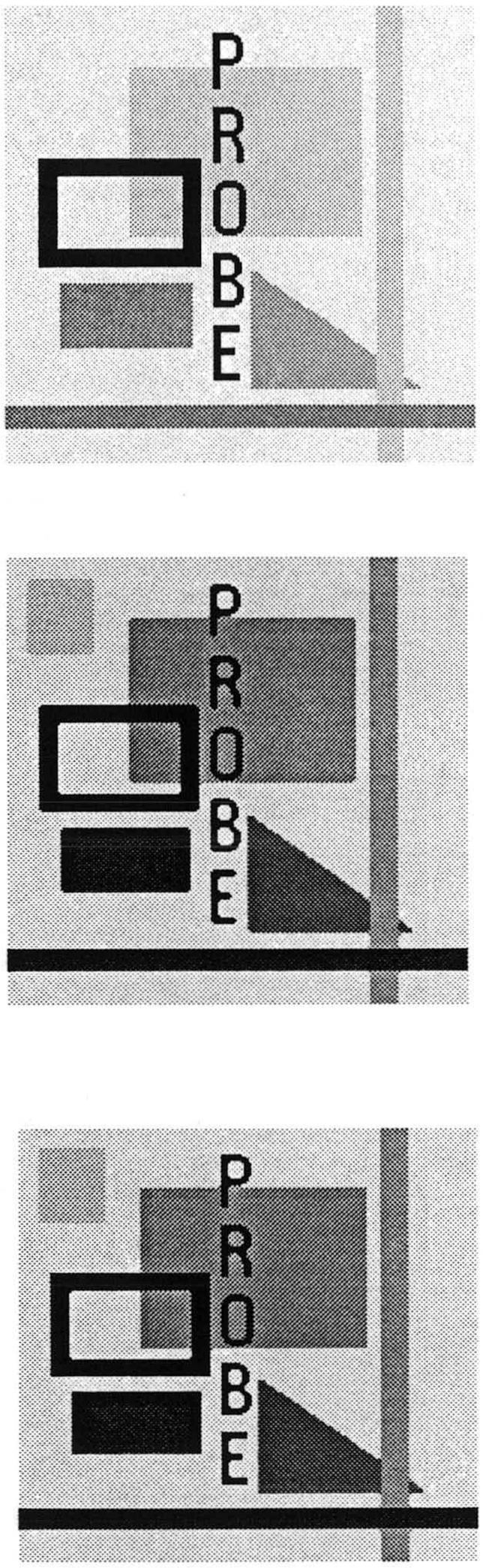


Figura 22: Imagen de motivos geométricos original sin ruido (parte superior) sometida a procesamiento mediante el filtro de mediana de tamaño de ventana 3×3 (centro) y con $MR_{4,6}$. Apréciese cómo este último conserva las esquinas mejor que el filtro de mediana tradicional.

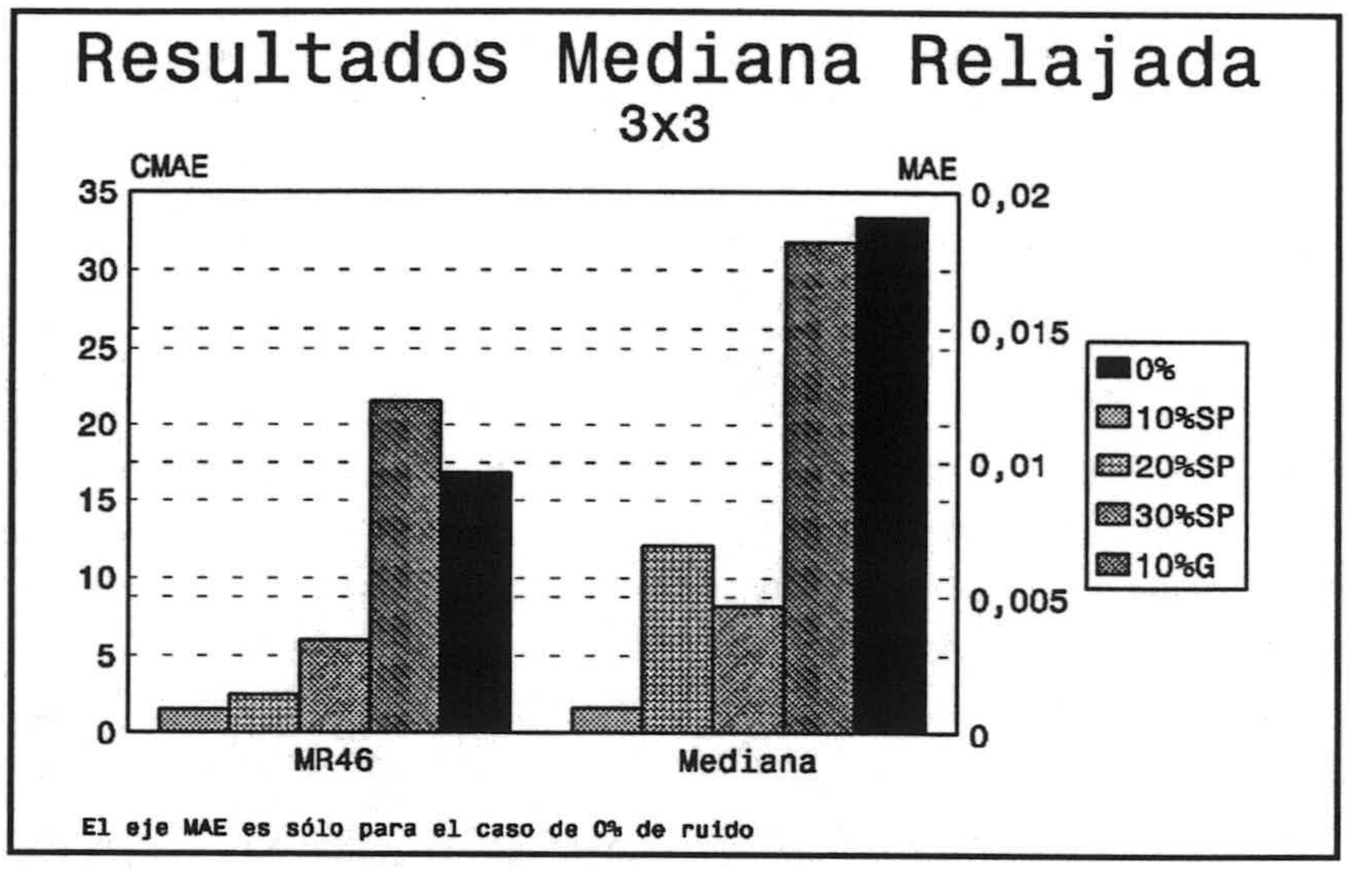


Figura 23: Para ventana 3×3 se ilustra el comportamiento de los filtros de mediana tradicional y relajada frente al procesado de una imagen de mostivos geométricos.

mente distribuido los resultados son muy similares entre el filtro de mediana y el MR de ventana 5×5 con intervalos de bondad moderados o pequeños.

De entre los mejores experimentos hemos escogido como ejemplo el $MR_{9,17}$ para ventana 5×5 y 10% de ruido sal y pimienta. Este ejemplo aparece en la figura 25. La limpieza de ruido es completa tanto para el filtro de mediana como para el MR, pero la distorsión de la imagen es menor en este último caso. Sin embargo, se observa en la tabla que a medida que el ruido aumenta, el filtro MR proporciona peores resultados. La razón de este comportamiento ya se apuntó al estudiar el punto de ruptura del filtro: a medida que el ruido se va haciendo más probable, es más fácil que entre en el intervalo de bondad y se de por bueno, dejándose sin filtrar y dando resultados espúreos.

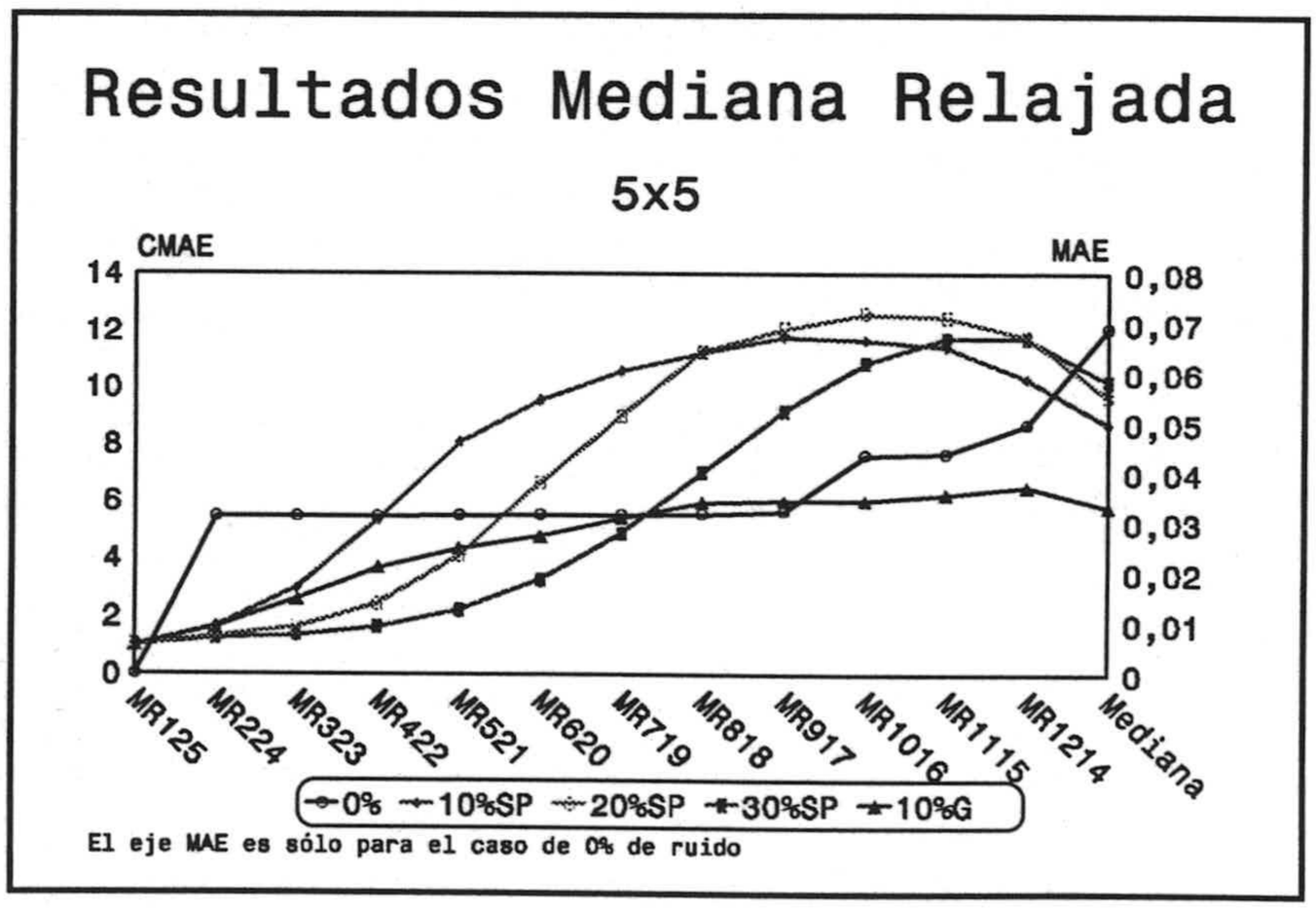


Figura 24: Se exponen gráficamente los datos correspondientes a los resultados de los experimentos llevados a cabo con el filtro de mediana y MR para una imagen de motivos geométricos.

Por último, los experimentos realizados para ruido de sal y pimienta se repitieron para ruido gris uniforme, obteniéndose los resultados que aparecen en la tabla 4. De nuevo el filtro MR se muestra más eficaz que el de mediana para ventana 5×5.

En algunos resultados de filtrado con mediana se ha observado que cuando existen dos líneas muy juntas en la imagen fuente S, éstas aparecen unidas en la imagen resultante R, en un fenómeno que hemos dado en denominar *capilaridad*. Ya el ejemplo numérico de la figura 15 mostraba este comportamiento. Sin embargo, para los filtros MR se minimiza tan indeseable efecto, e incluso se elimina, aunque los intervalos de bondad sean pequeños, como lo prueban los ejemplos visuales mostrados y la figura citada.

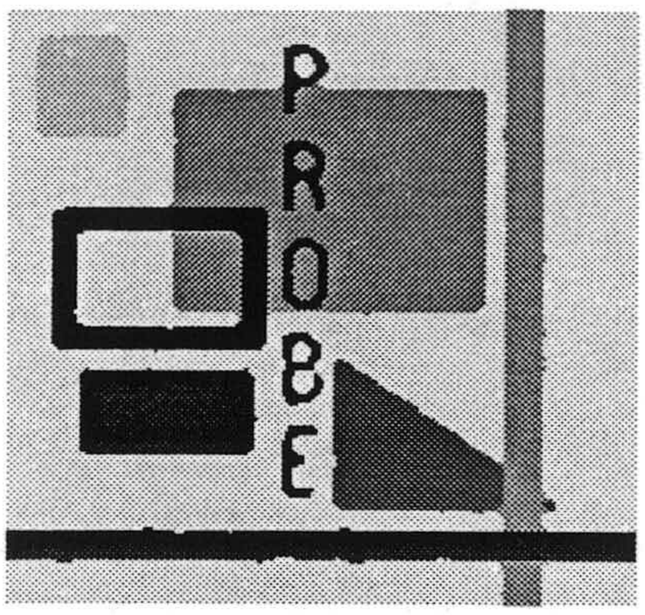
Respuesta a imágenes reales bidimensionales. Aún se han realizado nuevos experimentos para mostrar la bondad del filtro de mediana relajada para una imagen fuente no artificial: una imagen de texto de 200 × 200 píxeles y 64 niveles de gris, afectada por diferentes cantidades y tipos de ruido. Para el filtrado se ha empleado el algoritmo de mediana con ventana 3×3 y 5×5, para comparar con el MR_{46} y MR_{818} respectivamente. Los resultados se muestran en la tabla 5; mientras que en la figura 26 aparecen algunas imágenes resultantes. Aunque se observa que ambos filtros distorsionan la imagen, el MR muestra un mejor comportamiento, produciendo un resultado más legible que el del filtro de mediana. Además, la disminución del efecto de los *blotches* que se había previsto teóricamente se aprecia en este ejemplo, sobre todo en la zona correspondiente al papel que rodea al texto.

Se han representado los valores numéricos en la figura 27, tanto para ventana de tamaño 3×3 como para 5×5. Se ha estudiado el comportamiento frente a ruido impulsivo *sal y pimienta* y uniformemente distribuido, y también para ruido gaussiano de media nula y varianzas 5, 10 y 20. Como se observa, siempre que el ruido es menor del 30% los resultados deMR con ventana 3×3 son mejores que los de mediana de

Fuente	Ventana	Filtro	MAE(R-O)
Sin Ruido	3×3	$MR_{4.6}$ Mediana	**0.0096** *0.0191*
	5×5	$MR_{1.25}$ $MR_{2.24}$ $MR_{3.23}$ $MR_{4.22}$ $MR_{5.21}$ $MR_{6.20}$ $MR_{7.19}$ $MR_{8.18}$ $MR_{9.17}$ $MR_{10.16}$ $MR_{11.15}$ $MR_{12.14}$ Mediana	**0** 0.03108 0.03108 0.03108 0.03133 0.03153 0.03153 0.03173 0.03223 0.04355 0.04385 0.04965 *0.69000*
Sal y Pimienta 10% MAE(S-O)=0.7464	3×3	$MR_{4.6}$ Mediana	0.04820 ***0.04712***
	5×5	$MR_{1.25}$ $MR_{2.24}$ $MR_{3.23}$ $MR_{4.22}$ $MR_{5.21}$ $MR_{6.20}$ $MR_{7.19}$ $MR_{8.18}$ $MR_{9.17}$ $MR_{10.16}$ $MR_{11.15}$ $MR_{12.14}$ Mediana	0.7464 0.4492 0.2522 0.1404 0.0925 0.0783 0.0706 0.0665 **0.0633** 0.0640 0.0653 0.0724 *0.0858*
Sal y pimienta 20% MAE(S-O)=1.4964	3×3	$MR_{4.6}$ Mediana	0.6110 ***0.1247***
	5×5	$MR_{1.25}$ $MR_{2.24}$ $MR_{3.23}$ $MR_{4.22}$ $MR_{5.21}$ $MR_{6.20}$ $MR_{7.19}$ $MR_{8.18}$ $MR_{9.17}$ $MR_{10.16}$ $MR_{11.15}$ $MR_{12.14}$ Mediana	1.4964 1.1367 0.9257 0.6227 0.3675 0.2256 0.1665 0.1330 0.1240 **0.1189** 0.1201 0.1271 *0.1546*

Sal y Pimienta 30% MAE(S-O)=2.2327	3×3	MR4.6 Mediana	0.3735 ***0.2736***
	5×5	MR1.25 MR2.24 MR3.23 MR4.22 MR5.21 MR6.20 MR7.19 MR8.18 MR9.17 MR10.16 MR11.15 MR12.14 Mediana	2.2327 1.7936 1.6513 1.3762 1.0139 0.6880 0.4578 0.3186 0.2434 0.2057 **0.1903** **0.1903** *0.2189*
Gris UnifoMRe 10% MAE(S-O)=0.6268	3×3	MR4.6 Mediana	0.07033 ***0.06093***
	5×5	MR1.25 MR2.24 MR3.23 MR4.22 MR5.21 MR6.20 MR7.19 MR8.18 MR9.17 MR10.16 MR11.15 MR12.14 Mediana	0.6268 0.3894 0.2432 0.1711 0.1445 0.1315 0.1161 0.1053 0.1041 0.1041 0.1004 **0.0965** *0.1083*

Tabla 4: Datos correspondientes a los resultados de realizar filtrados de diferentes imágenes contaminadas con distintos tipos y cantidades de ruido. Se han empleado los algoritmos de mediana tradicional, con tamaños de ventana 3×3 y 5×5, y el filtro de mediana relajada de igual tamaño. En este último caso, se ha variado asímismo el intervalo de bondad con objeto de poner de manifiesto la evolución del comportamiento del filtro. Los mejores resultados se han resaltado en negrita.



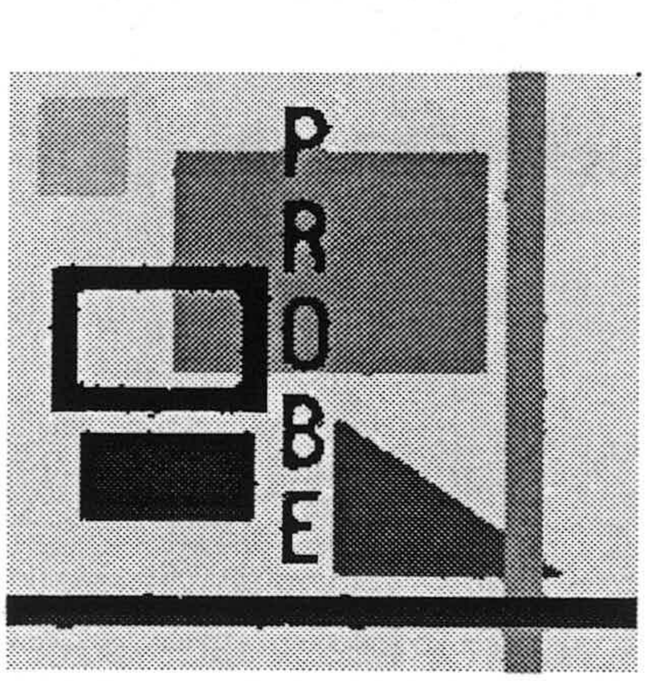


Figura 25: La imagen de motivos geométricos contaminada por 10% de ruido *sal y piemienta* (arriba) se ha procesado con el filtro de mediana de tamaño de ventana 5×5 (centro) y con $MR_{9,17}$ (abajo). Se observa cómo este último resultado respeta detalles que la mediana tradicional destruye, como por ejemplo el pico derecho del triángulo o las esquinas de los cuadrados.

igual tamaño, precisamente cuando la probabilidad de colapso es más parecida entre ambos tipos de filtro. En el caso de ventana de tamaño 5×5 la mejora se extiende incluso al 30% de probabilidad de aparición de pulsos. Cuando el ruido es uniformememnte distribuido el filtro de mediana y el MR son casi coincidentes para una ventana 3×3, mientras que para ruido gaussiano las mejoras son prácticamente iguales para todos los filtros, independientemente del tamaño de la ventana.

Es digno de mención el hecho de que las mayores mejoras que proporcionan los filtros se obtengan con ruido impulsivo *sal y pimienta*, seguidos del ruido uniformemente distribuido. El caso gaussiano es realmente malo, llegando incluso a empeorar la imagen cuando la varianza es de 5. En lo sucesivo, por lo tanto, no se considerá el problema de la supresión de ruido aditivo gaussiano debido al pobre comportamiento de los filtros. Por otra parte, debe notarse que para ventanas 3×3 las mejoras obtenidas son mayores que para 5×5. Por lo tanto debe inferirse que es preferible usar una ventana pequeña, que elimine menos ruido, pero que preserve mejor los detalles.

En todos los casos, las esquinas y otros detalles se han preservado mejor empleando nuestro algoritmo. Los resultados para MR son mejores incluso cuando se introduce una pequeña cantidad de ruido impulsivo, siempre que no sea demasiada cantidad. En definitiva, podemos decir que nuestro método mejora al filtro de mediana, sin coste computacional ni cambio en la filosofía subyacente.

TIPO DE RUIDO	MAE (S-O)	MAE(R-O)			
		MR818	MR46	MED(3x3)	MED(5x5)
10% SP	3,181025	2,428625	1,365975	1,772975	3,31135
20% SP	6,291075	2,40615	1,812325	2,09195	3,3702
30% SP	9,27985	2,84895	2,7213	2,638575	3,43905
10% G	2,569875	2,53915	1,445575	1,777925	3,282675
20% G	5,1038	2,727075	1,9596	2,18555	3,3364
30% G	7,5944	3,2526	2,934225	3,007325	3,583925
N(0,5)	3,607675	3,80575	2,888125	2,92615	3,942775
N(0,10)	6,268	4,828625	4,308375	4,25995	4,6589
N(0,20)	10,77833	6,568825	6,62455	6,452975	6,07505

Tabla 5: Se muestran aquí los valores de MAE correspondientes al filtrado de la imagen de texto mediante el algoritmo de mediana y el MR.

Figura 26: La imagen de texto real original (arriba a la izquierda) fue contaminada con 10% de ruido *sal y pimienta* (arriba a la derecha). El resultado de filtrar dicha imagen fuente con el algoritmo de mediana tradicional con ventana 5×5 puede verse en la imagen a la izquierda de la fila inferior. La imagen de abajo a la derecha muestra el resultado de emplear el filtro $MR_{8,18}$.

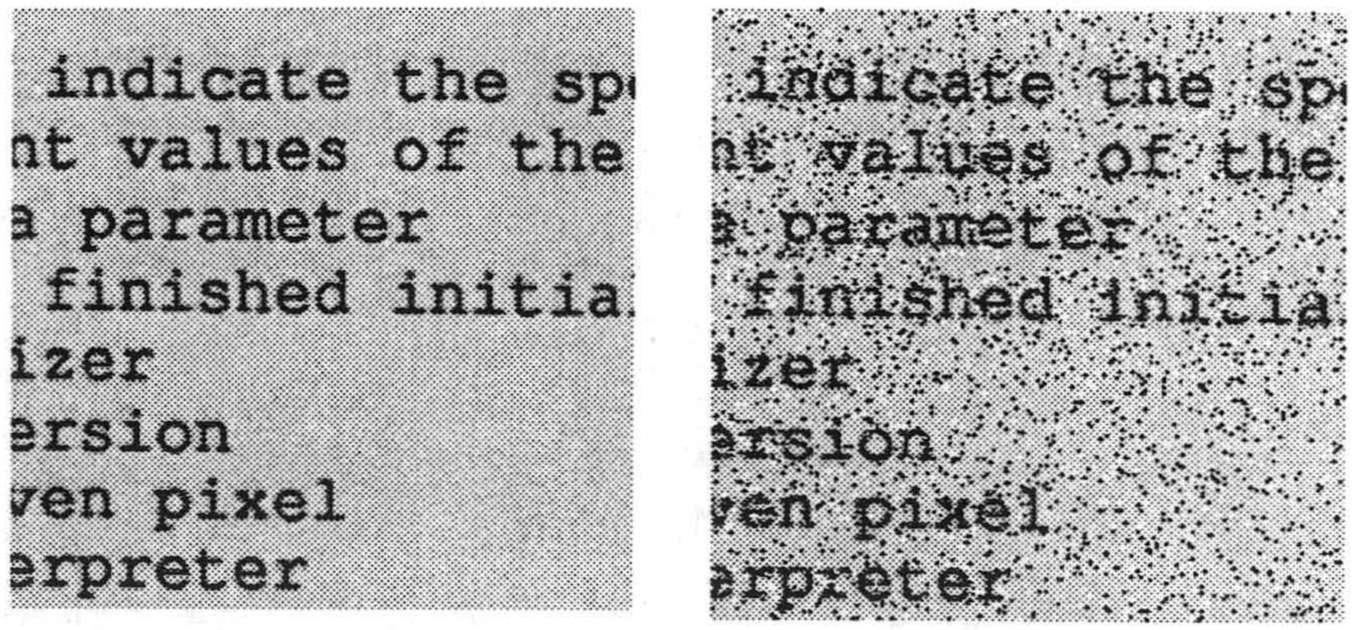
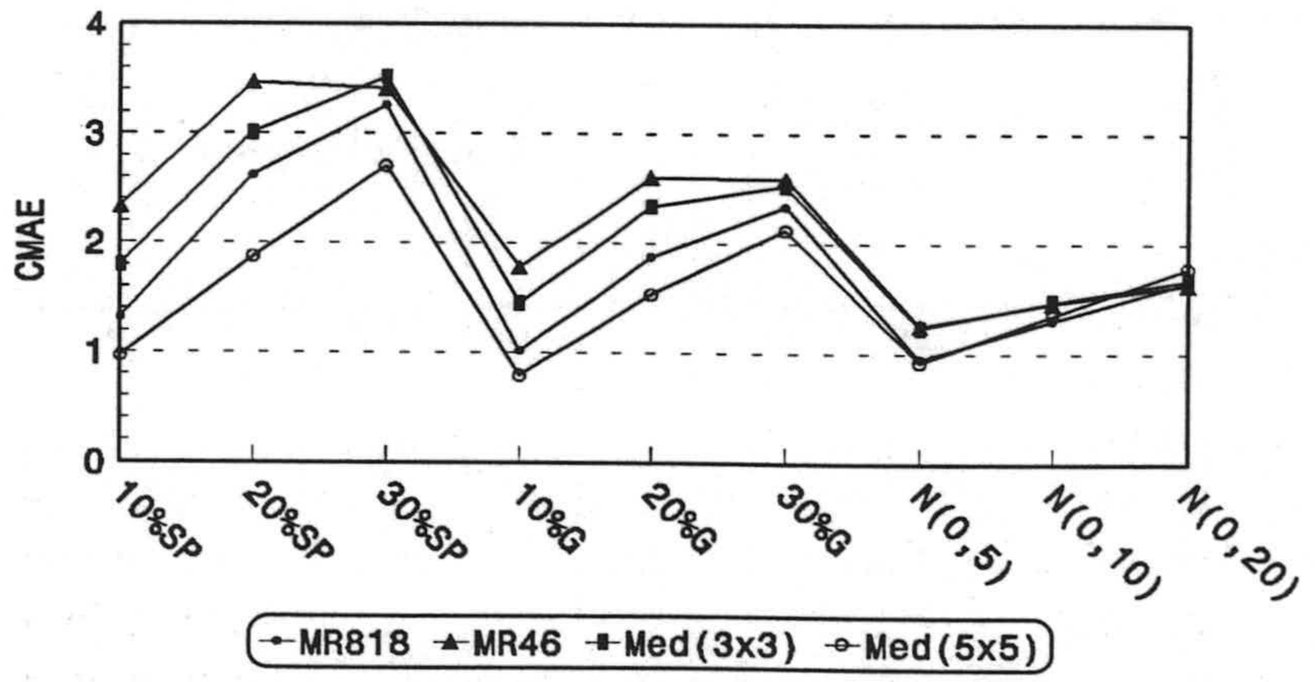


Figura 27: Comparación de los resultados obtenidos al procesar la imagen texto afectada con ruido mediante los filtros en estudio.



EL FILTRO DE MEDIANA CON VENTANA ADAPTATIVA

DESCRIPCIÓN

En una imagen, los bordes entre las regiones allí representadas contienen siempre una información esencial. Es más, nuestra percepción visual se basa en dicha información. Por tanto, cualquier filtrado debe tender a respetar los bordes de la imagen. Sin embargo, los filtros lineales distorsionan estos bordes, cambiándolos de sitio y difuminándolos [PRA78] [JAI89], mientras que los filtros no lineales tipo mediana, empleados para eliminar el ruido impulsivo, preserva la información contenida en los bordes, pero también produce un efecto negativo sobre ellos: los bordes pierden nitidez. De hecho, cuanto mayor es la ventana empleada, más ruido se elimina, pero mayor es la falta de nitidez de la imagen, pudiendo incluso eliminarse algunos bordes y detalles delgados [PIT90] [DAV88] [DAV89].

Para una imagen típica, cada región difiere de las otras lo suficiente como para que sea apropiada la descripción de la misma mediante un proceso bidimensional no estacionario [PRAT78]. Por tanto el filtrado de ruido será mejor llevado a cabo por métodos no estacionarios, o adaptativos a la señal. Como ya se apuntó, el empleo de esquemas adaptativos proporciona la manera más inteligente de trabajar, puesto que emplea toda la información apriorística acerca de la imagen, aunque, eso sí, a costa de un precio computacional. Dentro de este ámbito se engloba el filtro aquí propuesto, y que denominamos *restdiver*.

Se trata de un filtro adaptativo con tamaño de ventana variable. Tiene la ventaja de que elimina el ruido impulsivo, a la vez que preserva los bordes, mejor que el filtro de mediana tradicional, y en una sola pasada (sin reiteración del filtrado). El filtro propuesto es conceptualmente muy simple, y se basa en la adaptación del tamaño de la ventana del filtro de mediana a la presencia de bordes en la cercanía, detectados con un sistema nuevo de detección de bordes, de manera que si hay una región uniforme se emplea una ventana grande (con la máxima potencia limpiadora) pues tenemos la certeza de no estropear

ningún borde o detalle. Si alguna de estas características aparece, la ventana reduce su tamaño para minimizar el efecto de difuminado.

La idea fundamental es, pues, la siguiente: en la íntima vecindad de un borde, el respeto del mismo es más importante que la reducción de ruido, de acuerdo con el efecto de enmascaramiento que tiene lugar en la visión humana. Este efecto de enmascaramiento parte del hecho de que el ruido es percibido por el ojo de manera diferente, dependiendo de si está en una región uniforme o en un lugar cercano a un borde. El ojo es muy sensible al ruido, aunque sea escaso, en una región de gris uniforme, mientras que es capaz de tolerar una cantidad de ruido bastante mayor en la vecindad de un borde. A pesar de todo, con objeto de evitar cambios abruptos en la supresión de ruido, se realizará un filtrado.

A continuación expondré en primer lugar el esquema adaptativo que se emplea, detallando cómo varía el tamaño de la ventana. Después hablaremos del tipo de filtro que se emplea, y el porqué de su elección.

MECANISMO ADAPTATIVO DE SUPRESIÓN DE RUIDO

Recordemos lo dicho en el capítulo anterior acerca de los filtros adaptativos. En nuestro caso, el diseño del sistema adaptativo difiere ligeramente de la filosofía más general, expuesta en la sección citada. Mientras en la mayor parte de los filtros adaptativos el sistema de control se basa en información apriorística de la señal sin ruido original, y de la formación del ruido, en nuestro caso nos conformaremos con un conocimiento previo mucho más limitado, y más realista.

La imagen fuente S proporciona toda la información que requiere el sistema adaptativo. La situación de los bordes en la imagen se obtiene a partir de ella mediante un nuevo método de segmentación de imágenes desarrollado en nuestro Grupo de Investigación, utilizando la divergencia de Jensen-Shannon [BAR95]. Una vez que se tiene la imagen de bordes, B, resultado de la segmentación, el sistema de control asigna un valor al tamaño de la ventana adecuado para la supresión de ruido, pero manteniendo al máximo el respeto por los bordes en la imagen. Pasemos a ver en detalle el funcionamiento del filtro. En primer lugar mostraremos someramente el mecanismo de segmentación, para después detallar cómo opera *restdiver*.

SEGMENTACIÓN BASADA EN LA DIVERGENCIA JS

La segmentación y detección de bordes es una herramienta primordial en el procesamiento de imágenes. Algunos de estos mecanismos se basan en la búsqueda de diferencias de texturas entre dos regiones adyacentes para señalar el borde entre ellas, pero el problema dista de estar resuelto. Mientras los métodos más sencillos no aciertan a la hora de detectar los bordes entre regiones texturadas [PRAT78], existen procedimientos de regionalización que proporcionan características de la textura para reconocerla, pero son computacionalmente muy complejos.

En nuestro Grupo de Investigación se ha desarrollado un método para detectar bordes entre regiones de diferente textura (esto es, con histogramas de nivel de gris diferentes), basado en la divergencia de Jensen-Shannon, que es especialmente simple. No vamos a repetir aquí los fundamentos de la divergencia de Jensen-Shannon, que ya han sido expuestos en el anterior capítulo, por lo que pasaremos a detellar directamente el modo de operación del método propuesto.

En su versión más simple, una ventana deslizante W se divide en dos subventanas iguales, rectangulares y adyacentes, W_1 y W_2. Por tanto, los pesos de las distribuciones de probabilidad que aparecen en la expresión de la divergencia JS serán $\pi = (½, ½)$. Para cada localización de la ventana se calculan los histogramas normalizados $\mathcal{P}(W_1)$, $\mathcal{P}(W_2)$ y $\mathcal{P}(W) = ½[\mathcal{P}(W_1) + \mathcal{P}(W_2)]$, así como la divergencia de Jensen-Shannon. La ventana se desliza por toda la imagen con un factor de salto Δ. La divergencia JS varía con la localización de la ventana, dependiendo tanto del tamaño de la subventana $S_w = (X_w \times Y_w)$ como del tamaño del patrón de la textura, $S_t = (X_t \times Y_t)$.

Para texturas regulares y subventanas que comprendan un número entero de patrones de textura, la divergencia JS = 0 siempre que la ventana esté situada dentro de una misma región, independientemente de la textura que tenga. En cualquier otro caso, JS $\neq$ 0. Cuando una ventana cruza perpendicularmente el borde entre dos regiones a y b, puede demostrarse fácilmente que la máxima divergencia se obtiene si y sólo si cada subventana se sitúa enteramente dentro de una región diferente, y bajo la siguiente suposición simplificadora, no siempre verificada: el histograma $\mathcal{P}(W_i)$ se compone de $\mathcal{P}_a$ y $\mathcal{P}_b$ con las mismas proporciones (β, 1-β) con las que la superficie de la

UNIVERSIDAD DE GRANADA
15 JUL. 1996

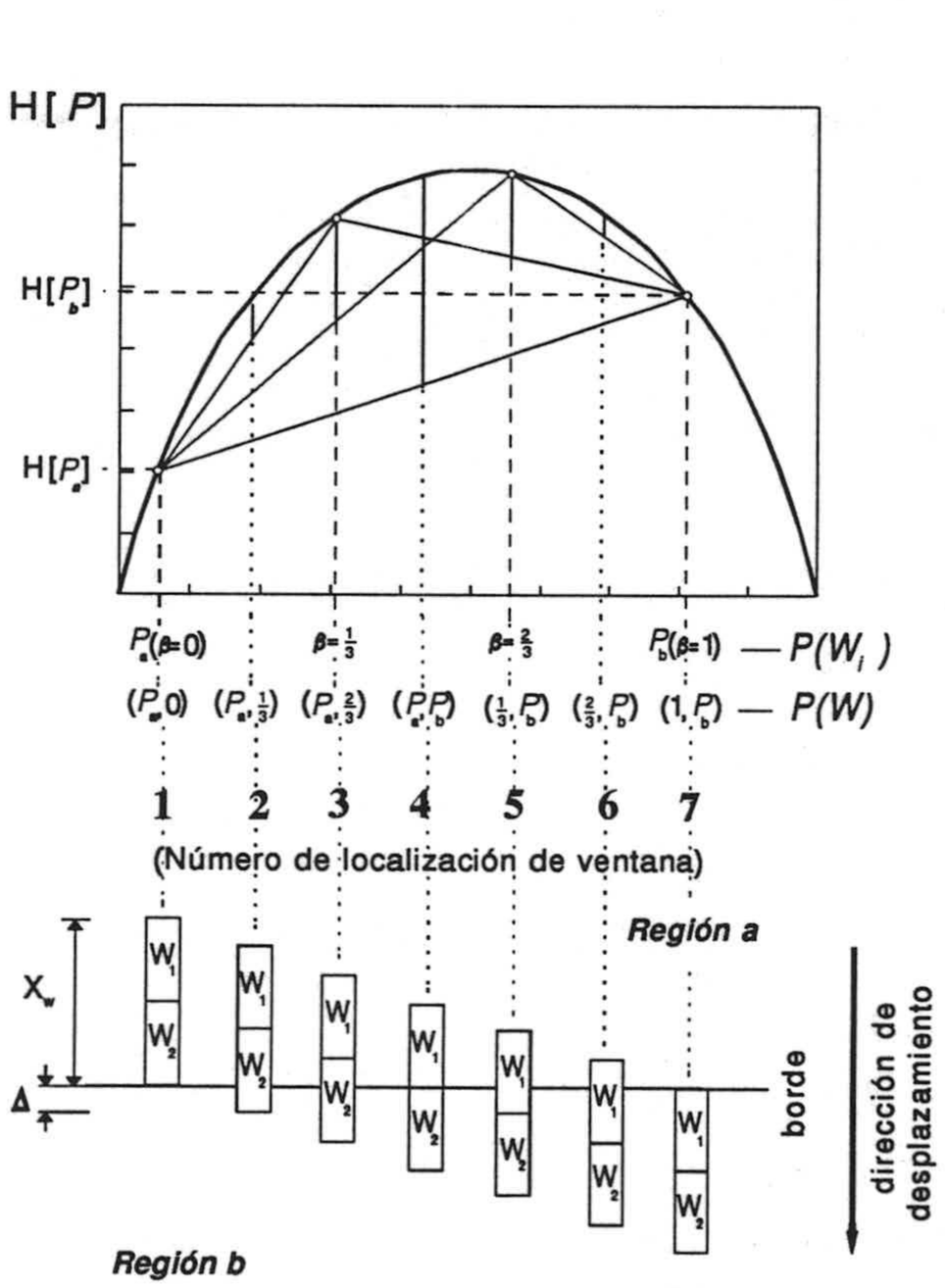
subventana W_i está compuesta por las regiones a y b. Este máximo sirve para determinar el borde entre las texturas.

Podemos intentar aclarar el modo de operación del método mediante un ejemplo sencillo, que aparece ilustrado en la figura 28. Se ha considerado una imagen binaria, con $X_w = 3$, $\Delta = 1$, y $X_t = 3$. La divergencia JS se muestra para todas las posiciones (1 a 7 en la figura). Los histogramas están representados en el eje horizontal como sigue: en una segunda fila se muestra $\rho(W) = \frac{1}{2}[\rho(W_1) + \rho(W_2)]$ correspondiente a la localización de la ventana que aparece debajo. Esta distribución viene dada en función de $[\rho(W_1) + \rho(W_2)]$, de modo que el número que aparece indica las proporciones $\hat{a}$ de la combinación lineal convexa de las distribuciones incluidas en la subventana W_i que cruza el borde:

$$\rho(W_i) = (1 - \beta_i)\,\rho_a + \beta_i\,\rho_b$$

Esta expresión está representada en la primera fila del eje de abscisas.

Figura 28: Se muestra cómo funciona el método de segmentación mediante el empleo de la divergencia de Jensen-Shannon.



En la práctica se realizan algunas variaciones respecto a lo dicho hasta aquí. Por ejemplo, se suelen emplear ventanas divididas en cuatro subventanas (cuadrantes) iguales y adyacentes. Puede calcularse la divergencia de Jensen-Shannon de los cuatro histogramas conjuntamente, o tomados de dos en dos, ya sea diagonalmente o en forma de cruz. En las aplicaciones que se mostrarán aquí, cada vez que desplazamos la ventana se calculan las divergencias entre los pares de subventanas (W_1, W_4) y (W_2, W_3). El mínimo valor de estas divergencias es tomado como la divergencia JS total.

Asignando al píxel central de la ventana el valor de la divergencia de Jensen-Shannon, para cada punto en la imagen fuente existirá un valor definido de la divergencia JS, obteniéndose por tanto una matriz de divergencias. Empleando la cota superior de JS, que es $log_2(n)$, siendo n el número de distribuciones de probabilidad, normalizamos la matriz. Acto seguido, mediante una

umbralización adecuada, se obtendrá por fin una imagen binaria que corresponde con la imagen segmentada de los bordes de la imagen B_{ij}.

De esta manera se han obtenido resultados mejores que os obtenidos con otros métodos de detección de bordes que abundan en la literatura, como el de gradiente gaussiano, el de campos aleatorios de Markov, el de Canny y algunos basados en redes neuronales [PIT90] [DAR93]. Pero sin duda la característica que nos ha movido a desarrollar este método de segmentación es su robustez frente al ruido.

Cuando trabajamos en procesamiento de imágenes se nos suministra una imagen fuente S afectada de ruido, y no disponemos de la imagen original O de la que proviene. Por tanto, el método de detección de bordes que habremos de emplear para nuestro esquema adaptativo habrá de funcionar incluso cuando la imagen fuente esté muy degradada debida al ruido. Y ello no sucede con los métodos generalmente empleados en programas comerciales, ni los propuestos en la bibliografía. Nuestro algoritmo detecta muy bien los bordes de la imagen, con bastante rapidez; y se ve relativamente poco afectado por la presencia de ruido. Es, por tanto, el candidato perfecto para nuestros propósitos.

Actualmente el Grupo de Investigación está trabajando en la mejora del método propuesto, con considerable éxito.

ESQUEMA ADAPTATIVO EMPLEADO

Acabamos de ver cómo el valor de la divergencia de Jensen-Shannon indica la existencia de bordes en una imagen. Del proceso completo de segmentación se tiene una imagen binaria de bordes B. Dicha imagen se pasa al algoritmo como parámetro, además de la imagen fuente S. Se trabaja con ventanas deslizantes, cuadradas y de lado impar, cuyos tamaños máximo $[2V_M+1]^2$ y mínimo $[2V_m+1]^2$ están predefinidos. Entonces el filtro opera como sigue.

En primer lugar se sitúa en la posición (i,j) de la imagen fuente, teniendo un tamaño inicial $[2(V_m+1)1]^2$. Si dentro de la ventana $W_{ij}(S)$ existe algún borde, al comparar con B_{ij}, la ventana se reduce de tamaño en una unidad hasta $[2V_m+1]$ y se filtra. Si no hay borde dentro de la ventana inicial, ésta se va aumen-

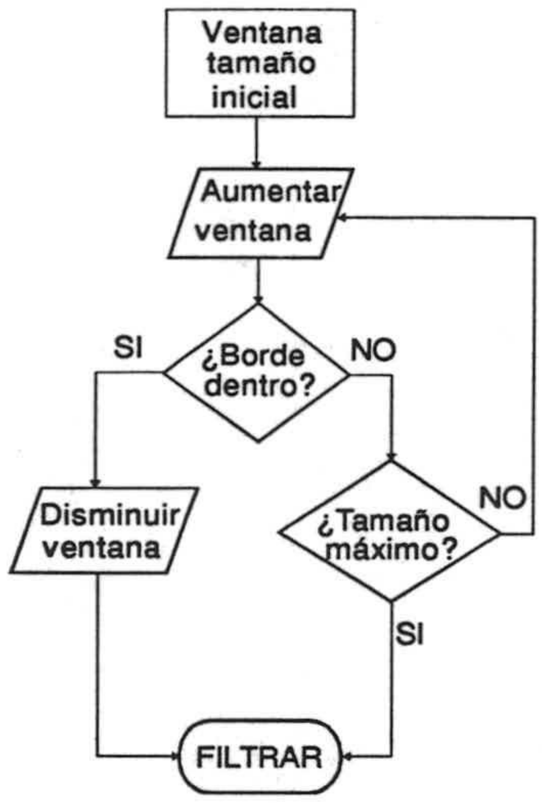


Figura 29: Organigrama del algoritmo del filtro de mediana con ventana adaptativa.

tando hasta que se produce una de estas dos cosas:

· Aparece un borde dentro de la nueva ventana.

· El tamaño de la ventana alcanza el valor máximo $[2V_M+1]^2$.

Entonces el tamaño se reduce en una unidad y se procede al filtrado. A continuación la ventana $W_{ij}(S)$ se desplaza a la posición adyacente, y se reiteran las operaciones. Este proceso queda más claro en el organigrama que aparece en la figura 29.

Vemos así que esta estrategia respeta nuestra filosofía inicial: se obtiene la máxima limpieza de ruido en las regiones uniformes, y una mínima limpieza, pero máximo respeto por los bordes, en regiones vecinas a éstos. Para ilustrar el mecanismo de filtrado adaptativo aquí expuesto, se propone un ejemplo.

La imagen utilizada es un gráfico sintético de imágenes geométricas generado por ordenador completamente limpio de ruido. Tiene un tamaño de 200 × 200 y una resolución de 8 bits por píxel. Esta imagen original se muestra en la figura 30. La fila 70 se ha marcado con una línea horizontal oscura. La imagen se ha sometido a un procedo de generación de ruido impulsivo de sal y pimienta afectando al 20% de los píxeles, dando lugar a la imagen fuente S que se aprecia en la figura 30. La segmentación de S se ilustra también en la figura 30. Para la fila indicada en la figura 30a, la señal sin ruido $O_{i=70,j}$ aparece gráficamente en la figura 31a, y la correspondiente señal ruidosa $S_{i=70,j}$ en la figura 31b. La posición de los bordes en esa fila detectados por el método de segmentación aparecen en la figura 31c. Se emplean como valores extremos del tamaño de la ventana $[2V_m+1] = 3$ y $[2V_M+1] = 11$, comenzando con una ventana de tamaño inicial 5×5. La evolución del tamaño de las ventanas se muestra en la figura 31d.

TIPO DE FILTRO EMPLEADO

Ya hemos dicho que el filtro de mediana elimina bien el ruido impulsivo y reduce el ruido aleatorio (especialmente cuando la densidad de probabilidad del ruido tiene colas largas), y

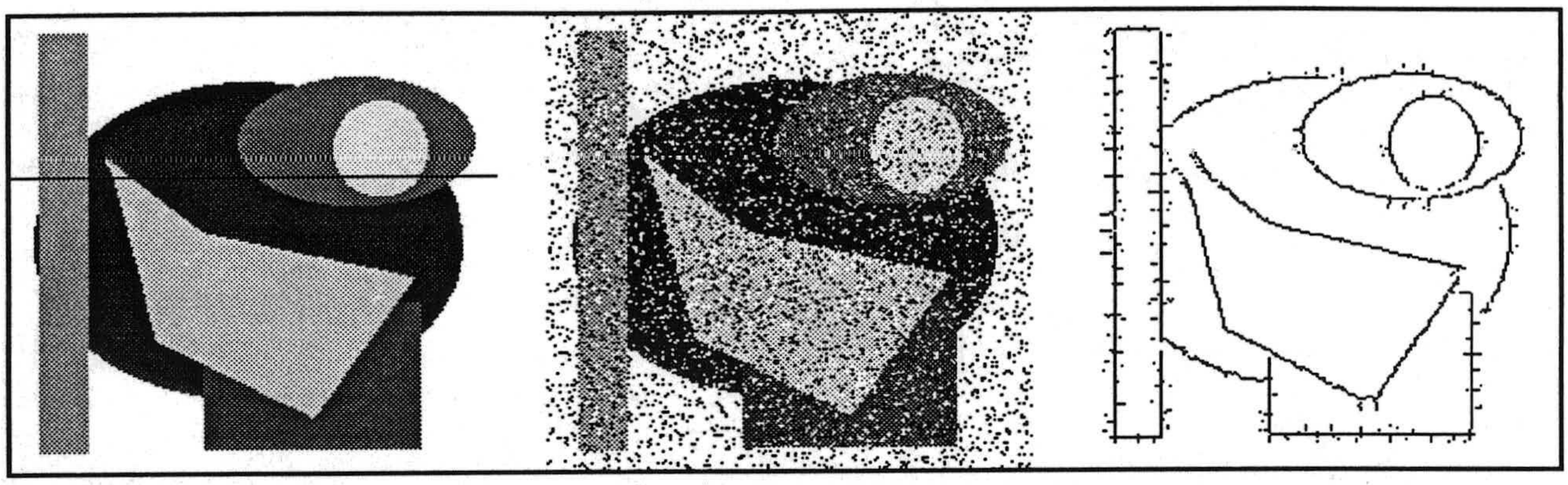


Figura 30: A la izquierda se muestra la imagen original de motivos geométricos. En el centro aparece la misma imagen contaminada por un 20% de ruido impulsivo *sal y pimienta* y a la derecha se expone el resultado de la detección de bordes con el método propuesto.

mantiene los bordes y tendencias monótonas de la señal [ARC83] por lo que se emplea ampliamente en procesado de imágenes. Además, es un filtro muy rápido y sencillo.

En cuanto al tamaño ($2N+1$) de la ventana que emplea, recordemos que se han de tener en cuenta dos efectos. La supresión de ruido es directamente proporcional al tamaño de la ventana; pero por otro lado la resolución espacial es nversamente proporcional al tamaño de la ventana. En particular, un filtro de mediana elimina todos los objetos sobre un fondo uniforme que contengan menos de ($N+1$) píxeles. Para solucionar este problema se suele emplear comúnmente una ventana pequeña (generalmente 3×3), pero iterando el proceso de filtrado hasta que el resultado sea suficientemente limpio, o hasta que se llegue a una raíz (señal inalterada por sucesivas iteraciones del filtro) [GAL81]. Estas raíces dependen de la forma y tamaño de la ventana empleada, y el estudio de la aparición y convergencia de las mismas es útil en el estudio del filtro de mediana [FIT85] y en codificación de imágenes. Sin embargo, la repetición del filtrado de mediana no implica necesariamente que se consiga una reducción de ruido suficiente [NOD82].

Por tanto, hemos elegido el filtro de mediana (no recursivo) para aplicarle el esquema adaptativo de ventana adaptable, por ajuste automático a la naturaleza de la parte procesada de la señal, esto es, si está o no en la vecindad de un borde.

Una vez descrito profusamente el método de funcionamiento del algoritmo del filtro de mediana con ventana adaptativa, se procederá a exponer sus principales propiedades, siempre comparadas con las del filtro de mediana tradicional.

Figura 31.a: Gráfica representando los niveles de gris de la imagen original de motivos geométricos de la figura 30 a lo largo de la fila j = 70.

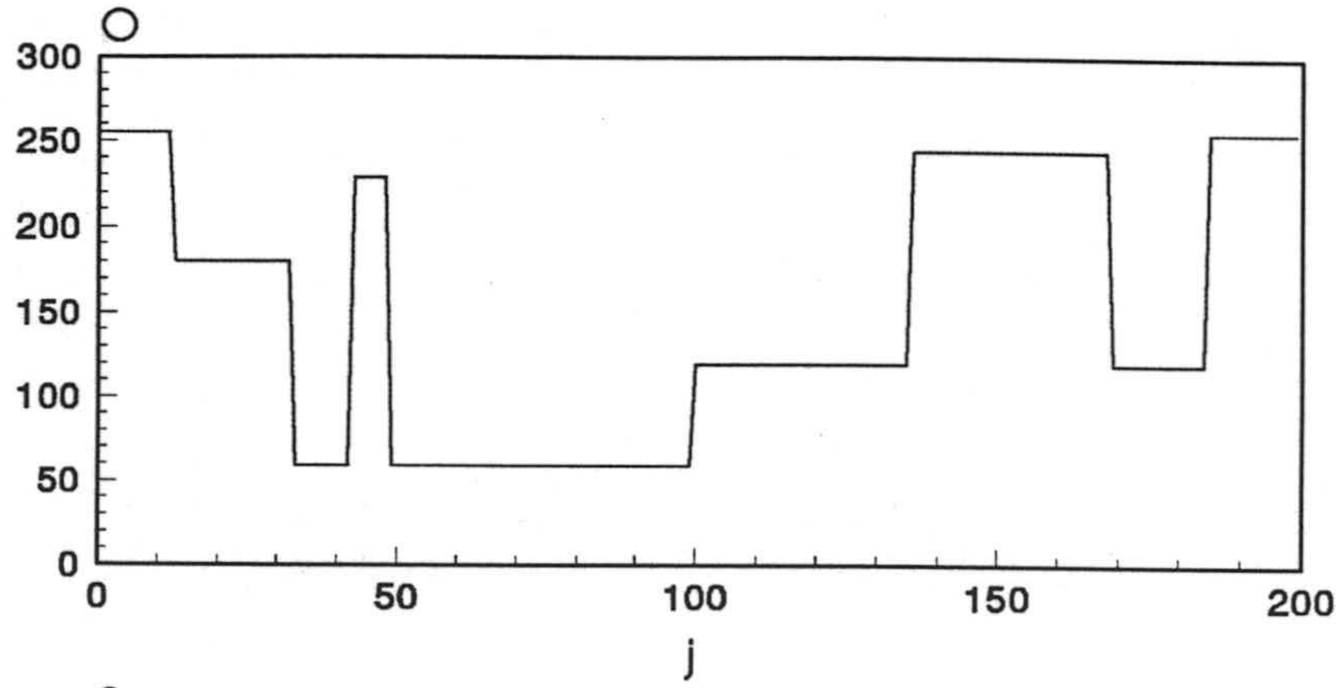


Figura 31.b: La misma fila j = 10, pero de la imagen fuente de la figura 30 (centro).

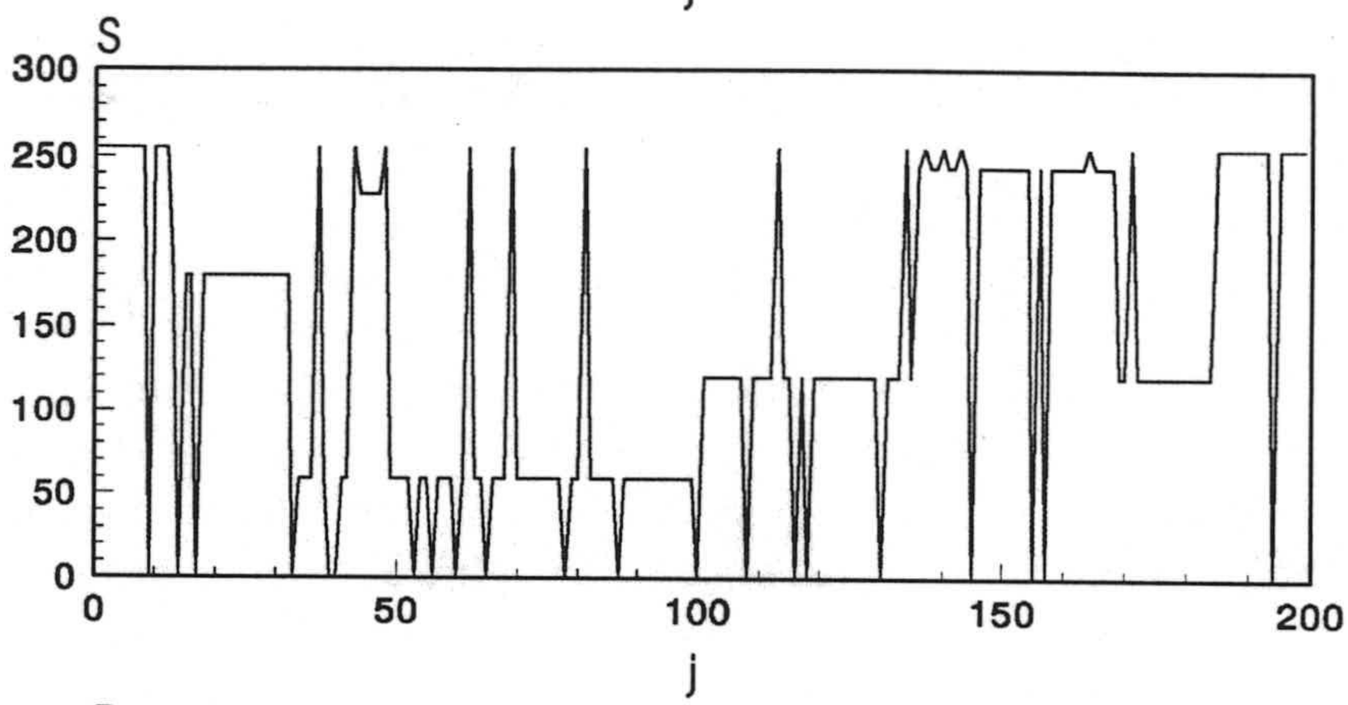


Figura 31.c: Imagen de bordes correspondiente a la imagen fuente de motivos geométricos, calculada mediante el nuevo método de detección de bordes.

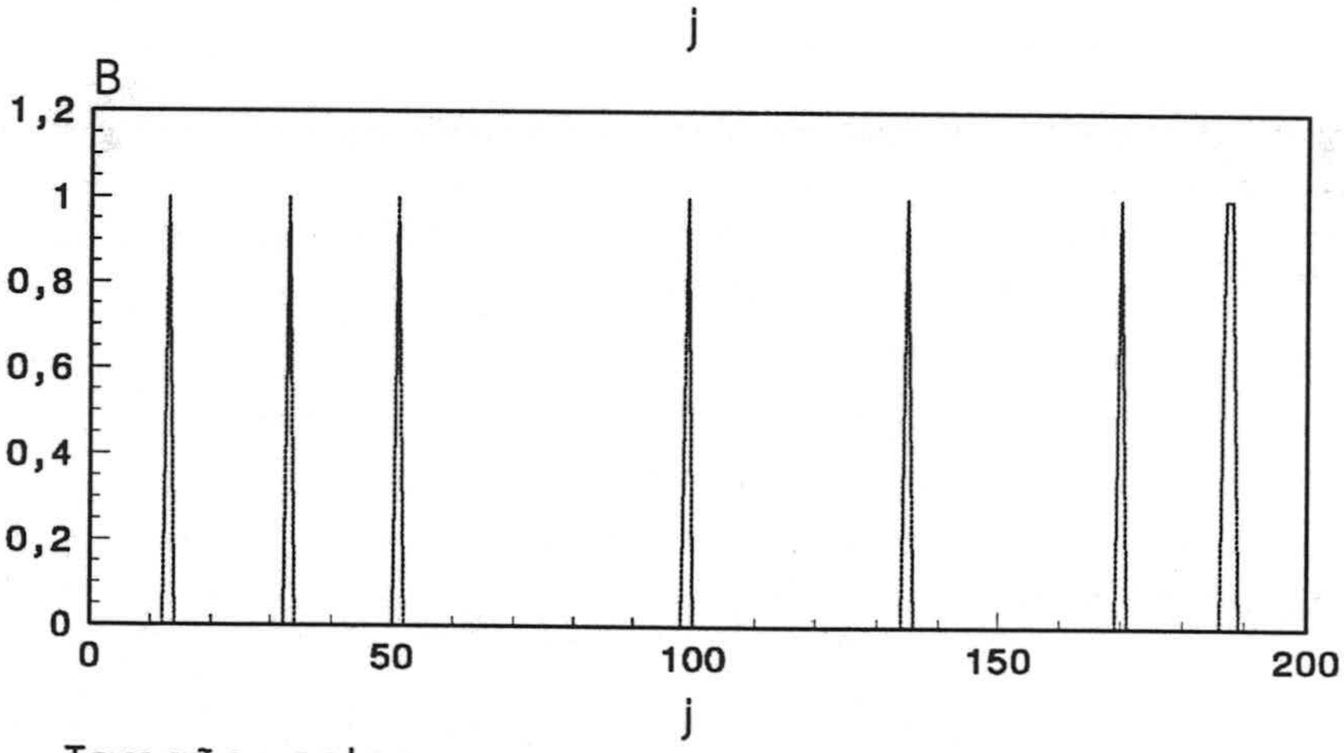
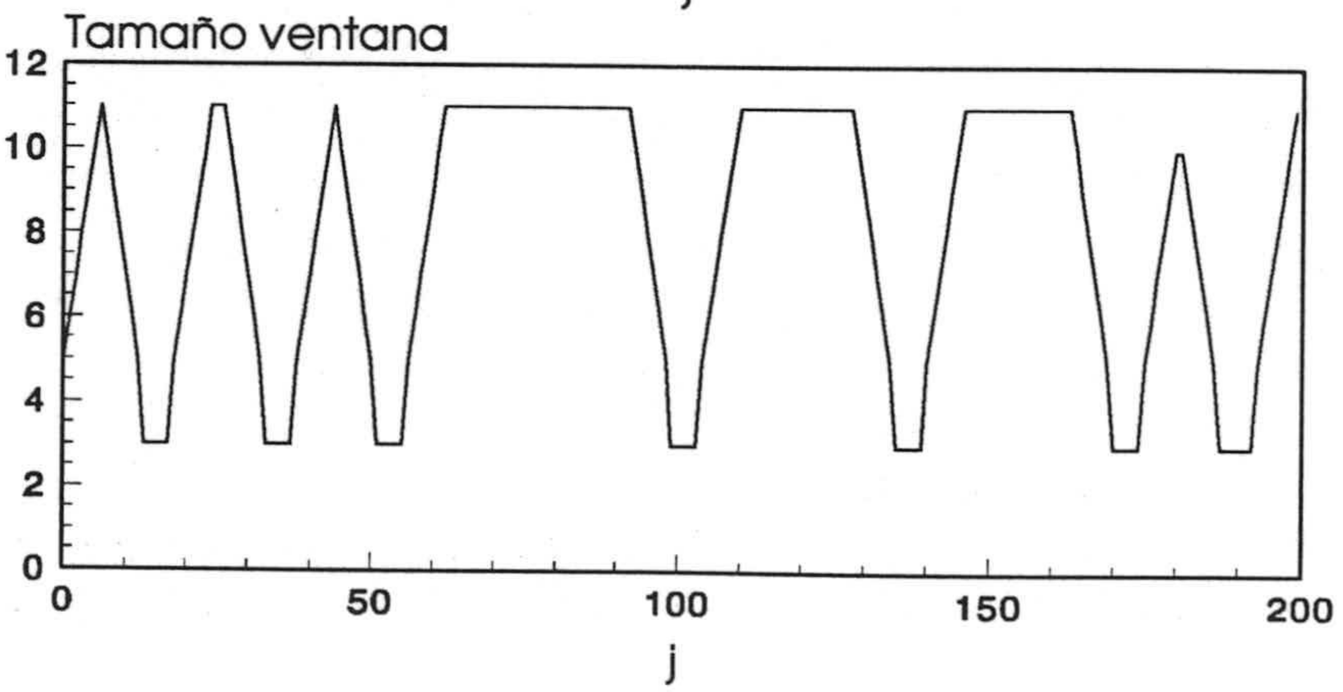


Figura 31.d: Tamaño de la ventana empleada por el filtro de mediana con ventana adaptativa, en función de la posición de los bordes en la imagen.



PROPIEDADES DEL FILTRO RESTDIVER

El filtro de ventana adaptativa que estamos presentando aquí posee una complejidad intrínseca que hace inviable un estudio estadístico detallado como el que se ha realizado para el filtro de mediana relajada. De hecho, incluso en un algoritmo tan sencillo como el anterior, la dificultad del análisis estadístico ha sido notable. Tanto más cuando el proceso de filtrado implica el empleo de un método de segmentación basado en una medida de cohesión de la complejidad como la de Jensen-Shannon. Por tanto, habremos de conformarnos con un estudio más reducido, fundamentado en el comportamiento del filtro de mediana, que es el núcleo de *restdiver*. Investigaremos las características de nuestro filtro frente a supresión de ruido y preservación de bordes y detalles. Lo primero se realizará bajo condiciones de ruido impulsivo, en dónde se calcularán las probabilidades de colapso y de ruido gaussiano. Las propiedades de conservación de bordes y detalles se examinarán considerando imágenes fuente geométricas.

Para realizar este estudio estadístico emplearemos un resultado simple, pero que facilitará enormemente la tarea. Fijémonos que *restdiver* no es más que un filtro de mediana, pero con diferente tamaño dependiendo de la cercanía de los detalles de la imagen. Por tanto, nuestro algoritmo se comportará en cada píxel como una mediana tradicional de tamaño $(2N+1)\times(2N+1)$ con N variable. Como las propiedades del filtro de mediana son bien conocidas [JUS81] [GAL81] [ARC82] [ARC83] [FIT85] [DAV88], podremos estimar el comportamiento de *restdiver* si conociéramos el número de veces que se utiliza cada tamaño de ventana.

Lema 4. *Sea el tamaño del filtro* $(2N+1)\times(2N+1)$, *con* N *variable. Sea* $\Pi_m(N)$ *una propiedad del filtro de mediana correspondiente a dicho tamaño. Sea card*(N) *el número de ventanas de tamaño* N *empleadas por restdiver en el procesamiento completo de la imagen. Entonces la correspondiente propiedad del filtro restdiver será:*

$$\Pi_{\text{restdiver}} = \sum_{N=N_{min}}^{N_{max}} card(N)\Pi(N)$$

Demostración.

Se sigue de la independencia estadística entre las diferentes realizaciones del filtro de mediana tradicional que realiza *restdiver* durante un filtrado. ■

SUPRESIÓN DE RUIDO Y PRESERVACIÓN DE DETALLES

Cuando la imagen está contaminada con ruido impulsivo, el filtro suele eliminarlo. Puede suceder que, cuando la cantidad de ruido (esto es, la probabilidad de aparición de pulsos) es demasiado elevada, la salida del algoritmo proporcione a su vez un valor ruidoso. La probabilidad de que aparezca un pulso en la salida de un filtro se determina mediante la probabilidad de colapso. La tabla 6 muestra dicha probabilidad para el filtro de mediana para diferentes tamaños de ventana y distintas probabilidades de ruido, calculadas a partir de la distribución de probabilidad de salida del mismo [DAV88] [PIT90], suponiendo una distribución binomial de entrada.

Tabla 6: Se muestran los valores de la probabilidad de colapso para el filtro de mediana tradicional. Se emplean diferentes tamaños de ventana y distintas probabilidades de aparición de ruido impulsivo.

p	TAMAÑO DE VENTANA				
	3x3	5x5	7x7	9x9	11x11
0,05	0,00725	0,001158	0,000194	0,000033	0,000006
0,625	0,01123	0,00222	0,00046	0,0001	0,00002
0,1	0,028	0,00856	0,002728	0,000891	0,000296
0,125	0,04297	0,01605	0,00624	0,00248	0,001
0,2	0,104	0,05792	0,033344	0,01958	0,011654
0,25	0,156	0,103	0,07	0,048	0,034
0,3	0,216	0,16308	0,123036	0,098809	0,078225
0,375	0,316	0,275	0,243	0,216	0,194
0,4	0,352	0,31744	0,289792	0,266568	0,246502
0,5	0,5	0,5	0,5	0,5	0,5

Aplicando el lema 4, el cálculo de la probabilidad de colapso es inmediata. Basta multiplicar la probabilidad de colapso del filtro de mediana para cada tamaño de ventana por el número de ventanas de ese tamaño que ha empleado *restiver*, y normalizar. En esencia, la probabilidad de colapso del filtro *restdiver* es un promedio ponderado de los valores de dicha probabilidad, para diferentes tamaños de ventana, correspondientes al filtro de mediana. Es decir, que en puridad sólo podemos hablar de probabilidad promedio de colapso en el caso del filtro que nos ocupa. Los resultados obtenidos aparecen en la tabla 7. Comparando ésta con la tabla 6 se deduce que las probabilidades de colapso para *restdiver* son muy similares a las de la mediana de ventana 5×5 (mayores para 5% y 10% de proba-

bilidad de aparición de ruido; menores para 20% y 30%). Por tanto el filtro de mediana con ventana adaptativa se comporta en promedio como un filtro de mediana tradicional de ventana 5×5 a efectos de robustez frente a ruido impulsivo, lo cual constituye una buena propiedad.

FUENTE DEL BORDE	NUMERO DE VENTANAS EMPLEADAS					PROBAB. COLAPSO
	3x3	5x5	7x7	9x9	11x11	
Geométrica 5% SP	12402	4446	3732	3007	16353	0,002399
Geométrica 10% SP	12143	4588	3819	3281	16169	0,00973
Geométrica 20% SP	10726	4306	3682	3319	17967	0,044052
Geométrica 30% SP	11058	4074	3519	3215	18134	0,130816
Geométrica 50% SP	10730	4189	3563	3269	18249	0,5

Tabla 7: Valores de la probabilidad de colapso para el filtro de mediana con ventana adaptativa, calculados mediante el lema 4, una vez conocido el número de ventanas de cada tamaño empleadas por el filtro.

Del mismo modo podemos calcular la varianza de salida del filtro *restdiver* para una imagen fuente S afectada de ruido blanco gaussiano $N(0,1)$, con sólo aplicar el lema 4 a los valores conocidos de tal varianza para el filtro de mediana. En la tabla 8 se plasman los resultados de dicha operación, para diferentes tamaños de ventana. La varianza promedio de salida para el filtro *restdiver* es de 0.058865, muy parecido valor de nuevo al correspondiente al filtro de mediana tradicional de ventana 5×5. Podemos apreciarlo gráficamente en la figura 32.

	FILTRO DE MEDIANA					RESTDIVER
	Tamaño de la Ventana					
	3x3	5x5	7x7	9x9	11x11	
Varianza	0,166101	0,061746	0,031775	0,01929	0,012936	0,058865
NºVentanas	10445	3279	3110	2949	20241	

Tabla 8: Valor de la varianza local promedio para el filtro de mediana con ventana variable.

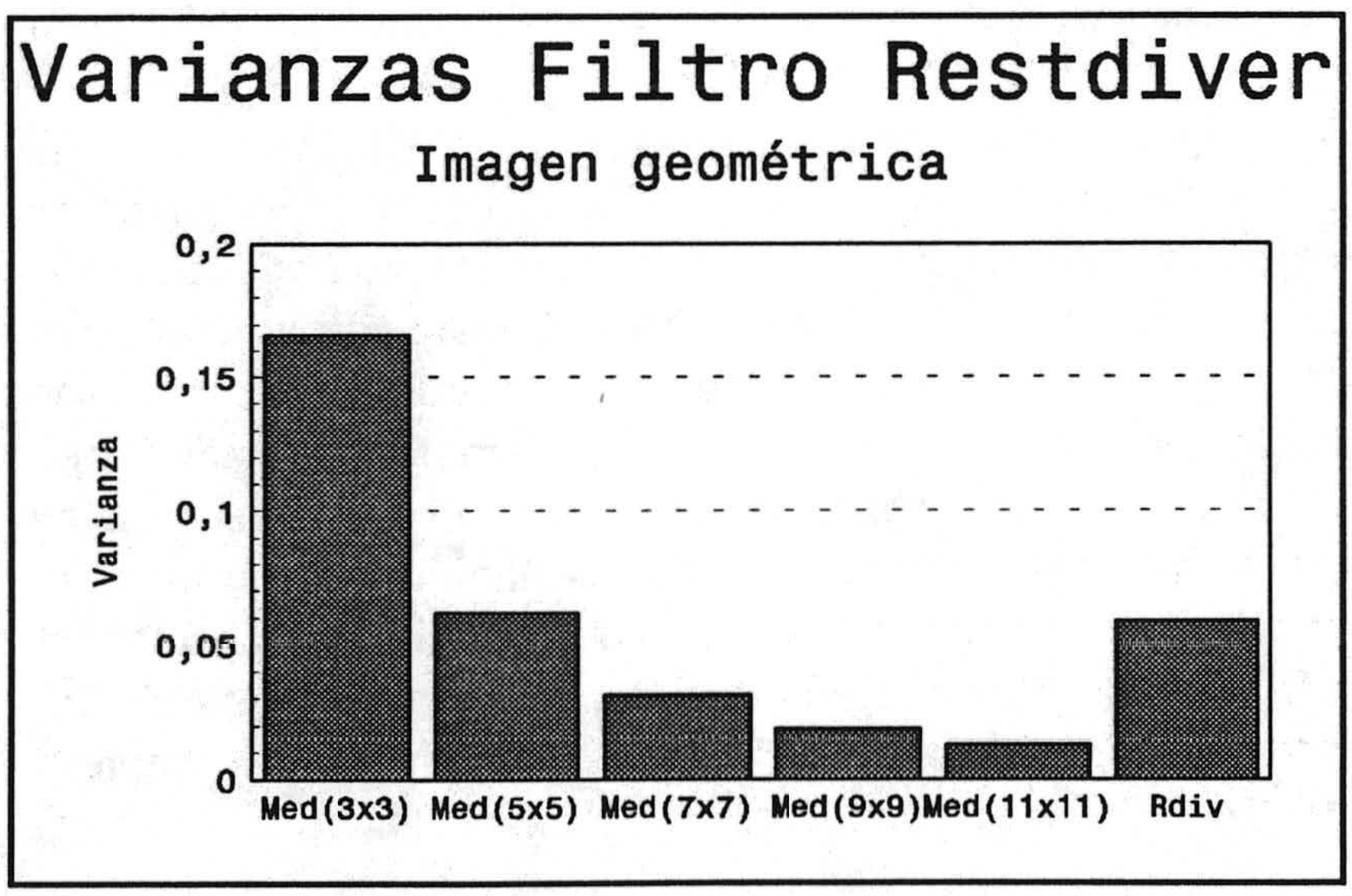


Figura 32: Comparación de las varianzas locales de los filtros de mediana tradicional y con ventana adaptativa (*restdiver*, en la figura se denota por *Rdiv*).

APLICACIÓN A IMÁGENES RUIDOSAS

Con objeto de comprobar los resultados obtenidos acerca del comportamiento teórico del filtro frente al ruido y bordes, y para obtener nuevos datos sobre el mismo, realizaremos una serie de experimentos de filtrado empleando imágenes sintéticas y naturales contaminadas con diferentes cantidades y tipos de ruido.

Como el objetivo principal de *restdiver* es la preservación de bordes, manteniendo la capacidad de supresión de ruido, hemos elegido como originales imágenes que presentan bordes bien contrastados. Así, se emplearán en los experimentos las siguientes, todas ellas con escala de píxel de 64 niveles de gris y tamaño de 200×200 píxeles:

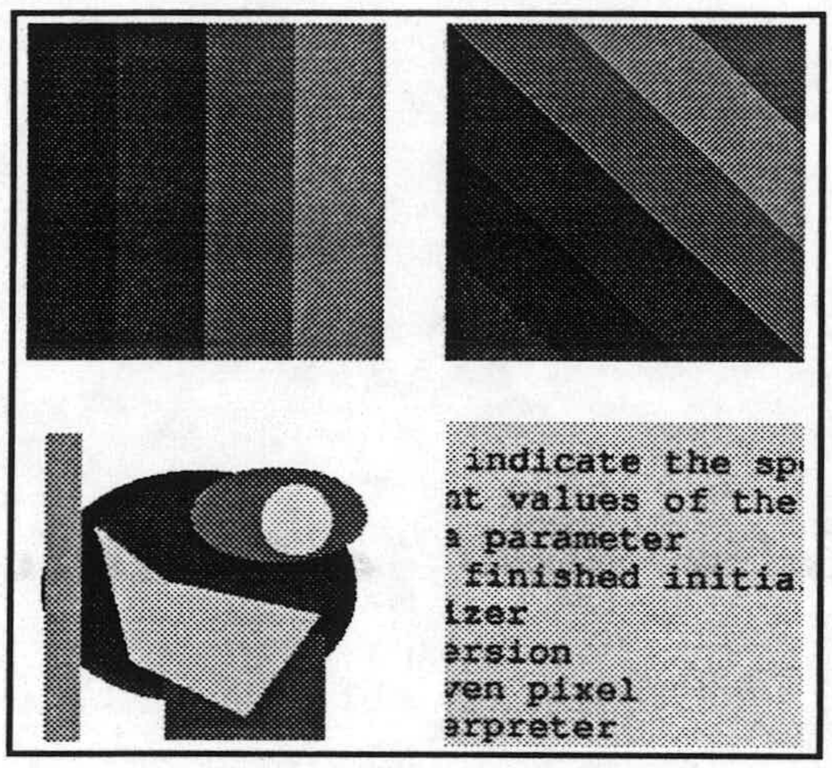


Figura 33: Imágenes originales empleadas para estudiar las propiedades del filtro de mediana con ventana adaptativa.

·*Imagen de bandas verticales:* Representa una serie de 4 bandas verticales de igual anchura, cada una de las cuales tiene un nivel de gris que difiere de la vecina en 10 unidades.

·*Imagen diagonal:* Conformada por 6 bandas diagonales de igual anchura, con niveles de gris diferentes. El salto de una banda a la siguiente es de 10 niveles de gris.

·*Imagen geométrica:* Constituida por diferentes formas geométricas, con varios perfiles y niveles de gris. Permite el estudio del comportamiento frente a bordes y líneas con diferentes orientaciones.

·*Imagen de texto:* Se trata de una imagen de texto mecanografiado, obtenida mediante digitalización de una página real.

Todas ellas pueden verse en la figura 33. Cada una de ellas ha sido contaminada con ruido impulsivo, tanto *sal y pimienta* como uniformemente distribuido, con diversas probabilidades de aparición: 5%, 10%, 20%, 30% y 50%. Se han filtrado entonces las imágenes fuente tanto con el *restdiver* como con el filtro de mediana tradicional, con ventanas de tamaño 3×3 y 5×5, con objeto de comparar comportamientos. El procesamiento se ha iterado para poner de manifiesto la evolución del resultado. En las tablas 9 y 10 se plasman los datos obtenidos de la realización de tales experiencias. En particular, la tabla 9 muestra los valores correspondientes al filtrado de las imágenes fuente *S* afectadas de ruido impulsivo *sal y pimienta*. En la columna

EXPERIMENTO		MAE (S-O)	MAE (R-O) RESTDIVER				MAE(R-O) MEDIANA			
			PERF, 1ª	PERF, MEJOR	JS, 1ª	JS, MEJOR	3x3, 1ª	3x3, MEJOR	5x5, 1ª	5x5, MEJOR
DIAGONAL	5	6,51235	0,09575	0,091525[2]	0,09585	0,091650[3]	0,09575	0,091525[2]	0,1041	0,104100[1]
	10	12,947925	0,121275	0,108550[2]	0,1223	0,108500[2]	0,135475	0,108550[2]	0,125575	0,115600[2]
	20	25,577049	0,223675	0,162425[3]	0,21835	0,162925[3]	0,44675	0,162425[3]	0,195425	0,156275[3]
	30	38,525425	0,5196	0,245325[3]	0,515825	0,242650[3]	1,982375	0,272700[3]	0,325325	0,236450[2]
	50	63,944275	2,3902	0,640850[3]	2,595475	0,676350[3]	13,13908	1,968375[3]	1,6744	0,393925[3]
VERTICAL	5	6,40075	0,011575	0,007875[3]	0,012825	0,008825[c]	0,011875	0,007875[3]	0,0039	0,003900[c]
	10	12,7221	0,041925	0,028375[3]	0,041925	0,028375[3]	0,0538	0,028375[3]	0,0287	0,017625[c]
	20	25,627001	0,1018	0,073075[c]	0,111625	0,073100[3]	0,34465	0,073075[3]	0,08435	0,050425[3]
	30	38,087399	0,35675	0,176225[3]	0,433	0,186025[3]	1,660975	0,226025[3]	0,20565	0,121625[3]
	50	63,14415	1,206375	0,370775[3]	2,051125	0,366875[3]	12,54518	1,494350[3]	1,31405	0,352825[3]
GEOMÉTRICO	5	6,40955	0,27575	0,275750[1]	0,2815	0,281500[1]	0,27575	0,275750[1]	0,630525	0,630525[1]
	10	12,5007	0,57655	0,526975[2]	0,5789	0,543875[2]	0,58225	0,526975[2]	0,742725	0,742725[1]
	20	25,263451	1,1095	0,997925[2]	1,163075	1,045075[2]	1,3096	0,997925[2]	1,41025	1,285675[2]
	30	37,4338	1,8613	1,449525[2]	1,873625	1,431875[2]	2,9215	1,478425[3]	1,9242	1,679275[2]
	50	64,261826	5,547325	2,893300[3]	5,793725	2,894000[3]	15,27443	4,192950[3]	4,2497	2,683900[3]
TEXTO	5	1,57195	1,680525	1,680525[1]	1,807525	1,807525[1]	1,645475	1,645475[1]	3,312725	3,312725[1]
	10	3,181025	1,800175	1,800175[1]	1,915275	1,915275[1]	1,772975	1,772975[1]	3,31135	3,31135[1]
	20	6,291075	2,076575	2,076575[1]	2,179025	2,179025[1]	2,09195	2,091950[1]	3,3702	3,37020[1]
	30	9,27985	2,449725	2,449725[1]	2,5963	2,596300[1]	2,638575	2,483075[2]	3,43905	3,439050[1]
	50	15,82598	4,00835	3,310950[2]	4,424225	3,441450[3]	5,9058	3,519675[3]	4,090175	3,722150[2]

Tabla 9: Datos correspondientes al procesamiento de las imágenes propuestas afectadas de ruido impulsivo *sal y pimienta* mediante los filtros de mediana tradicional y con ventana adaptativa. Entre corchetes, la iteración en la que se produjo el resultado mostrado.

MAE(S-O) aparece el error medio absoluto medido entre la imagen fuente y la original, mientras que MAE(R-O) contiene ese error, pero entre la imagen resultado y la original. Se aprecia también la existencia de columnas encabezadas por *Perf.*, que corresponden con el filtrado de *restdiver* pero empleando una imagen de bordes perfecta, esto es, obtenida a partir de la imagen original sin ruido. El objeto de haber incluido tales columnas no es otro que el de comprobar si el filtrado empleando imágenes de borde detectados a partir de imágenes ruidosas, que aparecen en las columnas encabezadas por *JS*, altera apreciablemente el resultado. El índice *1ª* significa que el resultado mostrado corresponde a la primera iteración del filtro, mientras que *mejor* indica que se plasma el mejor valor obtenido en la serie de iteraciones; entre corchetes aparece la iteración en la

que se produjo dicho valor. La tabla 10 muestra los mismos resultados de filtrado, pero para ruido impulsivo uniformemente distribuido (denotado por *G*), con un formato análogo al de la tabla anterior.

Evidentemente, el filtro *restdiver* se comportará como mínimo igual que el filtro de mediana 3×3 por construcción del algoritmo, dónde la ventana mínima es la 3×3. Si el ruido es muy abundante, o hay bordes muy separados en la imagen, el *restdiver* mejorará al filtro de mediana 3×3, pues puede aumentar el tamaño de su ventana adaptable, mejorando las cualidades de eliminación de ruido. Por supuesto, la capacidad de preservación de bordes y detalles siempre será la misma que el de este filtro de mediana, si la detección de bordes ha sido buena. Si dicha operación resultase defectuosa es posible que empeore

Tabla 10: Resultados de filtrar las imágenes propuestas, contaminadas con ruido impulsivo uniformemente distribuído. Se han empleado los algoritmos de mediana tradicional y *restdiver*.

EXPERIMENTO		MAE (S-O)	MAE (R-O) RESTDIVER				MAE(R-O) MEDIANA			
			PERF, 1ª	PERF, MEJOR	JS, 1ª	JS, MEJOR	3x3, 1ª	3x3, MEJOR	5x5, 1ª	5x5, MEJOR
DIAGONAL	5	3,59025	0,0963	0,09292[2]	0,0963	0,092900[2]	0,096775	0,0929[2]	0,108575	0,108575[1]
	10	7,208925	0,11235	0,105425[3]	0,11235	0,105425[3]	0,118275	0,105425[3]	0,122425	0,117425[2]
	20	14,9392	0,25125	0,177125[3]	0.257775	0,176925[3]	0,390575	0,178275[3]	0,23295	0,197500[3]
	30	21,81465	0,4317	0,297975[3]	0,4504	0,297050[3]	1,07615	0,312425[3]	0,4097	0,359100[3]
	50	36,536049	1,390675	0,739425[3]	1,61955	0,798425[3]	6,38865	2,077625[3]	1,98865	0,930675[3]
VERTICAL	5	3,856725	0,013575	0,01105[3]	0,013575	0,011050[3]	0,013575	0,01105[3]	0,00425	0,001175[c]
	10	7,549125	0,034775	0,029825[3]	0,034775	0,029825[3]	0,042852	0,029825[3]	0,030025	0,021600[3]
	20	14,9662	0,13305	0,10925[3]	0,150325	0,109325[3]	0,3302	0,110850[3]	0,168725	0,157875[2]
	30	22,2272	0,26045	0,183325[3]	0,275375	0,181650[3]	1,106075	0,224550[3]	0,3035	0,267600[3]
	50	37,77515	1,2466	0,6535[3]	1,380625	0,672500[3]	7,741175	2,967475[3]	2,785275	1,244575[3]
GEOMÉTRICO	5	5,356575	0,333925	0,333925[1]	0,333925	0,333925[1]	0,340825	0,340825[1]	0,70405	0,70405[1]
	10	10,1433	0,572475	0,540000[2]	0,588875	0,556675[2]	0,584825	0,540000[2]	1,0029	1,0029[1]
	20	20,549049	1,40405	1,247500[2]	1,404625	1,236825[2]	1,885525	1,265750[3]	1,901275	1,843825[2]
	30	30,417101	2,4654	2,081175[3]	2,48395	2,114125[2]	4,843525	2,464150[3]	3,282	2,933325[2]
	50	49,927399	10,1587	7,702350[2]	10,5852	7,906900[3]	20,82995	14,23640[3]	14,4872	11,28425[3]
TEXTO	5	1,498	1,670475	1,670475[1]	1,832725	1,832725[1]	1,6446	1,644600[1]	3,265475	3,265475[1]
	10	2,569875	1,8039	1,8039[1]	1,968125	1,968125[1]	1,777925	1,777925[1]	3,282675	3,282675[1]
	20	5,1038	2,11475	2,11475[1]	2,29575	2,295750[1]	2,18555	2,185550[1]	3,3364	3,3364[1]
	30	7,5944	2,58215	2,58215[1]	2,77155	2,771550[1]	3,007325	3,007325[1]	3,583925	3,540750[2]
	50	12,728675	4,696825	4,395025[2]	5,266525	4,732325[3]	7,157325	5,773775[3]	6,118375	5,315000[3]

algo el resultado de la mediana. Estas propiedades se verán comprobadas a la vista de las gráficas que se han realizado a partir de los datos de las tablas.

En las figuras 34 y 35 se presenta la variación de la mejora del filtrado (representada mediante el CMAE) frente al tipo de imágenes para un nivel de ruido impulsivo, tanto *sal y pimienta* como uniformemente distribuido. En la figura 34 podemos comparar los resultados de filtrado de ruido iiimpulsivo *sal y pimienta*. Aquellos valores correspondientes a la primera iteración del procesamiento se representan mediante una línea continua, mientras que con trazo discontinuo se muestran los mejores resultados de todas las iteraciones. Observamos que el mayor grado de mejora se experimenta para la imagen de bandas verticales, como habíamos predicho, hasta el punto de que la mejora del resto de las imágenes apenas es perceptible. Sólo cuando el ruido es abundante (a partir del 20%) el CMAE del filtrado de la imagen de bandas verticales decrece y podemos discernir mejoras en las imágenes de bandas verticales y diagonales. En cualquier caso, se obtinen tanto mayores mejoras cuanto menor es el porcentaje de ruido.

El filtro de mediana de ventana 5×5 proporciona los mejores resultados si se itera el filtrado para todas las imágenes, aunque a medida que el ruido aumenta *restdiver* se le acerca; en cambio el mediana 3×3 en primera iteración es el que se comporta peor. De hecho, nuestro algoritmo filtra tanto mejor cuanto mayor es el ruido, como era de esperar, pues intenta eliminar el mayor ruido posible manteniendo la preservación de los bordes y líneas de la imagen. Si lo comparamos con el filtro de mediana 3×3, que es el que posee las mismas características de respeto de detalles que *restdiver*, el algoritmo adaptativo siempre se muestra más efectivo.

Respecto a los resultados de las primeras iteraciones, sólo difieren de los mejores a medida que aumenta el ruido. Dicha diferencia es mayor para los filtros de mediana que para nuestro algoritmo, lo cual indica que *restdiver* está especialmente indicado para proporcionar buenos resultados sin iterar el proceso, que si recordamos es para lo que fue diseñado. Por otro lado, el comportamiento del filtro adaptativo cuando se basa en una imagen de bordes B_{ij} perfecta (*Perf.* en los diagramas) es casi el mismo que tiene cuando emplea imágenes de borde obtenidas a partir de la imagen fuente (*JS* en las gráficas), y ello

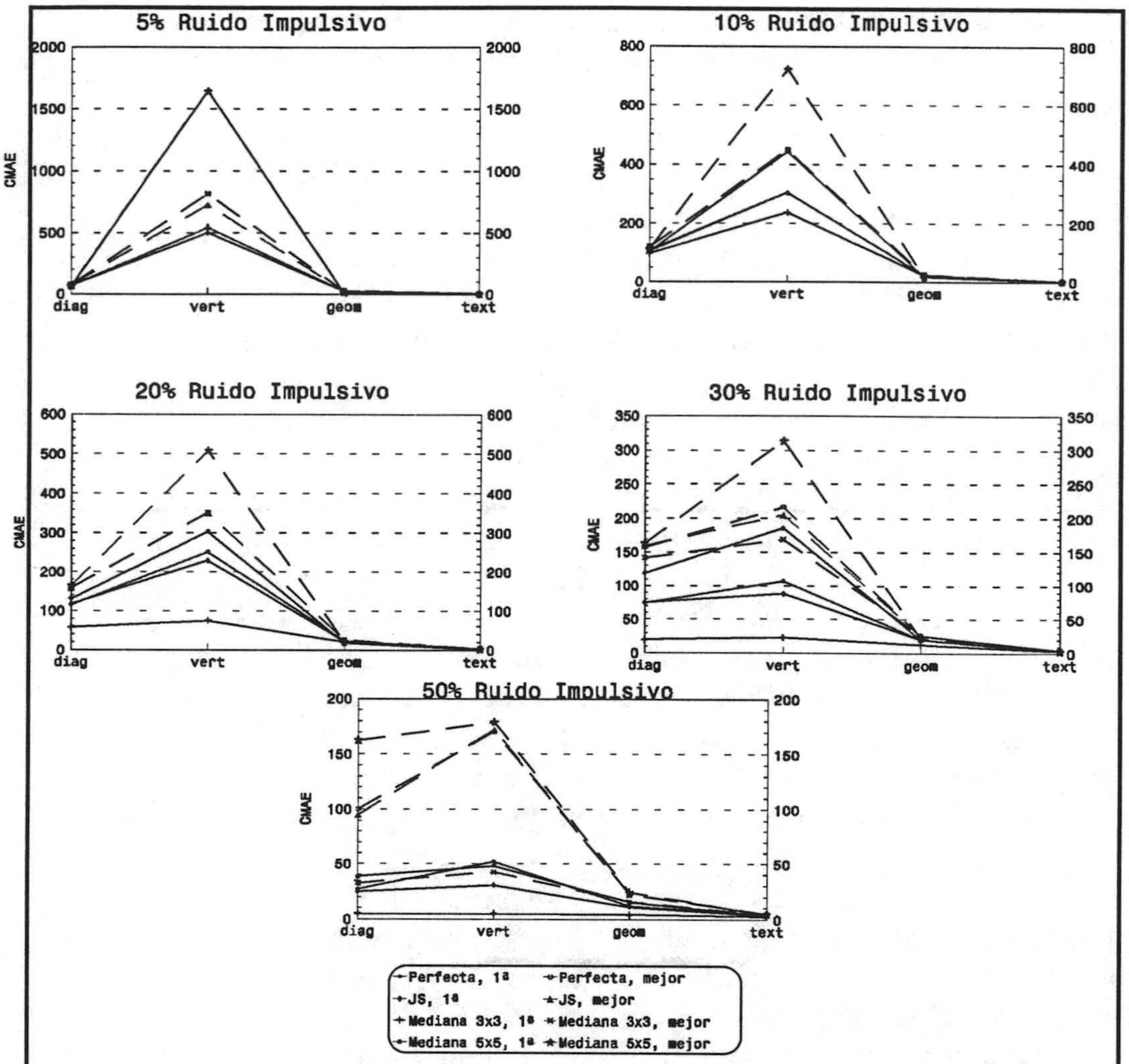


Figura 34: Representación gráfica de los resultados del filtrado de las imágenes escogidas mediante los algoritmos de mediana tradicional y con ventana adaptativa.

tanto para primera como para mejor iteración. Esta coincidencia habla en favor tanto de la robustez del algoritmo adaptativo como del método de detección de bordes empleado, el cual, incluso con un ruido tan intenso como el 50%, ha funcionado con una gran efectividad. En estas gráficas se aprecia también claramente el colapso de la mediana, especialmente con ventana pequeña, cuando el ruido es muy abundante.

Contemplemos a continuación la figura 35, dónde se pueden ver los resultados de filtrado de ruido impulsivo uniformemente distribuido en diversos porcentajes para la batería de imágenes con que venimos trabajando. Las características que de tal figura se desprenden son esencialmente análogas a las del caso de ruido *sal y pimienta*, aunque conviene destacar algunas diferencias. En este caso la mediana 5×5 se comporta bastante peor que antes cuando el ruido comienza a ser molesto. El filtro

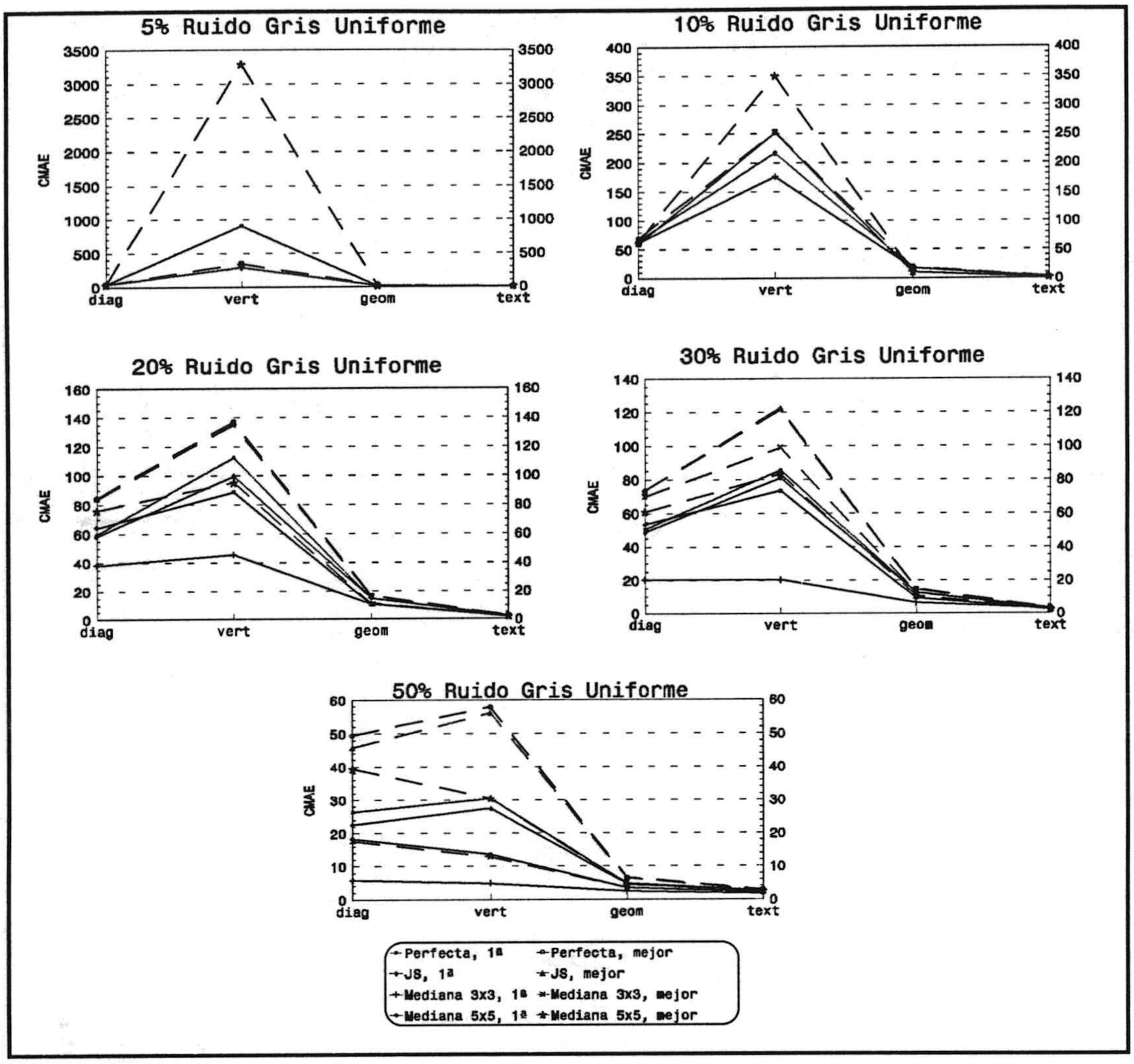


Figura 35: Resultados del procesamiento de las imágenes propuestas empleando los filtros de mediana tradicional y *restdiver.*

restdiver se muestra ahora más potente que en el caso precedente. Tanto para primera como para mejor iteración es mejor que el filtro de mediana. Es digno de mención el hecho de que incluso para un ruido tan extremadamente abundante como el 50% las diferencias entre el empleo de bordes perfectos y reales por parte del algoritmo adaptativo son mínimas.

También se han representado los valores numéricos de los experimentos agrupándolos por tipo de imagen, variando en cada gráfica la probabilidad de aparición de pulsos. Por ejemplo, en la figura 36 tenemos el comportamiento de los filtros frente a la imagen de bordes verticales. Para ruido impulsivo *restdiver* tiene el comportamiento más estable, el filtro de mediana 5×5 es el mejor y apenas se encuentran diferencias entre el uso de imagen de borde perfecta y real. El peor filtrado lo proporciona la mediana de ventana 3×3 en primera iteración. Si el ruido es uniformemente distribuido, se aprecia un notable

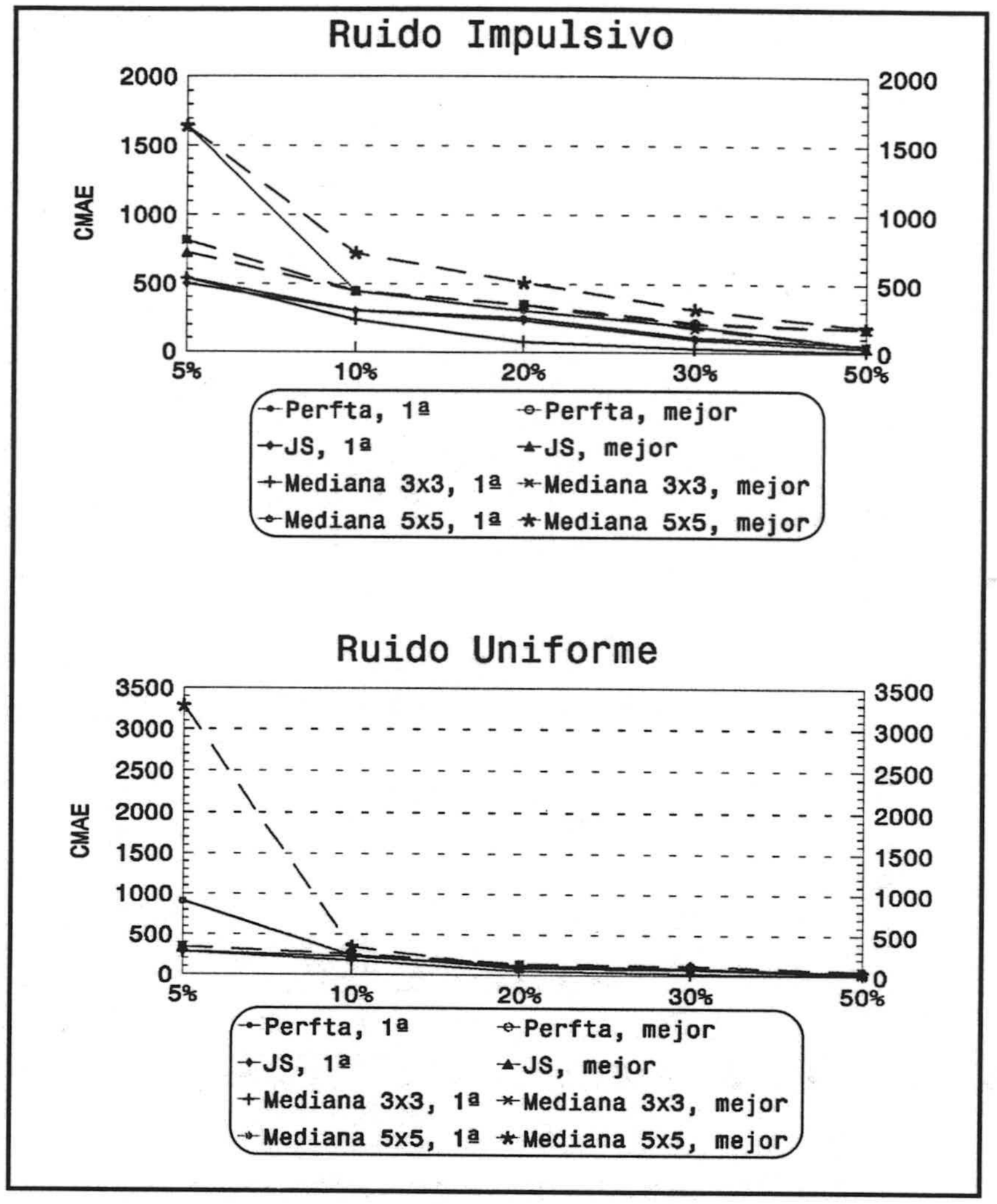


Figura 36: Comportamiento de los filtros de mediana tradicional y con ventana adaptativa frente al procesado de la imagen de bandas verticales contaminada de diferentes tipos y cantidades de ruido. Arriba vemos los resultados para ruido impulsivo *sal y pimienta* y abajo para ruido impulsivo uniformemente distribuido.

descenso en el CMAE del filtro de mediana de ventana grande al pasar del 5% de porcentaje de ruido en la imagen al 10%, a partir del cual todos los filtros tienen análogo comportamiento. Este hecho es achacable a que el filtro de mediana se satura en seguida con este tipo de ruido.

Fijémonos ahora en la figura 37, correspondiente a la imagen de bandas diagonales, la siguiente en cuanto a grado de mejora. De nuevo se observa que la influencia de la calidad de la imagen de bordes por parte de *restdiver* no es destacable, ni para ruido *sal y pimienta* ni para uniformemente distribuido. Este filtro mejora notablemente a la mediana de ventana 3×3 y a veces al de ventana mayor en ruido *sal y pimienta* y siempre en uniforme. En cuanto a la imagen de motivos geométricos, aparecen sus resultados en la figura 38. La mejora es mayor para ruidos altos en ruido *sal y pimienta* y para ruidos bajos en

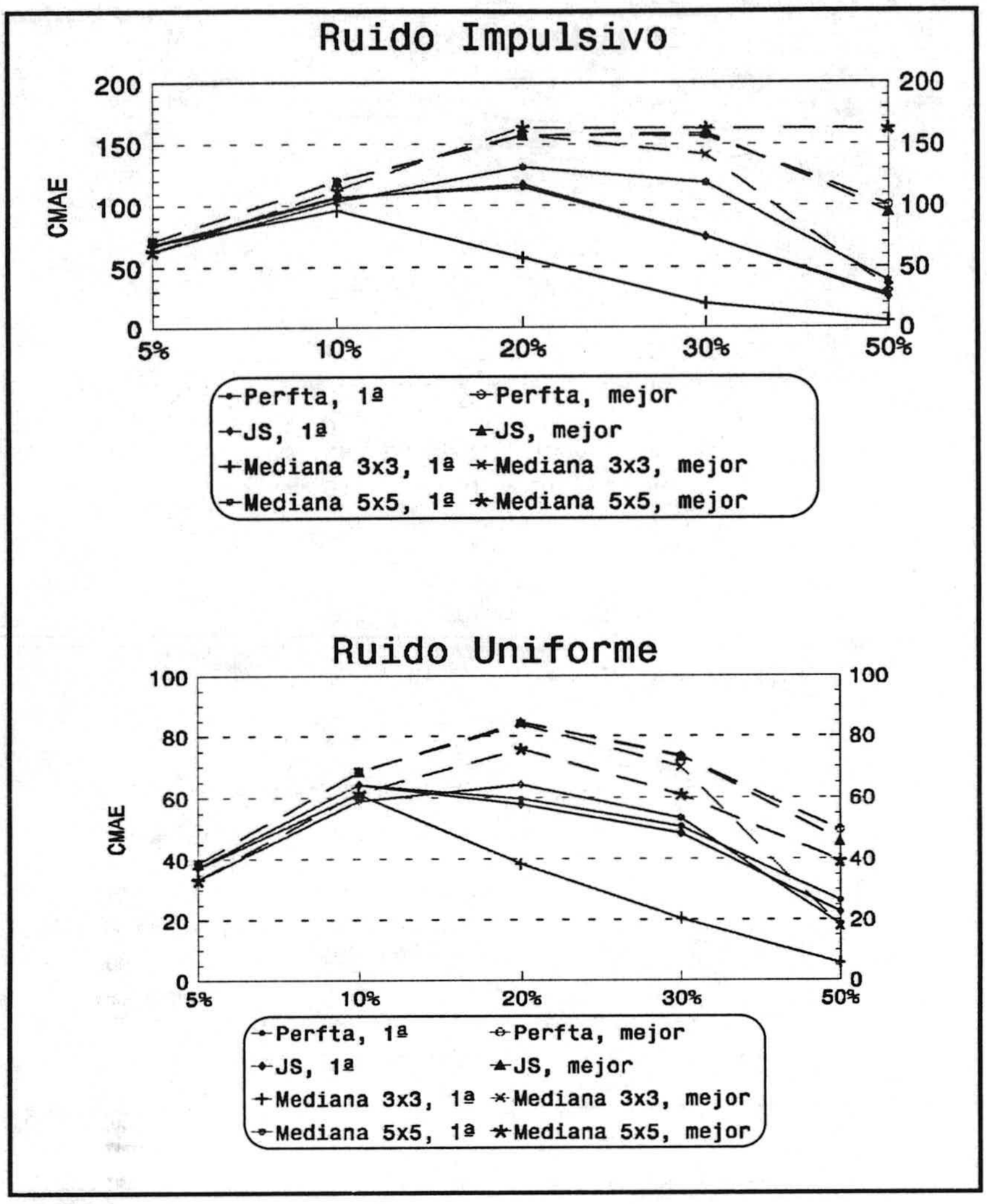


Figura 37: Resultados obtenidos con los filtros bajo estudio cuando se procesa la imagen de bandas diagonales afectada de ruido impulsivo de diferente naturaleza y abundancia. Los datos obtenidos de los experimentos con la imagen alterada con ruido *sal y pimienta* aparecen arriba. Abajo están los correspondientes al ruido impulsivo uniformemente distribuido.

ruido uniformemente distribuido. Aquí la mediana de ventana 5×5 no se comporta tan bien como en casos anteriores, al presentar la imagen estructuras más complejas. Ello permite a nuestro filtro mostrarse más efectivo, excepto para ruidos muy altos. La diferencia entre primera y mejor iteración aumenta con el nivel de ruido, aunque nunca es muy elevada. Se aprecia con claridad el declive al alcanzarse la probabilidad de colapso.

Veamos por último el comportamiento de los filtros frente a una imagen de texto real (figura 39), Aquí la diferencia entre el uso de imagen de bordes perfecta y real es nula hasta el 50% de procentaje de ruido. Entre los resultados de la primera y mejor iteración se aprecian levísimas discrepancias, que se mantienen constantes a pesar de que aumente el procentaje de ruido. De nuevo *restdiver* es el mejor, junto con la mediana de ventana pequeña, aunque ésta empieza a fallar cuando el ruido

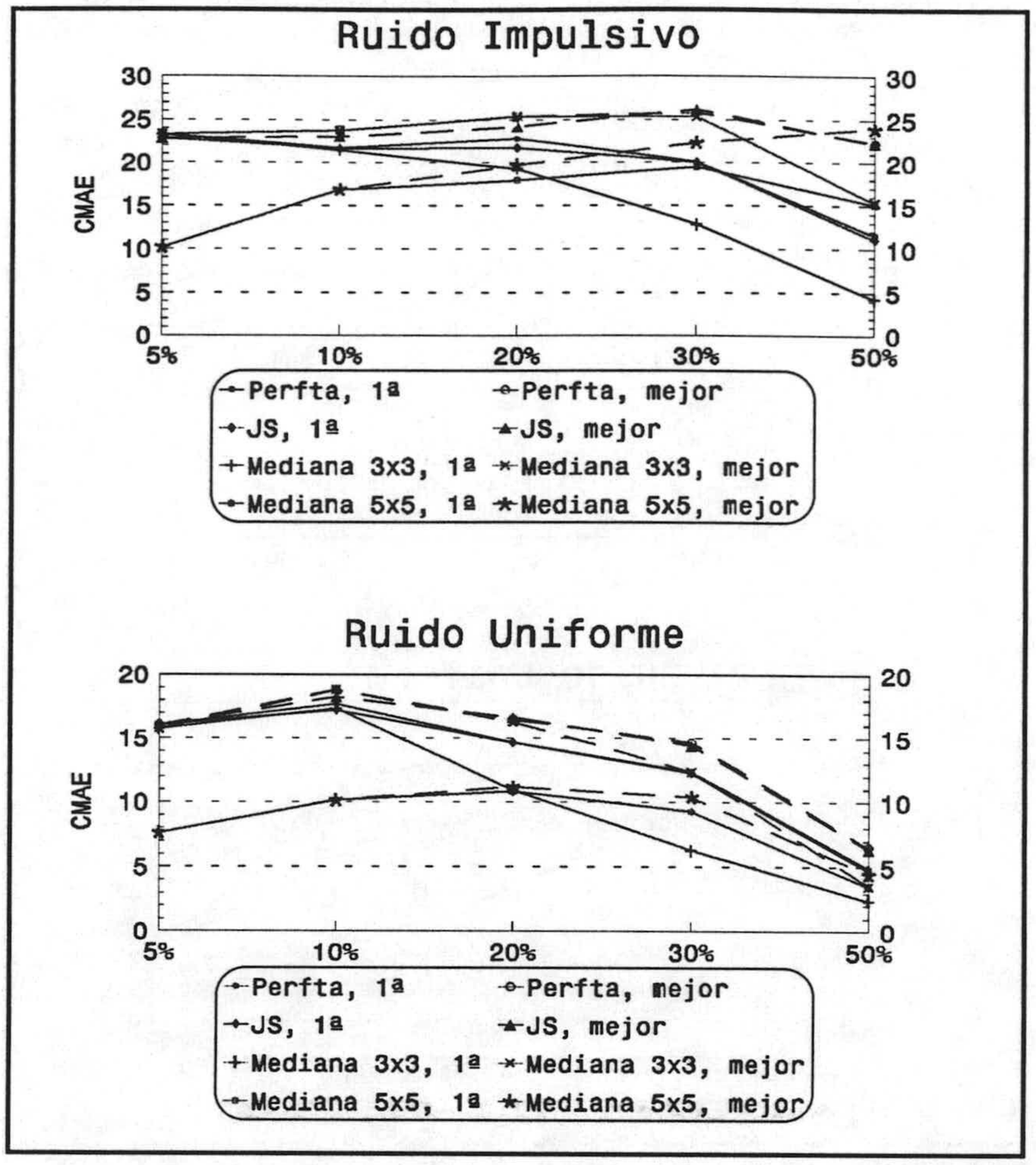


Figura 38: Gráficas representando los resultados de realizar filtrados a la imagen de mostivos geométricos afectada con distintas abundancias de ruido impulsivao *sal y pimienta* (arriba) y uniformemente distribuido (abajo).

es abundante. El filtro de mediana de ventana 5×5 es bastante lesivo con la imagen. La máxima mejora se obtiene para ruidos extremos, para ambos tipos de ruido.

En general puede decirse que los filtros se comportan mejor frente a ruido *sal y pimienta* que frente a ruido uniforme. Además, la mayor mejora se presenta con diferencia con la imagen de bandas verticales, seguida de la imagen de bandas diagonales, la imagen geométrica y la de texto, por ese orden. La razón principal para ello es que la imagen de bandas verticales presenta una estructura especialmente adecuada para un procesamiento con ventana deslizante, además de ser particularmente sencillo el detectar sus bordes. A medida que se complica la geometría, aparecen más bordes que impiden el crecimiento de las ventanas cuadradas deslizantes $W_{ij}(S)$, limitando la capacidad de eliminar ruido. Por lo tanto, podemos decir que el filtro de madiana con ventana adaptativa está especialmente indicado para procesar imágenes reales con ruido abundante, dón-

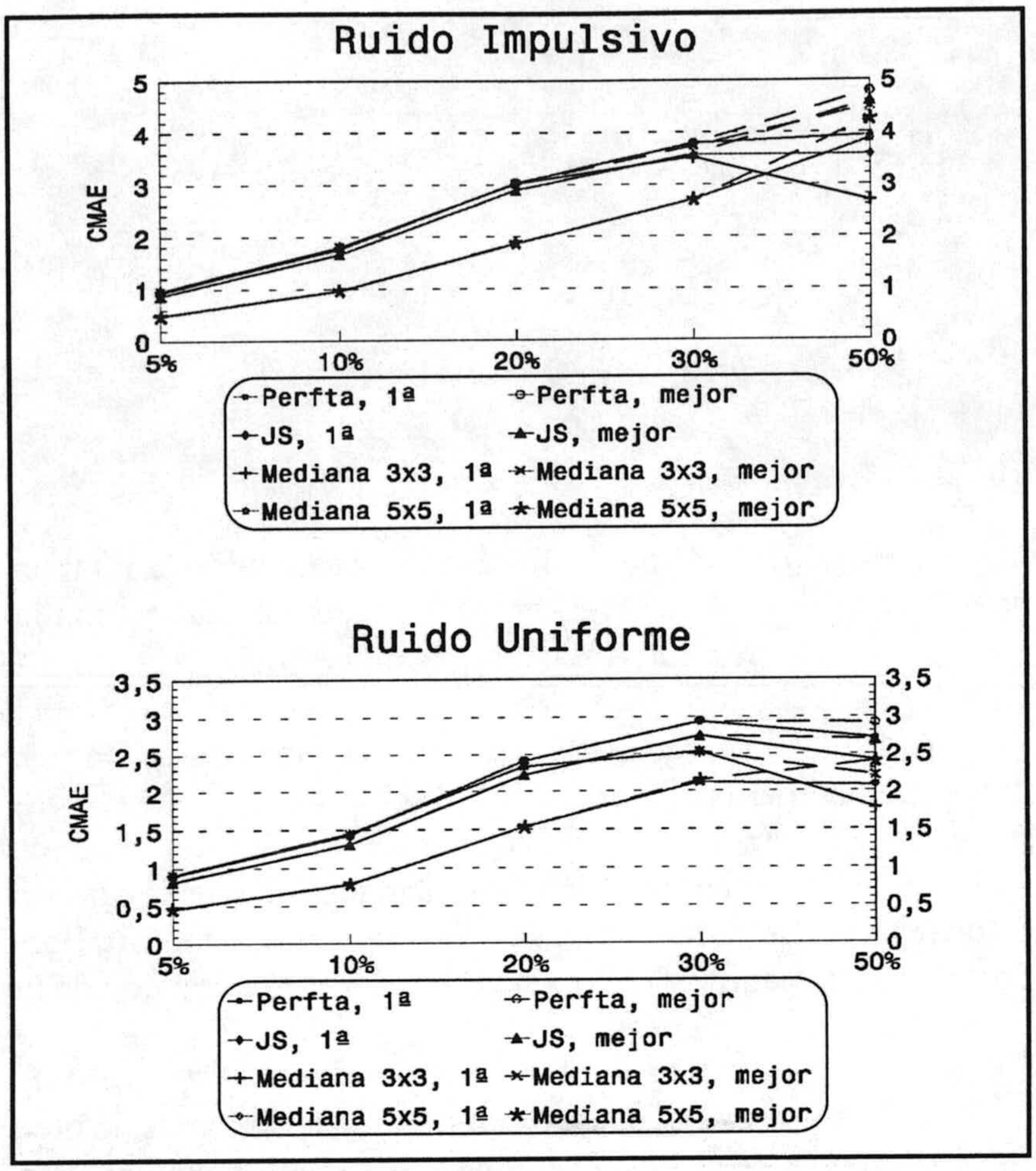
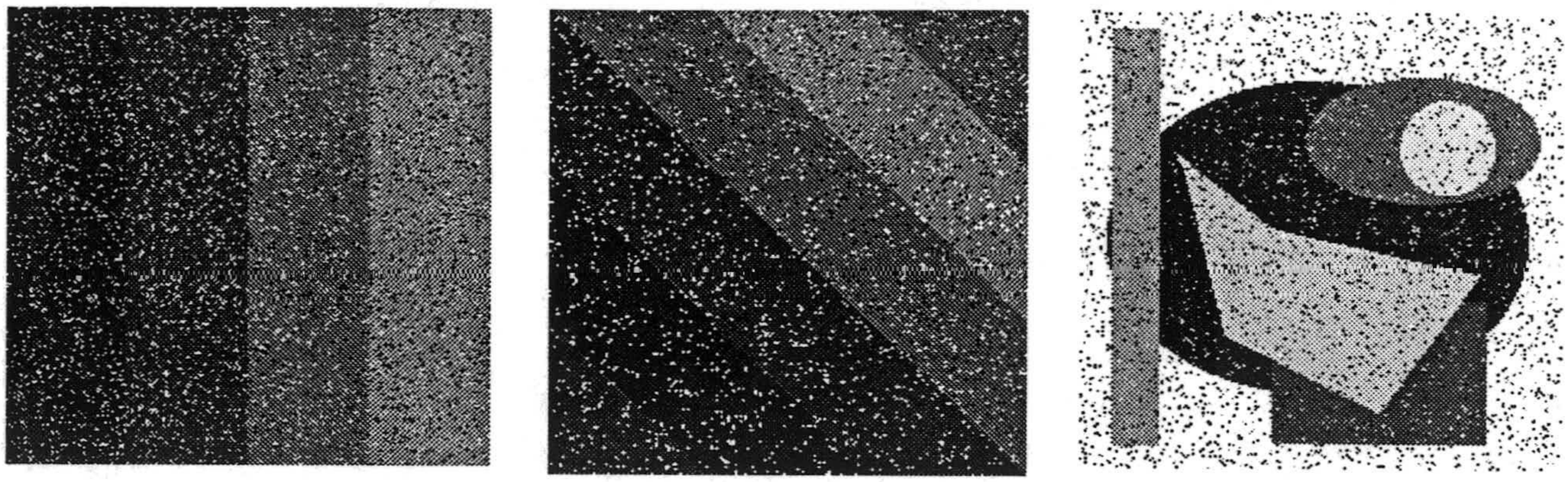


Figura 39: Representación gráfica de los datos obtenidos al realizar el filtrado de la imagen de texto contaminada por diferentes tipos y cantidades de ruido impulsivo (*sal y pimienta* arriba, y uniforme abajo).

de el filtro de mediana 3×3, que preserva los detalles igual que él, no es capaz de limpiar adecuadamente la imagen. Además, *restdiver* mejora al filtro de mediana cuando se realiza una única pasada de filtro, y no precisa el conocimiento de la imagen original para trazar la imagen de bordes B_{ij}, lo cual lo distingue de otros procesamientos adaptativos.

Para finalizar el capítulo, y antes de dar las conclusiones del mismo, conviene mostrar visualmente algunos de los resultados obtenidos en los experimentos arriba detallados. Así, en primer lugar se presentan (figura 40) las imágenes fuente empleadas, puesto que las originales ya han sido mostradas. En la figura 41 se pueden contemplar los resultados visuales del procesamiento de la imagen de bandas verticales contaminada de 20% de ruido uniforme, tanto con *restdiver* empleando imagen de borde real como con el filtro de mediana 5×5, el mejor de su clase. Se aprecia que la mediana elimina mejor el ruido en la primera iteración, aunque introduce más defectos en los bordes. Más acentuado se ve dicho efecto en los resultados de la



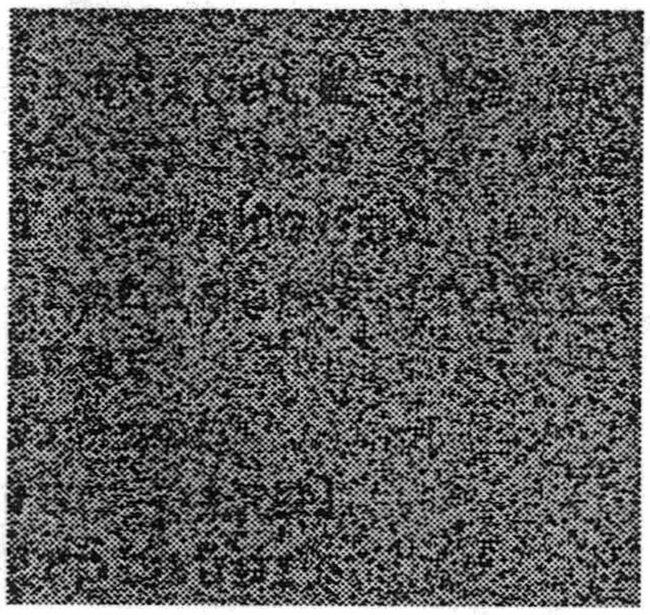


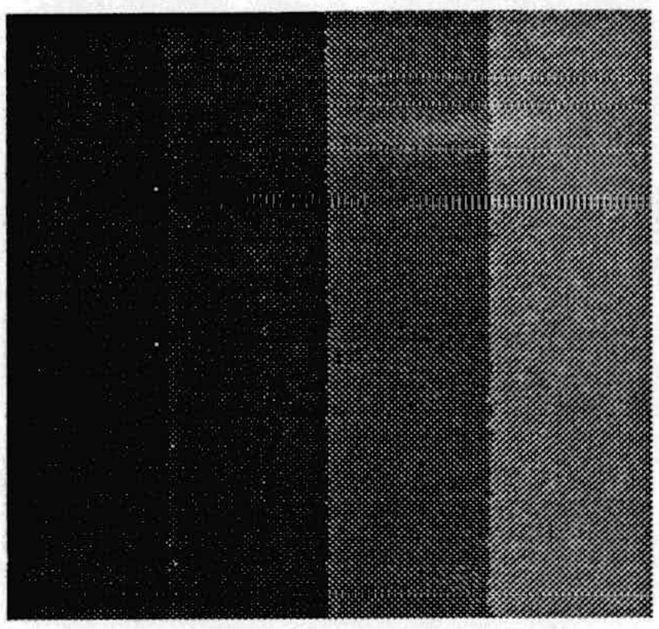
Figura 40: Ejemplo de imágenes fuente empleadas en los experimentos de filtrado. Arriba a la izquierda se observa la imagen de bandas verticales afectada por un 20% de ruido impulsivo uniformemente distribuido. En el centro, arriba, aparece la imagen de bandas diagonales contaminada con un 10% de ruido impulsivo *sal y pimienta*. A la derecha se muestra la imagen de motivos geométricos con un 10% de ruido impulsivo uniforme. Por último, abajo a la izquierda queda plasmada la imagen de texto alterada con un 50% de ruido impulsivo *sal y pimienta*.

mejor iteración. Con fines aclarativos se ha incluido la imagen de diferencias ecualizadas entre el resultado del filtrado con la mediana y *restdiver* con respecto a la imagen original.

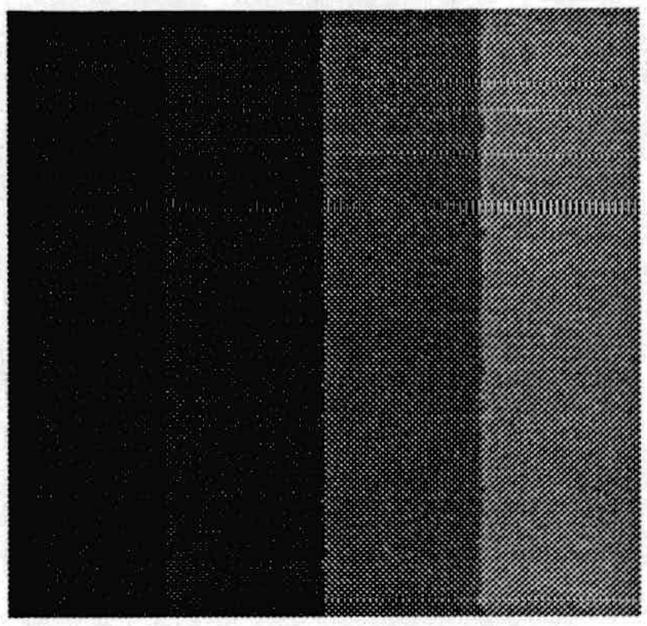
Las imágenes resultado correspondientes a la de bordes diagonales contaminada con 10% de ruido *sal y pimienta* aparecen en la figura 42. Los filtros comparados son los mismos que en la figura anterior. También se han incluido las imágenes de diferencia. Se observa que *restdiver* preserva mejor los bordes, mientras mantiene la capacidad de supresión de ruido de la mediana.

En la figura 43 se pueden comparar los resultados de procesar la imagen de motivos geométricos contaminada por 20% de ruido uniforme, de nuevo para los mismos filtros que en los casos precedentes. Podemos destacar algunos detalles que *restdiver* preserva mejor que el filtro de mediana. Por ejemplo, la delgada parte del óvalo que queda a la izquierda d ela imagen es casi eliminada por la mediana, no aí por el filtro adaptativo. También las esquinas del cuadrado sufren menos lesión en el caso de *restdiver*. Y el fenómeno de capilaridad observado entre el rectángulo y la forma irregular que muestra el filtro de mediana no aparece en el procesamiento con el filtro de ventana adaptativa.

Por otro lado, para la imagen de texto se ha escogido el procesamiento de la imagen fuente afectada de 50% de ruido *sal y pimienta* con *restdiver* con bordes reales y mediana 3×3, que es la que mejor se comporta en este caso (figura 44). El respeto de los bordes de las letras es el mismo, por construcción del algoritmo adaptativo. Pero la limpieza de la imagen es mucho más eficiente en el procesamiento de *restdiver*. El filtro de mediana 5×5 proporciona un resultado ilegible. Debe tenerse en cuenta la enorme cantidad de ruido empleada.

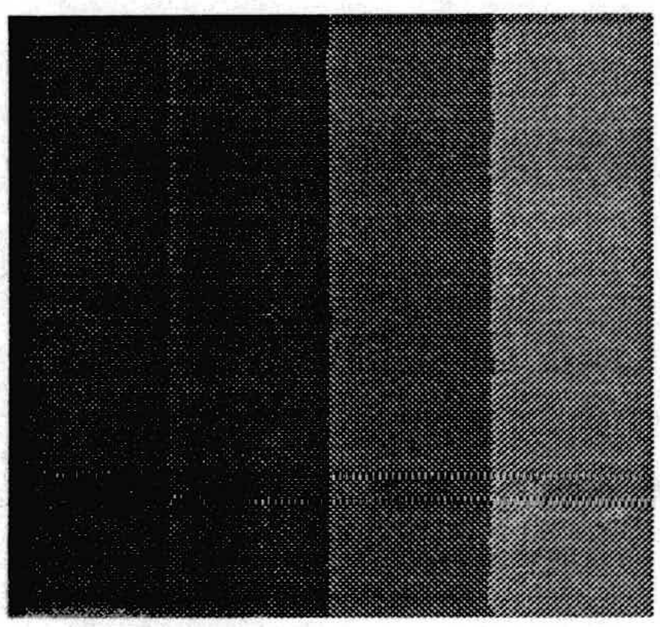


Restdiver, 1ª iteración

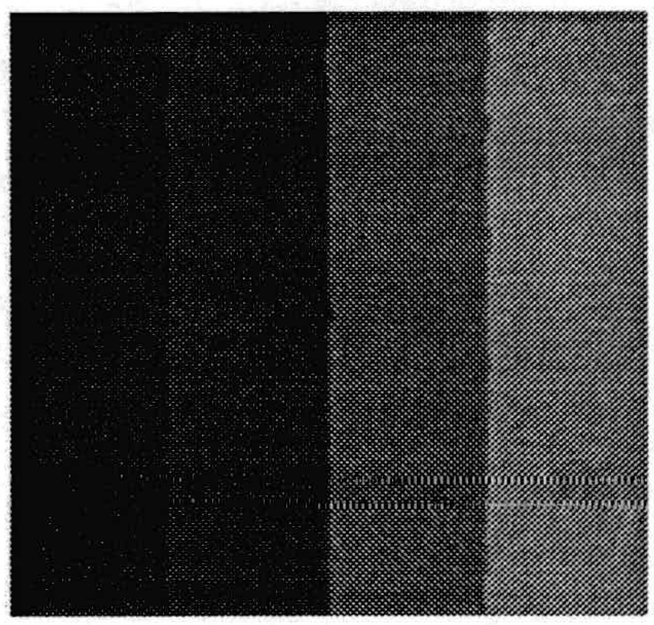


Restdiver, mejor iteración

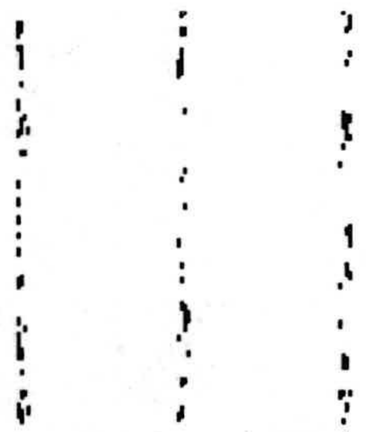
Figura 41: Imagen resultado de filtrar la imagen fuente de bandas verticales de la página anterior tanto con *restdiver* empleando imagen de bordes obtenida a partir de la imagen fuente como con el filtro de mediana tradicional, con ventana 5×5. Dada la similitud entre los resultados, se ha mostrado también la imagen de diferencia ecualizada entre el resultado y el original para ambos procesamientos. De su análisis se desprende que el filtro propuesto mejora sensiblemente el comportamiento de la mediana tradicional, sobre todo en el respeto de los bordes.



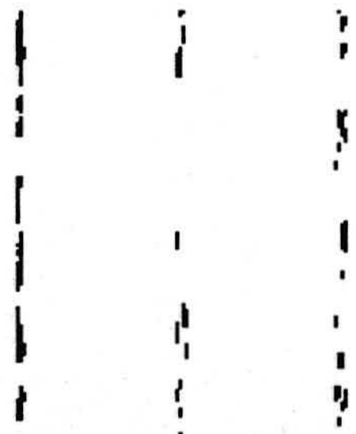
Mediana (5×5), 1ª iteración



Mediana (5×5), mejor iteración

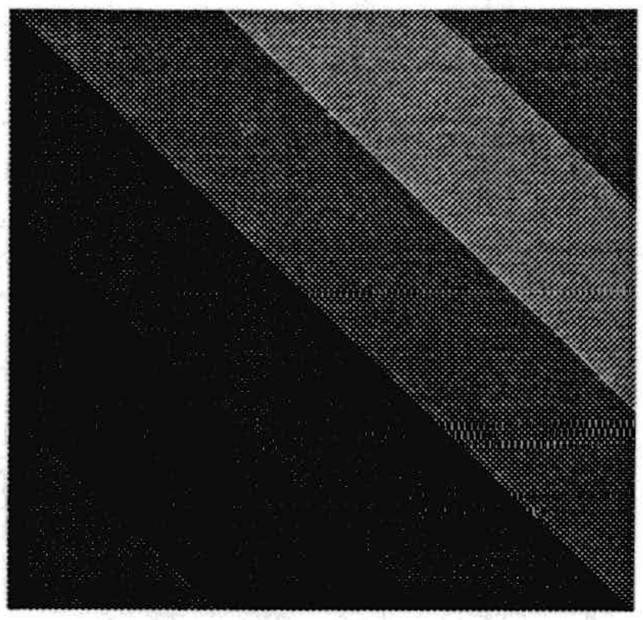


Diferencia resultado-original para el filtro *restdiver*.

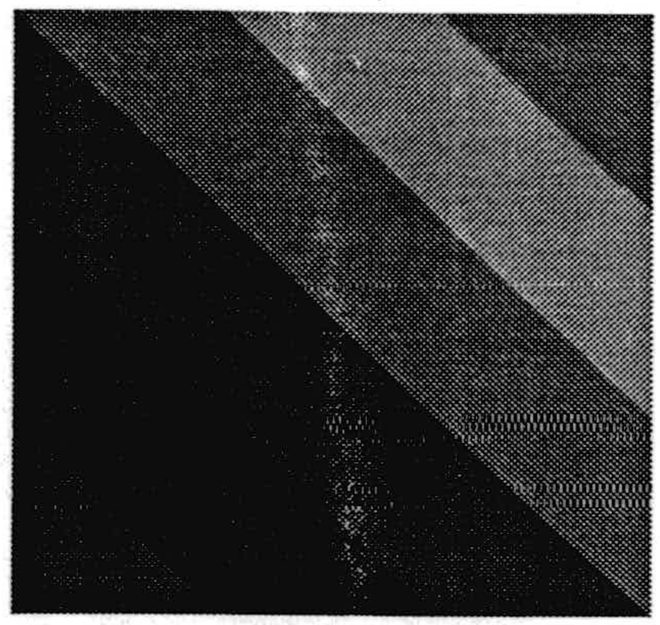


Diferencia resultado-original para el filtro de mediana

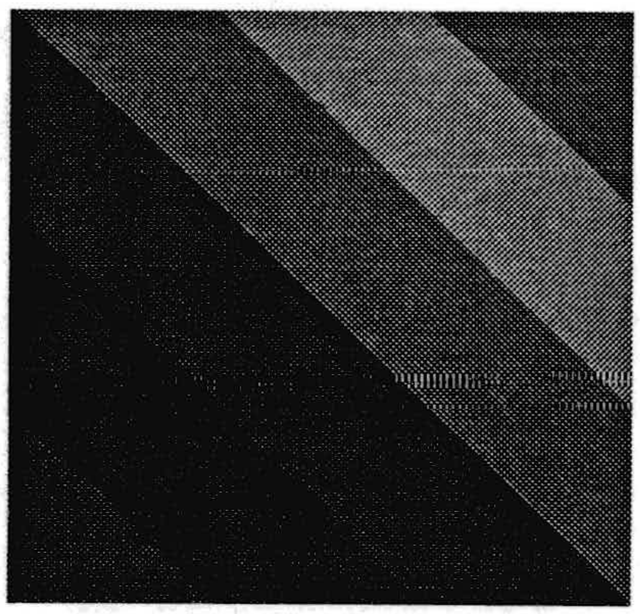
Figura 42: Imágenes de resultado obtenidas tras el procesamiento de la imagen fuente de bandas diagonales de la figura 40 empleando el filtro de mediana tradicional y con ventana adaptativa. De nuevo la similitud de los resultados aconseja incluir imágenes de diferencia ecualizadas resultado-original, de las que se desprende que el filtro *restdiver* es de nuevo el más respetuoso con los bordes de la imagen.



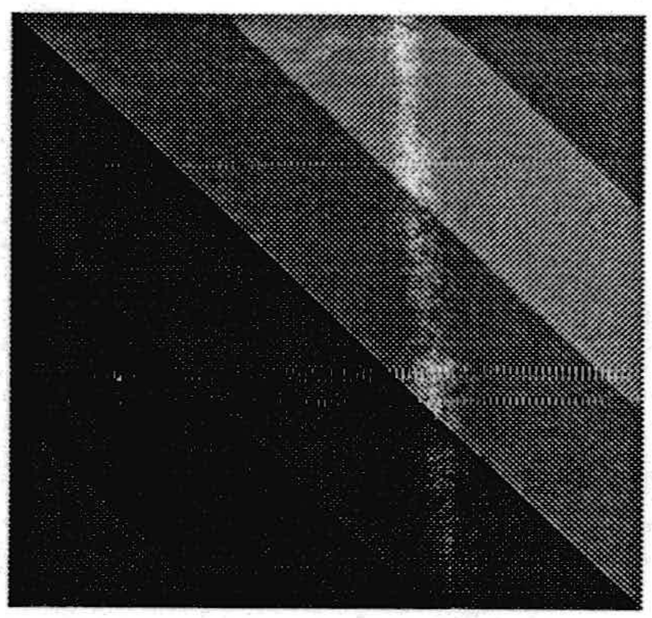
Restdiver, 1ª iteración



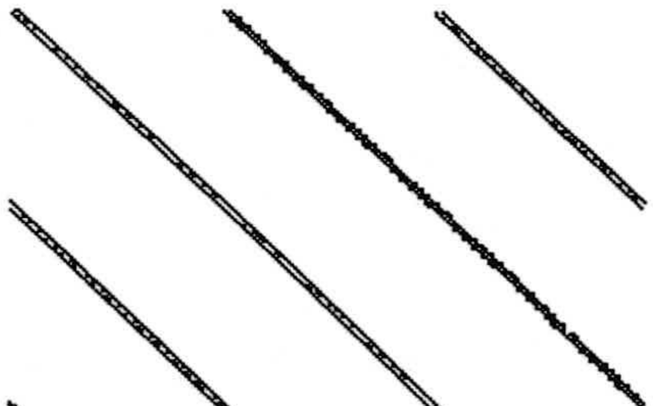
Restdiver, mejor iteración



Mediana (5×5), 1ª iteración

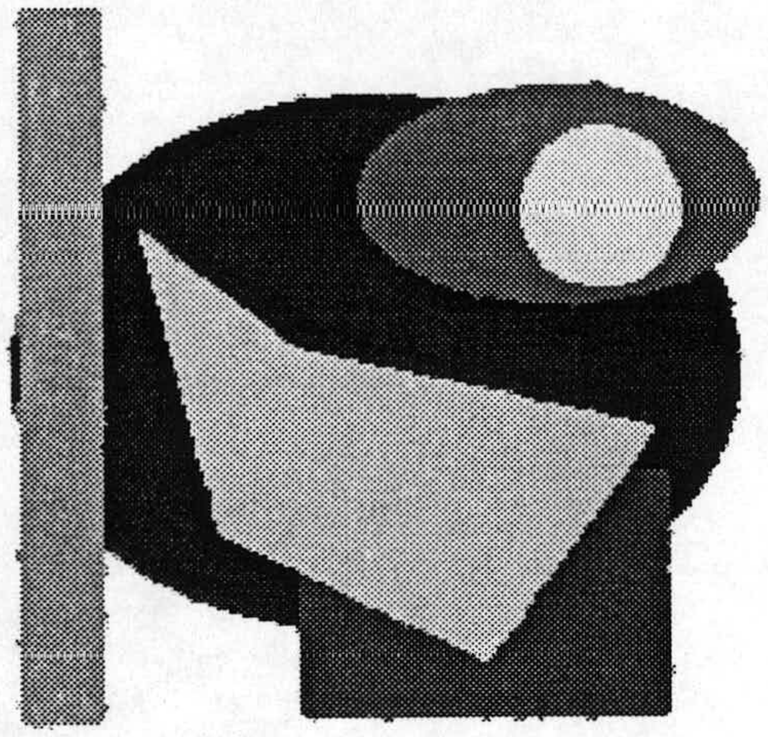


Mediana (5×5), mejor iteración

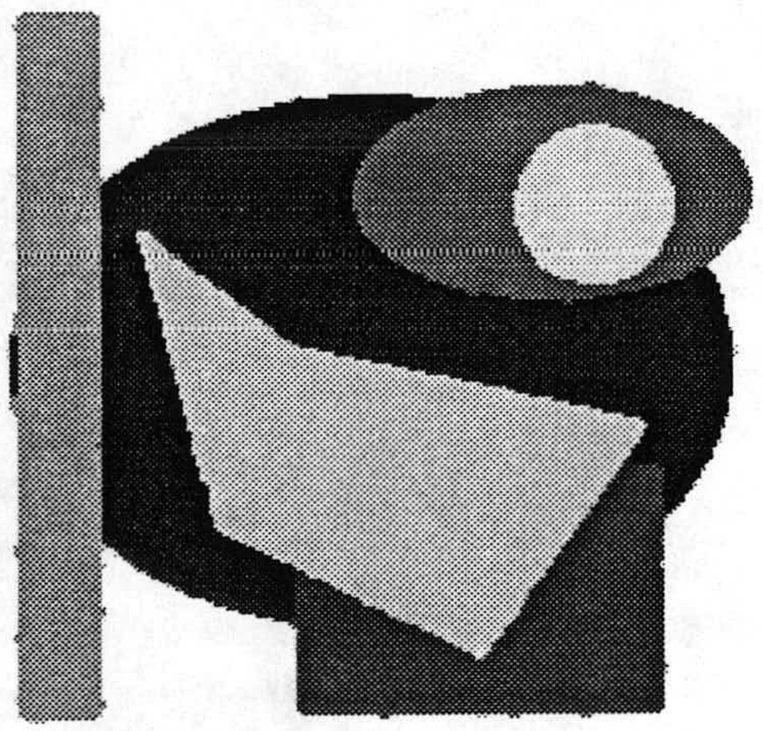


Diferencia resultado-original para el filtro *restdiver*.

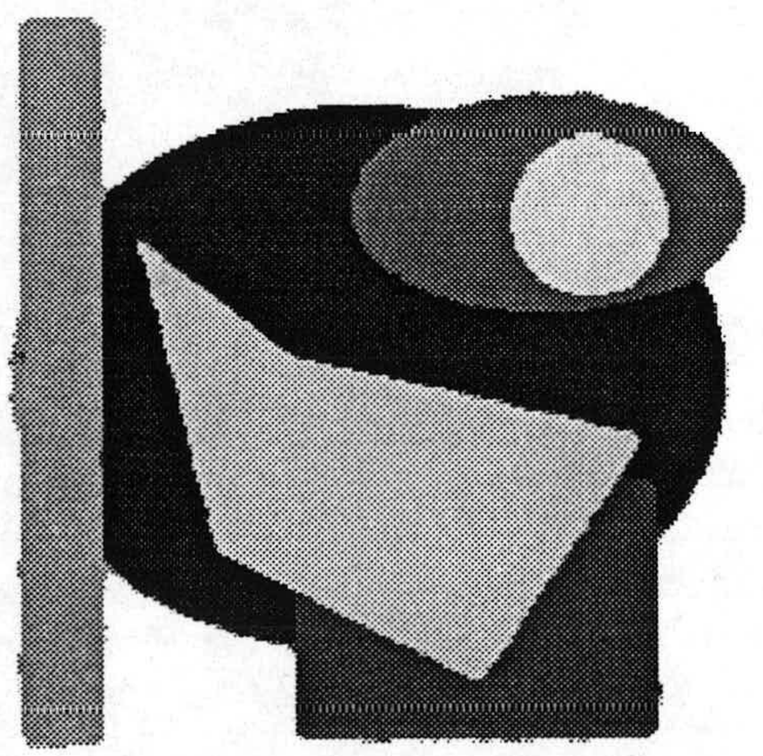
Diferencia resultado-original para el filtro de mediana



Restdiver, 1ª iteración

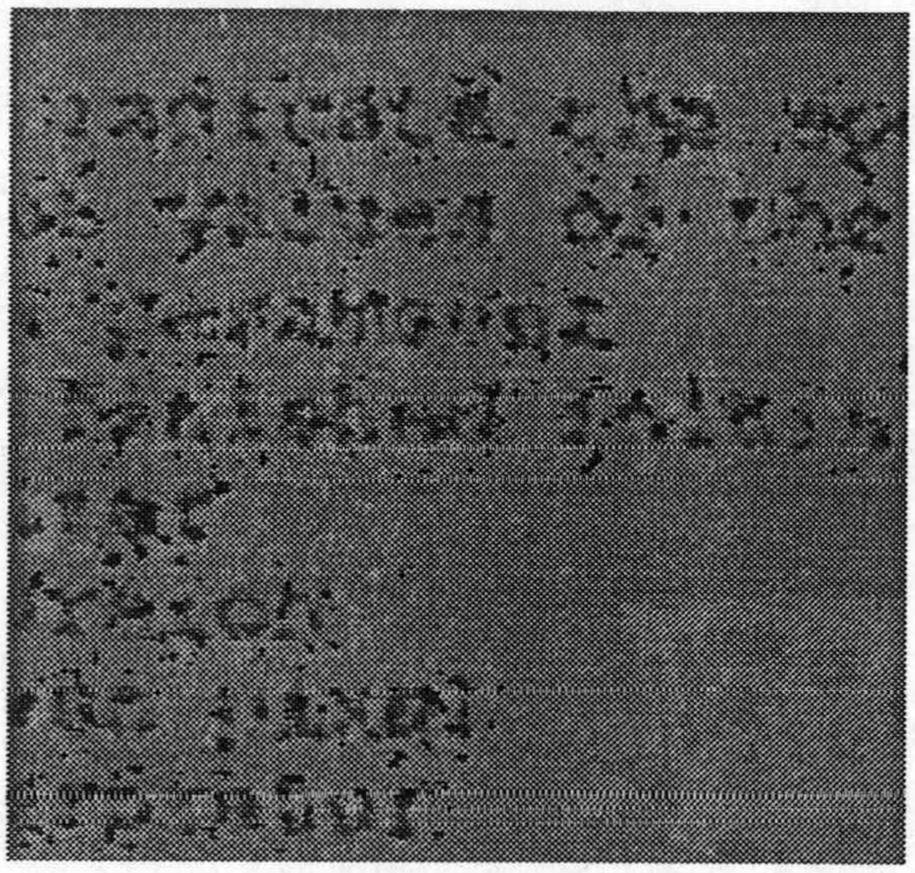


Restdiver, mejor iteración

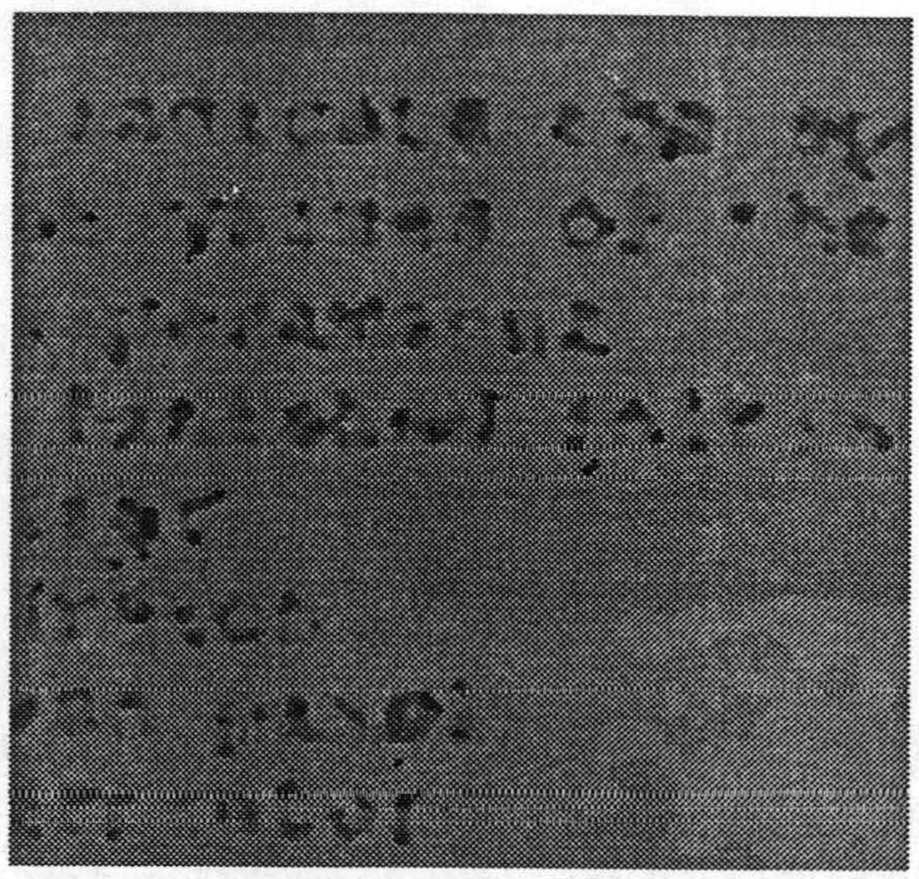


Mediana (5×5), 1ª y mejor iteración

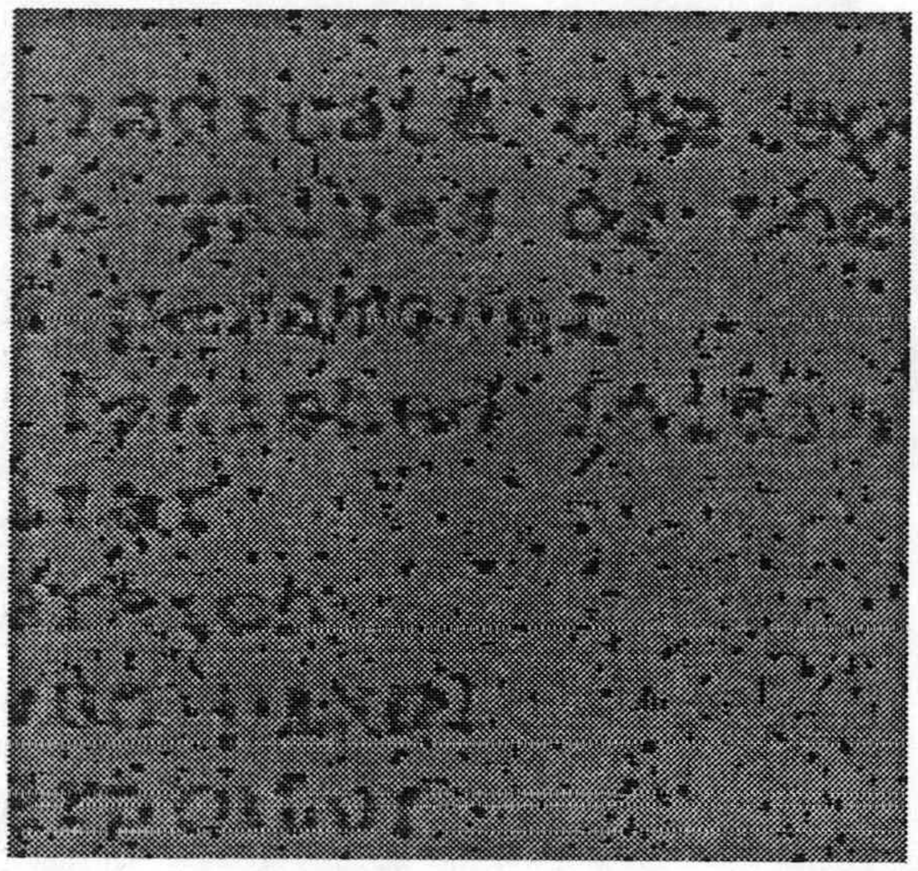
Figura 43: Se muestran las imágenes resultado del filtrado de la imagen fuente de motivos geométricos de la figura 40. Se han empleado tanto el filtro de mediana tradicional con ventana 5×5 como el propuesto, *restdiver*. En este caso la mejor iteración para el filtro de mediana coincide con la primera.



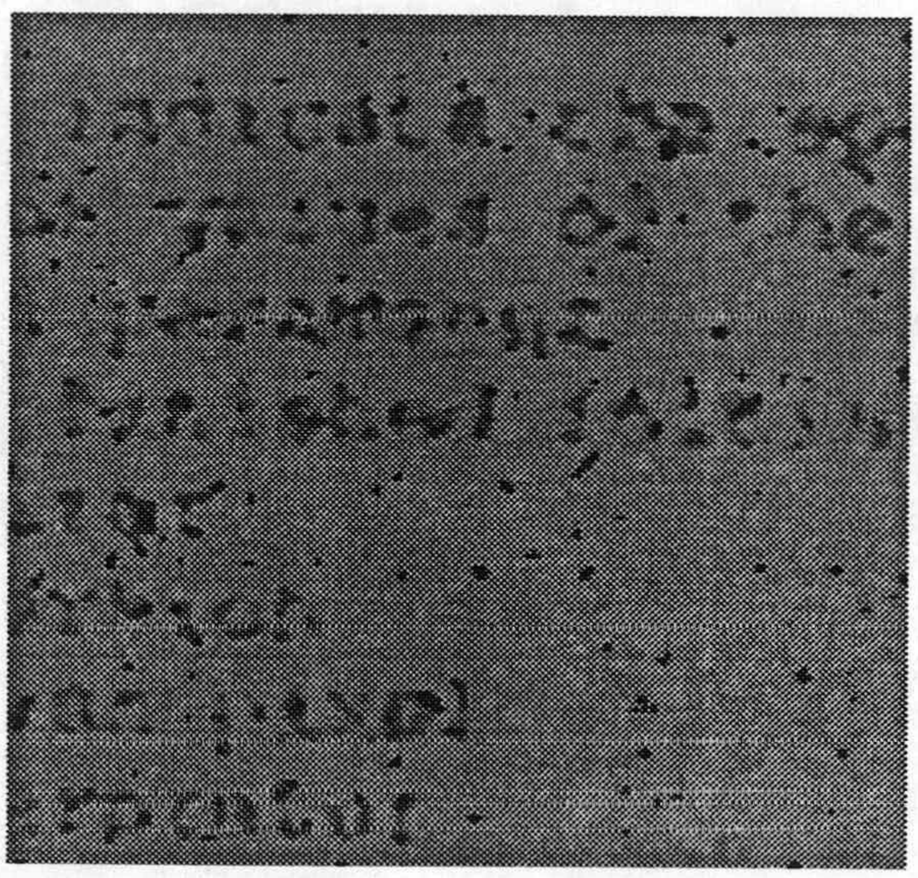
Restdiver, 1ª iteración



Restdiver, mejor iteración



Mediana (3×3), 1ª iteración



Mediana (3×3), mejor iteración

Figura 44: Imágenes resultado del procesamiento de la imagen fuente de texto que aparece en la figura 40. Obsérvese cómo, a pesar de la enorme cantidad de ruido, el filtro de ventana adaptativa ha sido capaz de conseguir una buena imagen de bordes, a partir de la cual elaborar el resultado que se muestra en la parte superior de la figura. Debe advertirse que los dos filtros puestos en comparación tienen el mismo comportamiento ante los bordes de la imagen, por lo que la mejora observada en el procesamiento con *restdiver* es achacable sobre todo a la capacidad limpiadora proporcionada por su capacidad de adaptar la ventana.

CONCLUSIONES

En este capítulo se han introducido dos nuevos algoritmos de filtrado basados en la mediana, el filtro de mediana relajada (MR) y el filtro de mediana con ventana adaptativa (*restdiver*). Ambos mejoran apreciablemente al filtro de mediana en imágenes reales. Sólo se ha considerado el ruido impulsivo, no sólo por ser el principal objetivo del presente estudio, sino porque el filtro de mediana, en el que se basan los filtros presentados, tiene un pobre comportamiento frente a dicho tipo de ruido.

El primero de ellos se basa en la idea de que los estadísticos de orden cercanos a la mediana pueden ser considerados suficientemente probables como para no pertenecer al ruido, y por tanto no tienen por qué ser procesados. De esta manera tan simple se ha desarrollado un algoritmo cuyas propiedades estadísticas de eliminación de ruido y preservación de bordes y líneas de la imagen han sido calculadas y comparadas con las del filtro de mediana tradicional. El filtro RM mejora la capacidad de preservación de detalles sin apenas disminuir las propiedades de supresión de ruido del filtro de mediana, siempre que el intervalo de bondad no sea demasiado amplio. Para comprobarlo se han realizado una serie de experimentos con imágenes sintéticas y reales que han demostrado el satisfactorio comportamiento del filtro de mediana relajada frente al filtro de mediana tradicional. La mejora en el respeto a los detalles de la imagen se produce tanto para imágenes limpias de ruido como a las contaminadas de él.

El filtro *restdiver*, por su parte, es un mecanismo adaptativo, que emplea un nuevo y prometedor método de detección de bordes basado en la divergencia de Jensen-Shannon. El fundamento del algoritmo es la ampliación de la ventana deslizante del filtro de mediana mientras no exista borde en la imagen, para aprovechar las excelentes cualidades de supresión de ruido impulsivo de dicho filtro. Se ha estudiado el comportamiento de *restdiver* de manera heurística, debido a la enorme complejidad de un estudio estadístico en este caso, y se ha comprobado la mejora de propiedades de eliminación de ruido y preservación de detalles. Para ello se han diseñado experimentos con imágenes sintéticas y reales alteradas con diferentes cantidades y tipos de ruido impulsivo. Para una sola pasada del

filtro, se muestra muy adecuado. La influencia de la calidad de la imagen de bordes que se le suministra al filtro es notoria si la imagen fuente se ve muy afectada por el ruido.

CAPÍTULO III

PRESELECCIÓN DE PÍXELES. MÁSCARA DE ESTIMACIÓN DE RUIDO

III.1 INTRODUCCIÓN.....217

III.2 PRESELECCIÓN BASADA EN MATRICES DE COMPOSICIÓN..222

III.3 CORRECCIÓN DE LA MÁSCARA.....235

III.3.1 El proceso iterativo.....235

III.3.2 El algoritmo del proceso iterativo..... 241

III.3.3 Medidas de calidad de la máscara.....243

III.3.4 Ejemplos de obtención de la máscara de estimación de ruido.....244

III.3.5 El problema del acceso a la LUT. La máscara gris..... 249

III.3.6 Influencia del umbral de selección sobre la máscara.....252

Influencia del umbral de selección en la corrección del nivel de gris de ventana de acceso a la LUT.....255

Influencia del umbral de selección en la localización del ruido.....272

III.3.7 Influencia de la calidad de la matriz de composición en la máscara de estimación de ruido.....279

III.3.8 CONCLUSIONES..... 280

INTRODUCCIÓN

Ya se ha indicado con insistencia que el objetivo de todo filtrado es la eliminación de la posibles distorsiones que aparezcan en la imagen observada, partiendo de determinados parámetros conocidos acerca del proceso de transformación de la imagen original a la fuente. En nuestro caso concreto, habremos de corregir la contaminación de ruido *impulsivo*, intentando que el resultado final sea una imagen tan parecida a la original como sea posible. Recordemos que ello entrañaba una doble dificultad, que surge de la ignorancia acerca de la imagen original: por un lado los datos acerca del proceso de alteración de la imagen son escasos, y por otro no podemos decidir si el resultado es aceptable. Por tanto las características del ruido y de la señal precisas para poder realizar un filtrado adecuado, habrán de extraerse directamente de la imagen fuente.

Por otro lado, existe una consideración de índole práctica. El ruido impulsivo no se manifiesta sobre la totalidad de los elementos de la imagen, sino en puntos espontáneamente localizados, dónde una variable aleatoria sustituye a la señal. Como el ruido afecta solo a muestras individuales de la imagen, al contrario de lo que sucede con, por ejemplo, el ruido aditivo, podemos distinguir dos estados en el algoritmo de filtrado:

- Preselección de píxeles erróneos (detección de pulsos)
- Estimación del nivel de gris original (corrección de la señal distorsionada).

Por tanto, a la hora de realizar un filtrado en una imagen fuente, afectada por ruido, parece lógico acudir directamente sobre los píxeles afectados por el ruido, dejando inalterados aquellos elementos de la imagen que no presentan contaminación. Esta idea surge de manera natural al observar la indeseable propiedad que

poseen los filtros digitales de suavizar cada píxel que es procesado, ya sea o no ruido. Obsérvese que el algoritmo de filtrado no suele ser inteligente, por lo que a medida que recorre la imagen, va procesando, sin importarle la naturaleza intrínseca del píxel. Esta característica de la mayoría de los filtros, como los de mediana, los hace especialmente inadecuados para la eliminación del ruido impulsivo, puesto que a cambio de corregir el nivel de gris de los píxeles afectados por el ruido, distorsionan aquellos que no lo están. Esta alteración se hace especialmente notoria en los bordes entre las diferentes regiones que componen la imagen, los cuales, como ya indicamos, contienen la mayor parte de la información que recoge el ojo humano.

Aparte de la estimación del nivel de gris original, cuyo estudio será abordado en el siguiente capítulo, lo que ha de intentar el diseñador de filtros que eliminen ruido impulsivo es *localizar* aquellos píxeles que hayan sido afectados por la distorsión, segregándolos de aquellos que no lo están. Procediendo de este modo, se puede discriminar perfectamente entre lo que es señal original y lo que es ruido, permitiendo estudiar ambos por separado. Entonces se pueden conocer las características del ruido y de la señal original a partir de la imagen fuente, sin disponer de la imagen original. Además, tenemos la ventaja añadida de poder realizar el procesamiento de filtrado solamente sobre aquellos píxeles que realmente lo necesiten, evitando el problema de la distorsión de los píxeles limpios inherente al empleo de filtros digitales. A este proceso de localización de ruido impulsivo dentro de la imagen fuente lo denominaremos ***preselección*** de los píxeles erróneos.

Este mecanismo se fundamenta en la verificación, mediante métodos estadísticos, de la pertenencia del píxel central de la vecindad —*i.e.* la ventana— a la muestra definida por los píxeles no ruidosos dentro de esa misma ventana. En caso negativo, se marca como píxel distorsionado, que deberá ser procesado. Al ser no lineal el mecanismo de introducción de ruido en la imagen, así lo será el método de detección del mismo. Es por esta razón por lo que los filtros no lineales son mejores para procesar ruido impulsivo que los lineales, incluso cuando éstos son adaptativos. El resultado de la preselección es, pues, una estimación de los píxeles que han sido sustituídos por píxeles en la imagen. Es decir:

$$\underline{M}_{ij} = \begin{cases} 1 \text{ (blanco)} & \text{si } x_{ij} \in \text{ vecindad no ruidosa de } x_{ij} \\ 0 \text{ (negro)} & \text{si } x_{ij} \notin \text{ vecindad no ruidosa de } x_{ij} \end{cases}$$

Ya quedaron profusamente recogidas en un anterior capítulo las metodologías existentes para la preselección en aquellos filtros que emplean cierta información acerca del tipo de señal que procesan, por lo que aquí se recordará lo más significativo, para centrarnos en el problema que nos ocupa en este apartado. Con ese objetivo, citemos de nuevo dos métodos básicos de detección que pueden implementarse con facilidad, basados en técnicas de estimación robusta, y que emplean una misma hipótesis acerca de la suavización de los detalles en una señal no distorsionada:

- *Predicción:* Se determina la diferencia entre el nivel de un determinado píxel y el que se estima a partir de sus vecinos ya filtrados, dentro de una secuencia de barrido de la imagen. Si dicha diferencia excede un determinado valor crítico, se considera como ruido. Obsérvese que tanto el método de filtrado como el valor crítico han de establecerse *a priori*. Dicha elección se basará, consciente o inconscientemente, en la información que se disponga del proceso de aparición de ruido y de la imagen original.

- *Democrático:* Durante el proceso de recubrimiento de la imagen, cada elemento se considera junto a sus $2N$ vecinos (n a la derecha y n a la izquierda) formando una muestra de $2N+1$ valores que se ordenan en magnitud. El valor del elemento en cuestión se compara con los k valores extremos de la secuencia ordenada. Si está en ese rango, se lo considera erróneo (por demasiado extremo). Como se observa, es parecido a la generalización del filtro demediana que propusimos con el filtro de mediana relajada, y se basa en la hipótesis de que el comportamiento *correcto* es localemente monótono, y que las desviaciones de esa monotonía son, si existieren, pequeñas. De este modo, la mayoría de los píxeles tienen aproximadamente el mismo valor, por lo que el elemento que posea un nivel de gris que destaque entre el de sus vecinos, será considerado anómalo, lo cual recuerda a la sociedad humana. Los valores de n y de k han de establecerse de antemano, de nuevo teniendo en cuenta consideraciones acerca de la naturaleza de la imagen original y del ruido que pueda inducir el usuario.

Estas dos metodologías se aplican en diferentes algoritmos de estimación de parámetros para el ruido impulsivo. Queda así patente que en un esquema adaptativo debe suceder que la decisión acerca de si un píxel es o no ruido se haga de acuerdo a valores locales de

la señal. En concreto, para detectar ruido impulsivo habrá que comparar el nivel de gris de cada de cada píxel con el de sus vecinos, y calcular una cierta función de error. Si ésta alcanza valores que superan un cierto umbral determinado—denominado ***umbral de selección***, el píxel se considera ruidoso.

La característica estadística básica del ruido impulsivo es la probabilidad de distorsión de las muestras de la señal, por lo que resulta lógico que el umbral de selección dependa de ella. De hecho, cuanto más probable sea la aparición de ruido, más exigente habrá de ser ese umbral. Sin embargo, la necesidad de utilizar este parámetro introduce un elemento de arbitrariedad del que debería estar exento el algoritmo adaptativo. Es por ello que algunos de estos filtros fijan automáticamente el umbral basándose en el histograma de distribución de los valores absolutos de las diferencias entre cada muestra de la imagen fuente y su valor predicho por el algoritmo a partir de la información local proporcionada por su vecindad:

$$\mathcal{P}[|S_{ij} - R_{ij}|]$$

Este histograma presenta dos partes características:

- La primera, definida por la distribución de la señal diferencia entre la muestra no afectada de ruido, y la estimada. Como los elementos vecinos suelen estar fuertemente correlacionados, como ya se dijo, y presentan por tanto niveles de gris similares, el valor de esta parte del histograma es pequeño.

- La segunda, determinada tanto por la distribución de error de la predicción de ruido, como por la diferencia entre el ruido y la estimación del nivel de gris original (siempre a partir de elementos no contaminados de la vecindad). Su valor decrece lentamente, pues los pulsos en el ruido impulsivo no están correlacionados.

Entonces el valor del umbral de selección se determina como el punto dónde se produce el cambio de tendencia entre ambas zonas del histograma, pudiendo así ser calculado automáticamente empleando el método de predicción. También pueden detectarse pulsos mediante métodos democráticos empleando vecindades adecuadas. En cualquier caso, debe notarse que el método de detección de los píxeles ruidosos está basado en un conocimiento previo implícito: el de que los pulsos tienen valores extremos, comparados

con los no ruidosos.

En otro orden de cosas, el tamaño de la vecindad de detección se definirá en función de la probabilidad de aparición de un pulso. En efecto, si ésta es elevada, debe emplearse una vecindad determinada por la ventana, grande, para que contenga suficiente información acerca de la señal no ruidosa. En el caso de que el ruido sea poco probable, se utilizarán ventanas pequeñas, pues serán suficientes para detectar el ruido cuando éste aparezca, ya que es poco probable que las distorsiones afecten a la mayor parte de la vecindad. Aprecie el lector en la anterior discusión un cierto paralelismo, deseable por otra parte, entre el modo de operación del ordenador y el mecanismo perceptual humano: cuanto mayor es el campo visual, tanto mejor se detectan las alteraciones provocadas por el ruido. De otro lado, si no hubiese conocimiento apriorístico alguno acerca de la probabilidad de aparición de pulsos, podría usarse el hecho reseñado de que los valores de la señal original en una vecindad pequeña tienen una dispersión pequeña de valores.

Pese a todo, los mecanismos propuestos en la literatura adolecen en general de cierto desequilibrio entre efectividad y complejidad; además de introducir parámetros arbitrarios escasamente justificables. A continuación se expone un método alternativo, basado en conceptos adaptativos y sustentado tanto por los resultados como por justificaciones teóricas.

PRESELECCIÓN BASADA EN MATRICES DE COMPOSICIÓN

Dentro del marco de los filtros adaptativos, hemos propuesto un método de filtrado de ruido impulsivo que se basa en la preselección de los píxeles que han de ser procesados. Para lograrlo, emplearemos la información contenida en las matrices de composición que ya han sido introducidas en un anterior capítulo. El objetivo último de la preselección es la construcción de una ***máscara de preselección*** $\underline{M}$ en la que aparezcan marcados los píxeles que habrán de ser procesados. Obsérvese que esta máscara verifica $\underline{M} \approx M$, siendo M la imagen definida en el capítulo II al introducir la notación.

En nuestra aproximación, cada imagen pertenece a una determinada clase C. De este modo, cada imagen fuente, pese a estar contaminada de ruido impulsivo, sigue teniendo características que nos recuerdan a qué clase de imágenes pertenece la imagen original de la que procede. Por ejemplo, podemos reconocer si la imagen fuente que vamos a procesar en un determinado experimento es una radiografía, una ecografía, un texto escrito ... Pues bien, cada clase de imágenes está representada por una determinada matriz de composición, también llamada modelo de composición. El modelo de composición que representa a la clase es una matriz bidimensional que, por definición, describe los histogramas proporcionados por una ventana deslizante que recorre todas las imágenes que componen esa clase. Definiendo el nivel de gris de ventana k' como la suma de los niveles de gris de todos los píxeles dentro de la ventana, cada columna $Q_{k'}$ de la matriz de composición es el histograma promedio de todas las ventana que tienen el mismo nivel k'. Por tanto, hay tantas columnas como niveles de gris de ventana, esto es, desde = a $G\,w$, siendo G el máximo nivel de gris de los píxeles (esto es, el máximo de la escala de grises), y w es el tamaño de la ventana, en píxeles. Podemos entonces escribir que:

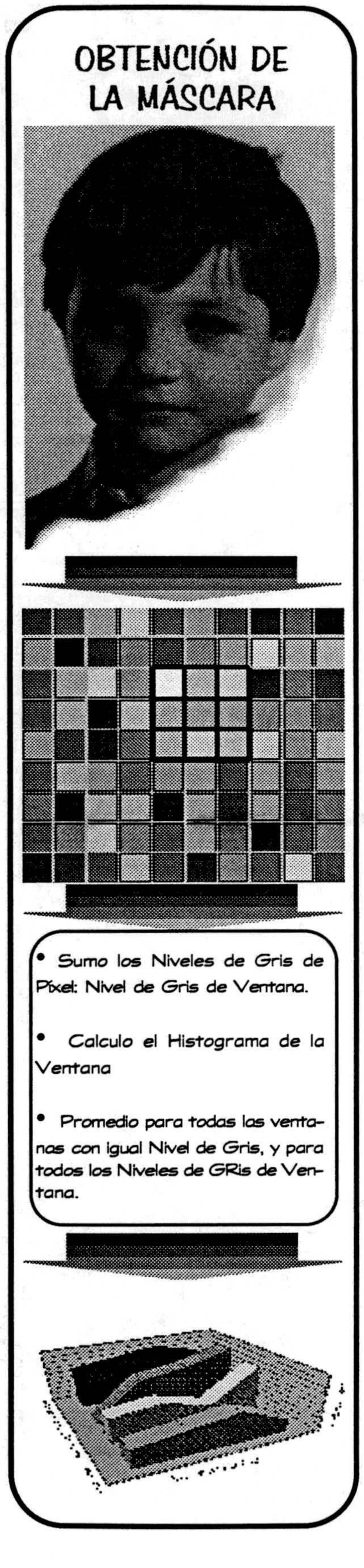
$$Q(C) = (Q_0\, Q_1\, Q_2 \cdots Q_{Gw})$$

Entonces, el *k-ésimo* elemento en $Q_{k'}$ es la frecuencia relativa del nivel de gris k para los píxeles dentro de ventanas deslizantes de nivel k'. En el caso de que no hubiese ventanas con nivel de gris k' entonces $Q_{k'}$ no existe, apareciendo una columna vacía en la matriz de composición $Q(C)$.

Por tanto, construyendo *a priori* el modelo de una determinada clase de imágenes (radiografías por ejemplo), dispondremos de un amplio caudal de información acerca de la manera en que se distribuyen en promedio los niveles de gris en dichas imágenes, dependiendo de la resolución a la que lo observemos. De esta manera nos será posible discernir si un determinado píxel tiene un nivel de gris discordante con el modelo, que será inmediatamente considerado como ruido.

Es evidente que no podemos trabajar con *todas* las imágenes que pertenecen a una misma clase para calcular así el modelo. Sin embargo, recordemos que una matriz de composición podía obtenerse de dos maneras. La primera, teórica, consistía en formular hipótesis acerca de la estructura interna de las regiones. Por ejemplo podríamos suponer que los niveles de gris dentro de las regiones de la imagen de un mismo tamaño y nivel k' se distribuyen siguiendo una hipergeométrica. La segunda manera de obtener un modelo es empleando mecanismos experimentales, mediante el cálculo de las columnas $Q_{k'}$ promediando los histogramas de todas las regiones en una muestra seleccionada a partir de un conjunto apropiado, representativo de una determinada clase de imágenes C. Repitiendo lo mismo para cada k' se obtiene finalmente la matriz $Q(C)$. Este modelo de composición podrá ser empleado para procesar cualquier imagen que pertenezca a la clase que representa. Este procedimiento se ilustra en el lateral de la presente página.

Una vez que disponemos del modelo de composición, representativo de la clase de imágenes a la que se supone que pertenece la imagen original de la que proviene la fuente, ¿cómo hay que proceder para construir la máscara de preselección, que me indique qué píxeles son ruido y cuáles no? El modelo va a proporcionar una gran cantidad de información relativa a la probabilidad de que aparezca un determinado nivel de gris k dentro de una región con nivel de gris k'. Por tanto, podemos saber si un determinado píxel tiene un nivel de gris probable con tan sólo mirar qué nivel k' tiene su vecindad, y consultando la columna Q_k del modelo. Si ello es así, el píxel no será ruido, pero si la probabilidad de aparición en una tal región es nula o despreciable, podemos tener casi la certeza



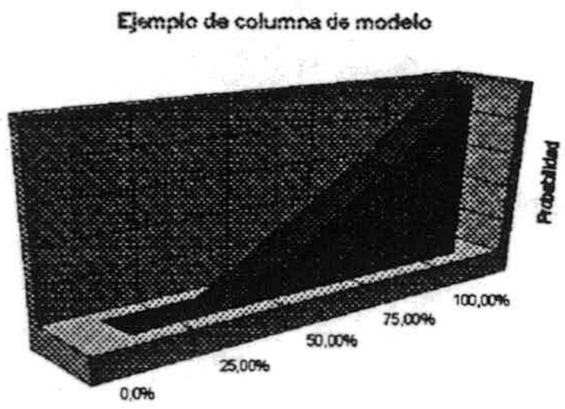


Figura 1: Columna ordenada de una matriz de composición cualquiera.

de estar ante un buen candidato a ser ruido. Definiremos como *selección* al proceso consistente en decidir si un píxel tiene un nivel de gris probable, de acuerdo con la matriz de composición.

Vemos que, de este modo, el resultado de utilizar el modelo de composición no es otra cosa que una estimación sobre si el píxel en cuestión es o no un ruido. Esta predicción dependerá en buena lógica de la calidad que tenga el modelo, la cual a su vez está relacionada con el tamaño de la muestra y con la homogeneidad del contenido de las ventanas en la muestra. Tal homogeneidad podía ser especificada, como vimos, en términos de una medida apropiada de cohesión entre los histogramas (que recordemos que no son otra cosa que distribuciones de probabilidad) de las ventanas observadas en la muestra de entrenamiento. Para este propósito hemos empleado la divergencia de Jensen-Shannon que ya definimos. Recordemos que cuanto menor era el valor de JS, mayor era la cohesión de la muestra, mejor definida estaba la clase, y por tanto mejor se comportaba el modelo correspondiente.

Pero aún no hemos especificado el método que hay que seguir para construir la máscara de preselección. El proceso de selección puede describirse como sigue:

1.- Se recorre la imagen fuente con una ventana deslizante de igual tamaño que aquélla que se empleó en la construcción del modelo que se va a utilizar. Por tanto, al comienzo del proceso de selección, se dispone de la imagen fuente y del modelo que representa la clase de imágenes a la que pertenece la imagen original de la que procede la imagen fuente.

2.- Para cada ventana, se computa el nivel de gris k' y se observa el nivel de gris k del píxel central.

3.- Se acude a la columna $Q_{k'}$ del modelo, que no es otra cosa que una distribución de probabilidad. La componente *k-ésima* del vector $Q_{k'}$ será la probabilidad promedio que tiene el nivel de gris k dentro de una región de nivel k' para imágenes de la clase C representada por el modelo.

4.- Se decide si dicha probabilidad de aparición del nivel k es suficientemente significativa para considerar dicho nivel como señal original, o si por el contrario indica que k es un buen candidato a ruido. Para ello se emplea un valor crítico

representado por el *umbral de selección.*

Obsérvese que este último paso contiene una dificultad intrínseca. No existe una manera unívoca de dirimir si la probabilidad es o no significativa. Podemos considerar al menos las siguientes elecciones:

ρ_1) Se escoge como parámetro de selección una determinada posición κ^* sobre la distribución $Q_{k'}$ *ordenada*, denotada por $(Q_{k'})_{[\cdot]}$ y mostrada genéricamente en la figura 1, a partir del cual los píxeles se considerarán no ruidosos. Para determinar esa posición κ^* se emplea un parámetro ρ_1 que representa un porcentaje de la distribución de niveles de gris de píxel. Por ejemplo, podemos elegir el 80% de la distribución como parámetro, de modo que se considerarían ruido aquellos píxeles cuyo nivel de gris quedase a la derecha de dicho percentil, por ser menos probables que el otro 20%. Es decir, este modo de operar implica que se considerará como ruido el 80% de los niveles de gris que aparezcan en la columna. Obsérvese que ello no quiere decir que se procese el 80% de los píxeles. Sin embargo, tenemos algunos inconvenientes:

- Es necesario ordenar la distribución, con el subsiguiente aumento de la complejidad.

- No conocemos de antemano la forma de la distribución $Q_{k'}$, por lo que la elección de ρ_1 es completamente arbitraria, al hacerse a ciegas. Entonces puede suceder la distribución sea muy uniforme, de modo que ese valor de 80% que supusimos en el ejemplo sea demasiado alto, y ningún nivel de gris lo supere. Estaríamos considerando como ruido la totalidad de los píxeles, lo cual es claramente erróneo.

- Por otra parte, cada columna del modelo representa una distribución de probabilidad diferente, por lo que un valor de 80% puede ser aceptable para algunas columnas, y demasiado alto para otras.

- Existen columnas en las que no aparecen muchos niveles de gris. Ello es lógico, pues en una región de nivel k' no pueden aparecer todos los k de la

escala, a no ser que $k' = G\,w$. Por tanto, al elegir un determinado percentil como parámetro de selección podría ocurrir que todos los píxeles considerados ruido tuviesen probabilidad nula, pero ¡algunos píxeles con igual probabilidad se darían por buenos! Si consideramos el ejemplo de la figura 1, y elegimos ρ_1=20%, observamos que el intervalo desde ese percentil hasta el 25% ha sido considerado no ruidoso, pero tiene la misma probabilidad que el intervalo en el que se han considerado los niveles de gris como constitutivos del ruido. Esta misma indeterminación sucede si tenemos *mesetas* de igual probabilidad en la distribución ordenada.

Para eludir esta dificultad, hicimos que el algoritmo situase el parámetro de selección en el último percentil que tuviese la misma probabilidad que la señalada inicialmente. En nuestro ejemplo, aunque hemos elegido $\rho_1 = 20\%$, el algoritmo situaría automáticamente $\rho_1 = 25\%$. De esta elegante manera hemos conseguido eliminar el problema, pero a costa de introducir un fuerte elemento de arbitrariedad, poco justificable por conductos teóricos. Siguiendo con nuestro ejemplo, daría exactamente lo mismo elegir $\rho_1 = 0\%$ que $\rho_1 = 25\%$. Y lo que es más grave, si tomásemos $\rho_1 = 30\%$ en realidad sólo consideraríamos un 5% de los niveles de gris.

Por tanto, parece que esta estrategia para elegir el parámetro de selección es bastante inadecuada.

ρ_2) De nuevo el umbral es un determinado lugar en la distribución ordenada, κ^*; pero en esta ocasión se determina mediante la elección de un porcentaje de área ρ_2 encerrada bajo la curva de $(Q_{k'})_{[\cdot]}$. Dado ρ_2, entonces se verifica una y sólo una de las siguientes condiciones:

1. $0 \le \rho_1 < (Q_{k'})_{[0]}$. En este caso trivial, no se filtra ningún píxel, por ser todos los niveles probables. Para evitar este problema, se introduce por convenio el lugar $\kappa = -1$ con $(Q_{k'})_{[-1]} = 0$.

2. Existe un único κ^* tal que:

$$\sum_{\kappa=-1}^{\kappa^*} (Q_{k'})_{[\kappa]} \leq \rho_2 < \sum_{\kappa=-1}^{\kappa^*+1} (Q_{k'})_{[\kappa]}$$

donde se ha empleado el convenio anterior. En este caso, sólo los píxeles con nivel de gris [0], [1], ...,[κ^*] son procesados.

3. $\Sigma(Q_{k'}) \leq \rho_2$. Esto es, todos los píxeles son filtrados (no hay preselección). Para ello, hay que permitir que $\kappa^* = G$.

Entonces, aquellos niveles de gris que tengan una probabilidad mayor que la que aparece cubierta por el área serían considerados no ruidosos. De esta manera, el 0% de área sería equivalente a decidir que los niveles de gris ruidosos son sólo aquellos cuya probabilidad es exactamente cero. Obsérvese que de esta manera hemos eludido la ambigüedad que aparecía en el criterio anterior (y mostrada en último lugar): la elección de un porcentaje de área define unívocamente la fracción de niveles de gris que van a ser filtrados. Surge en seguida la cuestión de interpretar físicamente el significado de ese porcentaje de área. Percátese el lector de que el área encerrada por la curva de distribución ordenada es exactamente la suma de todas las probabilidades de esa distribución. Por tanto, el área comprendida desde el percentil 0% hasta el percentil determinado por el usuario ρ_2%, no es más que la suma de las probabilidades, desde la nula hasta la p, correspondiente con el ρ_2% del área bajo la curva de distribución. Esta interpretación del nuevo parámetro de selección permite definirlo en términos de la distribución ordenada y *acumulada* como la probabilidad acumulada umbral que delimita el comportamiento ruidoso del no ruidoso.

A pesar del esfuerzo, seguimos arrastrando casi las mismas desventajas de que adolecía el anterior criterio de elección de parámetro de selección. uno parece insoluble: la indeterminación que surge al llevar a cabo la ordenación de la distribución cuando se nos presentan n niveles de gris con igual probabilidad. En ese caso no tenemos información que nospermita colocar a uno antes que a otro. Se presenta una especie de meseta en la distribución dónde el

ordenamiento se ha llevado a cabo de forma arbitraria. Esta dificultad no es grave a menos que nuestro percentil caiga exactamente en esas zonas. En tales casos, estaríamos dando por buenos niveles de gris con exactamente la misma probabilidad que otros considerados como ruido. La única manera de paliar en alguna medida el problema sería conociendo de antemano la posición de las mesetas, y no permitiendo que el percentil de área quedase justo en medio de tales zonas, obligando al mismo a irse al principio o al final de tales regiones. Sin embargo ello es poco práctico por dos razones: es costoso computacionalmente e introduce un nuevo elemento de arbitrariedad.

Recordemos que uno de los inconvenientes que presentaba el anterior criterio de selección era que se escogía un mismo parámetro para todas las columnas, independientemente del perfil que presentase la distribución, lo cual añadía, ante el desconocimiento previo de dichas distribuciones, una arbitrariedad poco justificable. En el presente caso la situación es parecida. Ahora el porcentaje de área será el mismo para todas las columnas.

Además, obsérvese que al elegir porcentajes de área necesitamos acumular probabilidades, lo cual añade mayor complejidad computacional al algoritmo.

ρ_3) Otra vez se escoge como umbral de selección una posición κ^* en la distribución ordenada. Pero en este caso se determina esa posición en virtud de una cierta probabilidad límite, dada por la matriz de composición. Por ejemplo, podemos decir que todos los píxeles cuyo nivel de gris supere una probabilidad de aparición, según el modelo, de 0.7 serán considerados como buenos, y los que queden por debajo se tomarán como ruido. Podemos enumerar una serie de ventajas e inconvenientes que presenta este procedimiento:

- Siguiendo este camino, obtenemos directamente la estimación del ruido, sin necesidad de ulteriores modificaciones del histograma (como ordenamiento, etc).

• No se conoce de antemano la forma de la distribución $Q_{k'}$. Se tienen entonces los mismos inconvenientes que en las anteriores posibilidades.

Sin embargo, parece que podemos encontrar una solución a esos inconvenientes. En principio, un buen parámetro de umbral sería el valor $1/G$, siendo G el máximo valor de la escala de grises de los píxeles. ¿Por qué ese valor? Porque esa es la probabilidad que tendrían los niveles de gris, *si no hubiese sesgo*, supuesta una distribución uniforme. De este modo, en primer lugar, aseguramos que siempre exiten niveles de gris con probabilidad por encima y por debajo de ese valor, independientemente de la forma de la distribución. Por otro lado, si un determinado píxel tiene un nivel de gris cuya probabilidad, según el modelo, es mayor que $1/G$ podemos asegurar que es más probable que el valor dado por la distribución uniforme, por lo que es razonable suponer que no es ruido.

Desgraciadamente existe un fuerte sesgo, una restricción muy fuerte introducido por la naturaleza discreta de la ventana deslizante. Si ésta tiene un tamaño $w = n \times n$ entonces se verificará que:

$$\mathcal{N}(k')(Q_{k'})_k = \frac{\alpha}{w}$$

donde $(Q_{k'})_k$ es la componente k de la columna $Q_{k'}$ del modelo; $\mathcal{N}(k')$ es el número total de ventanas con nivel de gris k' empleadas al construir la columna $Q_{k'}$ del modelo, cuyo valor está dentro del intervalo $[0,W/w]$, siendo W el tamaño de la imagen en píxeles; y $á$ es un número natural dentro del intervalo $[1,w\,\mathcal{M}]$. La relación anterior surge de considerar el mecanismo de construcción de $Q(C)$, promediando los histogramas de todas las ventanas con igual nivel de gris de ventana, y el tamaño de la ventana. Es, pues evidente el sesgo introducido por el empleo de una ventana deslizante finita, que hace que la distribución se aleje de la uniforme. Por esta razón, la hipótesis en la que se basa la elección del parámetro de selección anterior no se verifica, y dicho valor pierde su significado. Véase que en los experimentos que hemos realizado se ha empleado $G = 64$; $w = 9$ ó 25, con lo que entonces el parámetro de selección $1/G = 1/64 \lll 1/w = 1/9$ ó $1/25$. Es decir, apenas habría nive-

les de gris menos probables que $1/G$, por lo que no encontraríamos casi ruido empleando dicho valor como parámetro de selección.

Las diferencias entre las tres alternativas de definir el umbral de selección se ilustran en la figura 2. Por lo dicho hasta el momento parece no existir una elección que eluda los dos principales problemas que se han planteado: la arbitrariedad del parámetro al desconocer los perfiles de las columnas y el elemento de azar que supone escoger un mismo umbral para todas las columnas. Sería muy interesante poder definir un criterio de selección que tuviese en cuenta cierta información acerca de la distribución, de manera que tales arbitrariedades quedaran, cuanto menos, paliadas. Tal información puede provenir de la *entropía* de la columna, esto es, de la distribución de probabilidad, $H(Q_{k'})$. Esta función nos da una idea de cuán homogénea es la distribución de probabilidad. Si la entropía vale cero la distribución es degenerada, y posee por tanto un perfil muy asimétrico. En caso de acudir a una columna con entropía nula, la mayoría de los niveles de gris serían ruidosos, pues sólo algunos niveles tienen probabilidad no nula. Por otra parte, si la entropía alcanza su valor máximo, entonces la distribución es uniforme, por lo que la curva de distribución ordenada es una recta. Todos los niveles de gris tienen la misma probabilidad, por lo que o todos son ruido, o todos son no ruido.

Queda claro que la homogeneidad de la columna, es decir, la forma de la distribución, que antes no se tenía en cuenta, está caracterizada de alguna manera por la entropía de la distribución. De hecho, por la definición de entropía, resulta que $2^{H(Qk')}$ es el número de elementos (niveles de gris) significativos de la distribución $Q_{k'}$. Por lo tanto, $2^{-H(Qk')}$ será la probabilidad típica significativa en la distribución de probabilidad $Q_{k'}$. Parece adecuado que el umbral de selección dependa de la información proporcionada por $H(Q_{k'})$. La elección inmediata puede ser elegir el umbral κ^* definido por $\rho_3 = 2^{-H(Qk')}$, por ser el valor típico de la probabilidad en la distribución.

Sin embargo, este valor puede resultar demasiado restrictivo. Por ejemplo, si $w = n \times n$ es el tamaño de la ventana, entonces $2^{-H(Qk')} > 1/w \Leftrightarrow n > 2^{½H(Qk')}$ como puede comprobarse fácilmente. Esta desigualdad es conveniente, a la vista de los comentarios realizados al discutir el parámetro ρ_3, pero se presentará en pocas ocasiones, bajo circunstancias especiales. Si tenemos en cuenta que la

máxima entropía posible es la correspondiente a la distribución uniforme, que vale $log_2 G$, y que el mínimo valor de la entropía es cero, podemos acotar el valor de $2^{½H(Qk')} \in [1, 2^{½log\,G}]$. Como $G = 2^{\nu}$, siendo ν el número de bits de resolución, entonces $2^{½H(Qk')} \in [1, 2^{½\nu}]$. Dado que los valores más utilizados de ν son {2,4,6,8}, siendo los empleados en esta tesis {6,8}, entonces $2^{½H(Qk')} \in [1, 16]$ ó [1,8]. De ello se deduce que la condición $2^{-H(Qk')} > 1/w$ se verifica mejor cuanto mayor sea el tamaño de la ventana. Debemos tener en cuenta que además los valores del intervalo antes calculado no son exactos. Hay que considerar el hecho de que el modelo ha de ser consistente. Ello impone una restricción sobre las posibles distribuciones, por lo que la entropía máxima posible, al trabajar con imágenes tal como se ha mostrado a lo largo de esta memoria, es la llamada entropía de Gibbs [PEPE] que es siempre menor que la entropía de la distribución uniforme. Pero en todo caso, los resultados deducidos mantienen su validez, puesto que la anterior restricción no resta generalidad.

Por tanto, si cierto nivel de gris tiene probabilidad superior a ese valor podemos estar seguros de que aparece con suficiente frecuencia en la muestra como para ser considerado no ruidoso. Sin embargo, no tenemos la certeza de que un nivel de gris menos probable que $2^{-H(Qk')}$ sea realmente ruidoso. Para no perder generalidad, el umbral de selección será determinado en principio por $\rho_3 = A\,2^{-H(Qk')}$, siendo A una constante apropiada.

Este parece ser un parámetro de selección muy adecuado. Tiene en cuenta la información que posee cada columna, por lo que la selección ya no es la misma para todas las $Q_{k'}$. Y deja de tener el carácter arbitrario que tenían los otros candidatos. Sin embargo, aún puede ser mejorada susceptiblemente, considerando la influencia de la cohesión del modelo, esto es, de la calidad intrínseca de la matriz de composición. Si el modelo es muy cohesionado, tiene mucha calidad interna, la selección puede llevarse a cabo de manera muy restrictiva, en el siguiente sentido. Si según el modelo un determinado nivel de gris tiene realmente una probabilidad pequeña, podemos tener la segirdad de que realmente la tiene. Pero si el modelo no fuese de calidad, no podríamos fiarnos tanto del resultado de la selección. Lo sensato en tal caso sería elevar la probabilidad a partir de la cual se considera el nivel de gris no ruidososo, esto es, incrementar el valor de ρ_3.

Por lo tanto podemos multiplicar el parámetro propuesto que determina el umbral de selección por un factor que dependa

de la calidad intrínseca del modelo. Si ésta es elevada, el factor valdrá la unidad, indicando que el parámetro de selección no debe ser modificado. Si la calidad es pobre, el factor ha de ser mayor que la unidad, para incrementar el parámetro de selección, al ser este proceso ahora mucho menos fiable. Como la calidad de la matriz de selección se cuantificaba empleando la divergencia de Jensen-Shannon (denotada por JS), el factor que se sugiere habrá de depender de esta medida.

Un valor para dicho factor que verifique las condiciones requeridas sería por ejemplo 1 + JS. Pero parece mucho más adecuado, tanto por su significado como por su forma matemática la expresión 2^{JS}. Escogiendo esta opción el parámetro que define el umbral de selección adquiere su forma definitiva:

$$\begin{aligned}\rho_3 &= A \cdot 2^{-H(Qk')}\, 2^{JS} \\ &= A \cdot 2^{JS-H(Qk')} \\ &= A \cdot 2^{-\mathrm{MED}(H(Qk'))}\end{aligned}$$

donde MED($H(Q_{k'})$) es la media de las entropías de las columnas del modelo. Esta expresión se deduce directamente de la definición de la divergencia de Jensen-Shannon. Existe una interpretación física del significado del parámetro. Si $2^{H(Qk')}$ es el número de elementos (niveles de gris) significativos de la distribución $Q_{k'}$, el factor $2^{\mathrm{MED}(H(Qk'))}$ puede considerarse como el valor medio de dichos niveles de gris. Es una especie de promedio sobre todas las ventanas que intervinieron en la determinación de la columna $Q_{k'}$. Por tanto, su valor inverso significará la probabilidad típica promedio de la distribución. Obsérvese que entonces el empleo de este valor de ρ_3 relaciona de alguna manera las ventanas que intervinieron en la creación de la matriz de composición con las ventanas que se utilizarán en el filtrado.

Para que el parámetro ρ_3 sea práctico, su valor no debe constituir una incógnita acerca de su significado. Es decir, si elegimos un determinado valor del mismo debemos ser conscientes de lo que dicha elección implica a la hora de realizar el proceso de selección, pues en caso contrario estaríamos obrando al azar. Para lograr este objetivo, se ha diseñado un intervalo de valores del parámetro de selección entre 0 y 200, con el significado que se muestra a continuación:

ρ	0	100	200
Probabilidad	0	$2^{-\mathrm{MED}(H(Qk'))}$	1

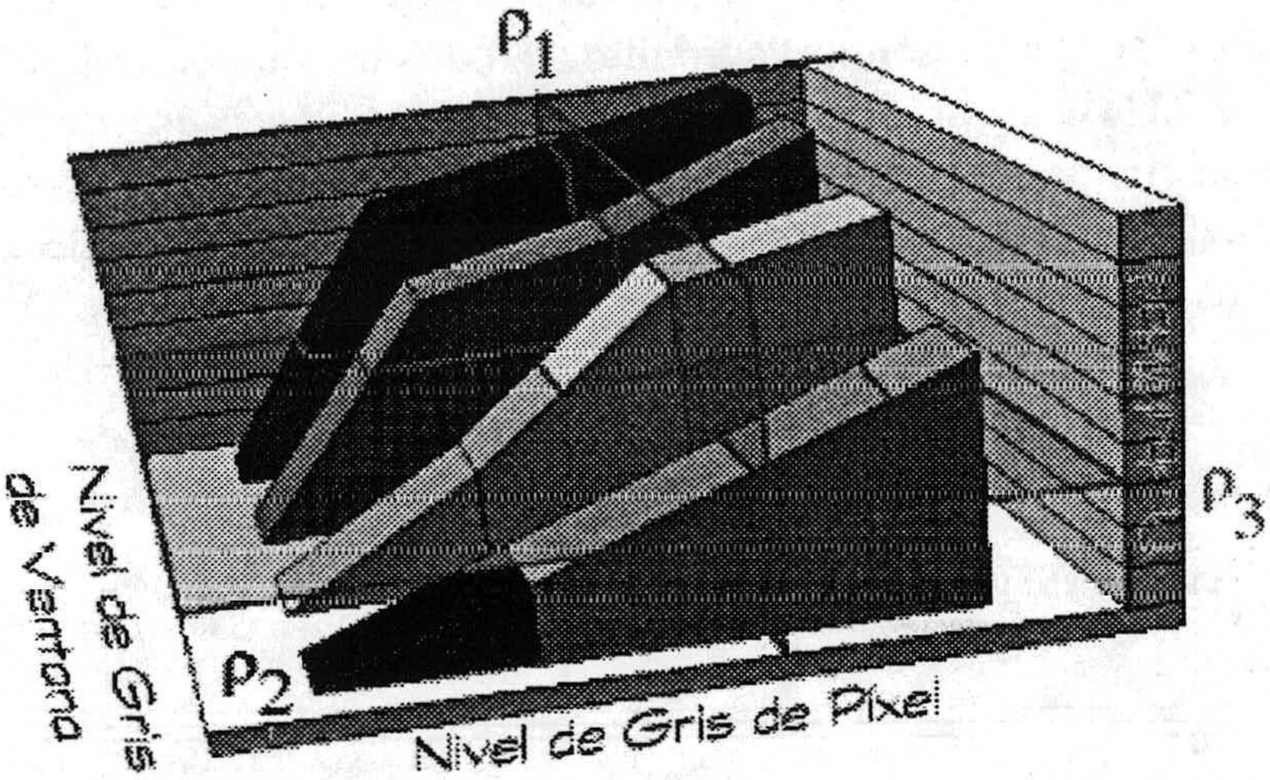


Figura 2: Aparece una representación gráfica de la matriz de composición, con sus columnas ordenadas, sobre las que se realiza la selección de niveles de gris de acuerdo a las tres estrategias expuestas en el texto..

Es decir, se ha modificado el significado del parámetro que define el umbral de selección. Ahora ya no es $\rho_3 = A \cdot 2^{-\mathrm{MED}(H(Qk'))}$, sino que $\rho_3 = A$. A la vista de la tabla, eligiendo un valor nulo para el parámetro, la probabilidad empleada en el proceso de selección es 0, con lo que se darían por malos todos aquellos niveles de gris cuya probabilidad fuere exactamente nula. Mientras, un valor de $\rho_3 = 200$ supone una probabilidad de 1, por lo que todos los píxeles se considerarían ruidosos. El caso $\rho_3 = 200$ implica seleccionar como ruidosos aquellos niveles de gris cuya probabilidad fuese menor que $2^{-\mathrm{MED}(H(Qk'))}$. Para valores intermedios de ρ_3 habrá valores intermedios de probabilidad de selección. Se ha escogido una interpolación lineal entre estos valores fijos, obligados por el hecho de que el uso de polinomios interpoladores de Newton-Lagrange o *splines* proporcionan valores negativos de la probabilidad para ρ_3 pequeños, o valores superiores a la unidad para ρ_3 elevados.

El conjunto de niveles de gris que superan el umbral de selección, y que por tanto no será filtrado, puede determinarse de antemano, antes del propio procesamiento, si el modelo se conoce. Procediendo así se consigue facilitar el manejo de la selección, que de otra manera resultaría penoso. Por tanto, puede elaborarse una *tabla de consulta* (LUT, o *Look Up Table*) cuya entrada es el nivel de gris de la ventana k', y cuya salida es el conjunto de niveles de gris que no han de procesarse. Esta LUT se define como:

$$\mathrm{LUT}[I, \rho, v_w, h_w] : [0, G \cdot v_w \cdot h_w] \rightarrow \mathscr{P}(K)$$
$$k' \rightarrow \mathrm{LUT}(k')$$

donde $\mathscr{P}(K)$ representa el conjunto de todos los subconjuntos de K, y $\mathrm{LUT}(k') = \{k \circledR K \mid (Q_{k'})_k \leq (Q_{k'})_{[\kappa^*]}\}$. Obsérvese que el parámetro ρ_3 coherciona el resultado final mediante la determinación del umbral de selección, el cual segregará los píxeles ruidosos de los limpios.

CORRECCIÓN DE LA MÁSCARA

EL PROCESO ITERATIVO

Cuando el ruido afecta a la imagen, el nivel de gris de algunos píxeles es alterado. Por lo tanto, el nivel de gris original de la ventana también se ve modificado. Al deslizar la ventana sobre la imagen para construir la máscara de preselección, el nivel de gris de la ventana que se le pasa a la tabla de consulta puede estar equivocado debido a la presencia de ruido. Como la tabla de consulta se elaboró a partir de la matriz de composición, que a su vez se construyó empleando imágenes originales limpoas de ruido, el proceso de selección se verá alterado notablemente. La máscara resultante no será de utilidad. Se impone introducir una corrección al nivel de gris de ventana k' que se pase a la tabla de consulta, para eliminar la posible alteración debida al ruido.

Supongamos que el ruido es de tipo impulsivo (sal y pimienta). La introducción de píxeles blancos y negros con igual probabilidad es equivalente a cambiar el nivel de gris de la ventana k' hacia un nivel de gris más central, entendiendo por central la parte del histograma más alejada de los extremos blanco y negro. En efecto, si la ventana fuera muy oscura (k' pequeño) y le introdujésemos ruido blanco y negro en idéntica proporción, el efecto neto es comparable a la adición de ruido gris medio. Por lo tanto hemos cambiado píxeles de nivel de gris k original oscuros por niveles de gris intermedios, esto es, de valor más alto. Evidentemente, el nivel de gris de la ventana k' crece también. Análoga tendencia hacia niveles de gris medios se observaría si la ventana fuese clara.

Queda, pues, patente, el efecto centrípeto que tiene el ruido impulsivo sal y pimineta sobre la imagen original. El nivel de gris de la ventana que se observa sobre la imagen fuente al construir la máscara de preprocesamiento es k_s', que no tiene por qué coincidir con el nivel de gris de la misma ventana en la

imagen original k'. El objetivo de la ***corrección*** será estimar el valor de k' a partir del nivel conocido k_s'.

Denotemos por r el número de píxeles que cambian debido al ruido dentro de la ventana de tamaño w; y sea Δ la desviación del nivel de gris de una ventana debido a la introducción de ruido impulsivo. Dentro de los r píxeles alterados habrá igual proporción de píxeles blancos y negros, si la equiprobabilidad del ruido sal y pimienta se mantuviese estadísticamente para todas las ventanas. Entonces por cada píxel de ruido blanco, el nivel de la ventana aumenta una cantidad ($-k'/w + G$), siendo G el máximo nivel de gris de la escala de píxel (el color blanco). Por cada píxel de ruido negro, el nivel de gris de ventana disminuye en $-k'/w$. Pues bien, si cambiamos r píxeles por ruido sla y pimienta dentro de la ventana de tamaño w entonces tenemos:

- $\frac{1}{2}r$ píxeles blancos, que aumentan el nivel de gris original de ventana k' en $G-k'/w$ cada uno.

- $\frac{1}{2}r$ píxeles negros, que hacen disminuir el valor de k' en $-k'/w$ cada uno.

Luego, en total, la variación de nivel de gris debida a la introducción de ruido impulsivo sal y pimienta será:

$$\Delta = \frac{r}{2}\left(G - \frac{k'}{w} - \frac{k'}{w}\right) = r\left(\frac{G}{2} - \frac{k'}{w}\right)$$

Pero fijémonos que $n = w \cdot p$, siendo p la probabilidad de aparición de ruido, y que puede ser estimado con facilidad. Además, el máximo de la escala de nivel de gris de ventana es $G' = w \cdot G$. por lo tanto:

$$\Delta = p\left(\frac{G'}{2} - k'\right)$$

Es fácil comprobar que esta expresión se comporta bien en los límites:

- Si $p = 1$ (100% de ruido) la imagen es toda de ruido, y se obtiene $\Delta = \frac{1}{2}G' - k'$, lo cual resulta lógico.

- Si $p = 0$ (no hay ruido), $\Delta = 0$.

Por otra parte, por definición $\Delta = k_s' - k'$. Igualando esta expre-

sión con la anterior, podemos despejar el valor de k_s' para obtener:

$$k_s' = k'(1-p) + p(½G')$$

que es una combinación lineal convexa, lo cual me garantiza que k_s' no va a alcanzar valores fuera de los límites de la escala de grises. El comportamiento en el límite vuelve a ser satisfactorio. Para $p = 1$, $k_s' = ½G'$; mientras para $p = 0$, el valor de $k_s' = k'$. Por lo tanto hemos llegado a una expresión que relaciona los niveles de gris k' y k_s' como era nuestro deseo. Por lo tanto:

$$k' = \frac{k_s' - p\frac{G'}{2}}{1-p}$$

Como los valores de k_s' y G' son conocidos, y el de p puede ser estimado (contando los píxeles ruidosos y dividiendo por el número total de píxeles), podemos estimar el valor de k' que **realmente** tiene la ventana mediante la anterior ecuación. Hemos de ser cautos, sin embargo, puesto que la anterior igualdad permite alcanzar valores de k' fuera de los límites de la escala, lo cual no es posible. No es que sea un error de la ecuación, sino que la misma expresa un proceso opuesto al comportamiento centrípeto de la adición de ruido impulsivo de sal y pimienta. Es decir, se centrifuga el nivel de gris k_s' previamente centrado por el ruido, para estimar k'. Pero en regiones próximas a los límites de la escala de niveles de gris de ventana, 0 y G', el centrifugado llevaría al k_s' fuera de la escala. Pues bien, cuando ello suceda, podemos solventar el problema al obligar al nivel de gris k' a irse al extremo de la escala correspondiente. Obsérvese que el razonamiento es aplicable asímismo a ruido impulsivo uniformemente distribuído, con igual resultado, debido a la simetría de la distribución de los niveles de gris de los pulsos..

Desgraciadamente, esta manera de proceder, si bien parece prometedora, presenta un grave inconveniente. Para que la expresión de corrección obtenida sea válida, hemos de partir del hecho de que en cada ventana la probabilidad de aparición de ruido blanco o negro sea la misma. Ello es cierto para la imagen completa, pero deja de ser verdad para subregiones de la misma. Y cuanto más pequeña es la subregión, menos fiable es la condición de equiprobabilidad que sustenta el anterior estudio. Se hace preciso atacar el problema desde un punto de

vista estadístico, muy complejo, que daría lugar a cierto número de ecuaciones que tendrían un rango de validez limitado.

En lugar de proceder de ese modo, se ha preferido abordar la cuestión desde un punto de vista más práctico, basado en teorías de cálculo numérico. La entrada k' de la tabla de consulta puede corregirse mediante iteraciones, bajo la suposición previa de que tal proceso será convergente. Como resultado, se obtendrá una imagen binaria: la máscara de preselección. Los píxeles negros de la máscara marcan la posición de los que se han considerado ruidosos en la imagen fuente, mientras que los no ruidosos quedan marcados por píxeles blancos.

Antes del filtrado, la imagen fuente S es recorrida por una ventana deslizante como primer paso de la iteración del proceso de construcción de la máscara. Cada vez que la ventana está en una determinada posición (i,j), su nivel de gris k' será utilizado como entrada de la tabla de consulta. Evidentemente, el tamaño de la ventana es el mismo de la que se empleó para construir la máscara, pues en caso contrario estaríamos empleando informaciones de distinta naturaleza, correspondientes a diferentes interresoluciones. El resultado es una primera imagen binaria que denominaremos *premáscara de selección*, $\underline{M}_0$. Esta imagen no es nada fiable como instrumento marcador de píxeles que deban ser procesados, debido al efecto centrífugo del ruido ya explicado. Es entonces cuando comienza el proceso iterativo.

La imagen S se vuelve a recorrer con la ventana deslizante. Pero en esta ocasión el nivel de gris de la ventana se computa de una manera distinta, con objeto de corregir el efecto centrífugo ya mencionado. Tan sólo los niveles de gris de los píxeles no ruidosos según $\underline{M}_0$ son acumulados: Se divide la suma de esos niveles de gris por el número de píxeles que han intervenido y se multiplica el resultado por el tamaño de la ventana. Se obtiene de esta manera el *nivel de gris de ventana estimado*, $\underline{k}'$:

$$\underline{k'} = \underline{\Sigma}\left[W_{ij}(S), \underline{M}\right] = \Sigma\left[W_{ij}(S \cdot \underline{M})\right] \frac{(2v_w + 1) \times (2h_w + 1)}{\Sigma\left[W_{ij}(\underline{M})\right]}$$

Este nivel de gris estimado corresponde a una entrada más adecuada de la tabla de consulta que k_s'. Repitiendo este proceso, empleando cada máscara previa $\underline{M}_i$ para obtener una nueva $\underline{M}_{i+1}$, puede alcanzarse un valor convergente en el límite que será la máscara de preselección $\underline{M}$. Obsérvese que el nivel de gris de

ventana estimado $\underline{k'}$ puede estar indeterminado, en el caso de que no haya ningún píxel no ruidoso en la ventana ($\Sigma[W_{ij}(\underline{M})] = 0$). En tal caso, por precaución, se considerará el píxel central como ruidoso.

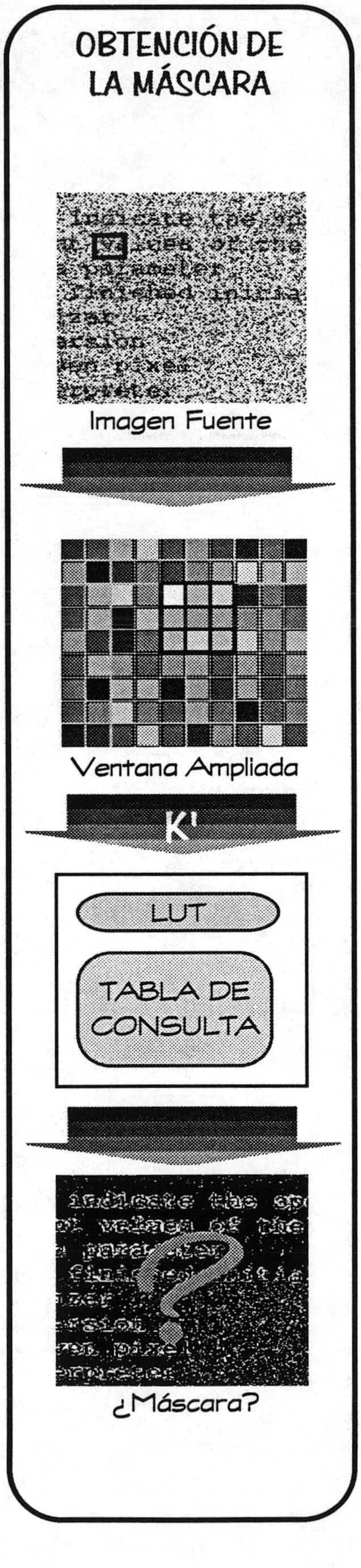
No debe olvidarse que para poder realizar este método iterativo hemos supuesto que el proceso era convergente. Desgraciadamente, no podemos asegurar que ello (condición de Lipschitz) sea cierto en todas las ocasiones. Las iteraciones podrían no converger en el límite, alcanzándose entonces soluciones oscilantes. Para evitarlo, se ha ideado un mecanismo de parada del algoritmo que fuerza la convergencia. Se basa en la aplicación de dos nuevos conceptos, denominados *descartes* y *persistencia*. La persistencia ρ representa el número de veces que un cierto píxel permanecería en el mismo estado (ruidoso o no) a lo largo de las iteraciones, a partir de un determinado paso inicial que queda determinado por los descartes d, que son un determinado número de pasos iniciales del proceso iterativo que se desechan. Los descartes tienen sentido porque los primeros pasos de la iteracón que da lugar a la máscara $\underline{M}$ están muy alejados de la convergencia (son muy diferentes entre sí), por lo que pueden ser ignorados. Como resultado, si un píxel alcanza el número p de iteraciones, una vez descartados los d primeros, y no ha cambiado su comportamiento (ruidoso o no), entonces se le considerará en ese estado para siempre.

De esta manera, si elegimos previamente los valores de ρ y de d, el proceso de iteración de máscara puede representarse como sigue:

$$\underline{M_0} \rightarrow \underline{M_1} \rightarrow \ldots \rightarrow \underline{M_d} \rightarrow \underline{M_{d+1}} \rightarrow \ldots \rightarrow \underline{M_{d+\rho}}$$

El número de pasos necesarios para la convergencia, que indican cuán rápido o lento es el proceso iterativo, se ha hecho coincidir con la suma de los valores de los descartes y la persistencia. Cuando este paso es alcanzado, la persistencia nos informa sobre el comportamiento de la mayoría de los píxeles. Cuando alguno de ellos ha variado su estado de ruidoso o no ruidoso mientras alcanzaba el último paso, se le considerará ruidoso por precaución. La máscara final es la que se alcanza en la iteración $(d + \rho)$-ésima:

$$\underline{M_{d+\rho}} \rightarrow \underline{M}$$



Los valores de d y μ que mejor resultado ofrecen son los siguientes:

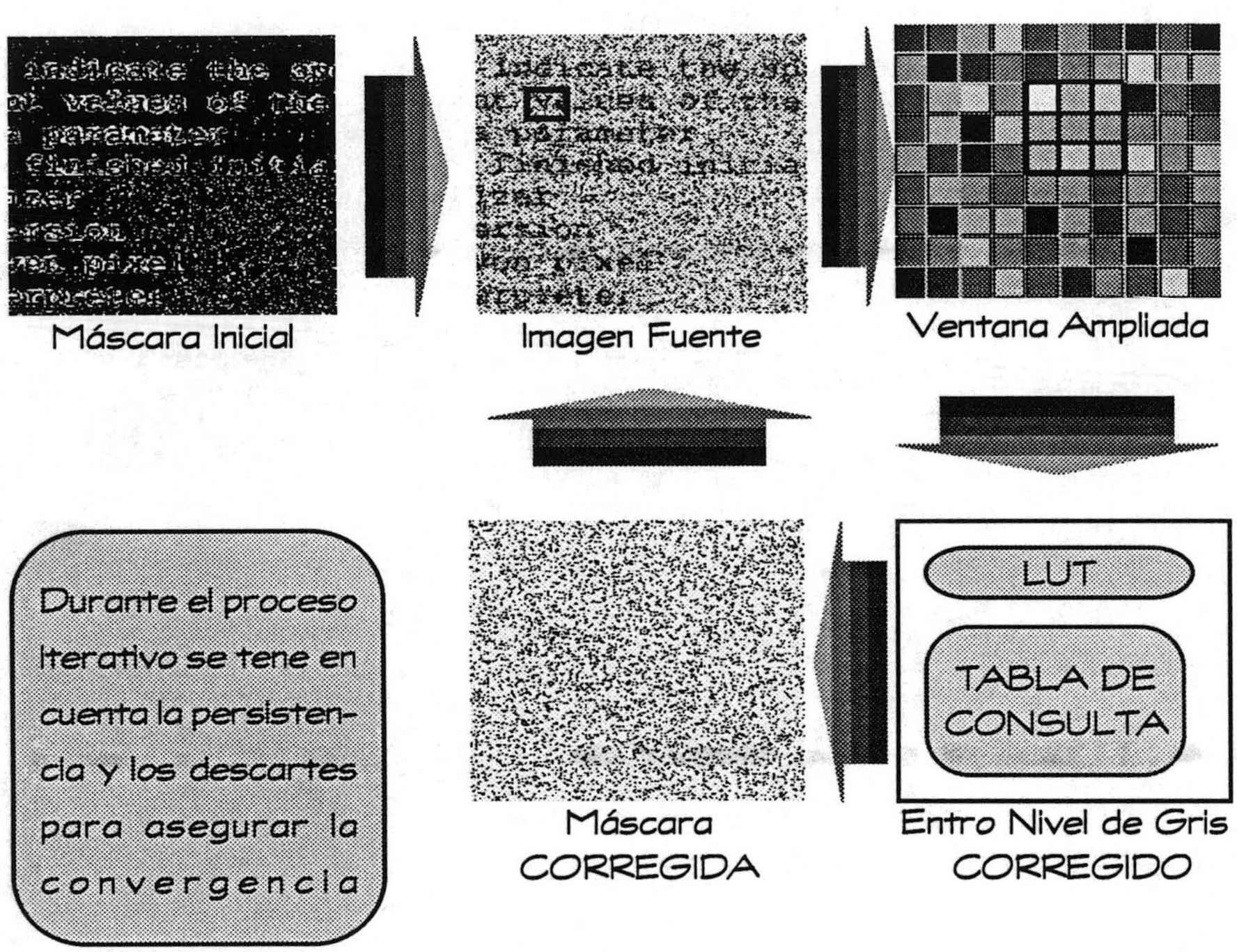
- $d = 5$
- $\mu = 4$

De este modo, se realizan siempre 9 iteraciones como máximo a la hora de construir la máscara de estimación de ruido.

Conviene distinguir aquí entre las *iteraciones de máscara* y las *iteraciones de filtro*. Mientras las primeras se refieren a los pasos del método de construcción de $\underline{M}$ explicado en esta sección, las últimas hacen mención al resultado de repetir el filtrado sobre la imagen resultado del procesamiento anterior. Esta diferenciación viene a colación porque la obtención de la máscara de estimación de ruido se realiza de diferente manera según estemos en la primera iteración de filtro, o en otra posterior. En efecto, en el caso de realizar la primera iteración de filtro, se llevan a cabo todas las iteraciones de máscara, como se ha explicado. Pero si repetimos el filtrado, la nueva máscara se calcula a partir de la máscara obtenida en la anterior iteración de filtro. Para ello se tiene en cuenta que aquellos píxeles que se dieron por limpios en la primera iteración de filtro seguirán siendo limpios una vez procesados: sólo aquellos píxeles considerados ruido podrán haber sido modificados por el algoritmo de filtrado. Es por ello por lo que la nueva máscara ($\underline{M}_{nueva}$) se determina a partir de la precedente ($\underline{M}_{antigüa}$) de la siguiente manera:

- Si $\underline{M}_{antigüa} = 1$ (píxel limpio) $\Rightarrow \underline{M}_{nueva} = 1$ (píxel limpio)
- Si $\underline{M}_{antigüa} = 0$ (píxel ruido) $\Rightarrow \underline{M}_{nueva}$ = calculado

Para calcular el valor de los píxeles de la nueva máscara no se realiza, por tanto, la totalidad de las iteraciones de máscara: tan sólo la primera, pero sólo sobre los píxeles dados por malos en la máscara precedente. Por lo tanto, en el caso de iteraciones de filtro posteriores a la primera, los valores de descartes y persistencia no son los expuestos arriba, sino que $d = 1; \mu = 0$.



El procedimiento que calcula la máscara se muestra a continuación:

ALGORITMO DEL PROCESO ITERATIVO

```
Algorithm Calcula_nivel_gris_ventana(ventana,
imagen, máscara): k';
{Este algoritmo devuelve el nivel de gris de una
ventana sobre una imagen}
begin
     k' := (suma niveles de gris de píxeles no
     ruidosos en ventana)*(tamaño de la ventana)/
     (número de píxeles no ruidosos en ventana);
end;
```

```
Algorithm Máscara(imagen_fuente, máscarantigüa):
máscaractual;
{Calcula una imagen binaria que marca los píxeles
ruidosos y no ruidosos}
begin
  for todas las posiciones (i,j) de ventana sobre
  imagen_fuente do
    if no hay píxeles no ruidosos en la ventana
    then máscaractual_ij := Malo;
    else
       begin
         k' := Calcula_nivel_gris_ventana(ventana,
         imagen_fuente, máscarantigüa);
         conjunto_píxeles_buenos := LUT[k'];
         {se usa la tabla de consulta, LUT, defini-
         da en el texto}
         if nivel_gris de píxel central de ventana
         ∈ conjunto_píxeles_buenos
             then máscaractual_ij := Bueno;
             else máscaractual_ij := Malo;
    end;
end;
```

```
Algorithm Iteración_Máscara(imagen_fuente):
máscara_definitiva;
{Obtiene la máscara definitiva como valor conver-
gente de esquema iterativo}
begin
      {Inicialización}
      máscarantigüa_ij := Bueno ∀ posición (i,j);
      máscara_definitiva_ij := Bueno ∀ posición (i,j);

      {Fase de DESCARTE}
      repeat número de descartes
          begin
            máscaractual:=Máscara(imagen_fuente,
            máscarantigüa);
            máscarantigüa := máscaractual;
          end;

      {Fase de PERSISTENCIA}
      repeat número de persistencia
          begin
            máscaractual:= Máscara(imagen_fuente,
            máscarantigüa);
            for ∀ (i,j) en imagen_fuente do
            if máscaractual_ij ≠ máscarantigüa_ij
              then máscara_definitva_ij := Malo;
              else máscarantigüa_ij := máscaractual_ij;
      end;
```

```
        for ∀ (i,j) en imagen_fuente do
            if máscaractual_ij = Malo
              then máscara:definitiva_ij := Malo;
end;
```

MEDIDAS DE CALIDAD DE MÁSCARA

Los errores característicos de la limpieza de ruido impulsivo son:

- Los de falsa detección, en los cuales píxeles limpios se consideran ruidosos. Resultan en un emborronamiento indeseable de detalles y perfiles en el procesamiento posterior.

- Los provenientes de la no detección de ruido, que da lugar a pulsos no eliminados en la imagen resultado.

Ambos tipos de error dependen del umbral de detección. En efecto, si disminuímos el umbral, el número de píxeles dados por limpios aumenta. Por tanto, hay mayor error por falsa detección, y menor de error producido por considerar píxeles ruidosos por limpios.

Por tanto, el grado de éxito en la estimación del ruido a través de esta máscara de preprocesamiento se puede medir mediante las proporciones de píxeles ruidosos y no ruidosos que $\underline{M}$ ha marcado adecuadamente. Estas medidas se definen a continuación:

Medida de Acierto	*Definición*
Acierto en píxeles no dañados (udr)	$\frac{\sum (M \cdot \underline{M})}{\sum M}$
Acierto en píxeles dañados (ddr)	$\frac{\sum [(1-M)\cdot(1-\underline{M})]}{\sum M}$

Recordemos que las operaciones entre imágenes se realizan píxel a píxel, y que las máscaras son imágenes binarias, en las cuales los píxeles con nivel 1 (blanco) son los considerados limpios; mientras que los ruidosos aparecen con nivel de gris 0 (negro). Desde luego, una máscara ideal sería capaz de detectar impecablemente ambos tipos de píxeles. Pero ello no siempre será alcanzable. Es posible diseñar una medida que aúne estos dos conceptos en uno solo. En principio podría pensarse que, cuando el ruido es difícil de detectar, como es el caso del impulsivo uniforme, sería mejor máscara aquella que detectase todos los píxeles erróneos, aunque para ello confundiese algunos píxeles limpios y los declarara como ruido. Esta decisión se justificaría por la escatología del proceso de filtrado: si se pretende eliminar el ruido, hay que detectar todo el ruido. Sin embargo, un análisi más detallado muestra que este aserto no es del todo válido. No cabe duda que la detección, —y posterior eliminación—, de aquellos píxeles cuyo nivel de gris destaca sobre el de sus vecinos dará lugar a una considerable mejora de la calidad de la imagen. Pero si se pretende suprimir la totalidad de los pulsos presenta un importante problema, en el caso del ruido impulsivo uniforme: al detectarlos a todos se incluirán, de forma irremediable, algunos píxeles limpios, puesto que existen niveles de gris de los pulsos que se parecen a los niveles originales. Por lo tanto, si la máscara localiza todos los impulsos, dará por buenos sólo aquellos pixeles cuyo nivel de gris es muy probable de acuerdo con el modelo. Es decir, quedarán sólo los muy buenos. Eso introducirá un notable sesgo en la LUT, que dará lugar a resultados imprevisibles. En definitiva, no hay que primar ni a *udr* ni a *ddr*. Se propone entonces la siguiente medida:

$$CM = (udr)\cdot(ddr)/100$$

denominada *CM, calidad de máscara*. Cuanto más cercano a 100, mejor es la máscara en cuanto a estimación de la posición de los pulsos.

EJEMPLOS DE OBTENCIÓN DE MÁSCARA DE ESTIMACIÓN DE RUIDO

A continuación se presentan unos ejemplos que ponen de manifiesto el satisfactorio comportamiento de la máscara de estimación de ruido, que es capaz de detectar los pulsos con gran precisión, incluso cuando el modelo es muy disperso o el ruido presenta distribución uniforme de niveles de gris.

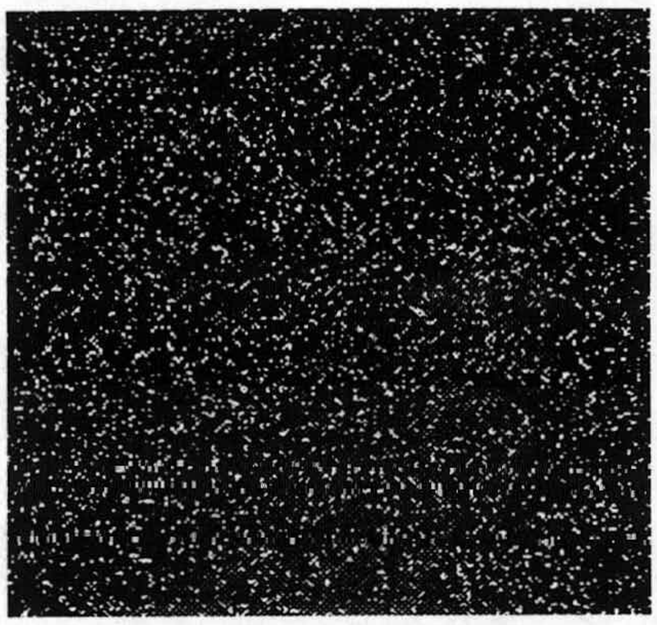
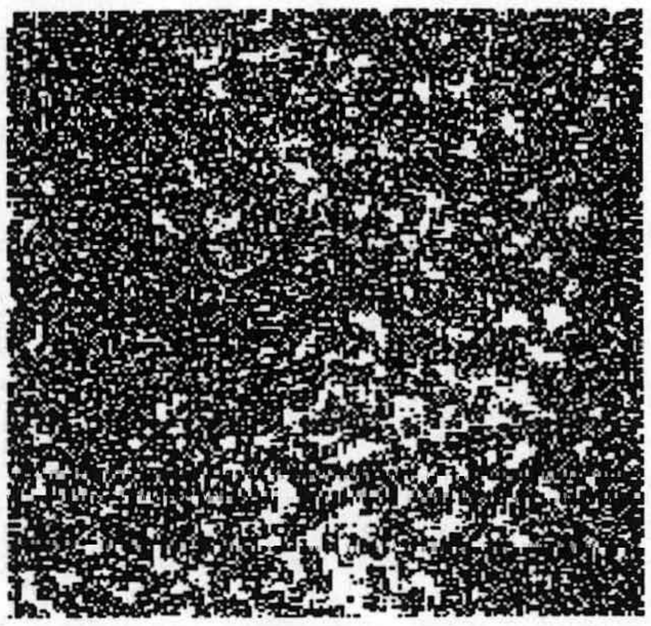
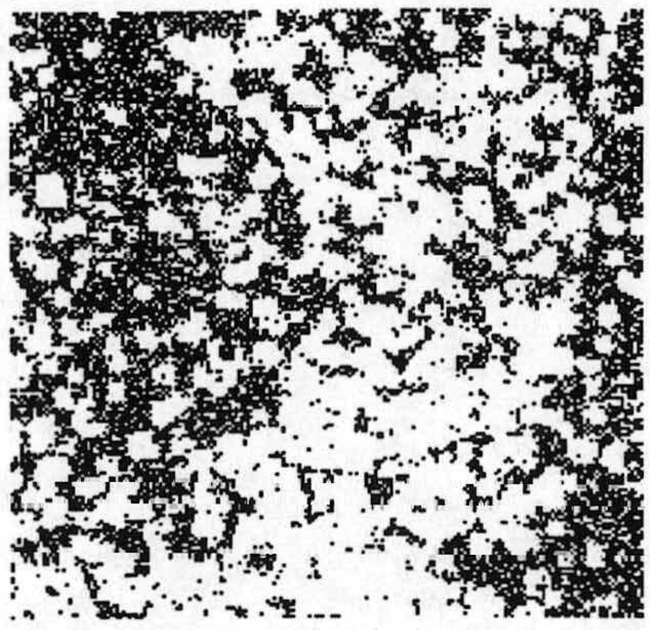


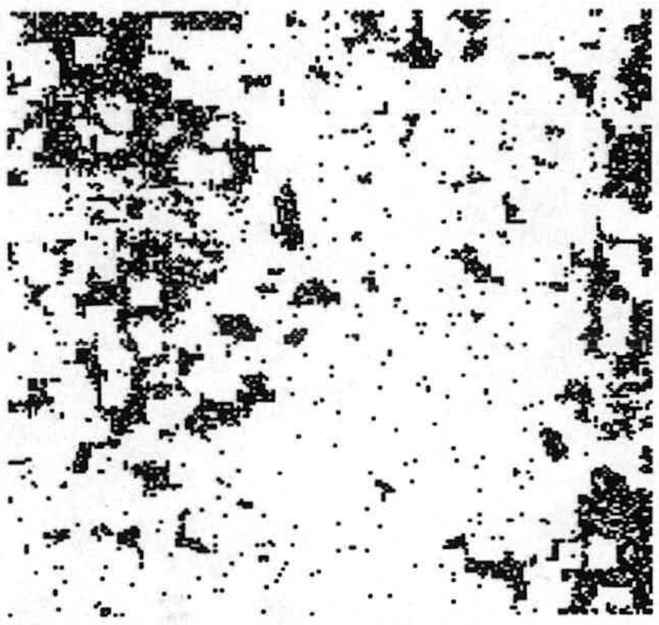
Imagen Fuente



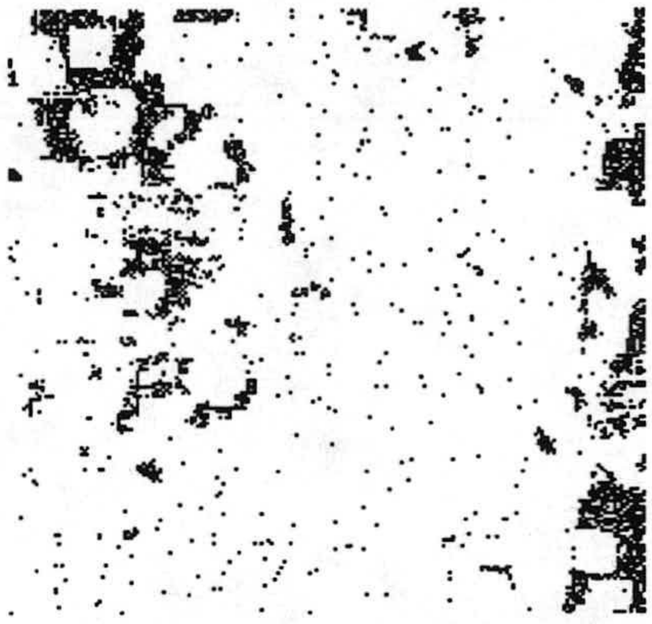
Δ_0



Δ_1



Δ_2



$\Delta_7 = |\underline{M} - M|$

Figura 3: Se muestra la convergencia estadística del proceso de construcción de máscara, para el caso de una imagen de la clase de radiografías afectada por 20% de ruido impulsivo uniformemente distribuído. Se ha utilizado una ventana 3×3. Las imágenes $\Delta_i = |\underline{M}_i - M|$ representan el error cometido al obtener la máscara en la iteración *i*. Los resultados son evidentes. También se muestran valores numéricos de calidad de máscara para el caso representado, empleando las medidas de calidad propuestas en el texto. De nuevo, se aprecia cómo la convergencia se asegura, manteniendo una calidad de máscara más que aceptable.

Iteración i	*udr*	*ddr*	*CM*	δ_i
0	15,34	93,11	14,28	
1	56,19	93,99	52,81	135544
2	80,03	94,64	75,74	7880
3	89,78	95,75	85,96	3317
4	94,36	96,13	90,71	1554
5	96,72	96,3	93,14	800
6	96,72	96,58	93,41	22
7	96,72	96,58	93,41	0

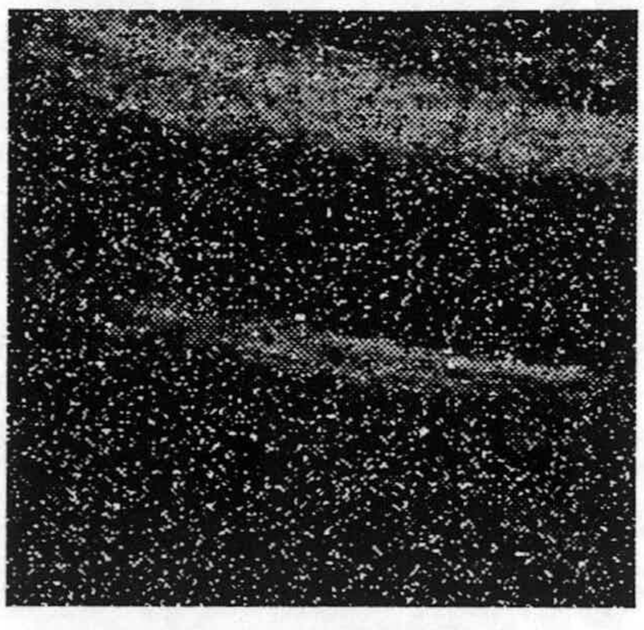
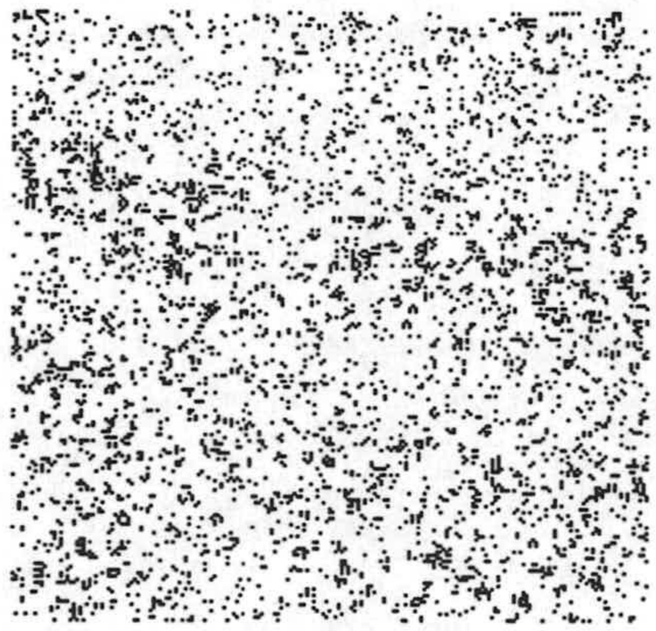
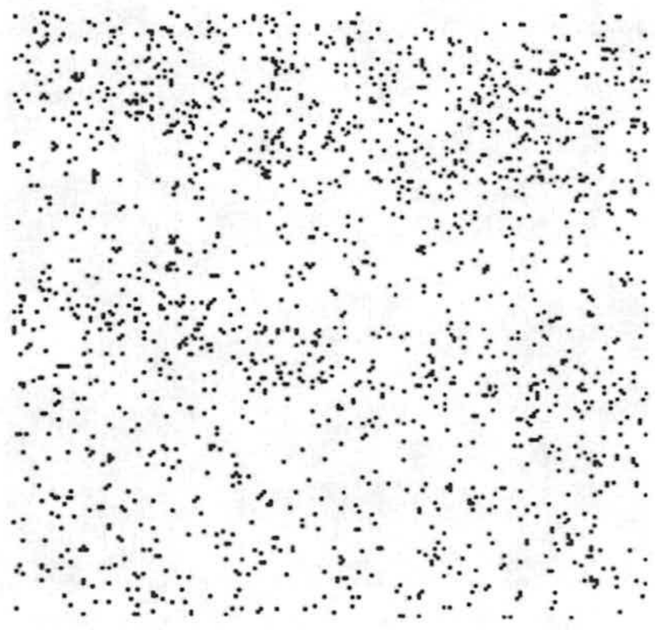


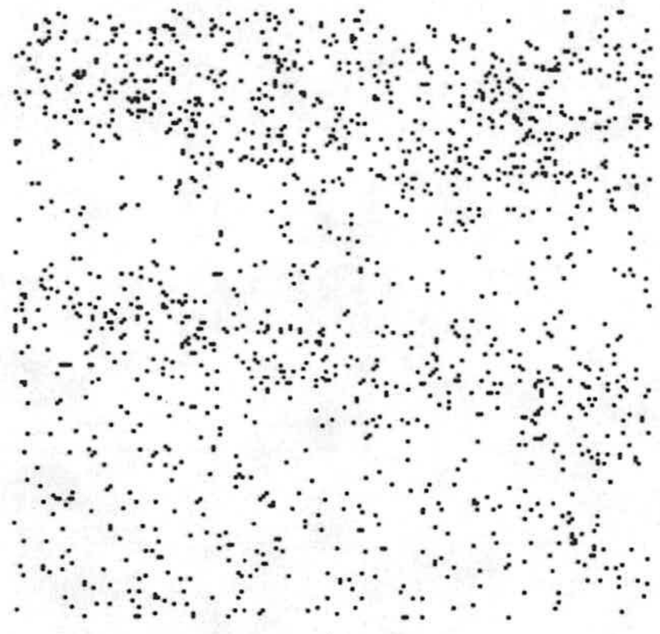
Imagen Fuente



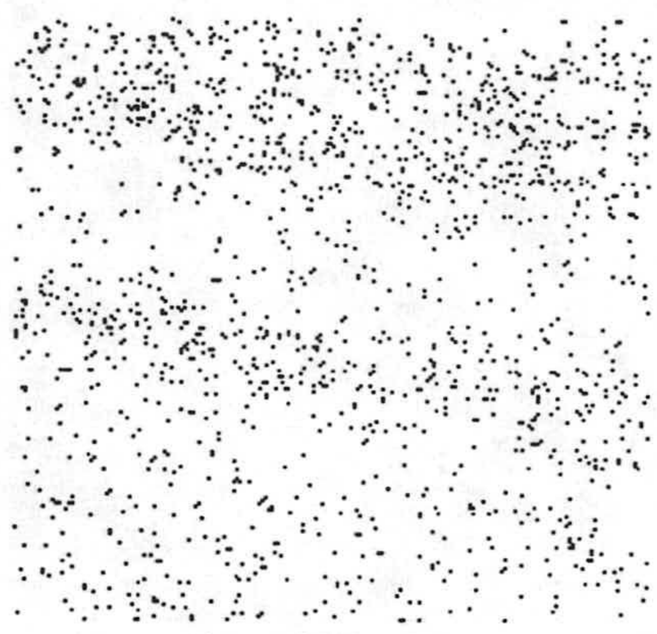
Δ_0



Δ_1



Δ_2



$\Delta_8 = |\underline{M} - M|$

Figura 4: Imagen fuente de la clase de ecografías (20% ruido impulsivo uniformemente distribuído) e imágenes diferencia entre máscara real y calculada para la iteración indicada en el subíndice. La convergencia es evidente. Se muestra también una tabla con resultados numéricos de calidad de máscara de acuerdo con las medidas propuestas en el texto.

Iteración i	*udr*	*ddr*	*CM*	δ_i
0	68,3	84,25	57,54	
1	94,15	87,51	82,39	9043
2	97,88	89,56	87,59	1586
3	98,62	89,99	88,75	392
4	98,82	90,18	89,12	160
5	98,85	90,17	89,13	102
6	98,75	90,36	89,23	47
7	98,74	90,37	89,23	2
8	98,74	90,37	89,23	0

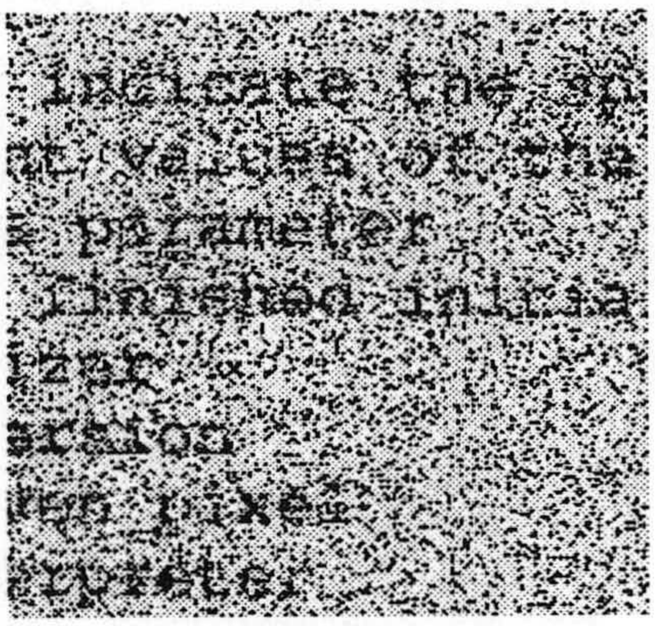
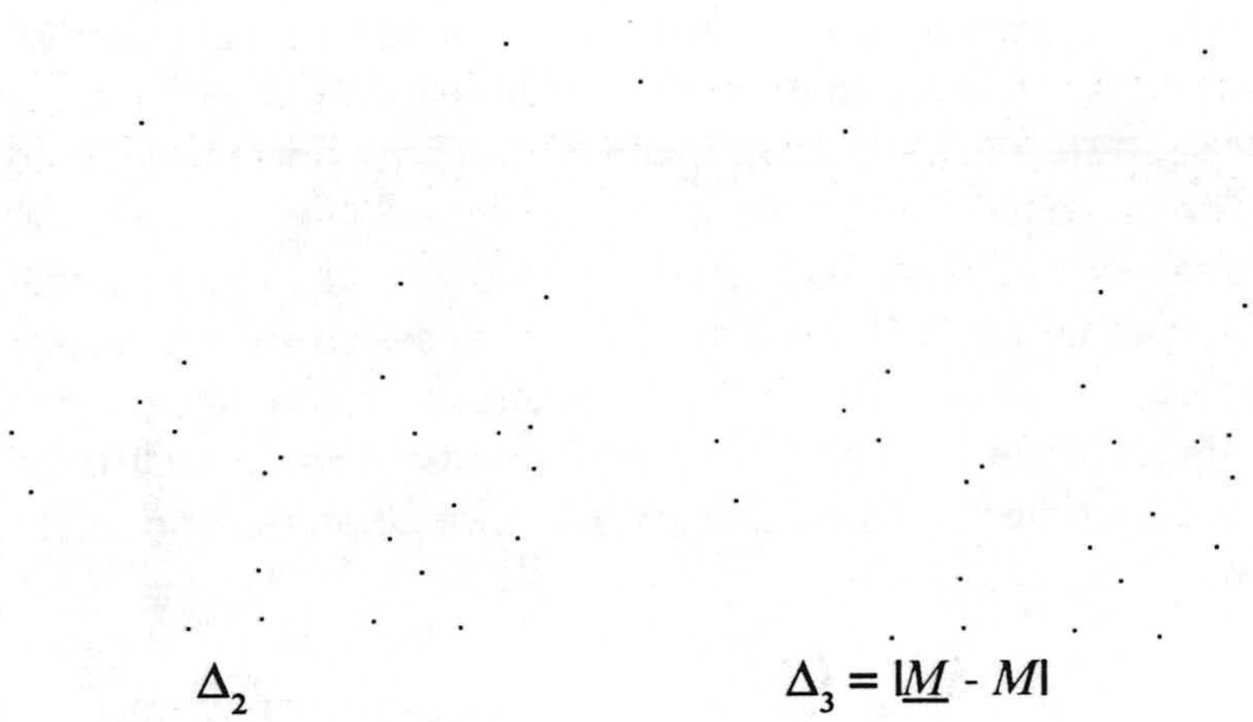


Imagen Fuente

Δ_0

Δ_1



Δ_2

$\Delta_3 = |\underline{M} - M|$

Figura 5: Imagen de texto mecanografiado afectado por 30% de ruido impulsivo *sal y pimienta* y las imágenes de diferencia entre máscara estimada y real para diferentes iteraciones. Obsérvese cómo la convergencia sucede en la 3 iteración, siendo el error minúsculo. Este hecho puede comprobarse a la vista de la tabla de datos numéricos de calidad de máscara.

Iteración i	*udr*	*ddr*	*CM*	δ_i
0	97,17	100	97,17	
1	99,91	100	99,91	258
2	99,91	100	99,91	1
3	99,91	100	99,91	0

La evolución del proceso iterativo que da lugar a la máscara $\underline{M}$ puede observarse en la tabla contenida en la figura 3 para una radiografía afectada por un 20% de ruido impulsivo uniforme, en la que se ha empleado una ventana de tamaño 5×5. Tanto *udr* como *ddr* alcanzan valores cercanos al 97% tras siete iteraciones. Obsérvese cómo los valores iniciales, sobre todo en la premáscara, están realmente alejados de la convergencia, lo

cual justifica el empleo de los descartes. La convergencia del proceso se observa con mayor claridad en la última columna de la tabla, dónde se han plasmado los valores de la diferencia $\delta_i = /\underline{M}_i - \underline{M}_{i-1}/$ entre dos iteraciones consecutivas.

Esta evolución está gráficamente mostrada en la figura 3 para el caso de una radiografía afectada de 20% de ruido impulsivo uniformemente distribuído, donde aparecen las imágenes de diferencia entre máscara real y calculada $\Delta_i = /\underline{M}_i - \underline{M}/$ para las iteraciones $i = 1, 2, 3, 7$.

En las figuras 4 y 5 se muestra también la evolución del proceso iterativo de construcción de máscara tanto para el caso de ecografías como para textos. En la primera de las figuras la ecografía ha sido contaminada por 20% de ruido impulsivo uniformemente distribuído, y se ha empleado una ventana de 3×3. Tras 8 iteraciones, la máscara converge a una máscara más que aceptable. En la figura 5 se presenta el caso de la evolución de la máscara construída para una imagen de texto mecanografiado afectada por 30% de ruido impulsivo *sal y pimienta*. En este cos la convergencia sucede ya en la tercera iteración, con un resultado excelente.

A la vista de los resultados expuestos podemos extraer algunas conclusiones. En primer lugar, el comportamiento general del método iterativo propuesto es más que aceptable. Obsérvese que para ruido impulsivo *sal y pimienta* la máscara definitiva se obtiene pronto y con una aproximación inmejorable a la máscara real. Sin embargo, lo realmente llamativo es que para ruido impulsivo uniformemente distribuído la máscara también acierta muy bien. Obsérvese que en este tipo de ruido los pulsos tienen niveles de gris aleatorios, por lo que es posible que adquieran el nivel de la vecindad, y puedan detectarse como buenos. Y también puede darse el caso contrario, en el que ciertos píxeles limpios se confundan con el ruido. Pues bien, aun en estos casos extremos el algoritmo se muestra robusto y presenta un muy buen comportamiento. Por otra parte, la introducción de los descartes y la persistencia asegura la convergencia, pero no introduce errores significativos. Además, la idea de los descartes se ve justificada por la gran diferencia entre la imagen calculada y la real en las priemras iteraciones del proceso. Queda, pues, validado el método de iteración para la construcción de la máscara de estimación de ruido, y será empleado en los futuros diseños de filtro que se expondrán en el siguiente capítulo.

EL PROBLEMA DEL ACCESO A LA LUT. LA MÁSCARA GRIS.

La máscara binaria de estimación de ruido que se ha presentado cumple la tarea de marcar los píxeles que han de ser procesados. Sin embargo, la máscara debe cumplir otra tarea igualmente importante si el proceso de filtrado emplea la información contenida en las matrices de composición. En efecto, el adecuado uso de los modelos consiste en adquirir conocimiento acerca de la probabilidad que un determinado nivel de gris de píxel tiene dentro de una región de un cierto nivel de gris de ventana. Así, hay que acudir a la columna del modelo adecuada en cada momento para conseguir esa información. Sin embargo, en la práctica la columna consultada suele ser muy diferente a la correcta. Ello se debe a la presencia de ruido impulsivo en la imagen fuente. Cuando la ventana deslizante la recorre, va delimitando regiones cuyo nivel de gris se calcula como suma de los niveles de los píxeles integrantes. Pero como algunos de estos píxeles son en realidad ruido impulsivo —tanto más cuanto mayor sea la cantidad de ruido introducido—, el nivel de gris de la ventana k_s' no es el correcto. El acudir a la columna correspondiente a k_s' nos proporcionará información errónea acerca del nivel de píxel k en consideración. Por lo tanto, el empleo de conocimiento presente en los modelos, cuando hay ruido presente, es inadecuado.

El estudio teórico realizado en este mismo capítulo proporciona una expresión de corrección válida, pero que precisa el conocimiento previo de lacantidad de ruido introducido en la imagen. Así pues, la máscara es la que debe solventar el problema. En efecto, si $\underline{M}$ tiene calidad suficiente, marcará la mayoría de los pulsos sin error. Entonces el nivel de gris de cualquier ventana se hará sumando los niveles de los píxeles limpios de la ventana, y extendiendo el valor a la totalidad de la ventana según se explicó anteriormente al introducir el concepto de nivel de gris estimado de ventana $\underline{k}'$. Entonces $Q_{k'} \approx Q_{\underline{k}'}$, por lo que la dificultad queda subsanada.

La bondad de este procedimiento depende de la calidad de la máscara, pero cuando ésta es buena, los resultados también lo son. Existe sin embargo la posibilidad de que la máscara no pueda ser lo suficientemente perfecta como para que el método de corrección propuesto sea eficaz. Tal sería el caso del ruido impulsivo uniformemente distribuído. Como el nivel de gris de los pulsos en este tipo de ruido puede ser cualquiera de la escala K, no sería extraño que algunos píxeles ruidosos quedasen sin detectar, e incluso que algunos píxeles limpios se die-

sen por malos. Esta dificultad intrínseca a la naturaleza de la distorsión no puede ser eliminada por métodos externos, como las iteraciones de máscara propuestas, por lo que la máscara resultante tendría una calidad limitada. En tal caso, el proceso de corrección propuesto no alcanzaría los resultados deseables.

En efecto, supongamos que la máscara es muy clara, esto es, casi todos los píxeles los considera limpios. Entonces, gran parte de los pulsos no han sido detectados. Por consiguiente, el error cometido al acceder a las columnas del modelo será notable. En el caso de que la máscara sea muy oscura, esto es, considera como ruido la mayor parte de la imagen, entonces la situacion es algo diferente. Los píxeles que esta máscara supone limpios son muy pocos. Se calcula entonces el nivel de gris de la ventana a partir de una muestra poco significativa. Como resultado, la LUT presenta un fuerte sesgo, que dará origen a comportamientos anómalos. De nuevo, la corrección pierde efectividad.

Bajo este punto de vista, se diseñó un nuevo concepto de máscara de estimación de ruido. En lugar de la decisión tan estricta que llevaba a cabo la máscara binaria (buenos o malos), se tomaba una determinación más relajada acerca de la limpieza del píxel. No se trataba ya de una imagen binaria, sino una imagen de niveles de gris. En la nueva máscara, $\underline{M}_g$, cada píxel tendría un nivel de gris proporcional a la probabilidad de aparición en la región, determinada por la ventana, centrada en él. Dicha probabilidad provendría de la información proporcionada por el modelo. De esta manera, el nivel de gris de una determinada ventana se determinaría así:

$$\underline{k'} = \Sigma[W_{ij}(S \cdot \underline{M}_g)]/G''$$

donde G'' es el máximo nivel de gris de la máscara $\underline{M}_g$. De este modo, la LUT no sufre sesgo de ningún tipo,pues todos los píxeles intervienen.

La construcción de este nuevo tipo de máscara sigue el mismo esquema iterativo que el propuesto para la máscara binaria. La única diferencia es que el umbral de selección no determina ahora si el píxel es ruidoso o no, sino que es un punto de referencia. Se calcula el nivel de gris del pixel central, y se compara con el umbral. Si es mayor, se le asignará un color

claro (que significa que es más bien limpio), y tanto más cercano al blanco cuanto más alejado del umbral. Análogo pocedimiento se sigue para asignar el nivel del píxel si su probabilidad, según el modelo, es menor que el umbral. Obsérvese que en el caso de la máscara gris, la convergencia del proceso iterativo es más compleja, pues las diferencias entre máscara sucesivas son ahora mayores. Pese a ello, los conceptos de descartes y persistencia siguen siendo válidos y adecuados.

Sin embargo, el uso de la máscara gris exige una binarización posterior, a la hora de realizar el procesamiento. Y es que en ese momento lo que interesa es localizar los pulsos, y la máscara $\underline{M}_g$ no lo hace. Mediante una umbralización adecuada, la máscara gris se convierte en binaria, se consigue ese objetivo. Desgraciadamente, ese umbral es una arbitrariedad poco defendible, cuyo valor es completamente *ad hoc*. Además, la complejidad algorítmica aumenta notablemente. Otra dificultad surge en el momento de buscar una medida de calidad de máscara gris, puesto que depende tanto del comportamiento fente a la corrección como frente a la localización.

En caulquier caso, se realizaron experimentos con la citada $\underline{M}_g$. Los resultados,que se muestran en el siguiente epígrafe, demostrarán que las ventajas proporcionadas por su uso no superan las dificultades que introduce, por lo que ha sido deshechada en los filtros diseñados que se analizarán en el próximo capítulo.

INFLUENCIA DEL PARÁMETRO DE SELECCIÓN SOBRE LA MÁSCARA

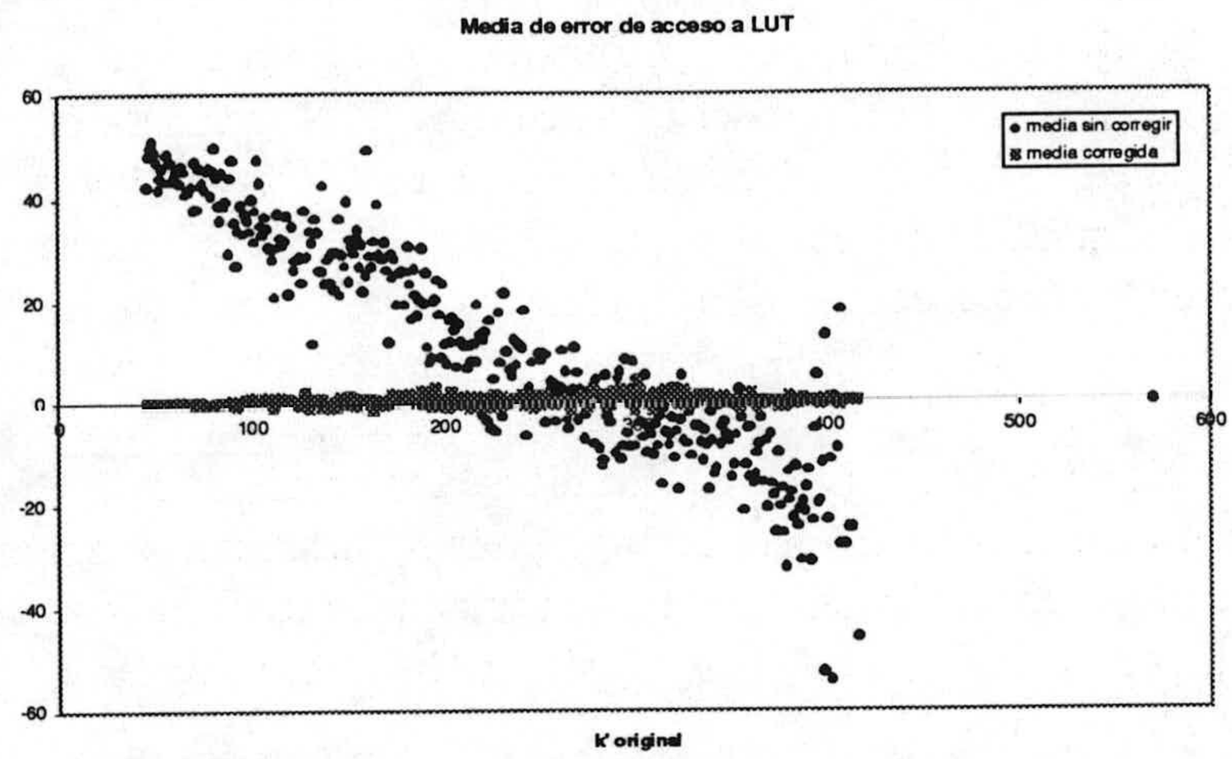
Una importante característica que puede determinar la calidad de la máscara $\underline{M}$ (y $\underline{M}_g$) es el umbral de selección definido por ρ —ρ_3 en nuestro caso—. Si éste posee valores bajos, entonces sólo se considerarán ruidosos aquellos píxeles cuyo nivel de gris tenga probabilidades de aparición realmente bajas. A medida que crece ρ, un mayor número de niveles de gris se tomarán como ruidosos. Por tanto, la calidad de la máscara puede incrementarse mediante una elección adecuada del umbral de selección mediante ρ. Puede apreciarse la veracidad de tal aseveración en los datos que aparecen en la tabla1. Fueron obtenido a partir de una imagen fuente de texto mecanografiado, afectada por un 20% de ruido impulsivo uniforme, utilizando una ventana de 5×5.

Tabla 1: Se plasman los resultados numéricos correspondientes al acierto en píxeles ruidosos y limpios por parte de la máscara, cuando se varía el umbral de selección de acuerdo con el cambio del parámetro ρ. Los resultados muestran una dependencia importante entre ambas cantidades, por lo que se realizará un estudio más profundo.

ρ	*udr*	*ddr*
0	99,93	46,55
1	99,54	61,41
2	98,85	67,96

Cabe realizar una llamada de atención acerca de la elección del parámetro ρ. Recuérdese que la existencia de ruido conducía a una alteración del valor del nivel de gris de ventana, por lo que se realizaba una consulta equivocada a la LUT, introduciendo un error en la determinación del carácter, ruidoso o no, de los píxeles en la imagen. Si se escoge un valor pequeño de ρ, entonces todos los píxeles marcados como ruidosos tendrán una probabilidad ciertamente pequeña de acuerdo con la entrada equivocada k_s' de la tabla de consulta. Pero, debido a que tal probabilidad es pequeña, es de esperar que dicho nivel de gris siga siendo considerado como ruidoso por entradas k' de la tabla de consulta vecinas a la incorrecta k_s'. Es decir, existe una cierta robustez frente a los errores producidos por el ruido si el parámetro de selección tiene un valor pequeño. Es posible que entonces la convergencia del proceso de construcción de la máscara sea más lenta, pero merece la pena, puesto que el incremento de calidad a través de las sucesivas iteraciones está garantizado.

En cualquier caso, se realizará a continuación un análisis más detallado de la influencia del umbral de selección en la calidad de la máscara, dada la importancia manifiesta de la relación



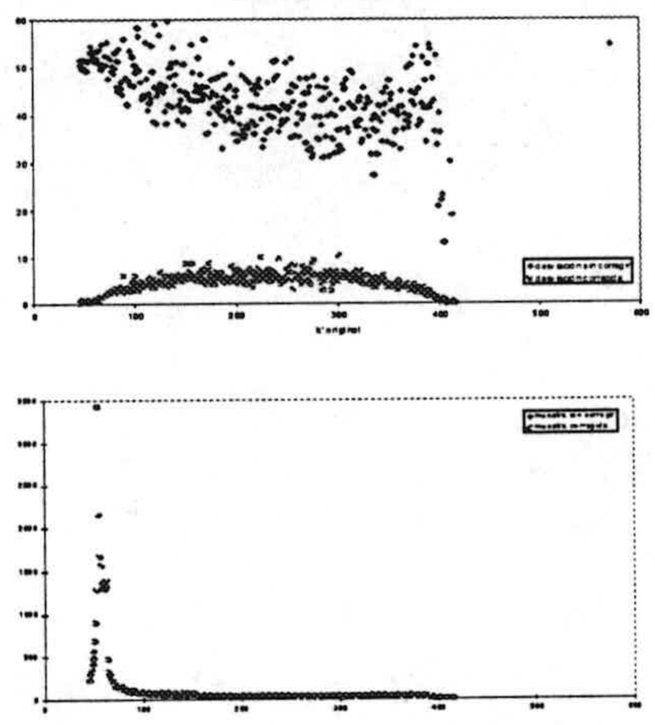
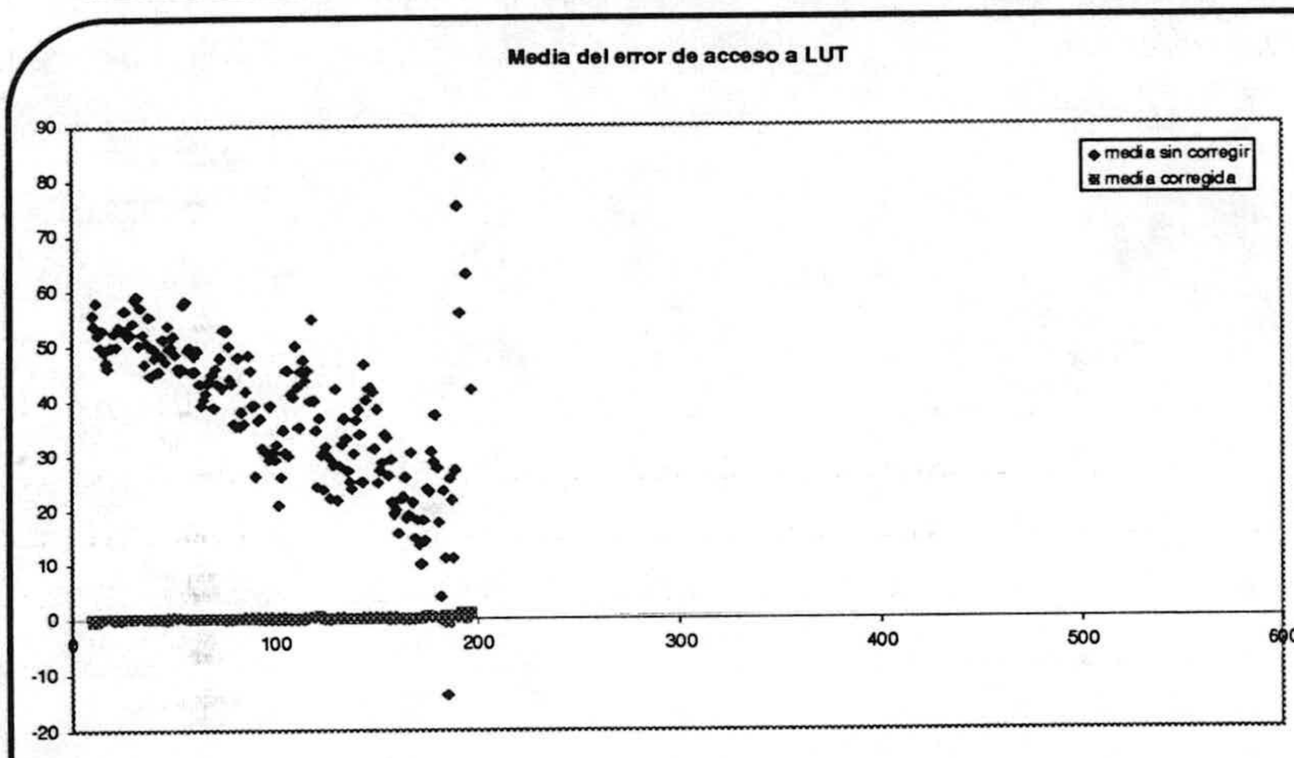


Figura 6: Media de los errores de acceso a la LUT. A la izquierda, arriba, la desviación típica de los errores. Abajo, el tamaño de la muestra usadas en las estadísticas. Imagen de ecografía con 20% ruido *sal y pimienta*. Tamaño de ventana 3×3. Se toma $\rho = 0$.



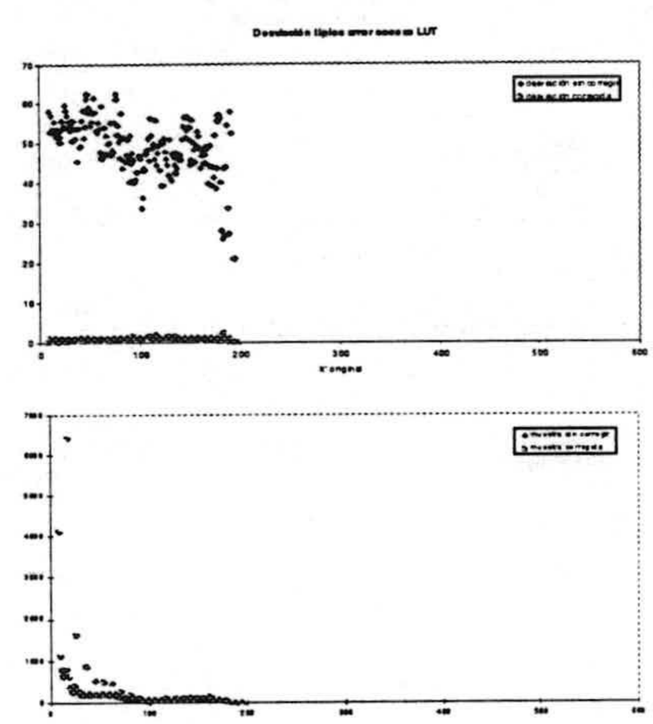
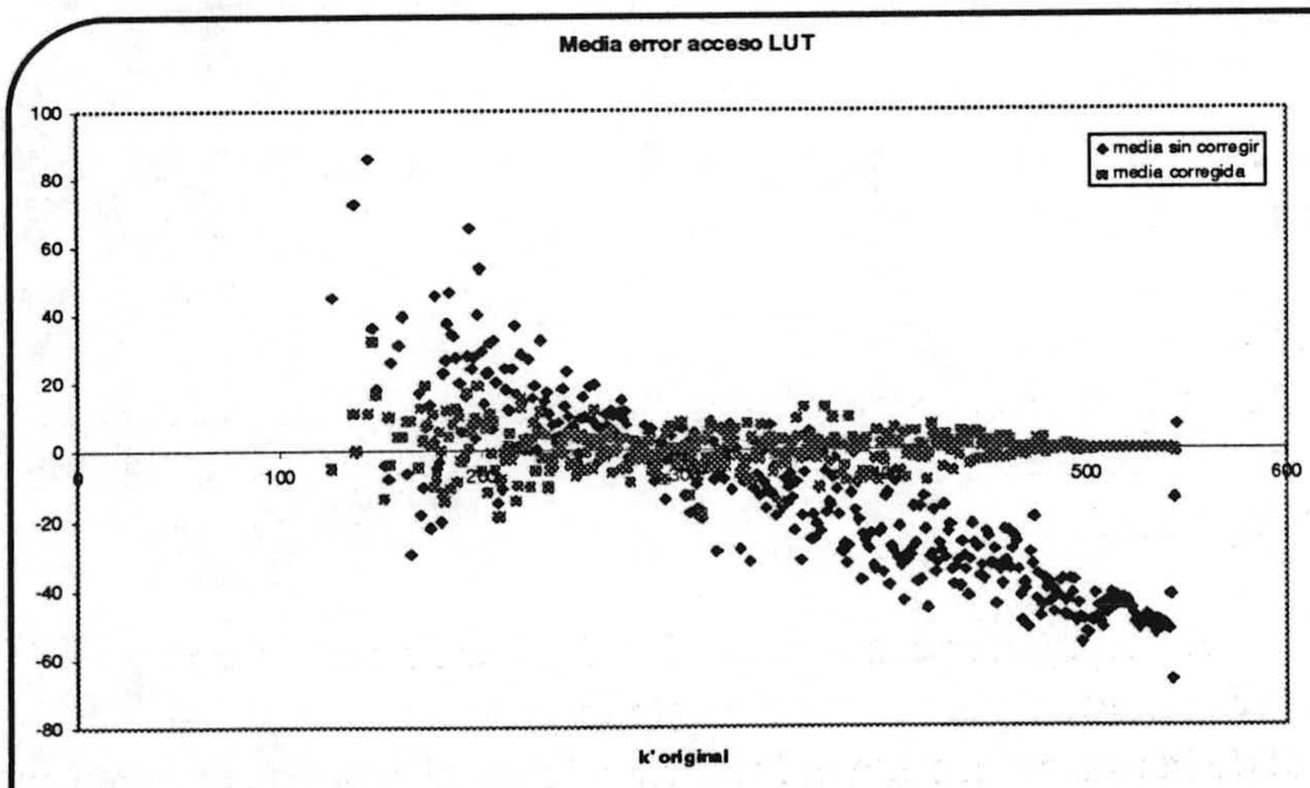


Figura 7: Igual que en la figura precedente, con igual ventana y ruido, pero para una imagen de radiografías. Aquí también $\rho = 0$.



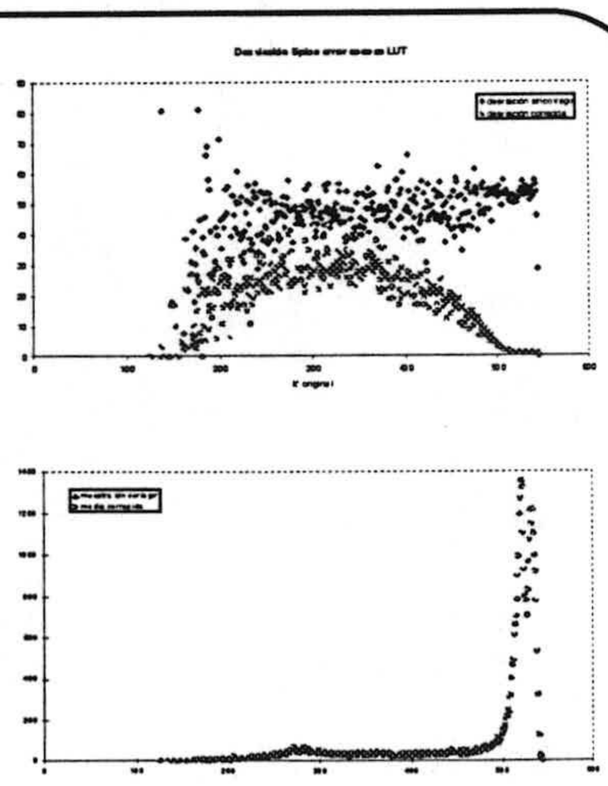


Figura 8: Se representa lo mismo que en las anteriores gráficas, pero para una imagen de texto y $\rho = 1$.

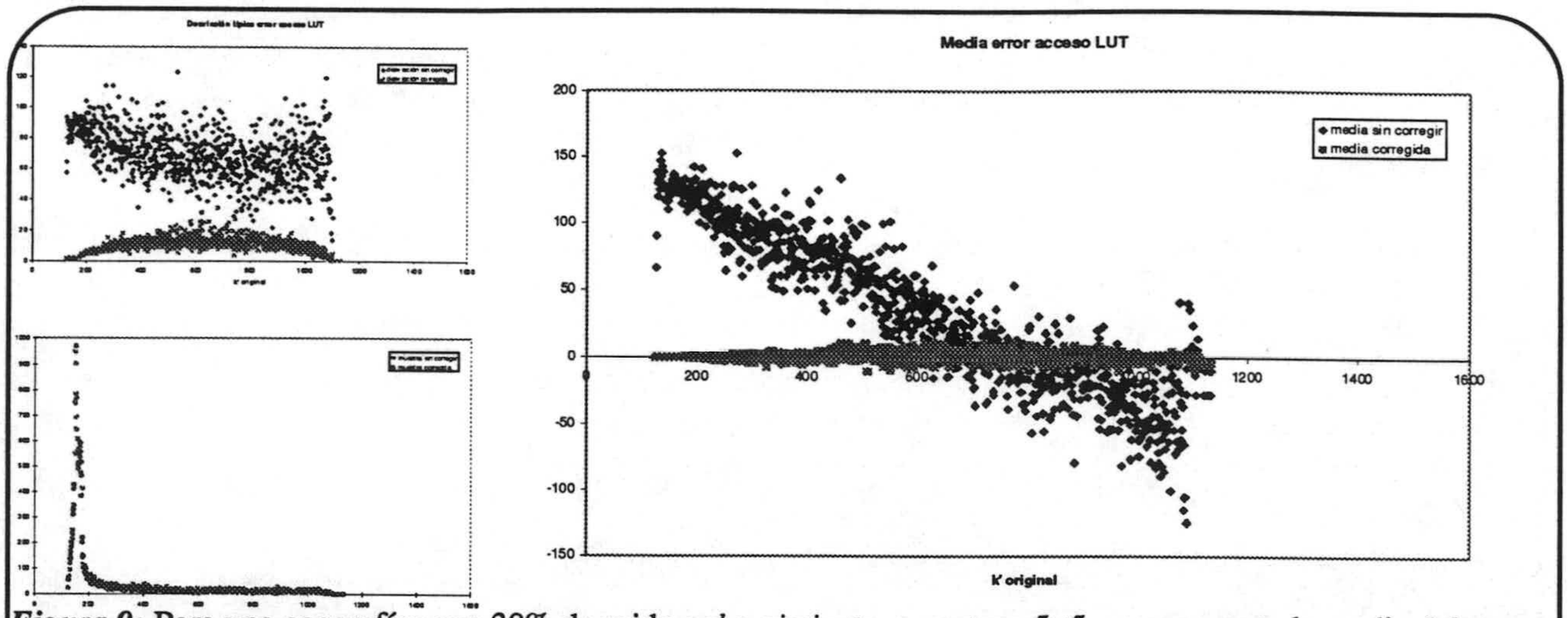


Figura 9: Para una ecografías con 30% de ruido *sal y pimienta*, y ventana 5×5, se representa la media del error de acceso a LUT (izq.), desviación típica de los errores (arriba, dcha.) y tamaño de la muestra. Aquí $\rho = 0$.

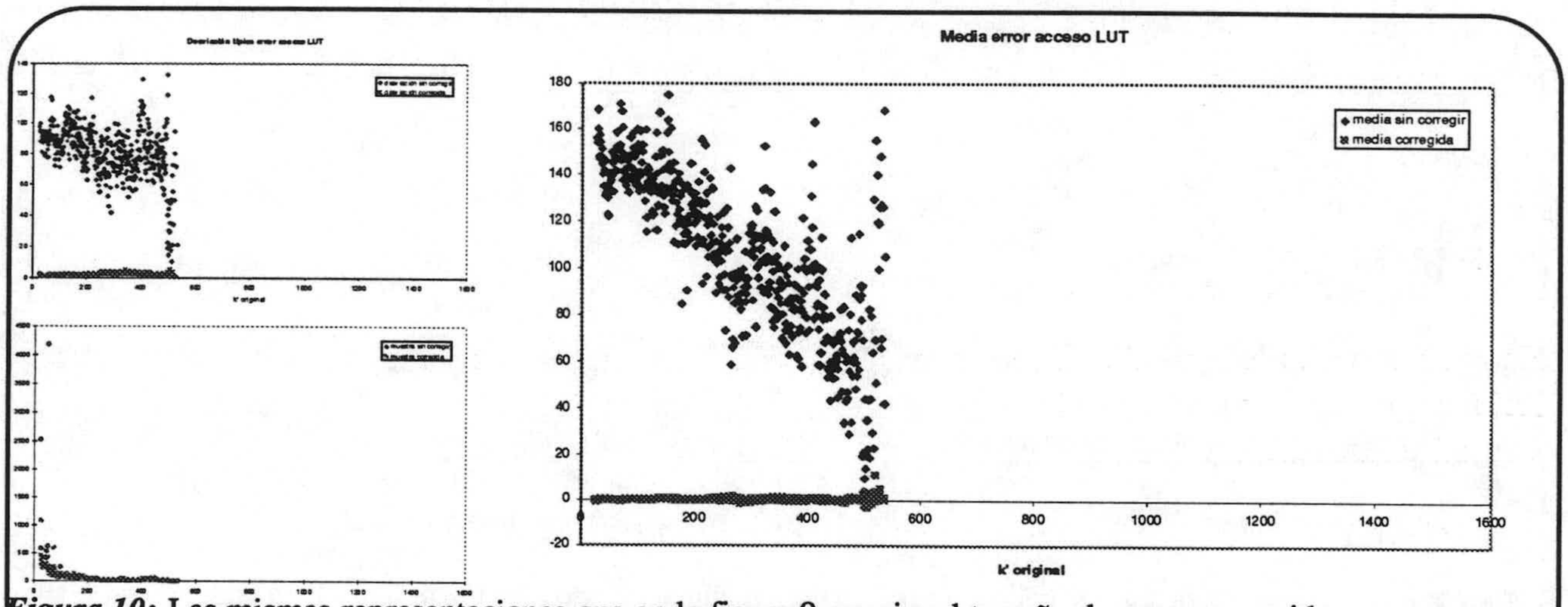


Figura 10: Las mismas representaciones que en la figura 9, con igual tamaño de ventana y ruido, pero para una imagen de radiografías. También $\rho = 0$.

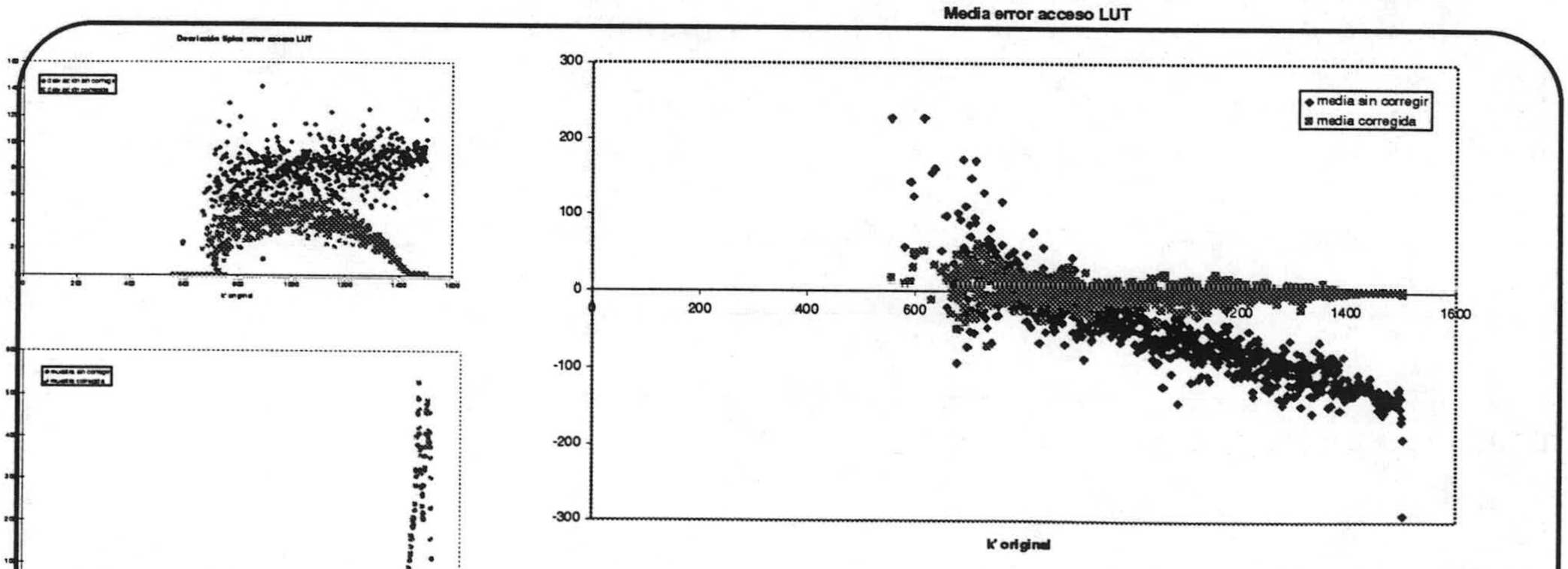


Figura 11: Igual que la figura 10, pero para un texto. Ahora $\rho = 1$.

entre ambas. De esta manera podremos concluir qué parámetros de selección pueden emplearse en el procesamiento, y cuáles no.

Influencia del Umbral de Selección en la Corrección del Nivel de Gris de Ventana de acceso a la LUT.

En primer lugar estudiaremos la influencia del umbral de selección en la capacidad de corrección del nivel de ventana por parte de la máscara. Para ello se han realizado una serie de experimentos en los que se ha calculado el error de acceso a la columna del modelo (en nuestro caso, a la LUT). Se muestra entonces tanto la media del error como la desviación típica del mismo. Se calculan las siguientes cantidades:

• *Media de la diferencia entre el nivel de gris de ventana en la imagen fuente y en la imagen original.*

$$\mu_{s.c.}(k') = \frac{\sum_{i=1}^{V}\sum_{j=1}^{H}\left(\Sigma\left[W_{ij}(S)\right] - \Sigma\left[W_{ij}(O)\right]\right)}{\chi_{s.c.}(k')}$$

Es decir, se calcula la media del error cometido en el acceso a la LUT cuando no se ha hecho la corrección, de ahí el subíndice *s.c.*. El valor de la muestra $\chi_{s.c.}$ es:

$$\chi_{s.c.}(k') = \sum_{i=1}^{V}\sum_{j=1}^{H}\delta\left(\Sigma\left[W_{ij}(O)\right] - k'\right)$$

donde $\delta(\cdot)$ es la función delta de Dirac.

• *Media de la diferencia entre el nivel de gris estimado de ventana y el nivel original.*

$$\mu_{c}(k') = \frac{\sum_{i=1}^{V}\sum_{j=1}^{H}\left(\Sigma\left[W_{ij}(S)\right] - \Sigma\left[W_{ij}(O)\right]\right)}{\chi_{c}(k')}$$

donde la muestra χ_c se calcula como:

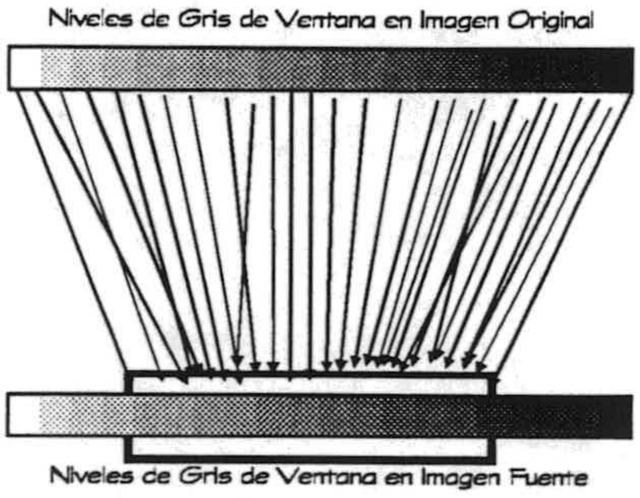


Figura 12: Ilustración del proceso que tiene lugar cuando se introduce ruido impulsivo a una imagen original. Los valores de la escala de grises de ventana se transforman en otros, al sumar el valor de los pulsos, que en promedio es igual a la mitad de la escala. Por eso el resultado es llevar el nivel de gris original a otro más centrado: es una especie de efecto *centrípeto*.

$$\chi_c(k') = \sum_{i=1}^{V}\sum_{j=1}^{H}\left\{\delta\left(\Sigma\left[W_{ij}(O)\right]-k'\right)\cdot\left[1-\delta\left(\sum_{u=1}^{V_w}\sum_{v=1}^{H_w}\underline{M}_{uv}-0\right)\right]\right\}$$

Es decir, se computa el error cometido al acceder a la LUT una vez realizada la corrección del nivel de gris de ventana.

• *Desviación típica de la diferencia entre el nivel de gris de ventana en la imagen fuente y en la imagen original.*

$$\sigma_{s.c.}(k') = \sqrt{\frac{\sum_{i=1}^{V}\sum_{j=1}^{H}\left(\Sigma\left[W_{ij}(S)\right]-\Sigma\left[W_{ij}(O)\right]\right)^2}{\chi_{s.c.}(k')}-\mu_{s.c.}^2(k')}$$

De esta manera se dispone de una medida de la dispersión de los errores de acceso a la LUT cuando no existe corrección del nivel de gris de ventana.

• *Desviación típica de la diferencia entre el nivel de gris estimado de ventana y el nivel original.*

$$\sigma_c(k') = \sqrt{\frac{\sum_{i=1}^{V}\sum_{j=1}^{H}\left(\underline{\Sigma}\left[W_{ij}(S)\right]-\Sigma\left[W_{ij}(O)\right]\right)^2}{\chi_c(k')}-\mu_c^2(k')}$$

Nos informa de la dispersión del valor de los errores de acceso ala LUT cuando se realiza corrección en los niveles de gris de ventana.

Una vez computadas las anteriores cantidades, para diferentes valores de ρ, se representan gráficamente. Por ejemplo, podemos observar las obtenidas en las figuras 6, 7, 8, 9, 10 y 11. En ellas se estudia el comportamiento, en cuanto a corrección del acceso a la LUT, de las máscaras binaria y gris frente a una imagen corrompida con un 30% de ruido impulsivo *sal y pimienta*. En abscisas se tiene el nivel de gris k' de ventana, medido sobre la imagen original O. En ordenadas aparece, bien la media del error cometido, bien la desviación típica, o bien el tamaño de la muestra empleada en el cálculo de las cantidades estadísticas anteriores. Se han observado los casos de corrección del acceso a la LUT mediante máscara gris y máscara binaria;

y también el caso en que no existe tal corrección. En todas las representaciones se ha escogido el parámetro ρ, que define el umbral de selección, que mejor resultado proporciona. Se muestran resultados para imágenes de tres clases diferentes: radiografías, ecografías y textos mecanografiados; y empleando dos tamaños de ventana distintos: 3×3 y 5×5.

De estas gráficas podemos extraer la conclusión de que, para el ruido *sal y pimienta*, la corrección llevada a cabo por las máscaras es realmente excelente para un valor de $\rho = 0$ (excepto para textos, en que $\rho = 1$). Es decir, basta con considerar como ruidosos aquéllos píxeles cuyo nivel de gris sea muy poco probable —e incluso nula—. Esta decisión viene justificada por el uso implícito de un conocimiento previo acerca del ruido en la imagen. Se sabe que el ruido *sal y pimienta* tiene niveles extremos, que son realmente improbables en una imagen típica. Por esta razón se emplea un parámetro ρ tan pequeño, y con tan buenos resultados. Se observa también que la dispersión de los errores disminuye al emplear la máscara en su faceta correctora. Las representaciones del tamaño de muestra ayudan a comprender qué parte de los datos son más significativos. En el caso de los textos, la zona más clara de la imagen contribuye más significativamente a los resultados de media y desviación de los errores, debido a que poseen un fondo casi blanco. Para radiografías y ecografías, sucede lo contrario, dado que ambas son imágenes generalmente oscuras.

Además, las figuras numeradas de la 6 a la 11 sirven para corroborar un resultado deducido al comienzo del presente capítulo. En efecto, la recta que aparece en las gráficas, correspondiente al caso en que no existe corrección de la entrada a la LUT, ya había sido predicha cuando se introdujo el concepto de corrección debida al efecto centrípeto que tenía el ruido impulsivo sobre los niveles de gris de ventana, k', de la imagen:

$$k' = \frac{k_s' - p\frac{G'}{2}}{1-p} \Rightarrow k'_s - k' = p \cdot \left(\frac{G'}{2} - k'\right)$$

Como se ve, la pendiente de la recta es negativa y el error es nulo en la parte media de la escala de nivel de gris de ventana G'.

Se ha realizado también un detallado estudio de los efec-

tos correctores de la máscara sobre imágenes fuente afectadas por 20% de ruido impulsivo con niveles uniformemente distribuídos. Aparecen en las figuras numeradas de la 13 a la 24.. Se ha trabajado de nuevo con las tres clases de imágenes —radiografías, texto, ecografías—, y con los dos tamaños de ventana. En este caso, para comprender mejor la evolución del comportamiento de las máscaras al variar el umbral de selección se muestra, junto con el valor de ρ que mejores resultados proporciona, otro más elevado. Debe hacerse una observación: en estas gráficas en abscisas se coloca el nivel de gris de ventana, k', medido sobre la imagen fuente, S, y no sobre O. La razón para este sutil cambio de escala es meramente práctica: el nivel de gris de ventana que procesa el algoritmo es el nivel de gris de ventana medido sobre la imagen fuente, y no sobre la original. En el caso anterior, con ruido *sal y pimienta*, se realizó con la otra escala por comprobar los resultados obtenidos al comienzo del capítulo.

En todas las gráficas en las que se muestra la media del error cometido al acceder a la LUT se aprecia una curva que corresponde con el caso en el que no se realiza corrección de dicho error. Podemos comprender cualitativamente la forma de dicha curva. Fijémonos en la figura 12; en ella se ilustra el proceso que tiene lugar cuando se introduce ruido impulsivo en una imagen original. Obsérvese cómo algunos elementos de la escala de grises de ventana en la imagen fuente proviene de la escala de la original, por el efecto *centrípeto* del ruido, que se expuso al comienzo del capítulo. Si nos fijamos en la escala de grises de la imagen fuente, advertimos que sólo exixte error de acceso a la LUT en la zona recuadrada, correspondiente con la escala sobre la imagen original. El error máximo se comete en los niveles situados en los extremos del recuadro, con signo distinto. En el nivel de gris central, no hay error. Teniendo en cuenta que el efecto *centrípeto* citado no es en realidad tal y como se explicó, sino que tiene ciertas desviaciones respecto de ese comportamiento ideal, la comprensión de la forma de las curvas en cuestión es ahora más sencilla.

En cualquier caso, podemos ver que el comportamiento de las máscaras, en cuanto a corrección de la entrada de la LUT, es muy bueno, incluso cuando se introdue ruido impulsivo uniformemente distribuído. Los mejores resultados se obtienen para las radiografías, seguida de las ecografías y de los textos. Este hecho demuestra, desde otra perspectiva, la influencia de la ca-

lidad del modelo en la calidad de la máscara, puesto que el modelo más cohesionado es el de radiografías, y el menos el de texto. También se comprueba esta dependencia entre calidades modelo-máscara en la escasa diferencia en los comportamientos de la máscara en el caso de radiografías, cuando el parámetro ρ varía. En efecto, el modelo de radiografías está muy cohesionado, poco disperso, lo cual significa que cada nivel de gris de ventana k' está compuesto de pocos valores diferentes de niveles de gris de píxel, k. Por tanto, la máscara detectará perfectamente los pulsos del ruido, puesto que es muy fácil que resalten de sus vecinos. De este modo, casi cualquier umbral de selección será suficientemente bueno como para conseguir una aceptable corrección del nivel de gris de acceso a la LUT.

El valor óptimo del parámetro ρ es algo mayor que para el caso de ruido *sal y pimienta*. Efectivamente, el ruido uniforme presenta pulsos de niveles de gris no necesariamente extremos, por lo que será necesario tomarlo un poco mayor que en el caso precedente, para así permitir que la detección de pulsos no sea para valores tan restringidos como antes (sólo blanco y negro). En otro orden de cosas, podemos apreciar que las máscaras construidas a partir de vantanas mayores suelen ser algo mejores que las obtenidas a partir de ventanas pequeñas, en concordancia con la similitud, ya establecida anteriormente, con el sistema perceptual humano, que detecta las distorsiones mejor si se amplia algo el campo de visión.

Cabe reseñarse que la mayor fiabilidad de resultados se deduce de las gráficas de tamaño de muestra. Cuanto mayor sea ésta, más han sido los elementos que han entradoa formar parde de la estadística que dió lugar a los valores de media de error o de desviación típica que se representaron. De nuevo, las zonas claras en los textos y las oscuras en radiografías y ecografías son las que tienen mayor representatividad.

A la vista de los resultados, la máscara gris se comporta siempre tan bien como la binaria. Sólo a veces lo hace un poco mejor: cuando el valor del parámetro ρ es elevado, la máscara de niveles de gris es más robusta. Sin embargo, para los valores óptimos de r apenas existe diferencia entre ambos tipos de máscaras. Quizá la mayor separación de comportamientos se observe en la desviación típica del error, dónde la máscara gris consigue, como era de esperar, una menor dispersión. En definitiva, podemos justificar el desuso de la máscara gris en lo sucesivo,

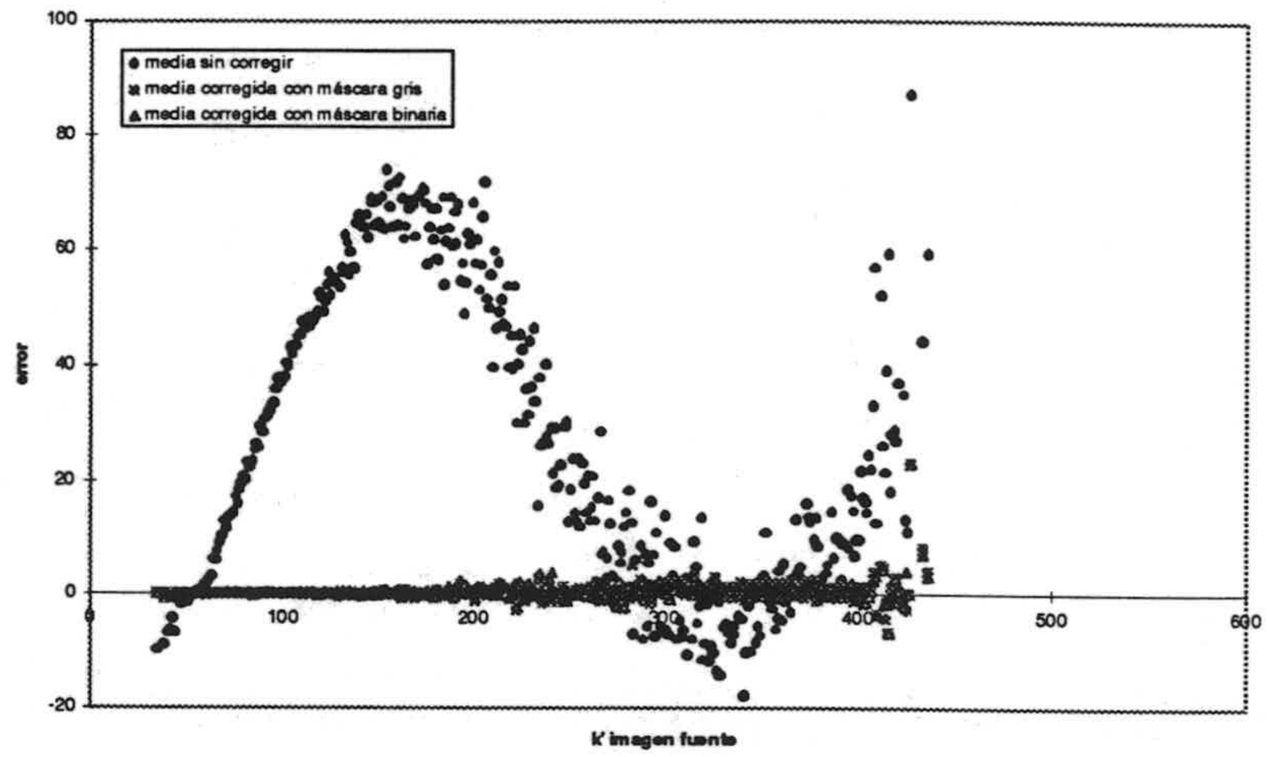


Figura 13a: Media del error de acceso a la LUT (diferencia entre el k' original y el observado en la imagen fuente) para el caso de una imagen de la clase de ecografías, contaminada por un 20% de ruido impulsivo uniformemente distribuído. El tamaño de la ventana utilizado es 3×3, y el parámetro $\rho = 10$. Se observa cómo tanto la máscara binaria como la gris corrigen casi totalmente el error en el nivel de gris de ventana introducido por el ruido. En abscisas se representa el nivel de gris de ventana medido en la imagen *fuente*; en ordenadas, las medias del error, con y sin corrección, $\mu_{s.c.}$ y $\mu_{c.}$.

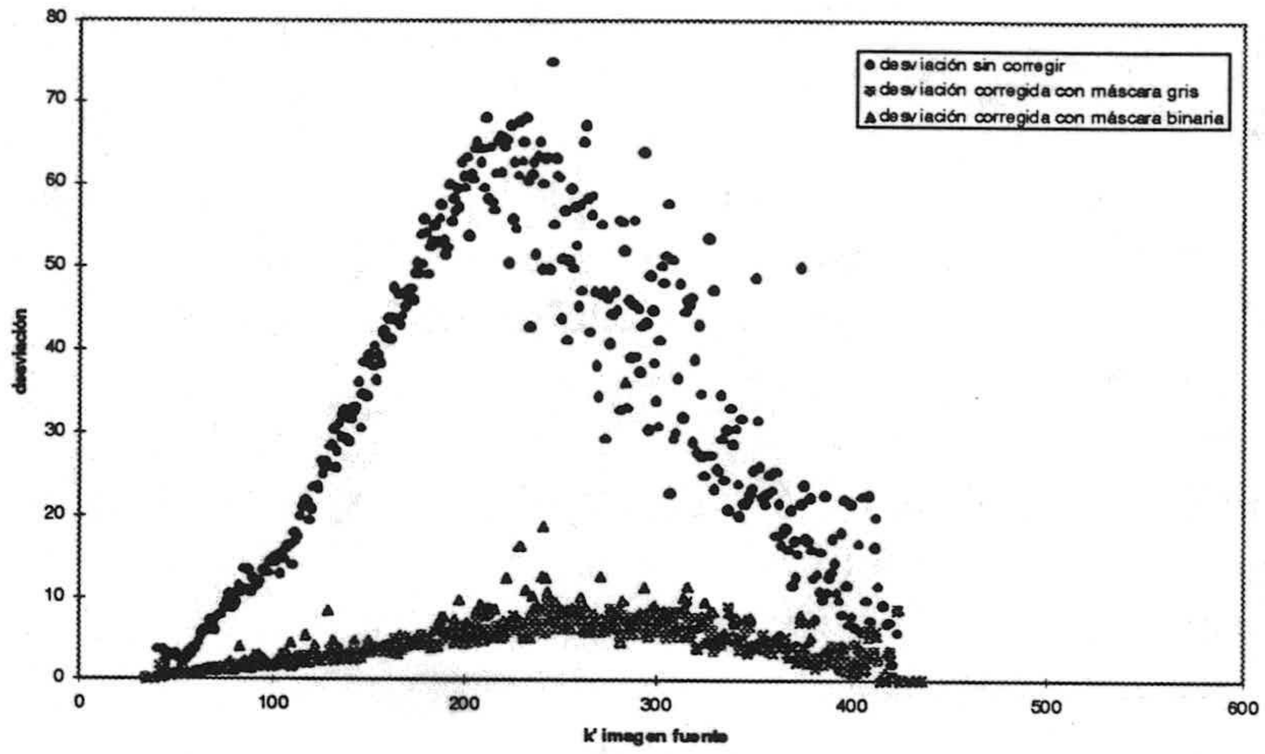


Figura 13b: Desviación típica de los errores de acceso a la LUT, con y sin corrección ($\sigma_{s.c.}$ y $\sigma_{c.}$). En abscisas, el nivel de gris de ventana k' medido sobre la imagen fuente S. El caso representado es el mismo que en la gráfica precedente. Se aprecia cómo tanto la máscara de estimación de ruido binaria como la gris disminuyen la dispersión del error.

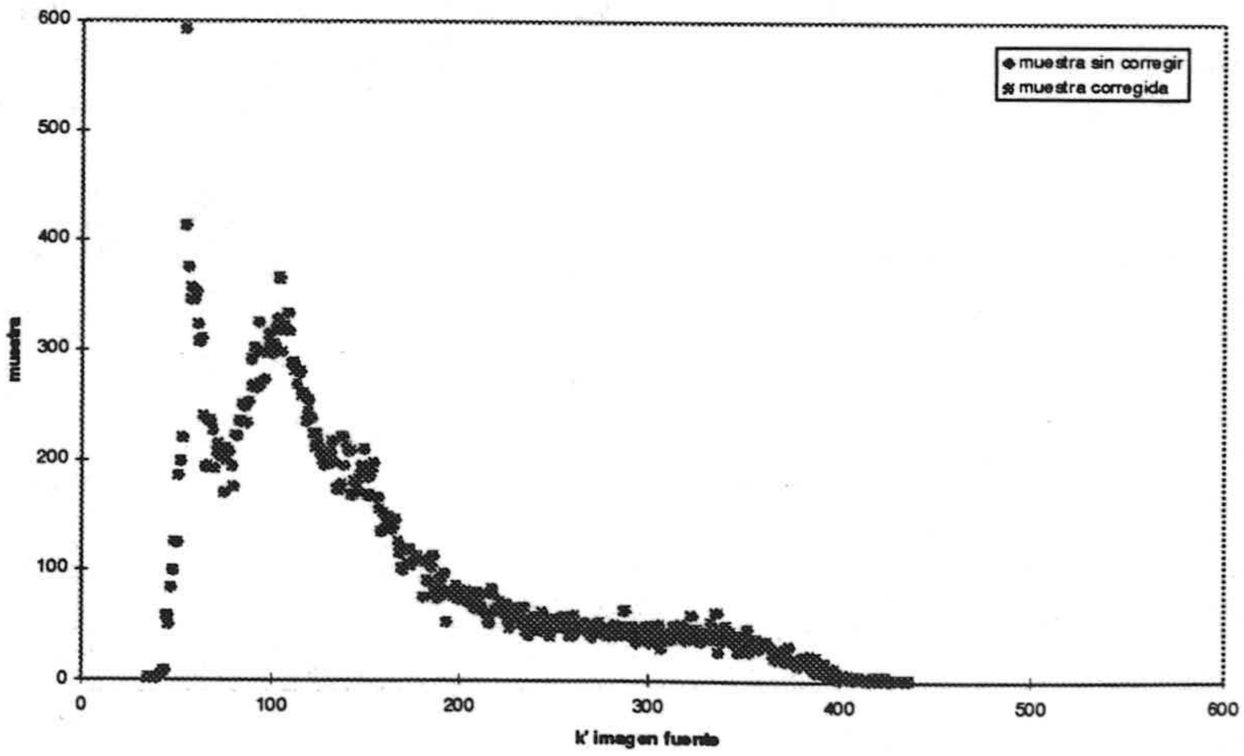


Figura 14c: Tamaño de la muestra empleada al hacer las medidas de error para el caso indicado de imagen de ecografías con 20% de ruido impusivo uniformemente distribuído; $\rho = 10$ y ventana 3×3. Se observa la coincidencia de las muestras para la estadística del caso en que existe corrección como en el caso contrario. El mayor número de medidas, y por tanto la mayor significación de los resultados, ocurre para niveles oscuros de k' medida sobre la imagen fuente.

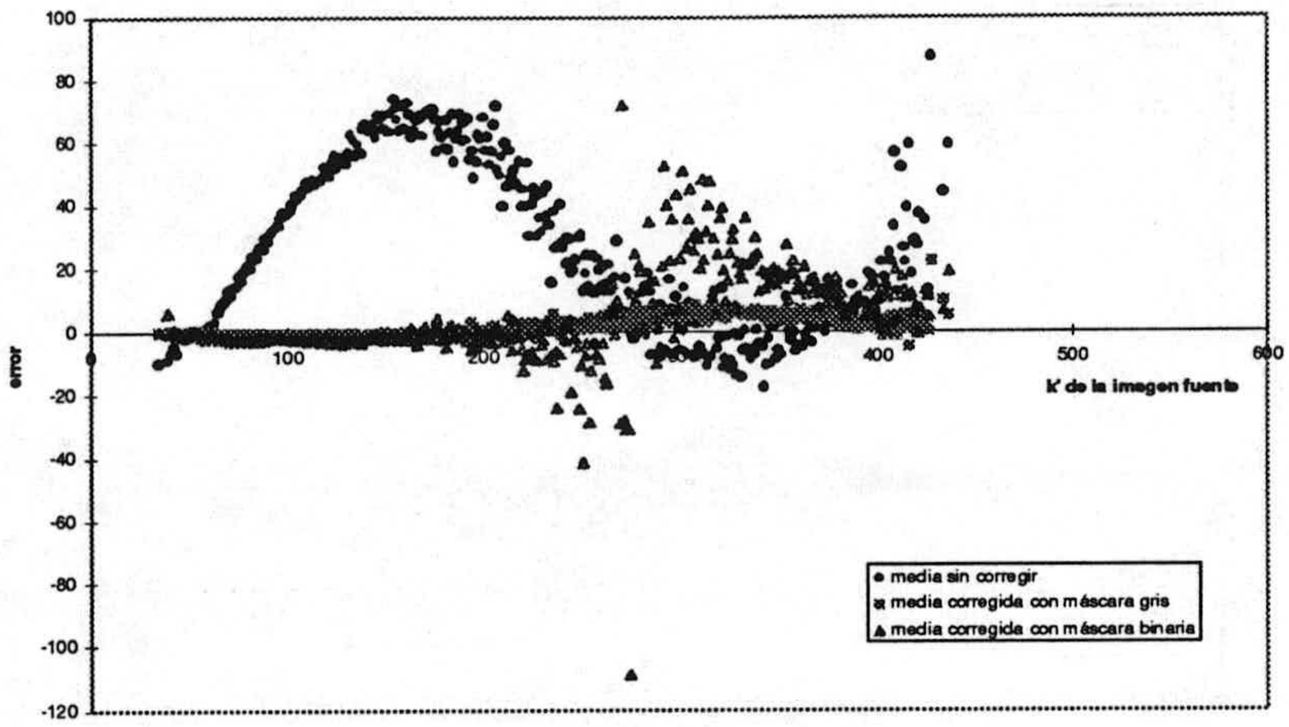


Figura 14a: Media del error cometido al acceder a la columna de la LUT, para una imagen de la lcase de ecografías afectada por 20% de ruido impulsivo uniformemente distribuído, con un parámetro definidor del umbral de selección ρ = 100. El tamaño de la ventana utilizada es 3×3. En abscisas, las *k'* sobre la imagen fuente. Se aprecia cómo la correeccción con máscara gris es más robusta que la binaria al aumentar el parámetro de selección.

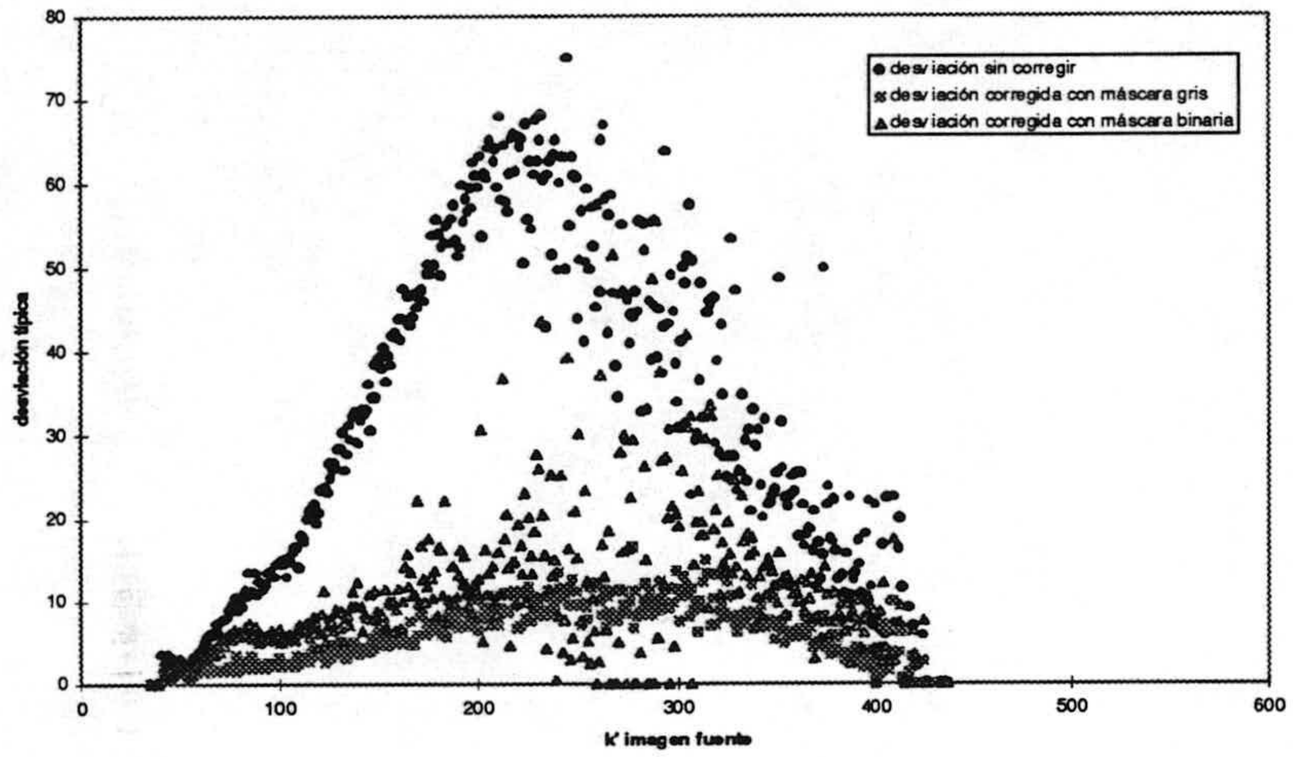


Figura 14b: Desviación típica de los errores anteriores, para el mismo caso. De nuevo los mejores resultados son los obtenidos con la máscara de niveles de gris. Sin embargo, recordemos que eligiendo adecuadamente el parámetro ρ (como en el caso de la página precedente) la máscara binaria es tan buena como la gris.

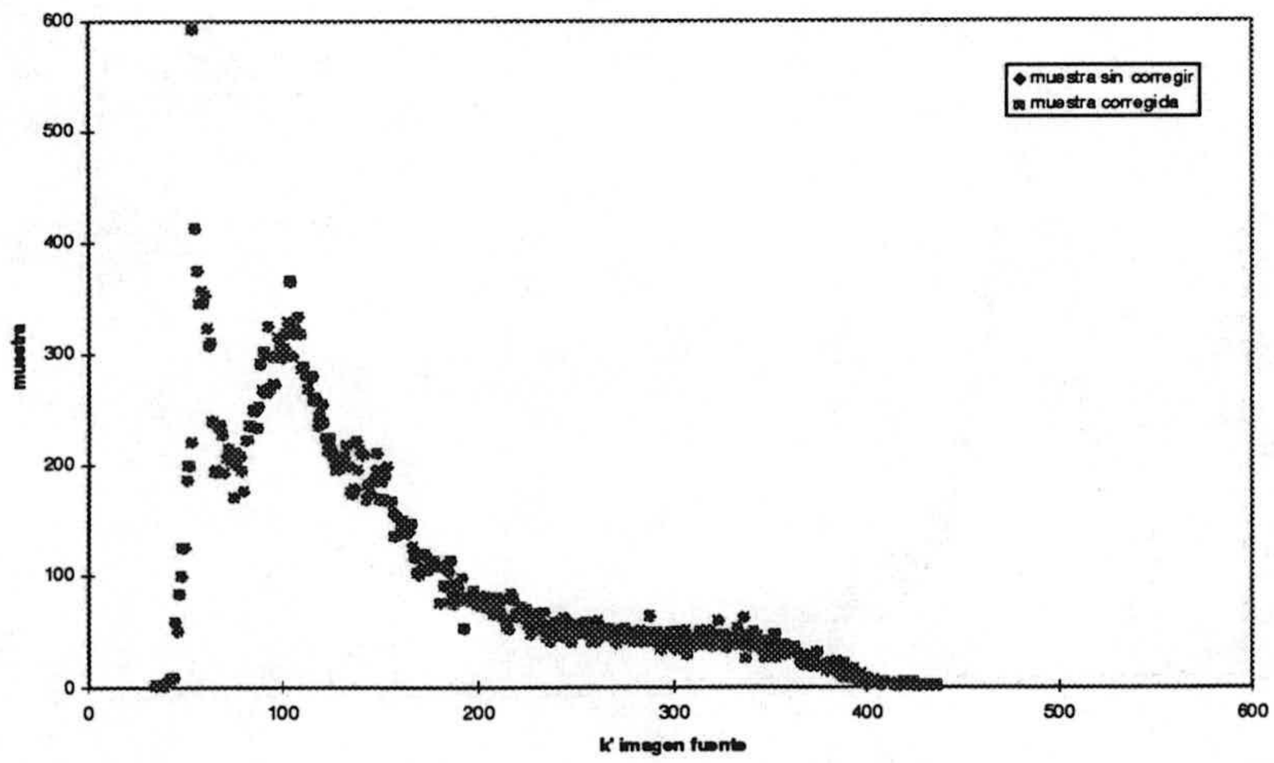


Figura 14c: Tamaño de la muestra al obtener las estadísticsa que aparecen en las anteriores gráficas. Se aprecia cómo los resultados son más significativos se encuentran en la zona oscura de la imagen. La muesta para el caso corregido y sin corregir es coincidente.

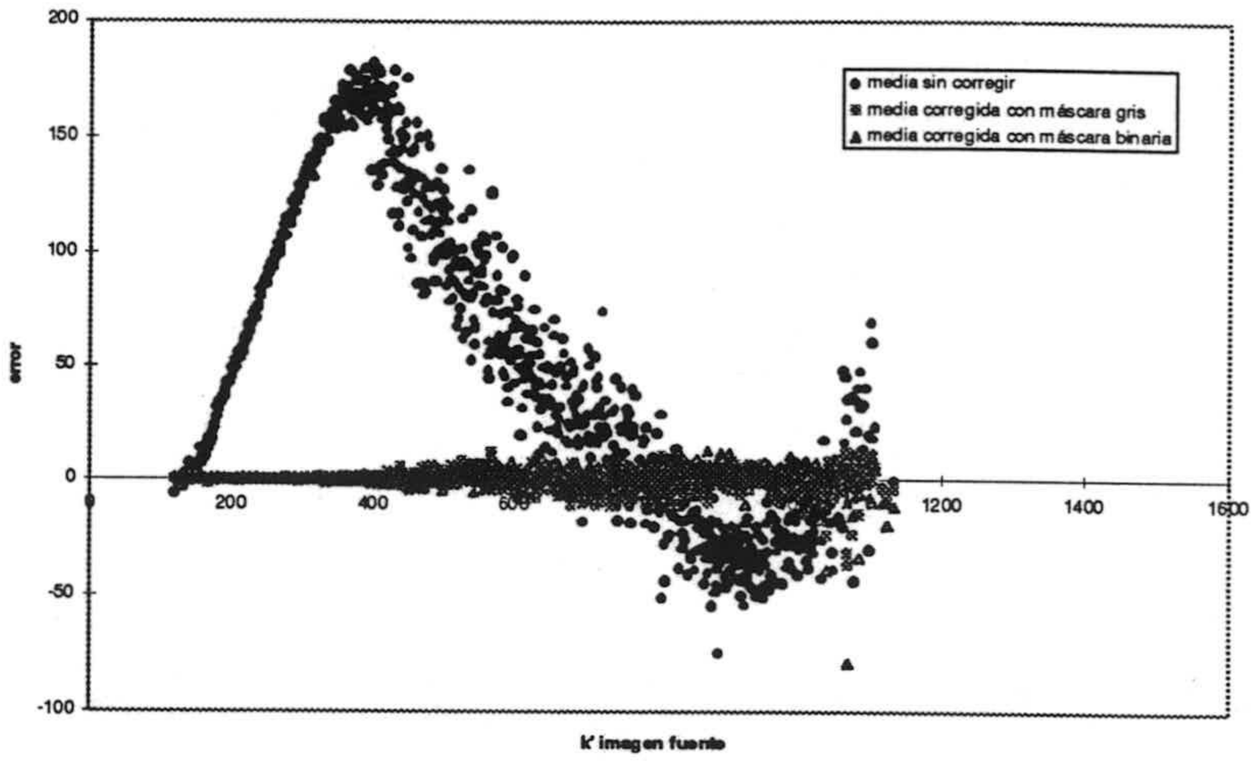


Figura 15a: Gráfica en al que aparece representada la media del error de acceso a la LUT frente a los k' calculados sobre la imagen fuente S. Se muestran tanto el caso en el que existe corrección como el que no ($\mu_{s.c.}$ y $\mu_{c.}$). La imagen es de la clase de ecografías, y el tamaño de la ventana es 5×5. El parámetro que define el umbral de selección vale $\rho = 10$. La corrección es excelente, tanto con máscara binaria como con máscara gris.

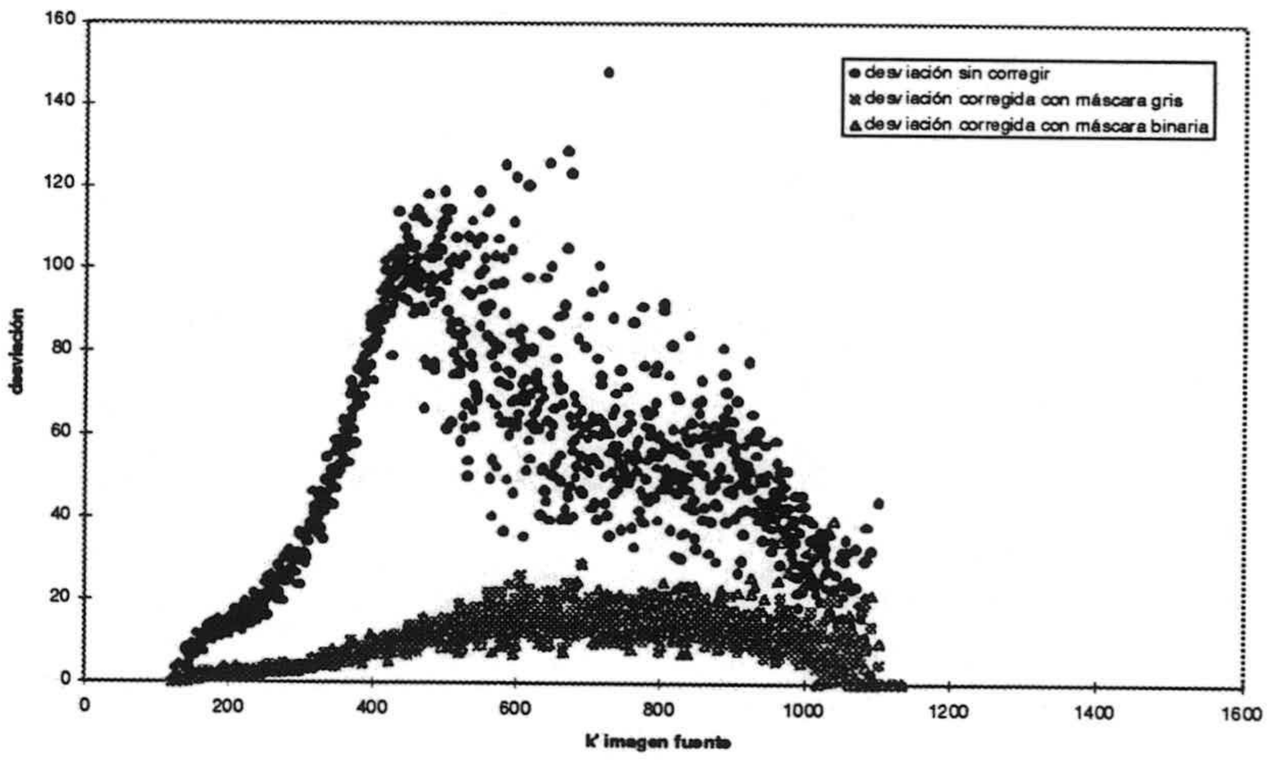


Figura 15b: Se representa la ´desviación típica de los errores de acceso a la LUT, con y sin corrección ($\sigma_{s.c.}$ y $\sigma_{c.}$). El caso mostrado es el mismo que el de la gráfica anterior. En abscisas, el nivel de gris de ventana k' medido sobre la imagen fuente S. Se aprecia cómo tanto la máscara de estimación de ruido binaria como la gris disminuyen la dispersión del error.

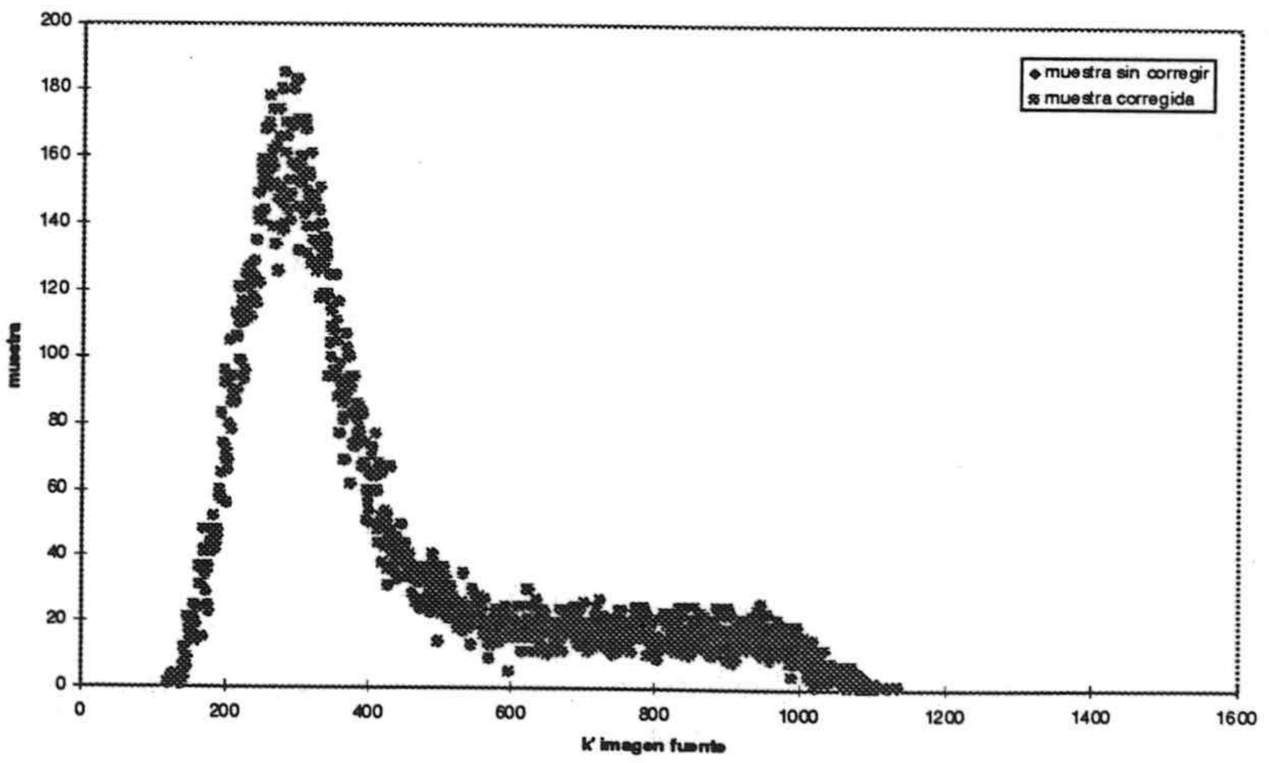


Figura 15c: El tamaño de la muestra que se empleó para la estadística mostrada en las gráficas precedentes, se plasma aquí en función del nivel k' medido sobre la imagen fuente S. Se aprecia tanto la coincidencia de casos cuando hay o no corrección , como la máxima representatividad de la xona oscura de la imagen fuente.

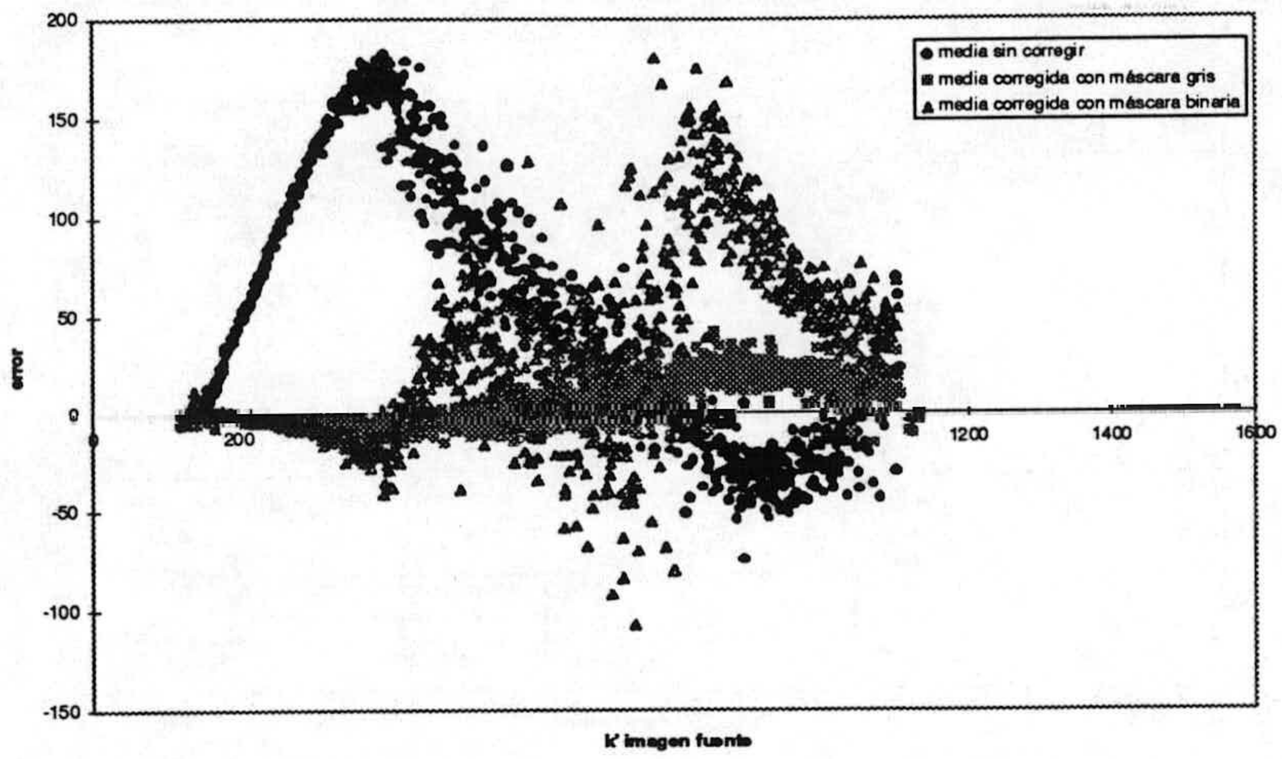


Figura 16a: Media del error cometido en el acceso a la columna de la LUT, para una imagen de la clase de ecografías con 20% de ruido impulsivo uniforme. El umbral de selección está definido por $\rho = 100$. El tamaño de la ventana utilizada es 5×5. En abscisas, las k' sobre la imagen fuente. Se aprecia cómo la correccción con máscara gris es más robusta que la binaria al aumentar el parámetro de selección. Con máscara binaria se producen errores notables debido a l sesgo introducido en la LUT.

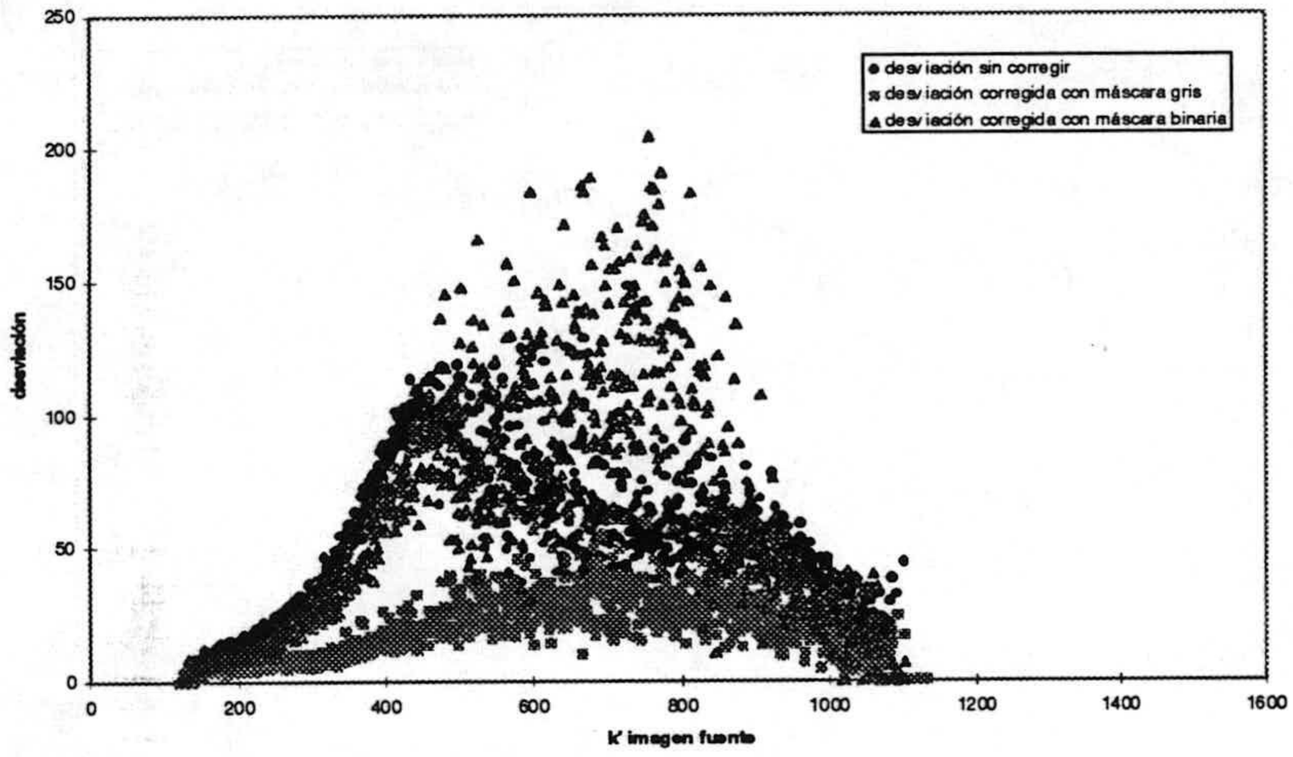


Figura 16b: Desviación típica de los errores anteriores, para el mismo caso. De nuevo los mejores resultados son los obtenidos con la máscara de niveles de gris.

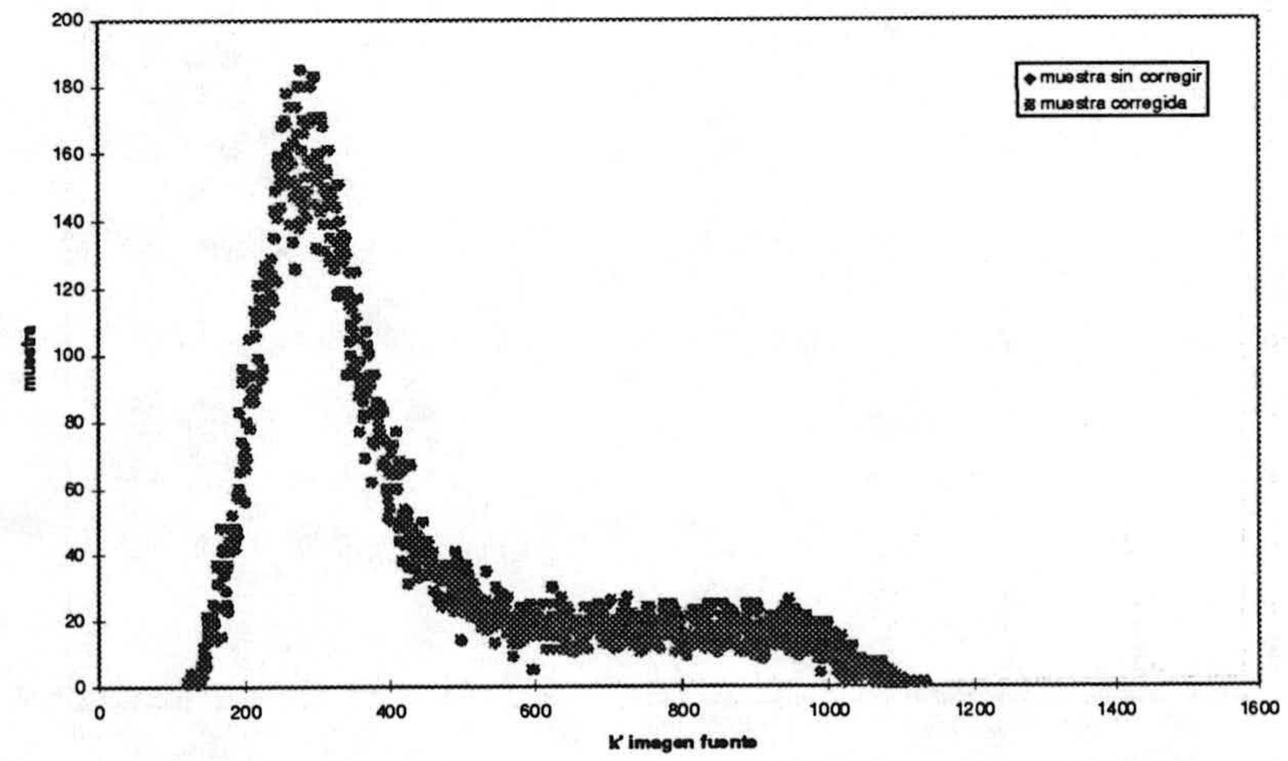


Figura 16c: Tamaño de la muestra al obtener las estadísticsa que aparecen en las anteriores gráficas. Se aprecia cómo los resultados son más significativos se encuentran en la zona oscura de la imagen. La muesta para el caso corregido y sin corregir es coincidente.

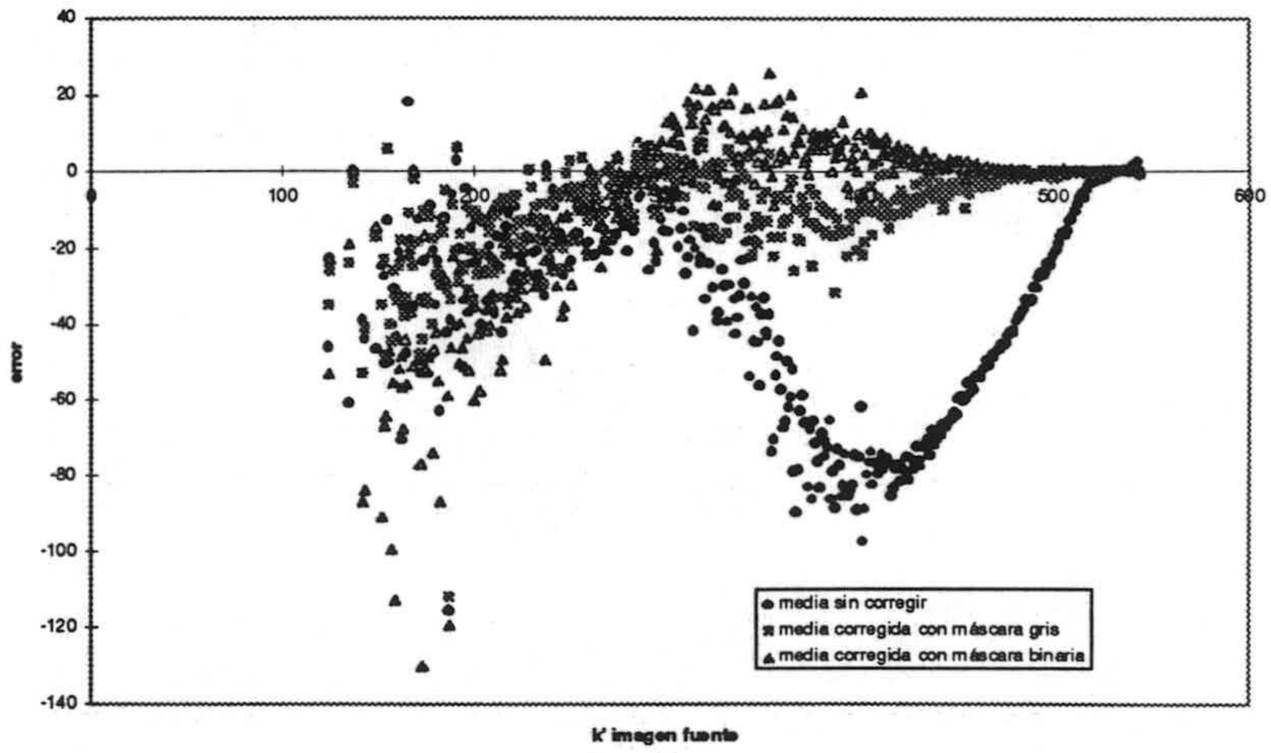


Figura 17a: Se representa la media del error cometido al acceder a la LUT, para una imagen de la clase de texto mecanografiado, con 20% de ruido impulsivo uniforme, ventana de tamaño 3×3 y umbral de selección determinado por el parámetro ρ = 10. Se aprecia una cierta mejora al realizar al corrección, si bien para las regiones claras se producen anomalías debidas a la falta de cohesión del modelo. Allá donde la muestra es más significativa (*cf.* apartado c), el error es menor. La máscara gris no introduce mejors sustanciales a la binaria.

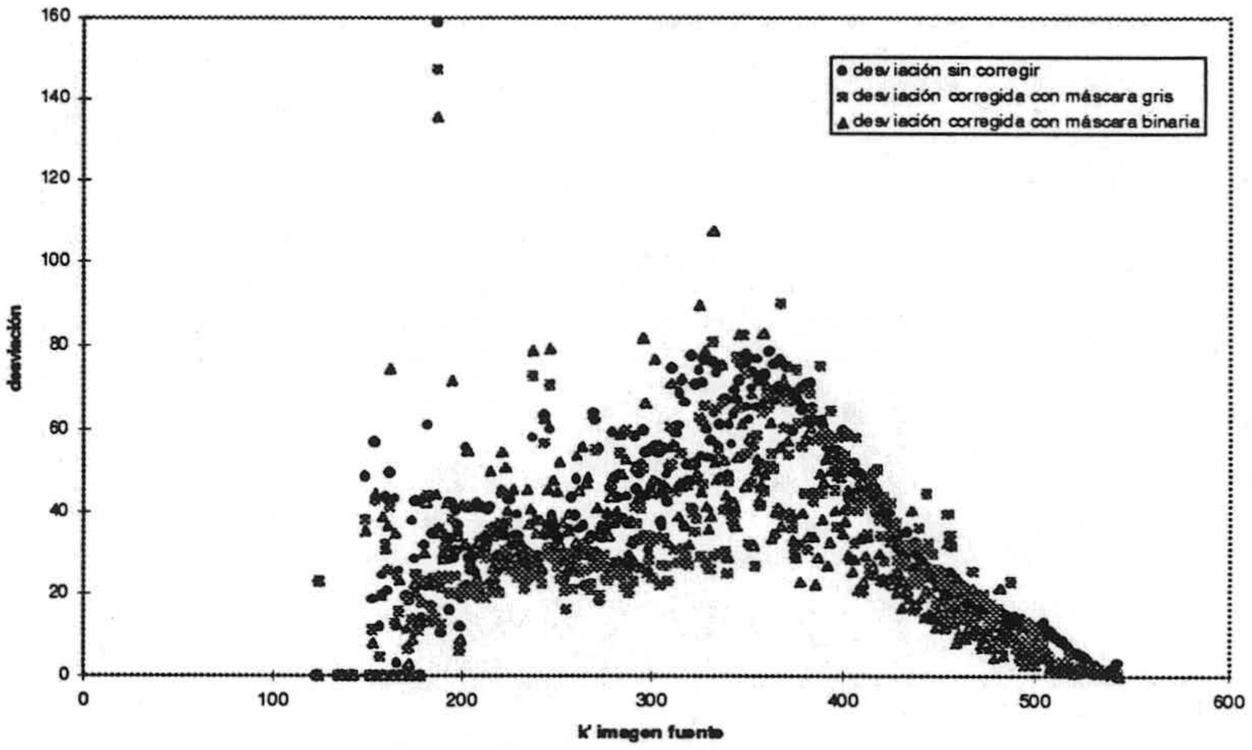


Figura 17b: Se muestran las desviaciones típicas de los errores de acceso a LUT para el mismo caso que en *(a)*. No hay mejora en la dispersión de los errores por el hecho de emplear corrección mediante máscara, sea gris o binaria, dada la escasa calidad del modelo.

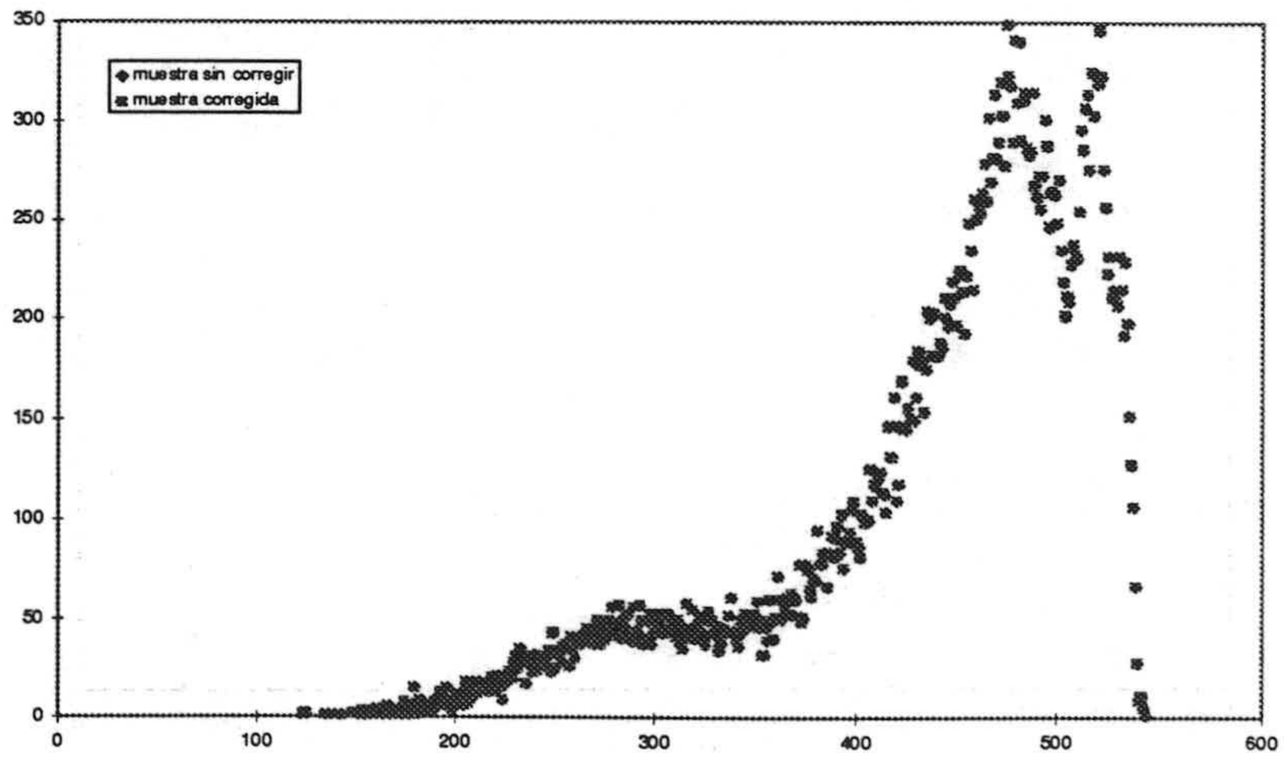


Figura 17c: La mayor concentración de muestra empleada en la realización de las estadísticas plasmadas anteriormente se localiza en la región clara de la imagen. La muestra concorrección y sin corrección es la misma.

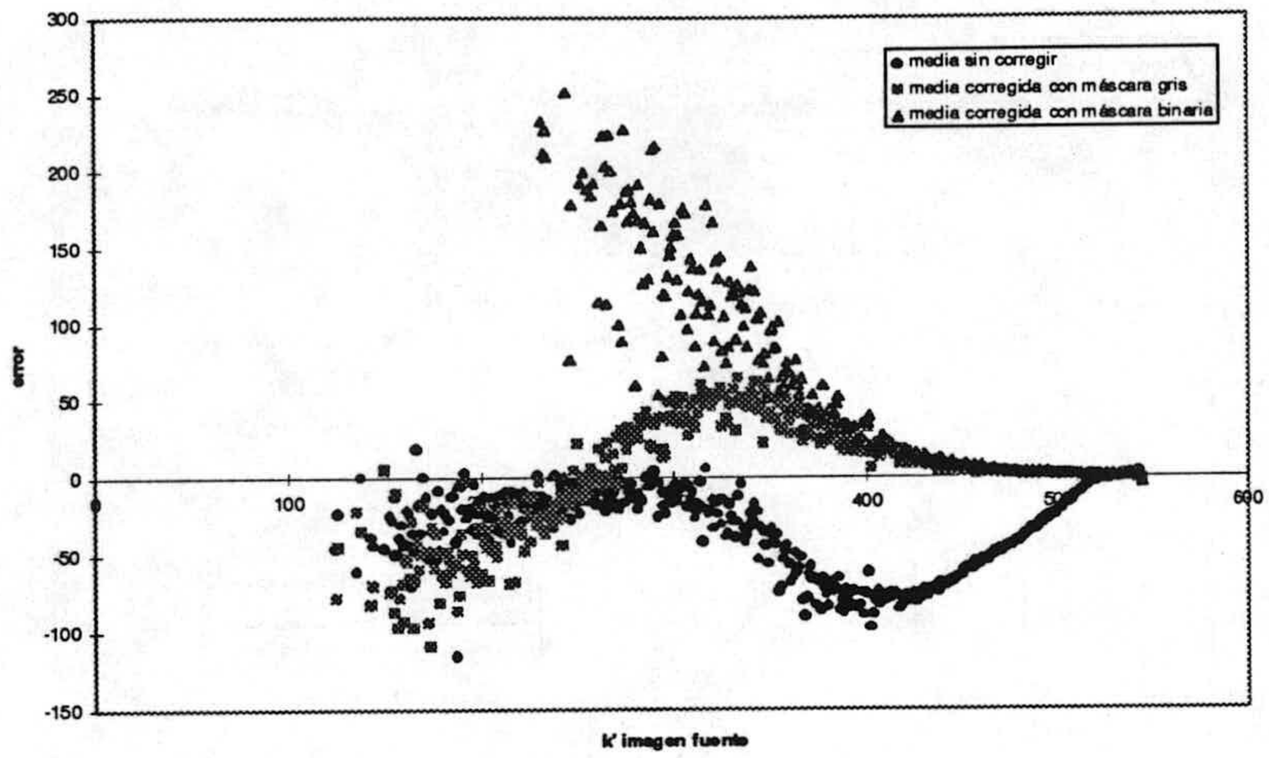


Figura 18a: Gráfica mostrando la media del error de acceso a la LUT para el caso de una imagen de la clase de texto mecanografiado, afectada de un 20% de ruido impulsivo uniforme. El umbral de selección se ha determinado mediante $\rho = 75$; y el tamaño de la ventana en 3×3. Se aprecia cómo la máscara gris arregla levemente el comportamiento de la imagen sin corregir; mientras la máscara binaria tiene un comportamiento anómalo, achacable a la falta de cohesión del modelo que provoca un sesgo notable en la LUT.

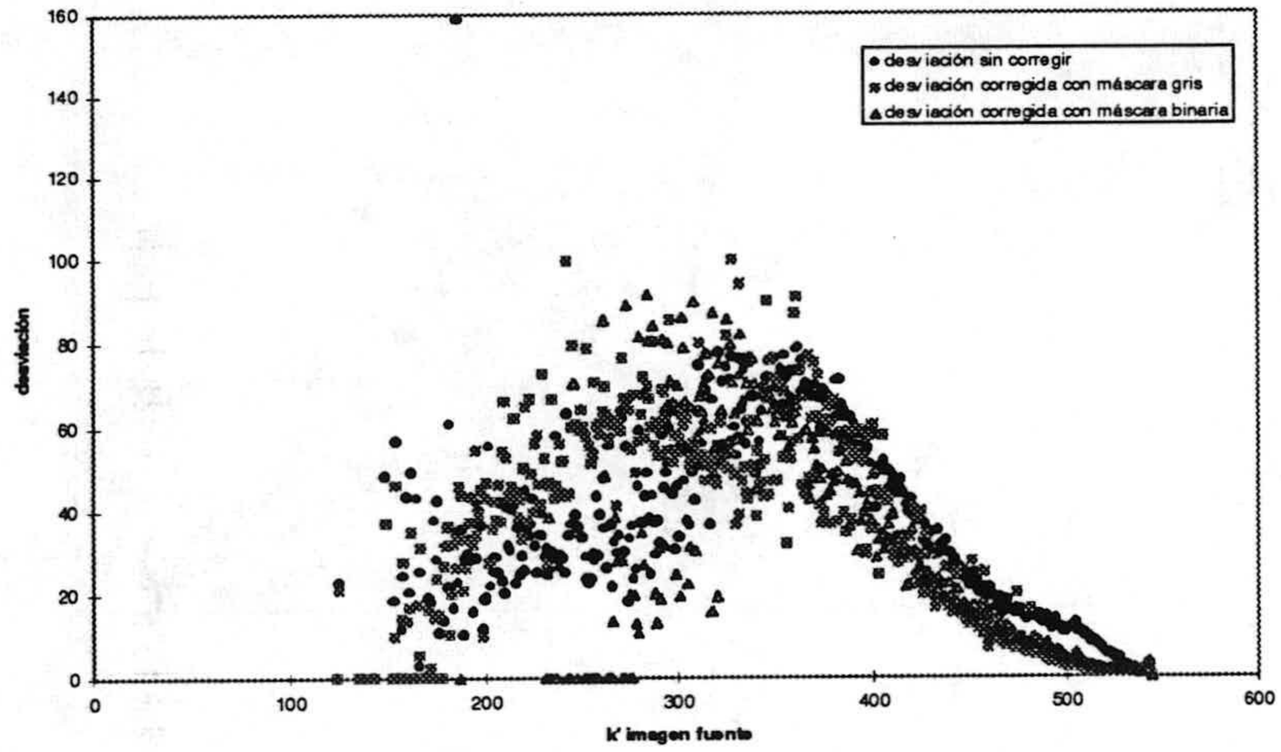


Figura 18b: Aparecen las desviaciones típicas de los errores anteriormente citados. No se observa gran mejoría en la reducción de la dispersión por el hecho de emplear máscaras.

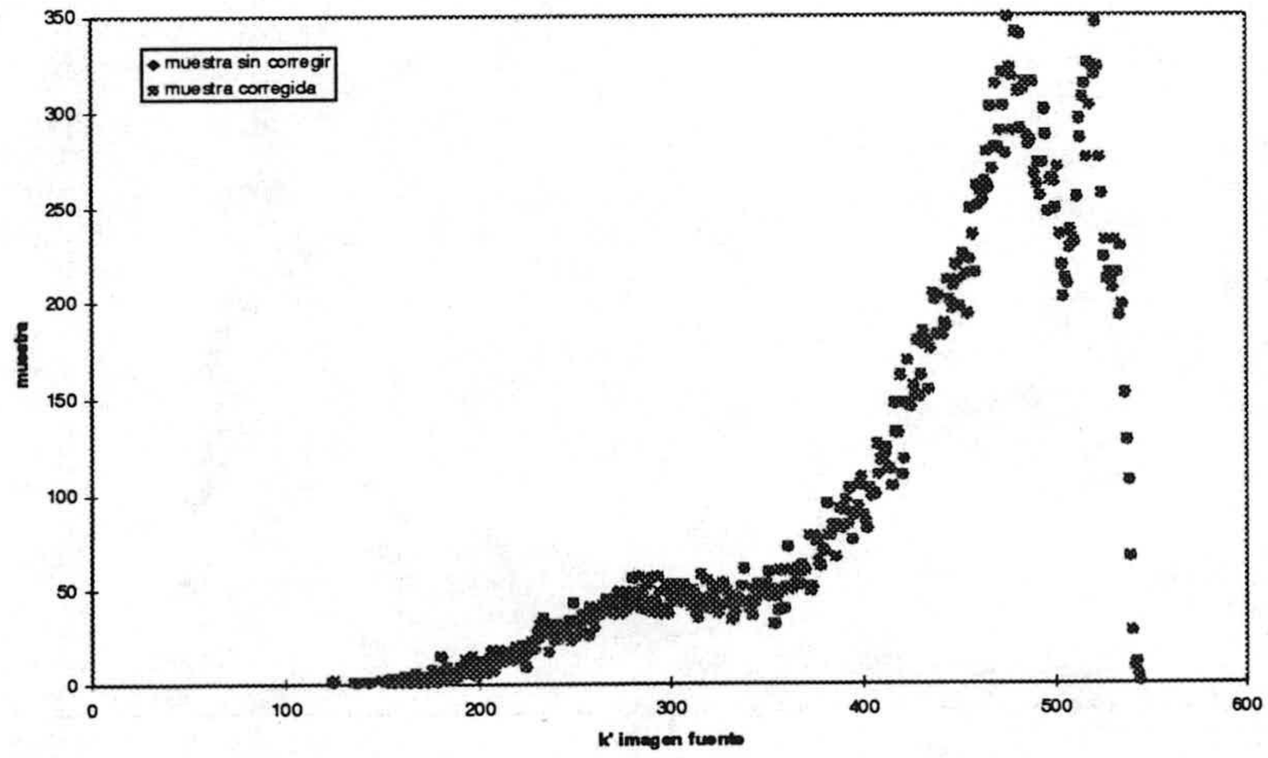


Figura 18c: Se muestran los tamaños de muestra empleados en las estadísticas. La mayor representatividad se produce en las zonas claras de la imagen. No hay diferencia entre el caso corregido y sin corregir.

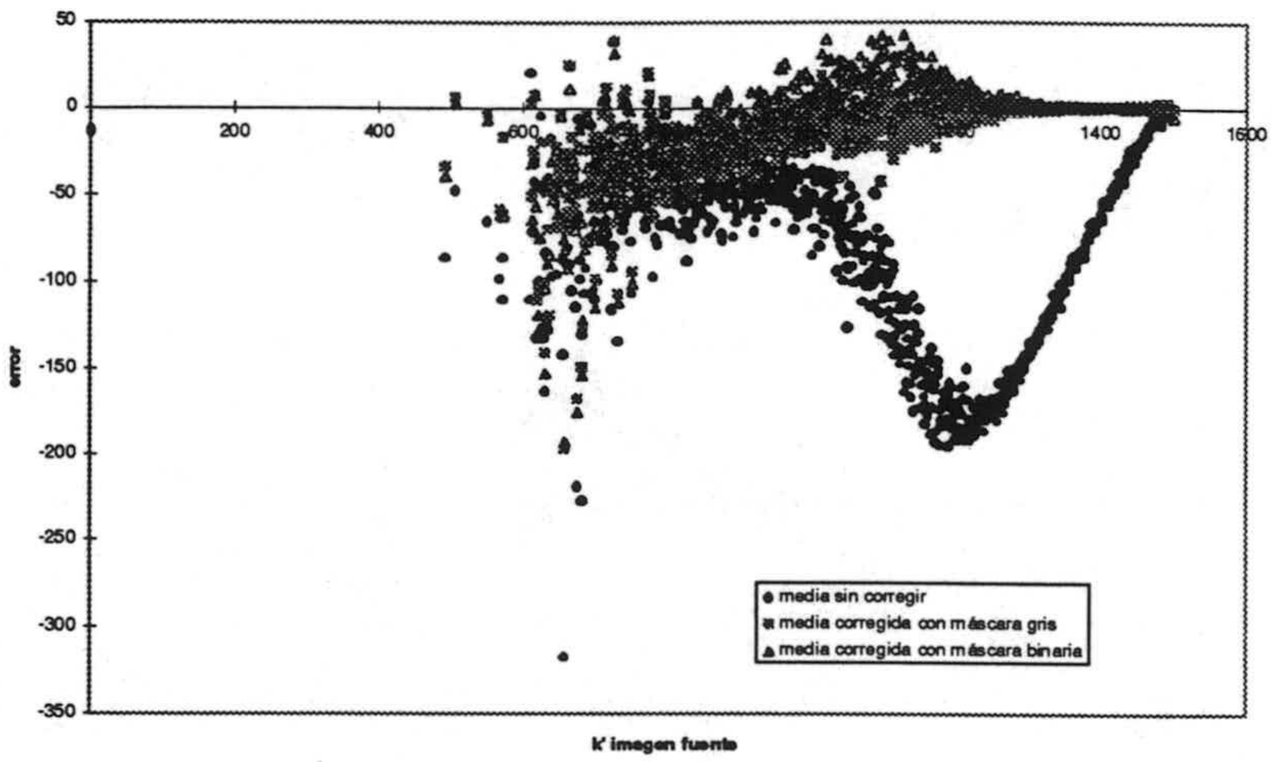


Figura 19a: Para cada k' sobre la imagen fuente se representa la media del error de acceso a la LUT para el caso de una imagen de la calse de texto mecanografiado, ventana de tamaño 3×3 y umbral de selección definido por el parámetro $\rho = 10$. Se observa cómo el empleo de máscara consigue corregir el error de acceso, pero no tan bien como en el caso de las ecografías o radiografías, debido a la dispersión interna del modelo. No se aprecian ventajas notables de la máscara gris frente a la binaria.

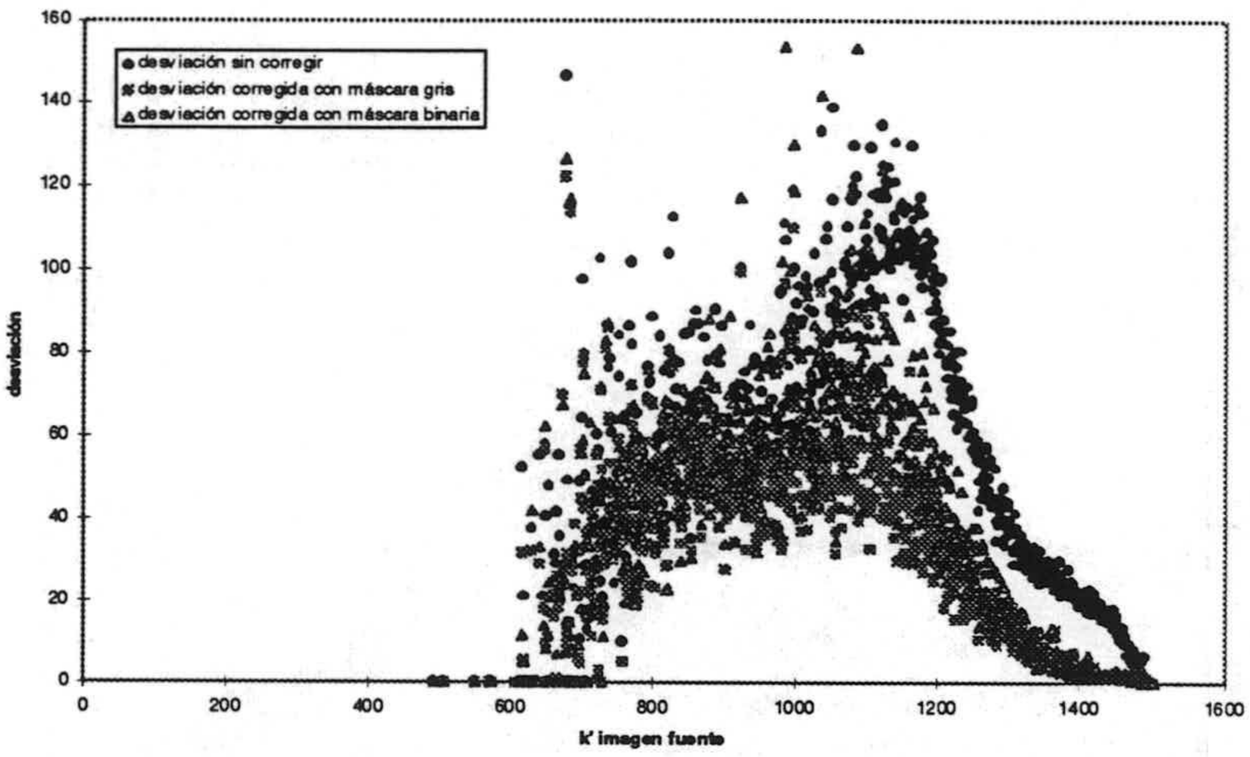


Figura 19b: Para el mismo caso de la gráfica precedente, se muestran aquí las desviaciónes típicas del error de acceso a LUT. El uso de máscaras no disminuye la dispersión de los mismos, como se desprende lde la gráfica.

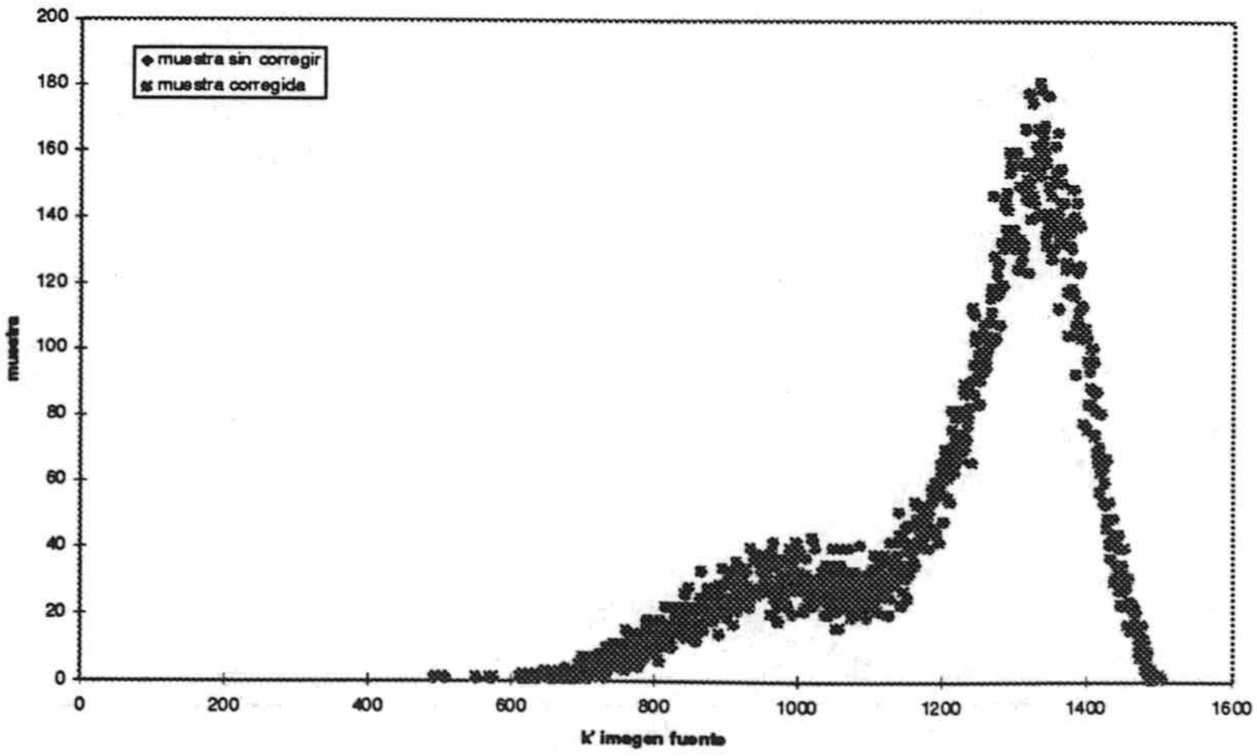


Figura 19c: Muestra que se ha utilizado para calcular las estadísticas plasmadas en las anteriores gráficas. Se observa la coincidencia de número de errores para el caso en que existe y no existe corrección. La mayor representatividad se alcanza en las regiones claras de la imagen.

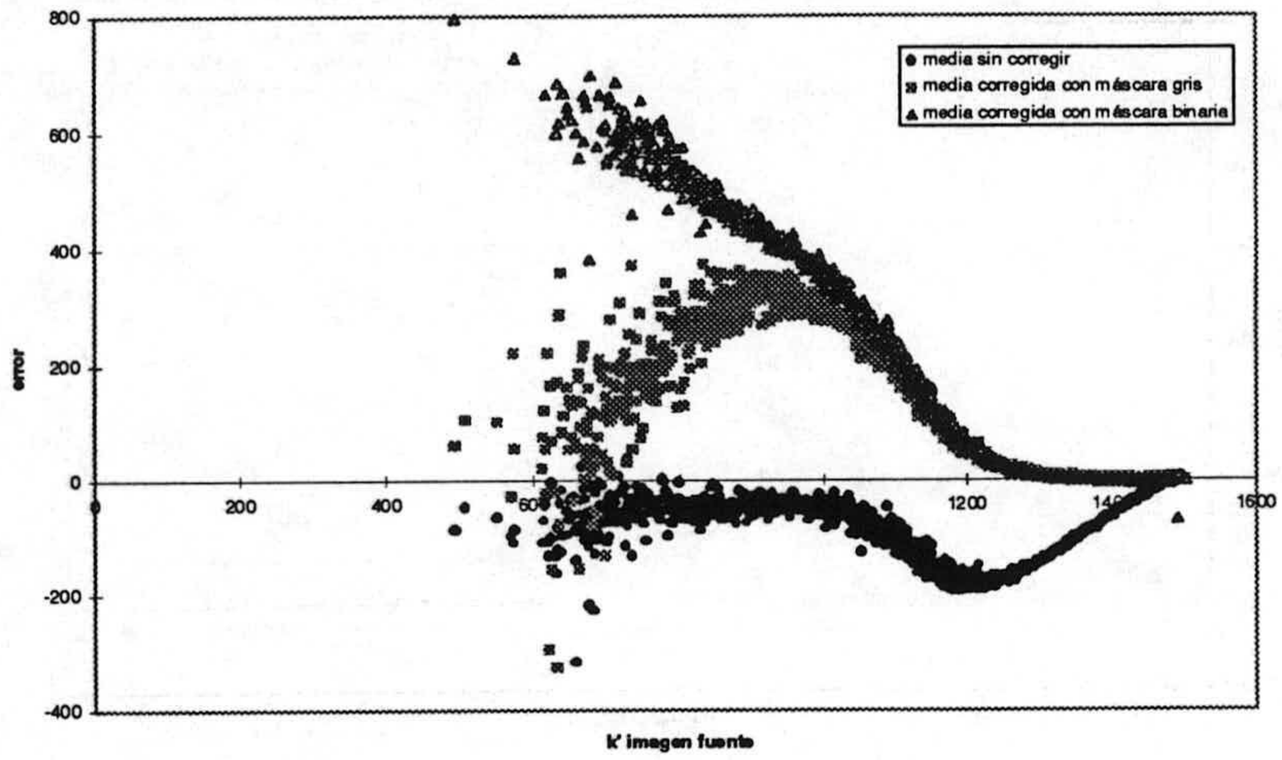


Figura 20a: Media del error cometido en el acceso a la columna de la LUT, para una imagen de la clase de texto mecanografiado con 20% de ruido impulsivo uniforme. El umbral de selección está definido por $\rho = 75$. El tamaño de la ventana utilizada es 5×5. En abscisas, las k' sobre la imagen fuente. Se aprecia cómo con máscara binaria se producen errores notables debido a l sesgo introducido en la LUT. La correccción con máscara gris es levemente más robusta que la binaria al aumentar el parámetro de selección. El mal comportamiento de ambas máscaras es achacable a la poca cohesión del modelo.

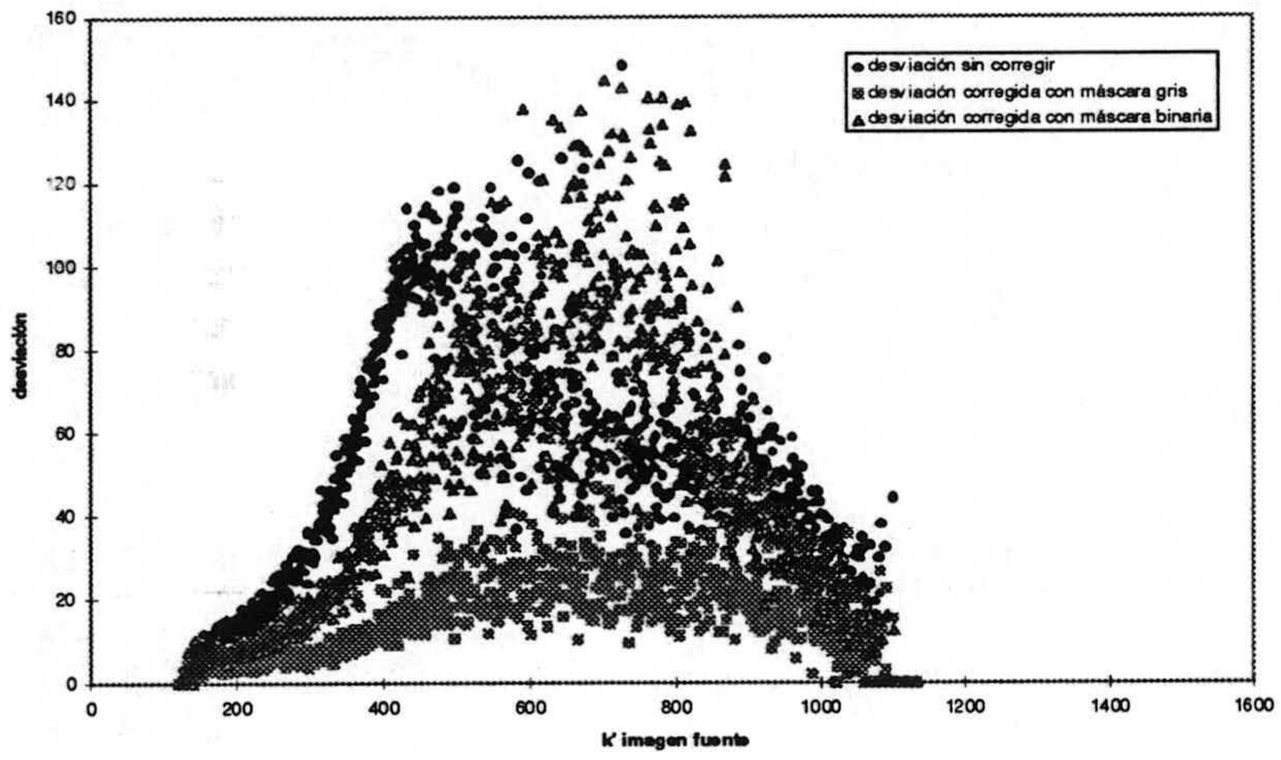


Figura 20b: Desviación típica de los errores anteriores, para el mismo caso. De nuevo los mejores resultados son los obtenidos con la máscara de niveles de gris.

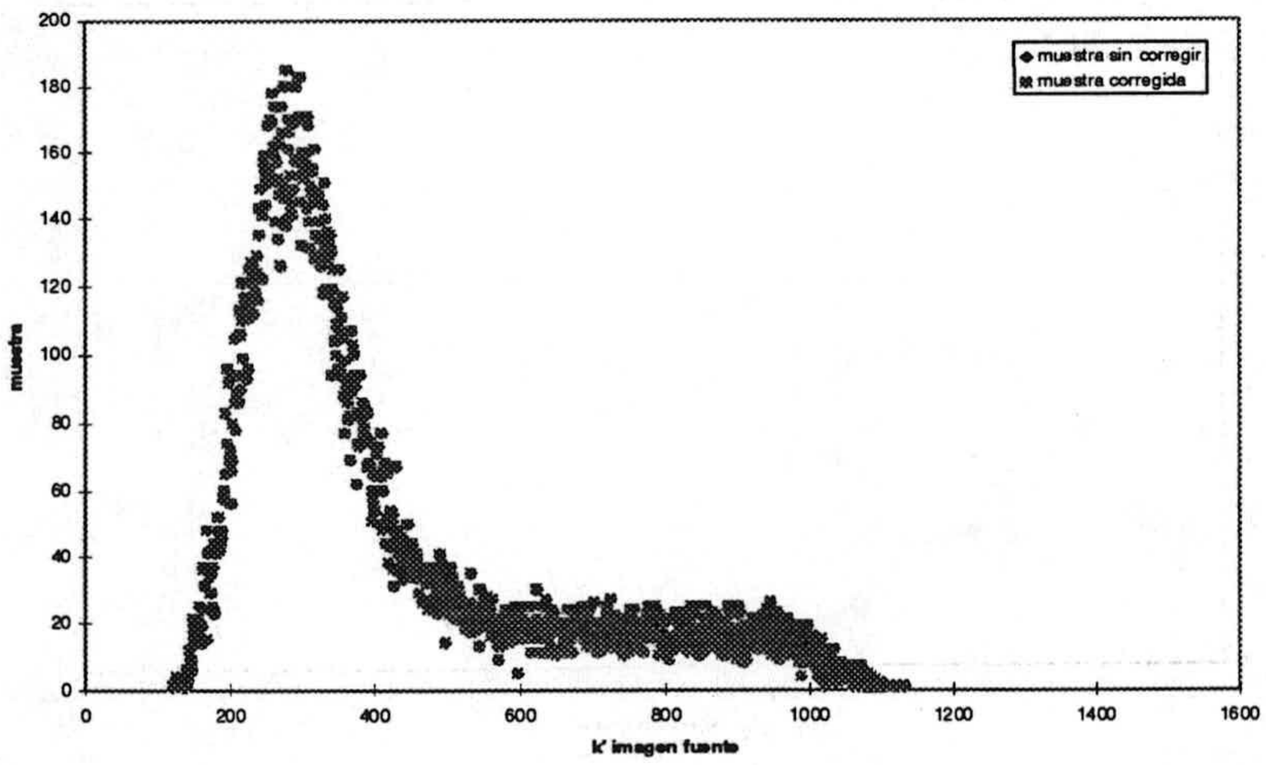


Figura 20c: Tamaño de la muestra al obtener las estadísticsa que aparecen en las anteriores gráficas. Se aprecia cómo los resultados son más significativos se encuentran en la zona oscura de la imagen. La muesta para el caso corregido y sin corregir es coincidente.

Figura 21a: Para una imagen de la clase de radiografías, con un 20% de ruido impulsivo uniformemnte distribuído y ventana 3×3 se representa la media de los errores de accseso a la LUT. El umbral de selección se determina mediante el parámetro $\rho = 1$. La corrección llevada a cabo por las máscaras es inmejorable. No se aprecia disferencia entre la máscaar gris y la binaria.

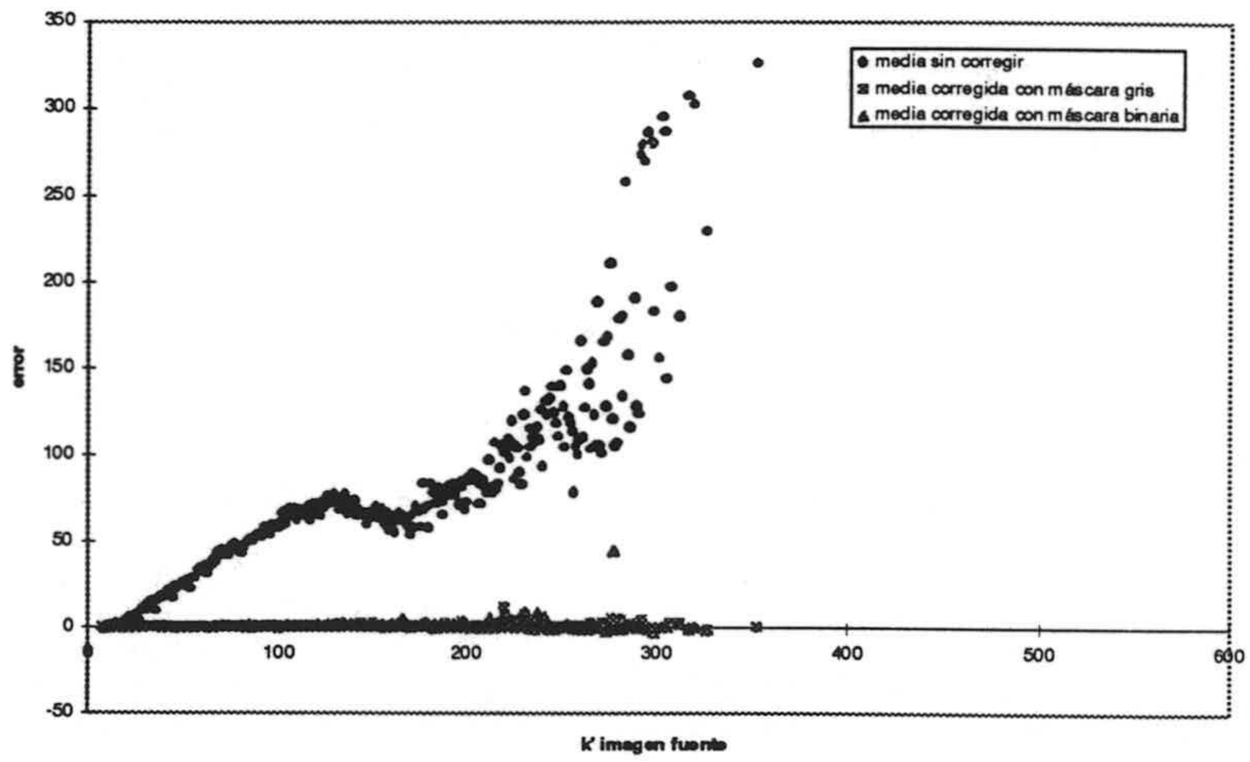


Figura 21b: Dispersión de los errores ya citados, medidos mediante la desviación típica. Aunque el mejor resultado lo proporciona la máscara gris, la binaria produce una aceptable disminución de la dispersión.

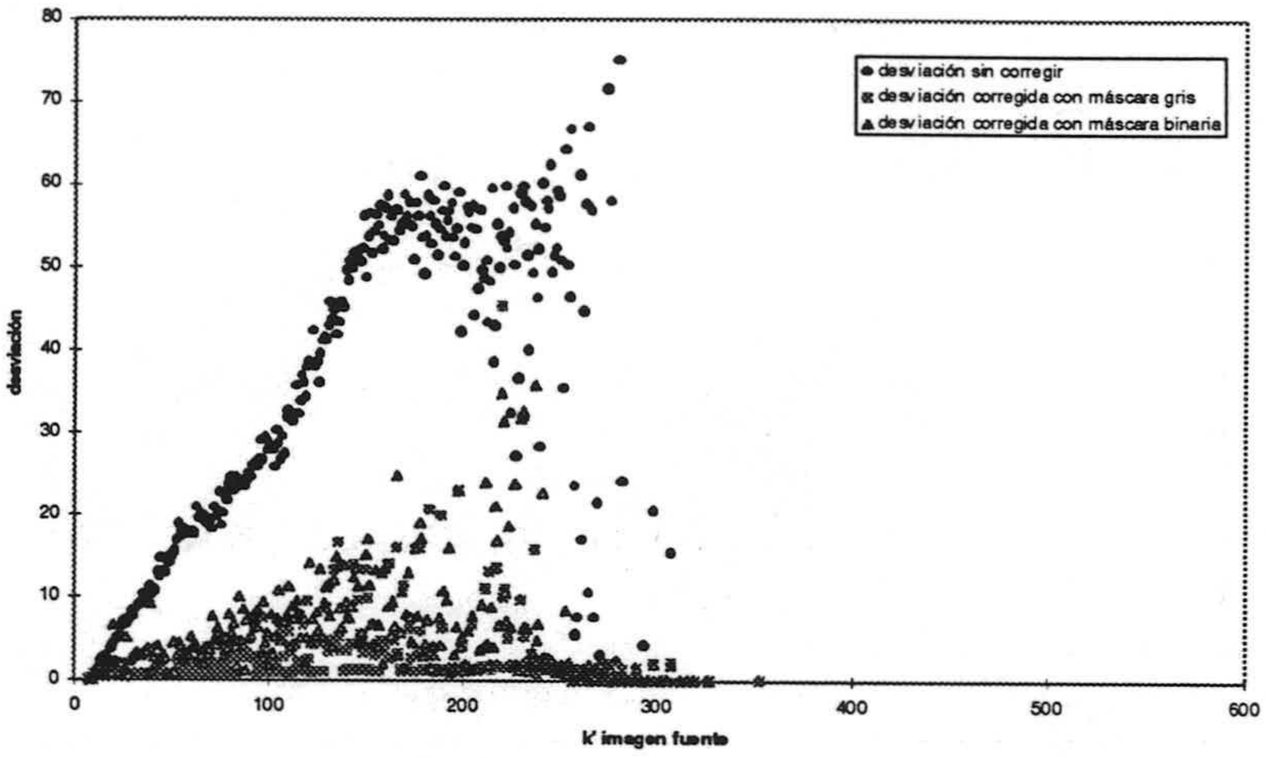
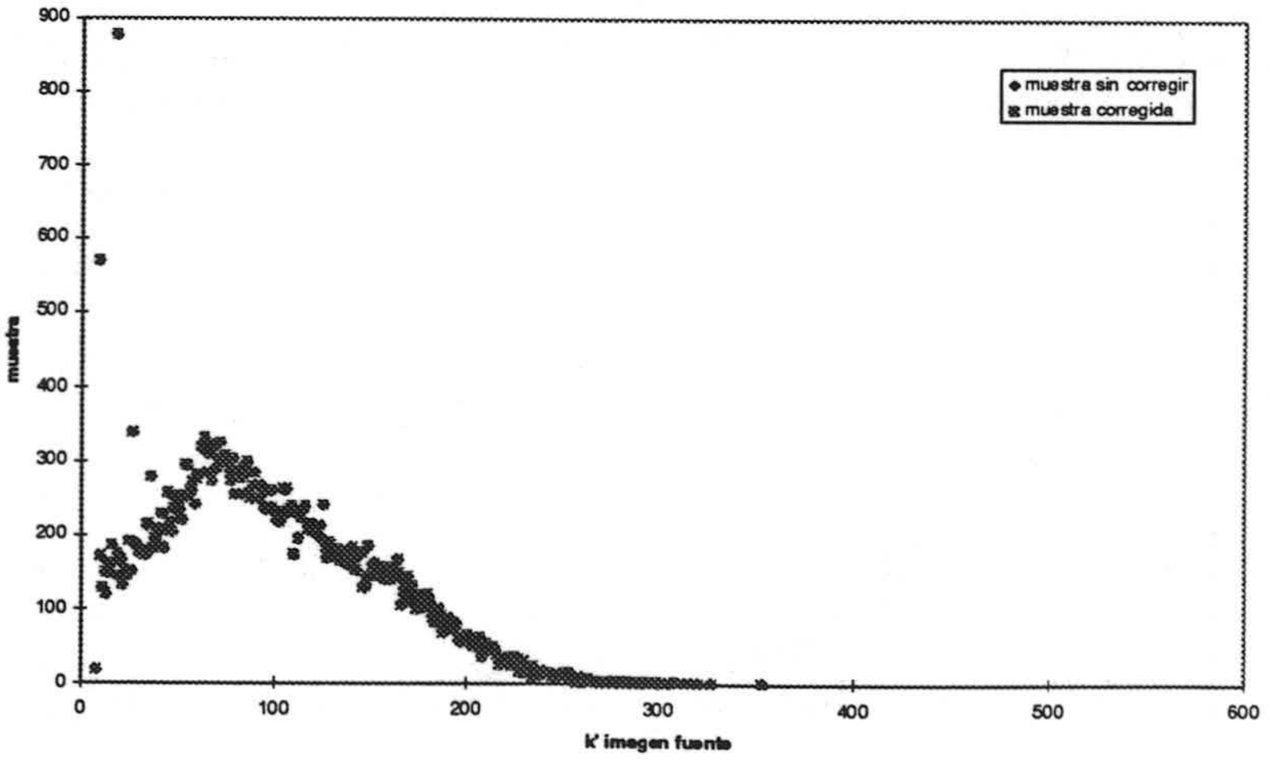


Figura 21c: Medida del tamaño de la muestra que entra a formar parte de las estadísticas. L aparte más oscura es la más representativa, como cabía esperar de una imagen de la clase de radiografías. Existe coincidencia en el número de regiones para los casos en que no existe corrección y sí la hay.



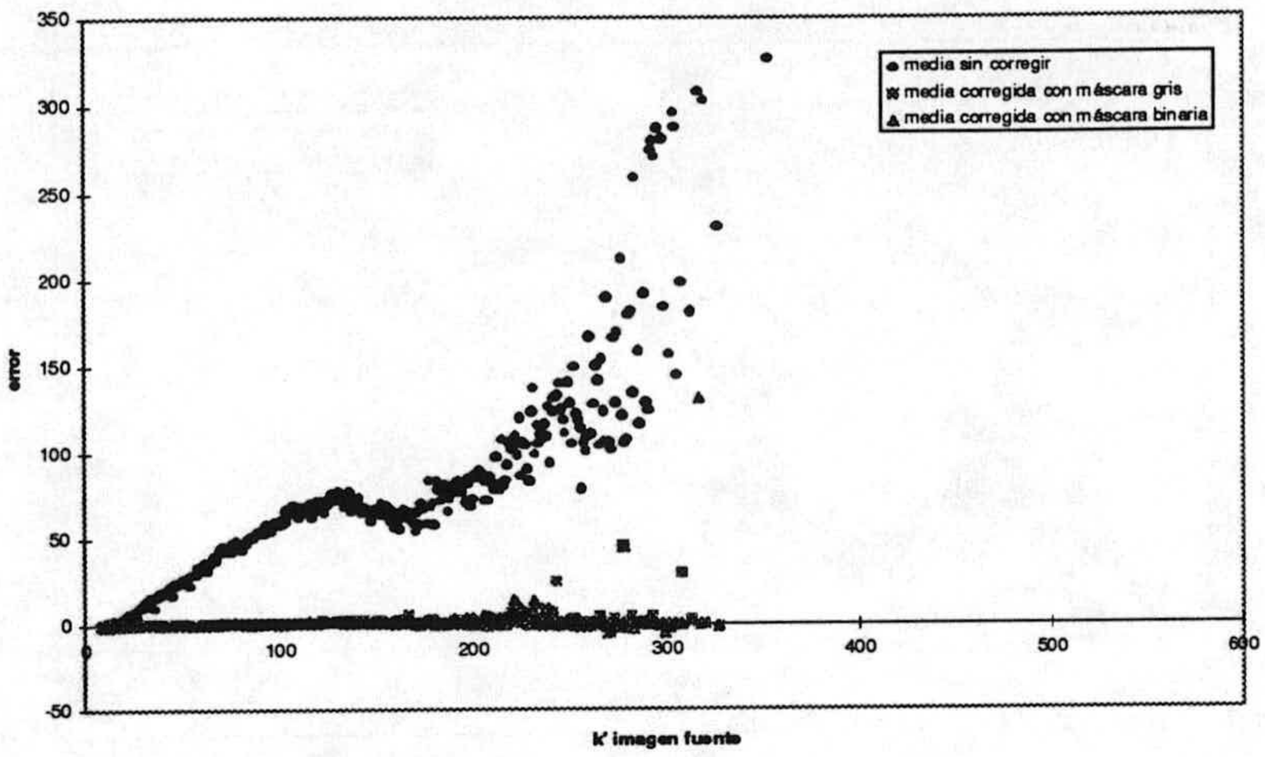


Figura 22a: Representación de la media de los errores cometidos al acceder a la LUT para una imagen de la clase de radiografías con 20% de ruido impulsivo uniformemente distribuído. El umbral de selección se ha fijado mediante el parámetro $\rho = 75$, y el tamaño de la ventana usado es 3×3. Se observa que tanto la máscara gris como la binaria dan resultados excelentes, dado que el modelo es muy cohesionado.

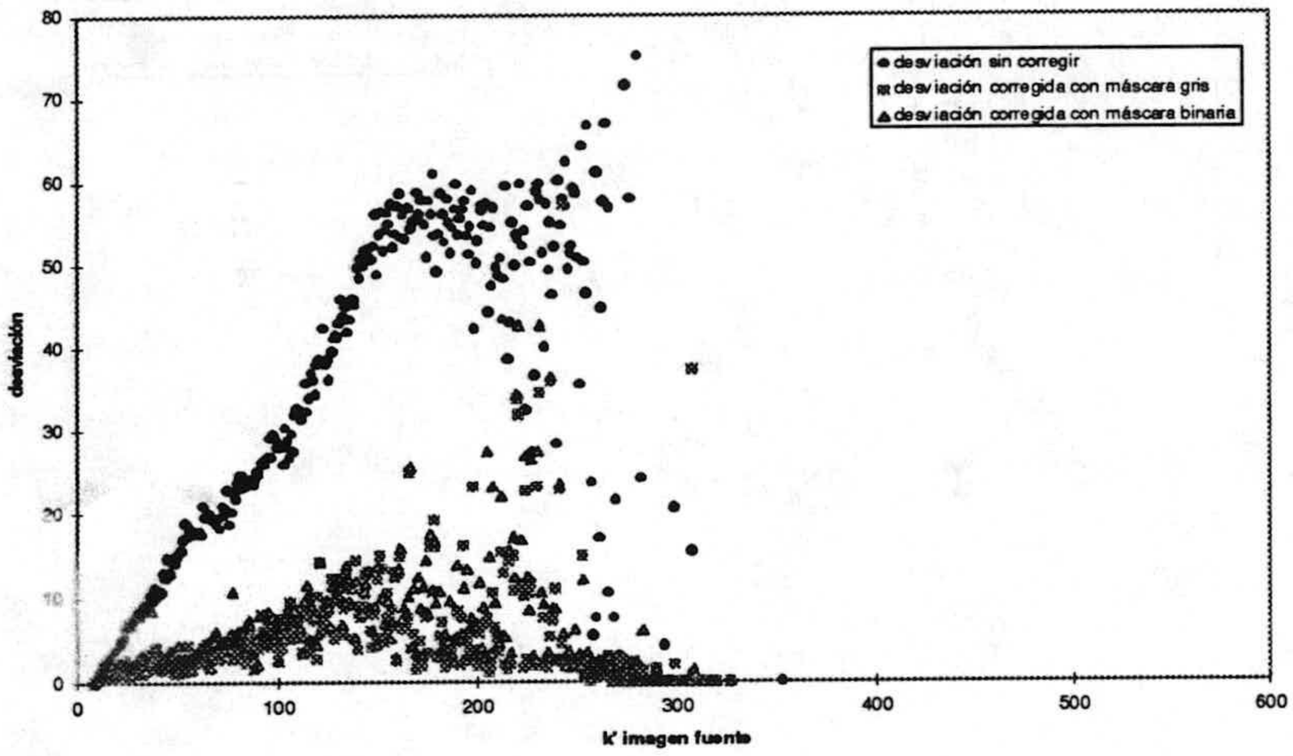


Figura 22b: Se plasman las desviaciones típicas de los errores anteriormente citados. Es apreciable la disminuciónde dispersión que aparece por el uso de las máscaras.

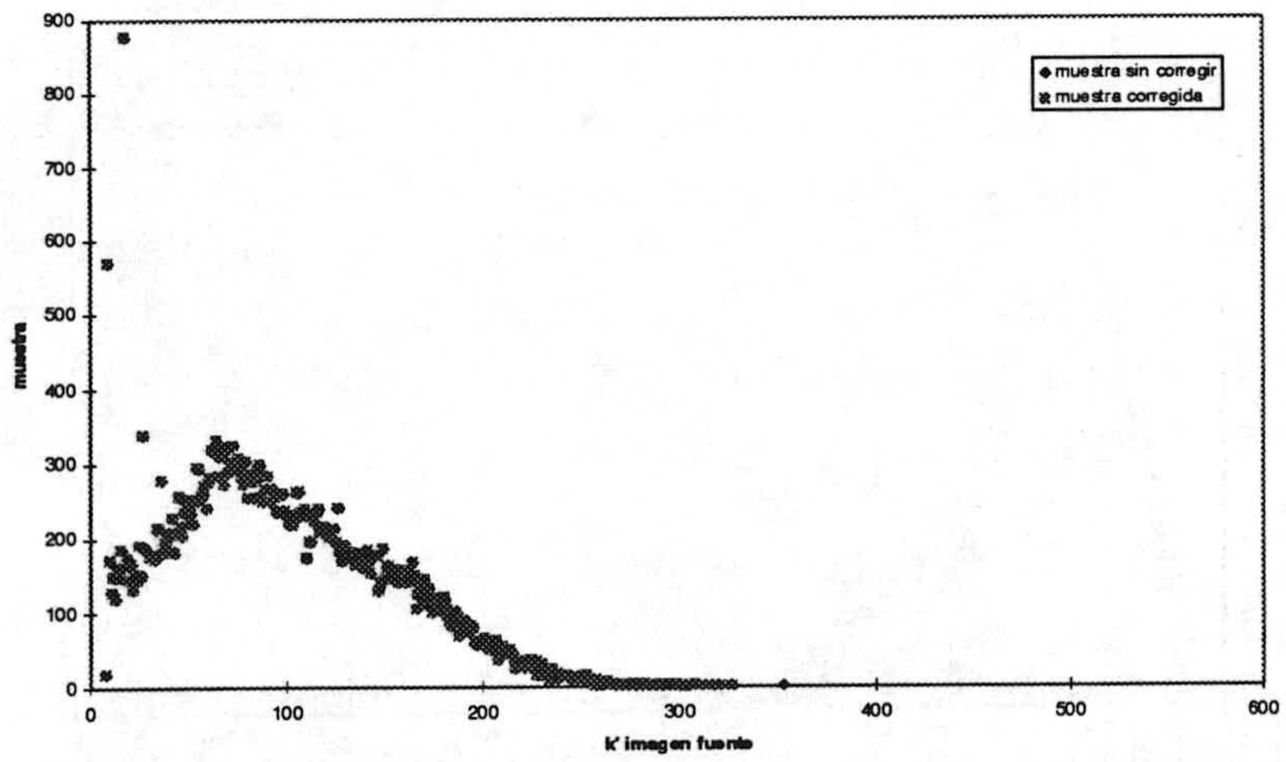


Figura 22c: Gráfica mostrando el tamaño de la muestra que se empleço en las estadísticas anteriormente utilizadas. La mayor representatividad se centra en las regiones oscuras de la imagen, y existe coincidencia en número de casos tanto para cuando existe corrección como cuando no la hay.

Figura 23a: Puede apreciarse la excelente capacidad de corrección del error de acceso a la LUT por parte de las máscaras, tanto gris como binaria. La imagen es de la clase de radiografías, contaminada por un 20% de ruido impulsivo uniforme. El umbral de detección se fija mediante el parámetro ρ = 1, y el tamaño de la ventana es 5×5. No se observa diferencia notable entre mácara gris y binaria.

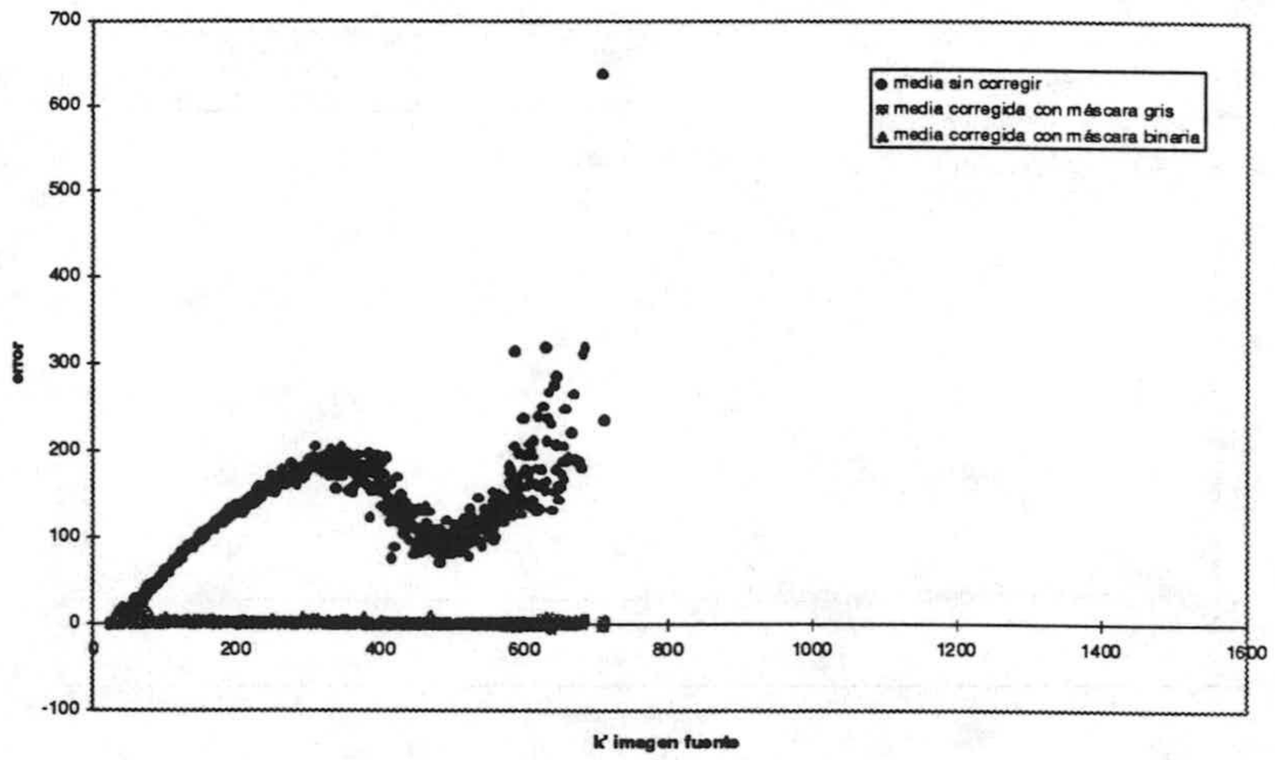


Figura 23b: Se aprecia la desviación típica de los errores representados en la gráfica precedente. Se aprecia cómo la mínima dispersión la produce el uso de la máscara de niveles de gris. En cualquier caso, la máscara binaria también presenta un buen comportamiento.

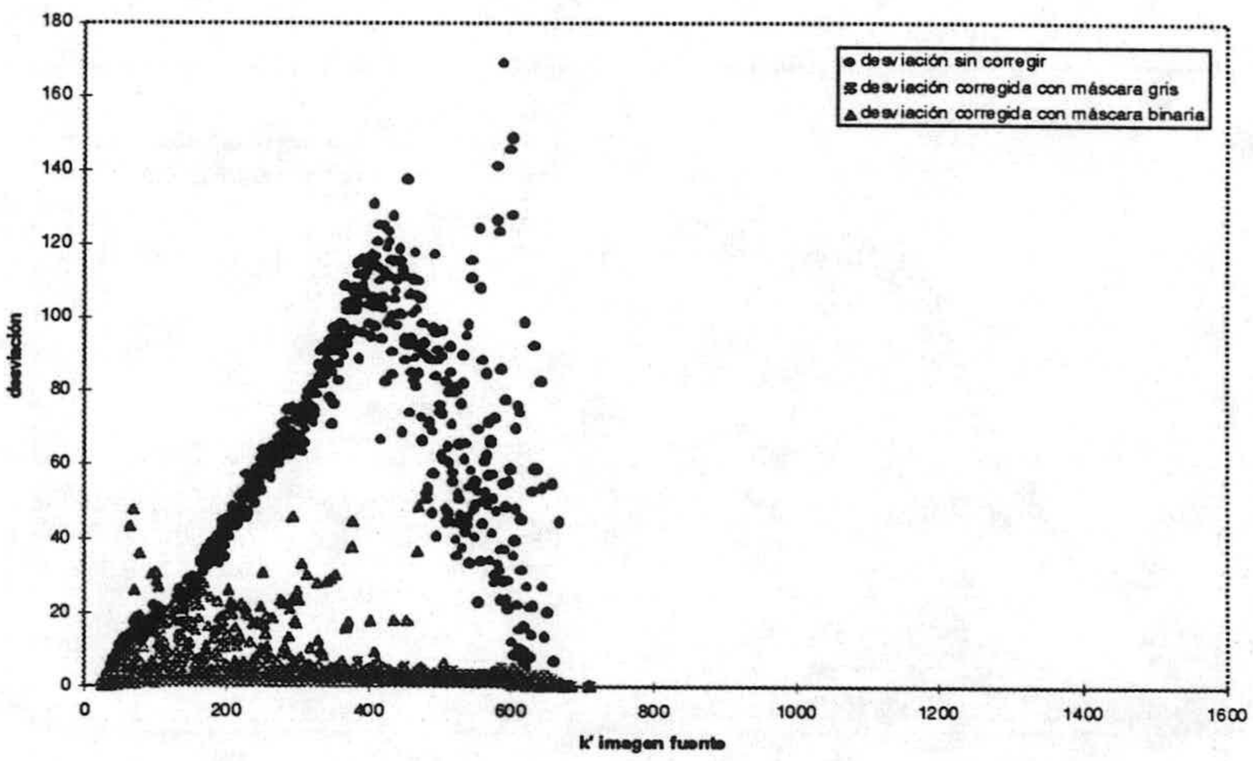
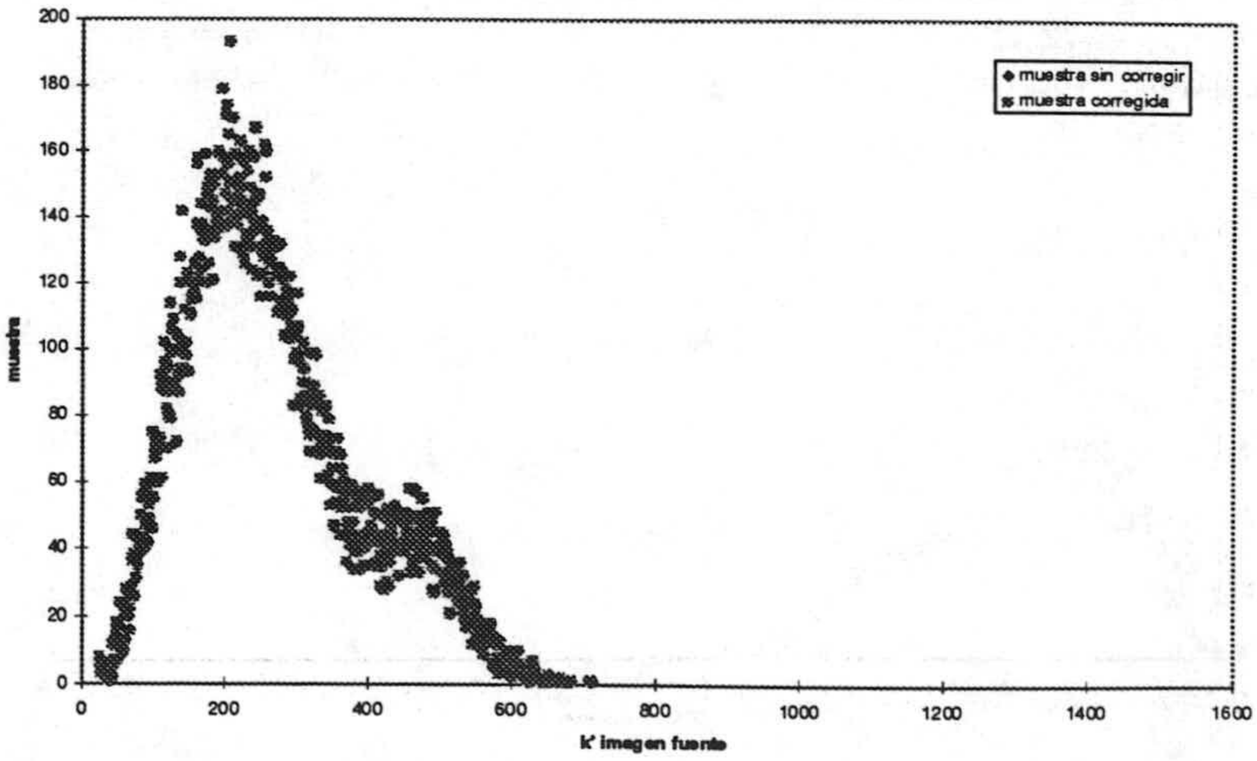


Figura 23c: Tamaño de la muestra que entra a formar parte de las estadísticas representadas. No hay diferencia entre el caso coregido y sin corregir. La parte más oscura de la imagen es la más representativa.



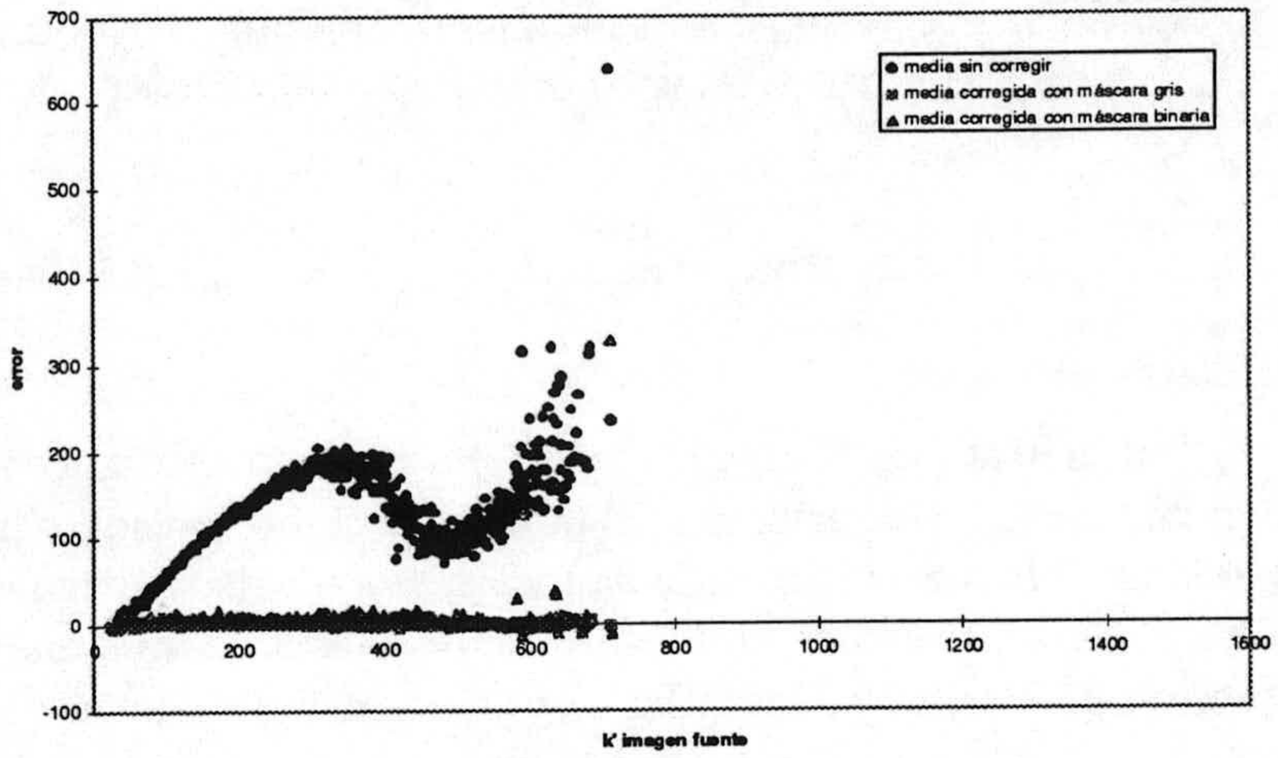


Figura 24a: Media de los errores cometidos al acceder a la LUT para una imagen de la clase de radiografías con 20% de ruido impulsivo uniformemente distribuído. El umbral de selección se ha fijado mediante el parámetro $\rho = 50$, y el tamaño de la ventana usado es 5×5. Se observa que tanto la máscara gris como la binaria dan resultados excelentes, dada la gran calidad de la matriz de composición..

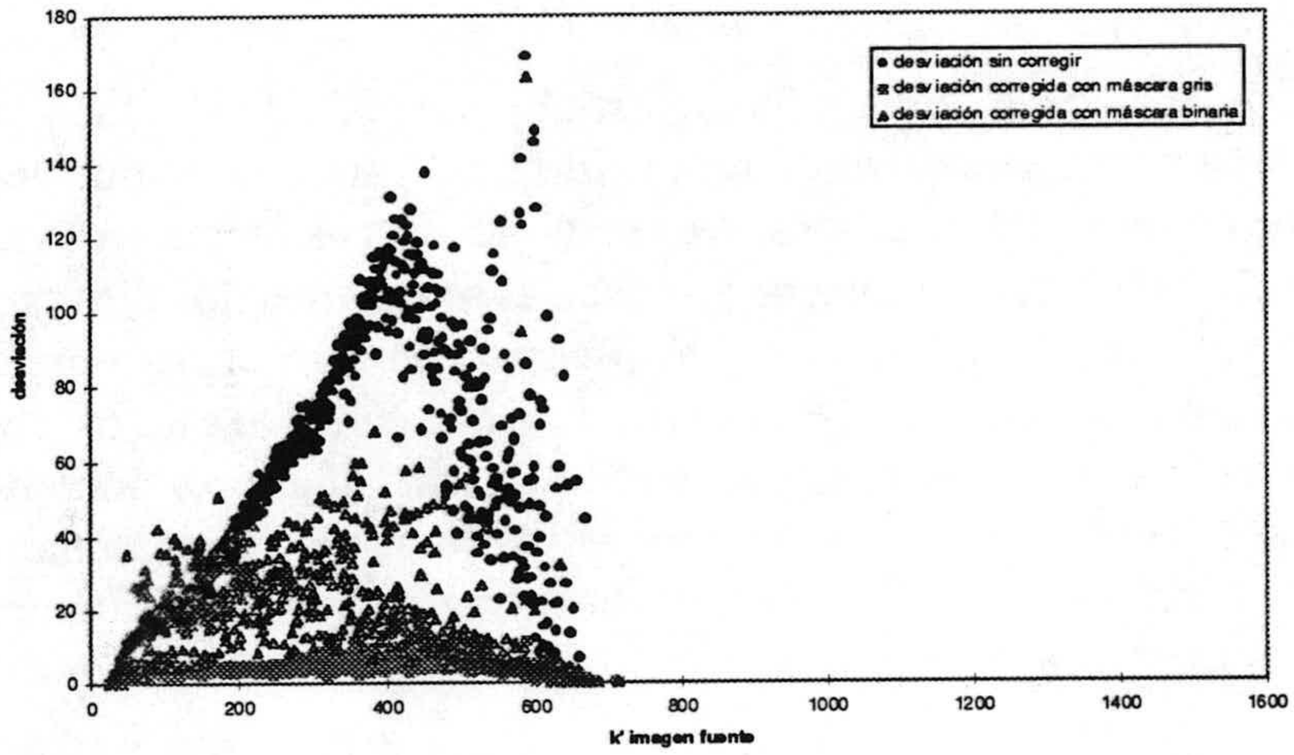


Figura 24b: Desviaciones típicas de los errores anteriormente citados. Es apreciable la disminuciónde dispersión que aparece por el uso de las máscaras. La máscara gris presenta el mejor comportamiento.

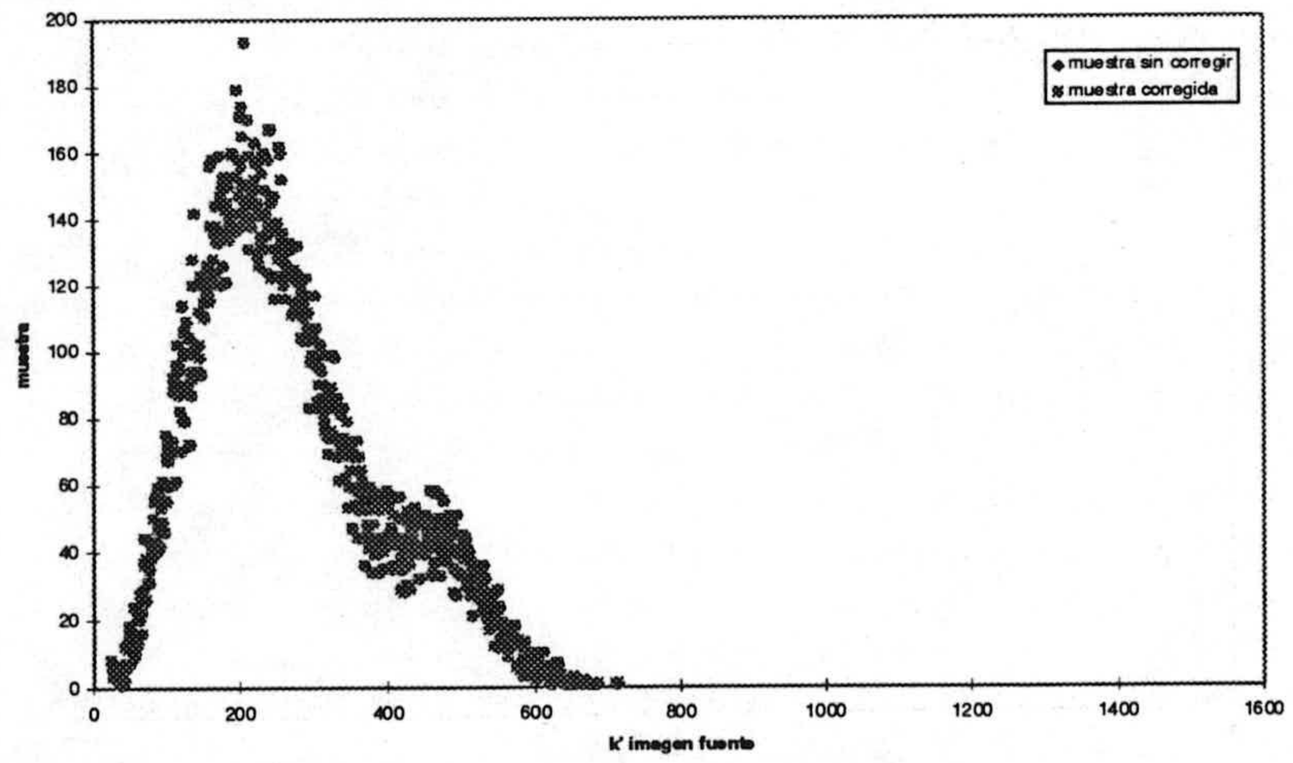


Figura 24c: Tamaño de la muestra que se empleó en las estadísticas anteriormente utilizadas. La mayor representatividad se centra en las regiones oscuras de la imagen, y existe coincidencia en número de casos tanto para cuando existe corrección como cuando no la hay.

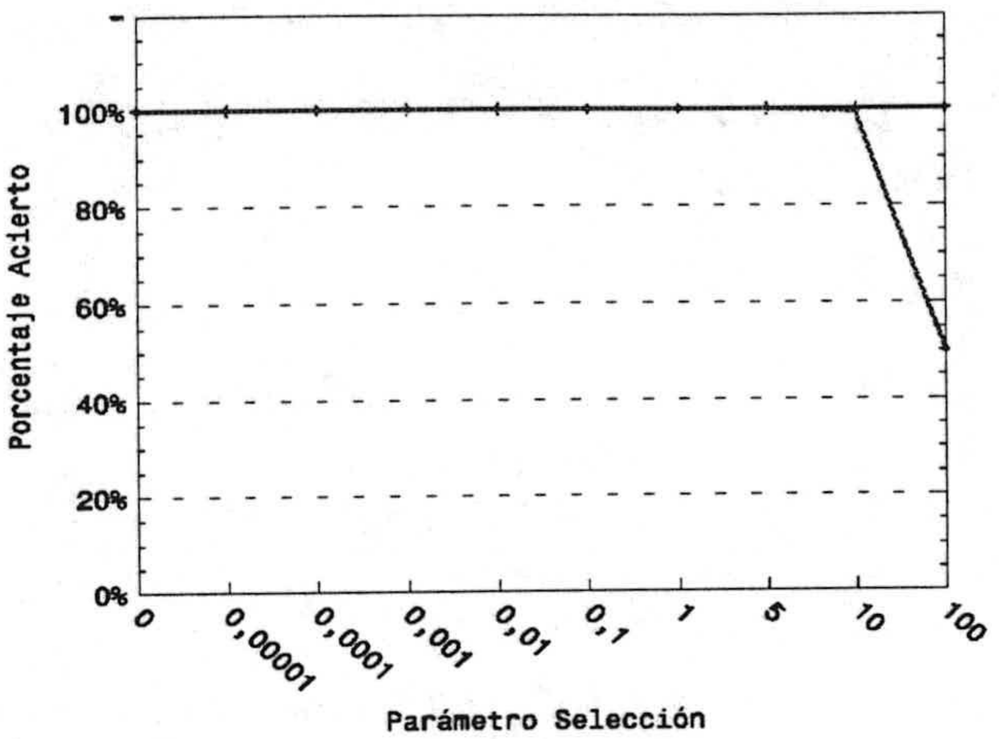
dado que no ofrece ventajas lo suficientemente notables como para que merezca la pena el costo computacional añadido.

Influencia del Umbral de Selección en la Localización del Ruido.

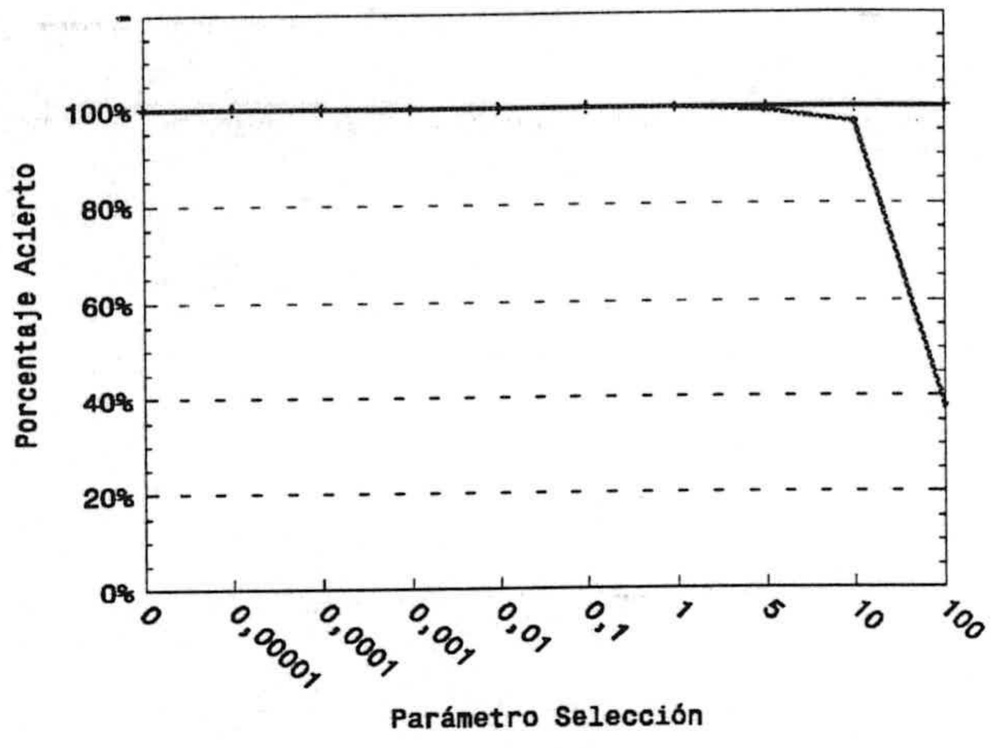
También se han realizado experiencias encaminadas a estudiar cómo afecta la variación del umbral de selección sobre la capacidad de la máscara binaria de localizar los pulsos adecuadamente. Para ello se han empleado las tres clases de imágenes —radiografías, texto y ecografías—, con dos tamaños de ventana 3×3 y 5×5. Una muestra de cada clase de imágenes ha sido contaminada con distintas cantidades y tipos de ruido impulsivo: 5%, 30% del tipo *sal y pimienta*; y 5% y 20% del uniformemente distribuido.

Para cada imagen fuente así obtenida se ha calculado la máscara, empleando diferentes valores del parámetro ρ que define el umbral de selección. Los resultados se han representado gráficamente, de forma que en abscisas se incluyen los diferentes valores de ρ utilizados; y en ordenadas aparece el procentaje de acierto y el factor *CM* de calidad de máscara. De este modo, se realiza una gráfica para cada clase de imagen, tipo y porcentaje de ruido, en la cual aparecen tres curvas: una, representando el *udr*; otra mostrando el *udr*; y la última, en trazo discontínuo, plasmando el valor de *CM*.

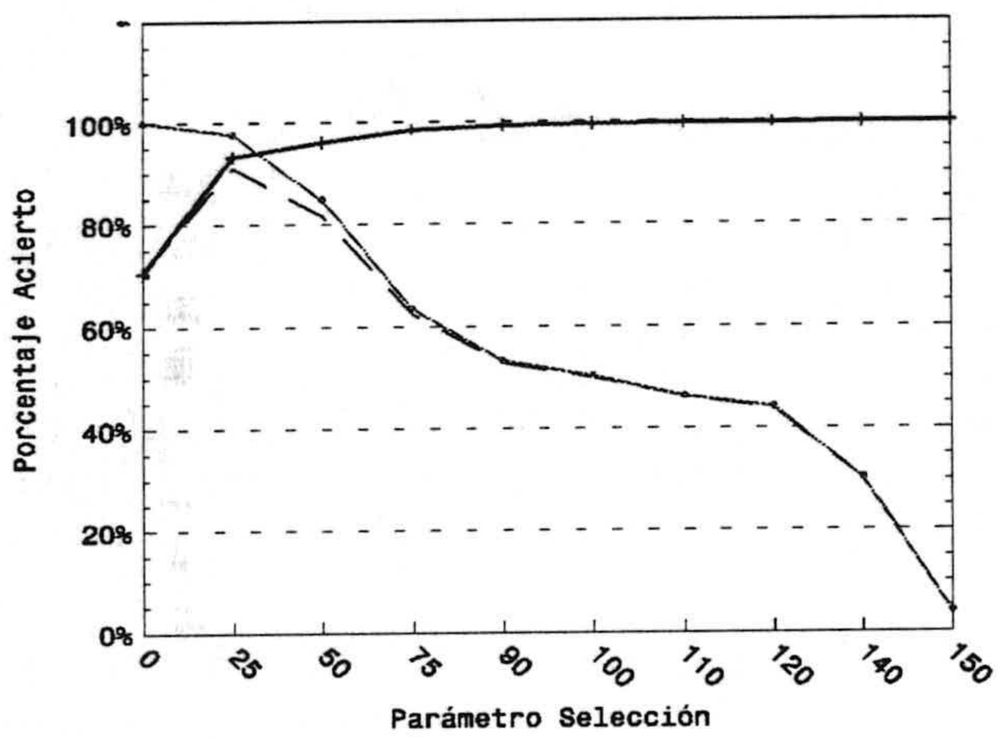
Estas ilustraciones aparecen en las figuras numeradas de la 25 a la 30. De las mismas observamos que para ruido *sal y pimienta* el comportamiento de la máscara *M* es excelente, localizando perfectamente la totalidad de los pulsos, para valores muy bajos del parámetro de selección ρ.



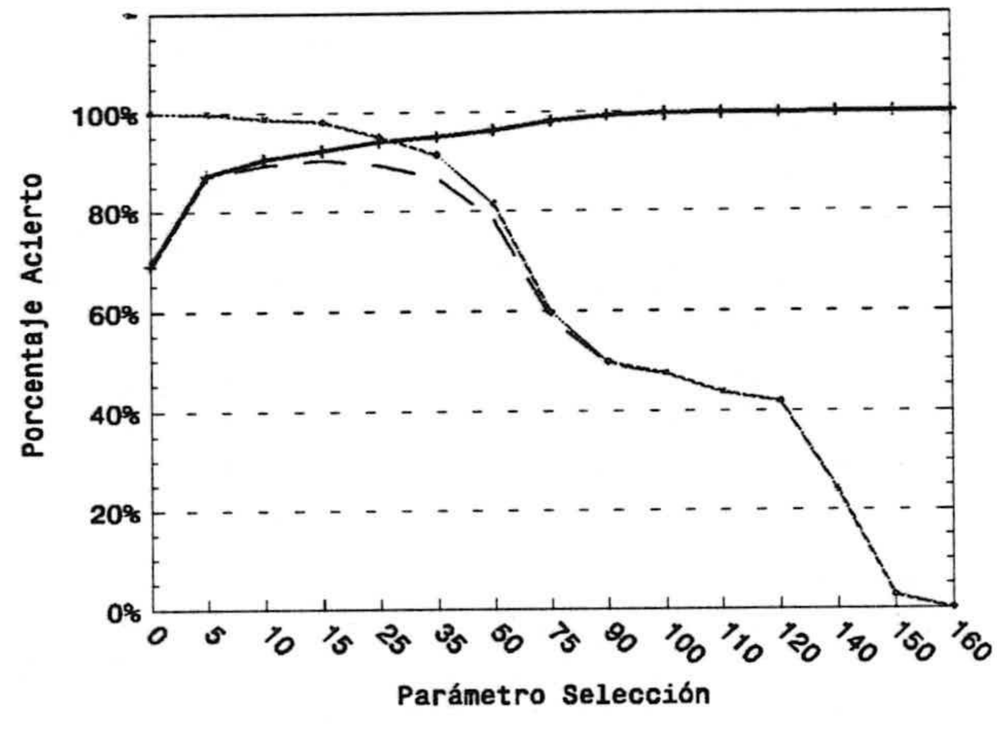
5 % Ruido Impulsivo *sal y pimienta*
[MAE(S-O) = 1.59400]



30 % Ruido Impulsivo *sal y pimienta*
[MAE(S-O) = 9.558975]



5 % Ruido Impulsivo *uniforme*
[MAE(S-O) = 1.14965]



20 % Ruido Impulsivo *uniforme*
[MAE(S-O) = 4.686675]

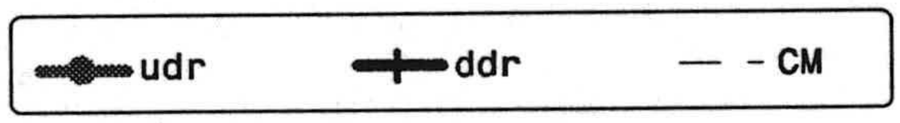
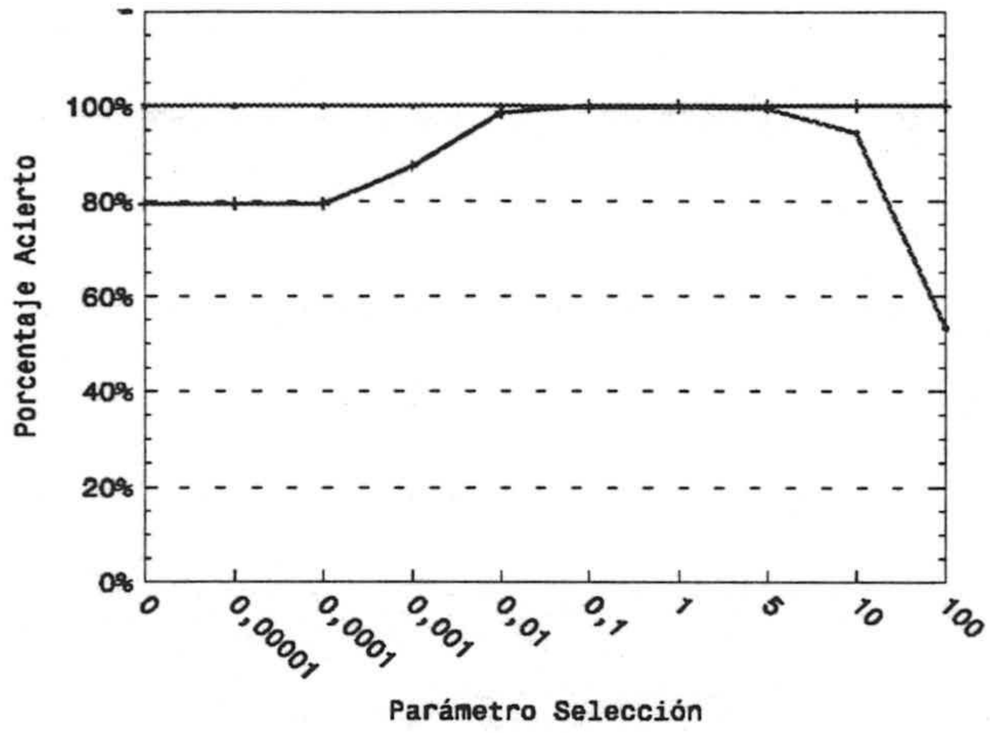
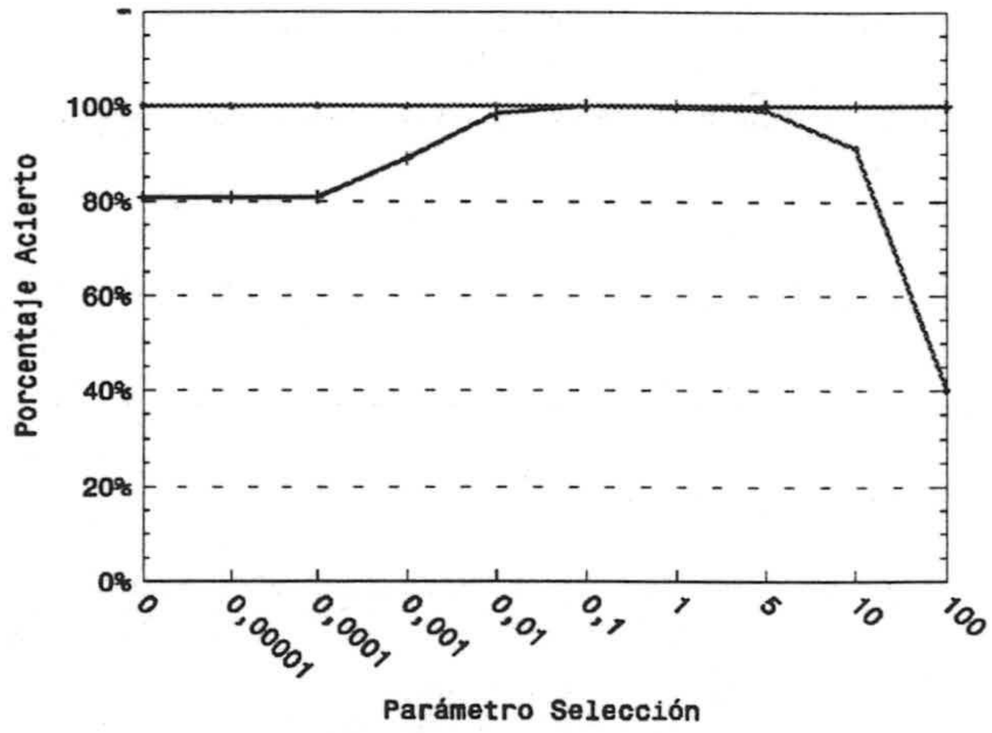


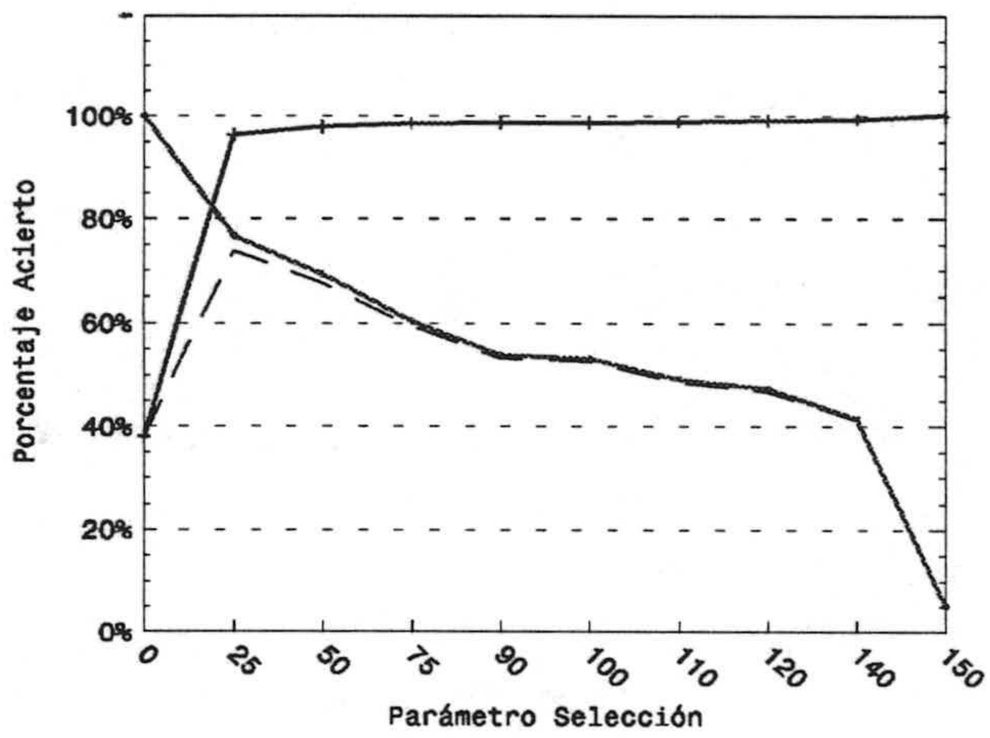
Figura 25: Aparecen gráficas que muestran la influencia del umbral de selección, determinado por el parámetro ρ, para el caso de una imagen de la clase de ecografías, en la que se ha empleado una ventana de tamaño 3×3. Las gráficas superiores ilustran el problema para ruido impulsivo *sal y pimienta*, de las cuales se desprende que un valor muy pequeño de ρ es lo más indicado para obtener una máscara de calidad. En las gráficas inferiores aparece el comportamiento frente a ruido imulsivo uniformemente distribuído. Cuando la imagen está poco afectada por la distorsión, la mejor máscara se obtiene para ρ en torno a 25. Pero si el ruido es más abundante, lo mejor es ρ = 0. En todos los casos los mejores resultados superan el 90% de acierto. Todas las máscaras empleadas son binarias.



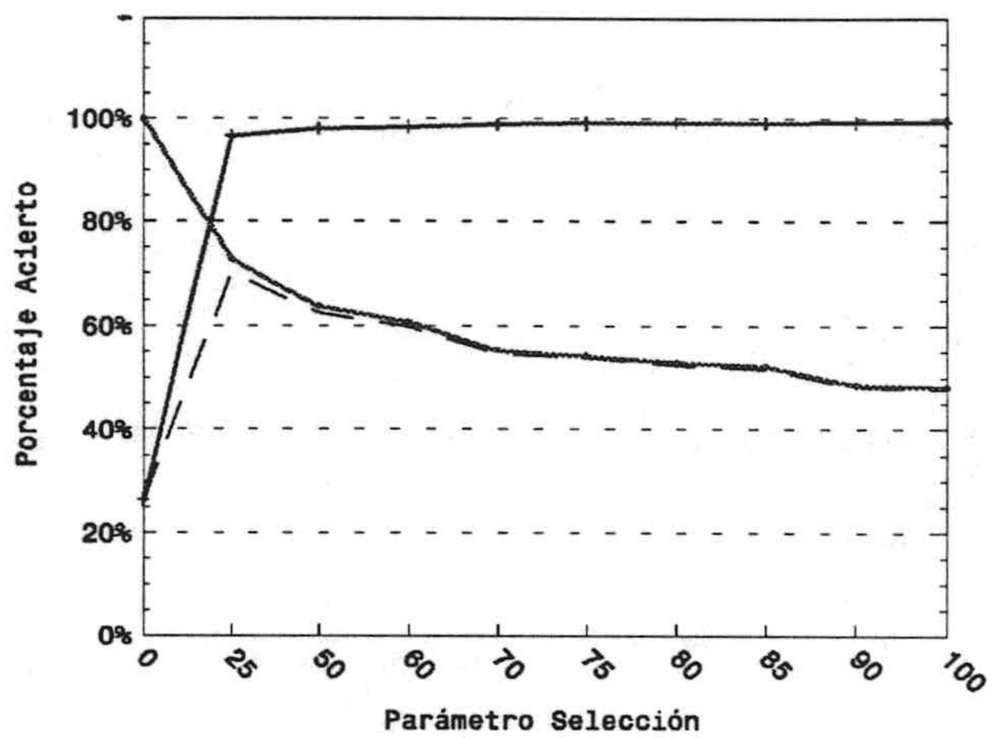
5 % Ruido Impulsivo *sal y pimienta*
[MAE(S-O) = 1.59400]



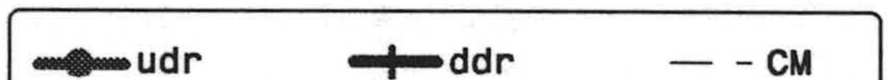
30 % Ruido Impulsivo *sal y pimienta*
[MAE(S-O) = 9.558975]



5 % Ruido Impulsivo *uniforme*
[MAE(S-O) = 1.14965]



20 % Ruido Impulsivo *uniforme*
[MAE(S-O) = 4.686675]



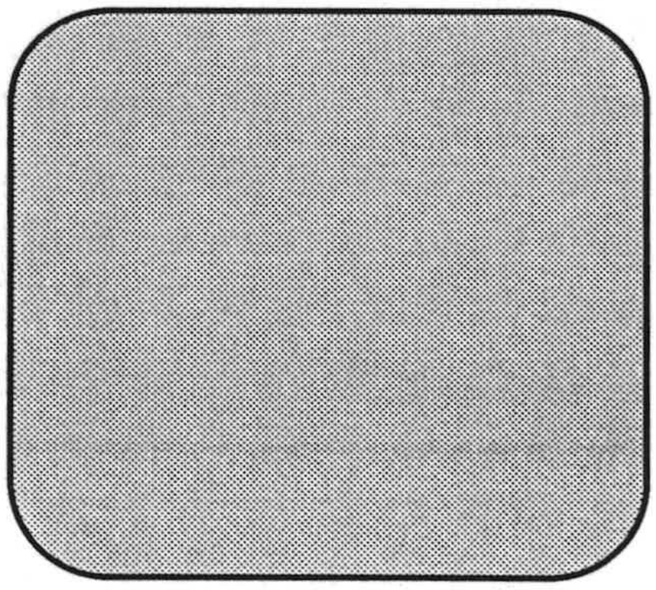
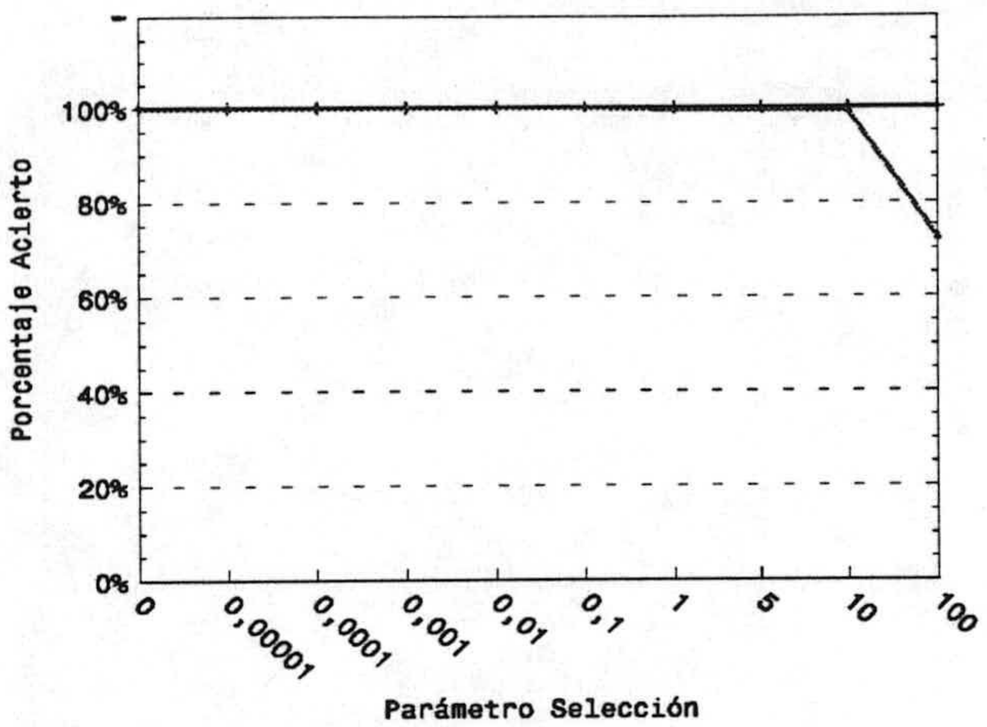
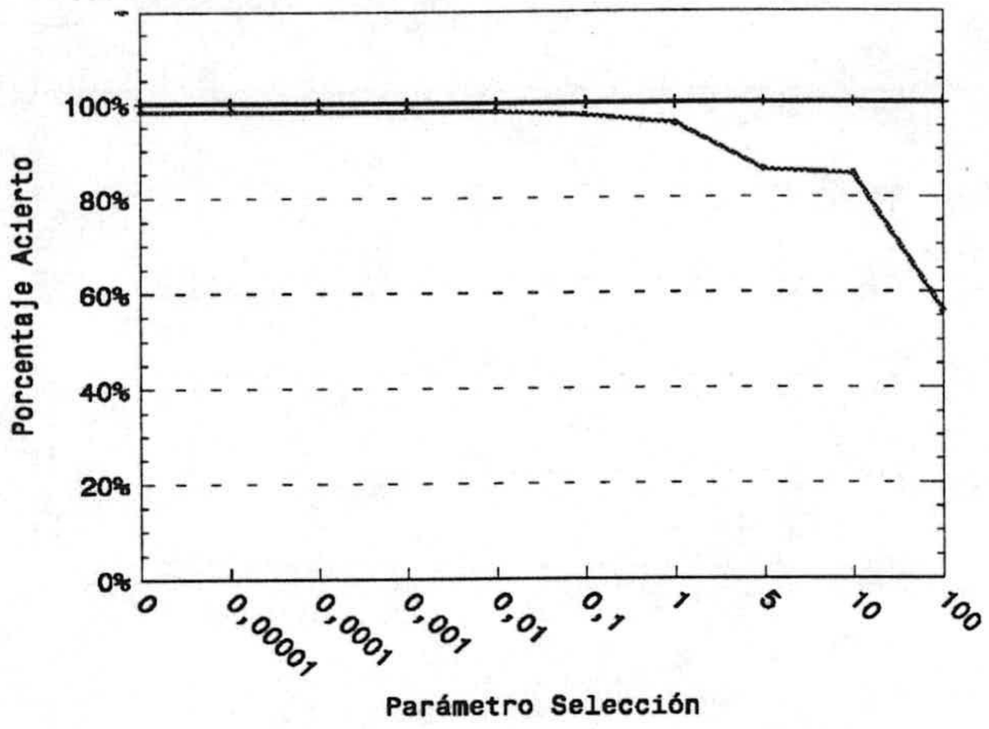


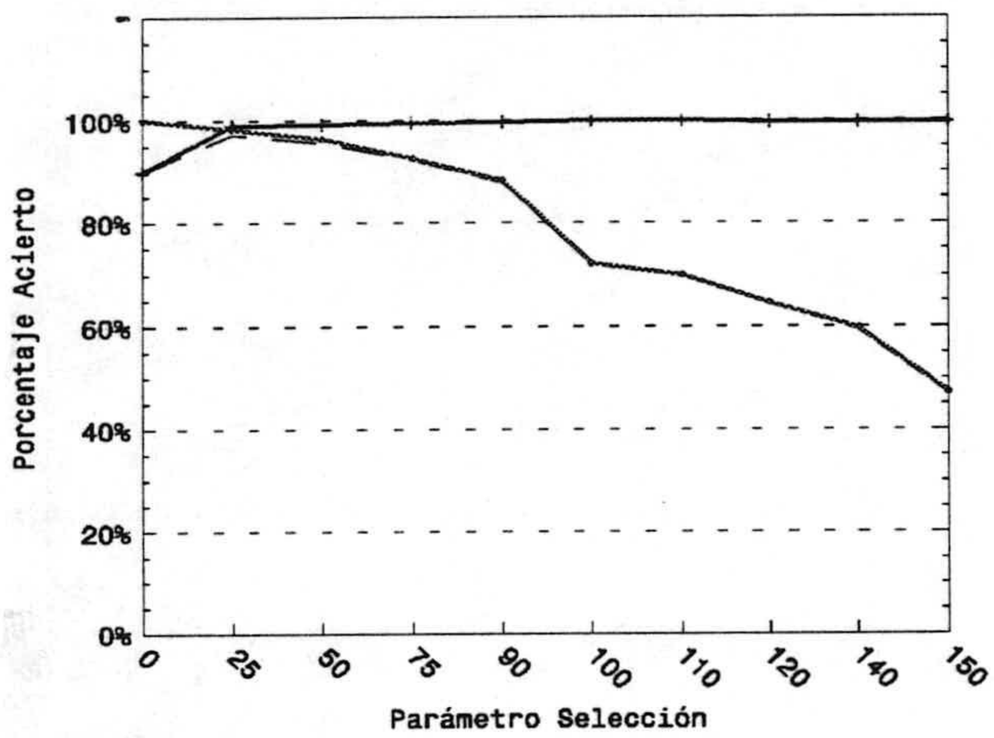
Figura 26: Se muestra la calidad de la máscara para diferentes valores de parámetros ρ (que determinan el umbral de selección) y para distintos tios y cantidades de ruido impulsivo. La imagen es de clase texto mecanografiado, y el tamaño de ventana utilizado es 3×3. Los resultados para ruido *sal y pimienta* son de un 100% de acierto para valores $0.01 \leq \rho \leq 1$, independientemente de la cantidad de distorsión. Para ruido impulsivo uniformemente distribuído, las mejores máscaras se obtienen para ρ en el entorno de 10, 20 ó 25. En todos los casos, la máscara empleada ha sido binaria. Los mejores resultados se acercan, en el peor de los casos, al 75% de acierto de máscara.



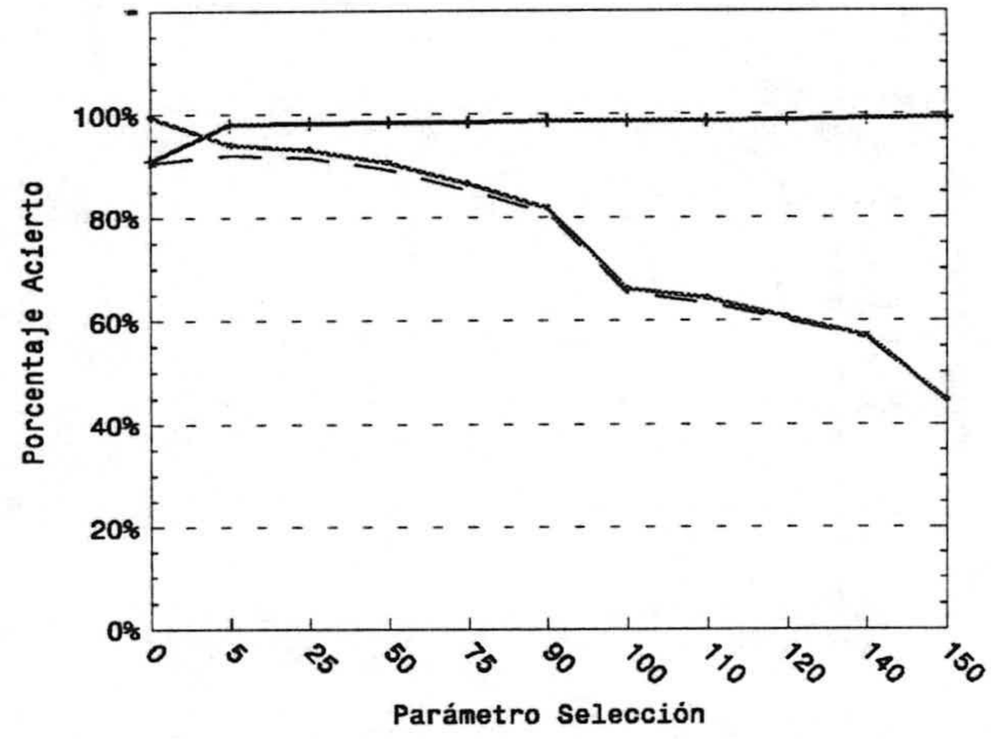
5 % Ruido Impulsivo *sal y pimienta*
[MAE(S-O) = 1.582125]



30 % Ruido Impulsivo *sal y pimienta*
[MAE(S-O) = 9.56045]



5 % Ruido Impulsivo *uniforme*
[MAE(S-O) = 1.30025]



20 % Ruido Impulsivo *uniforme*
[MAE(S-O) = 5.474025]

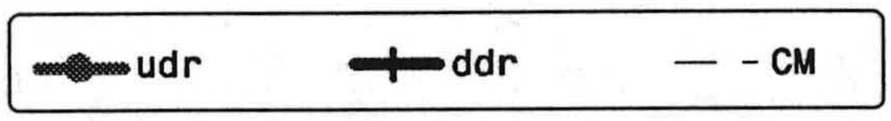
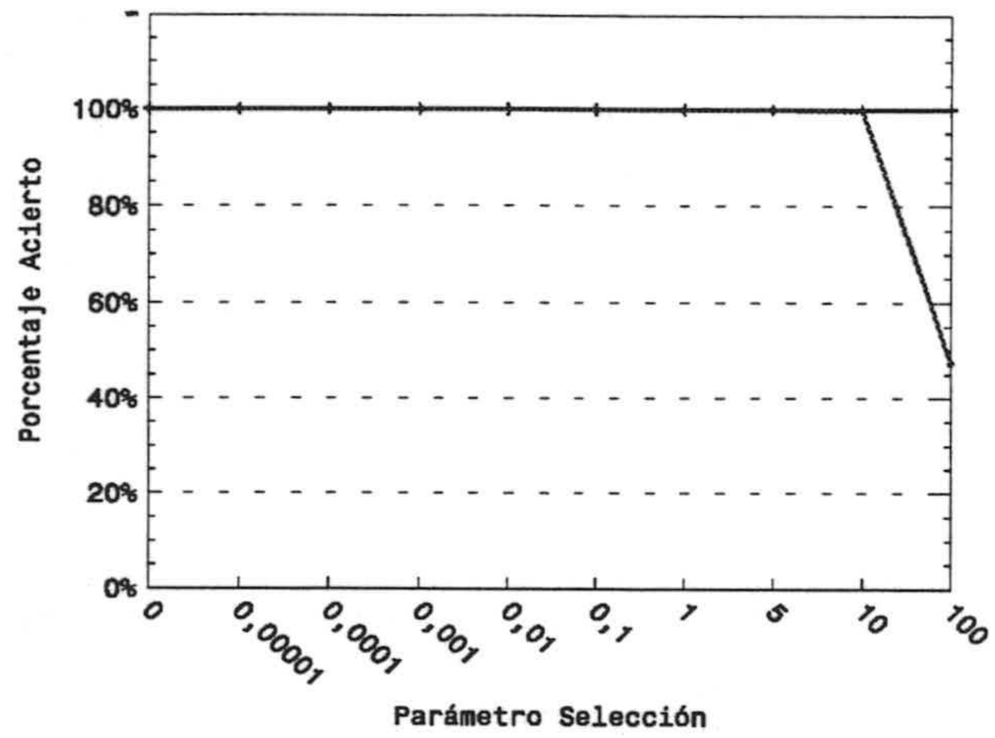
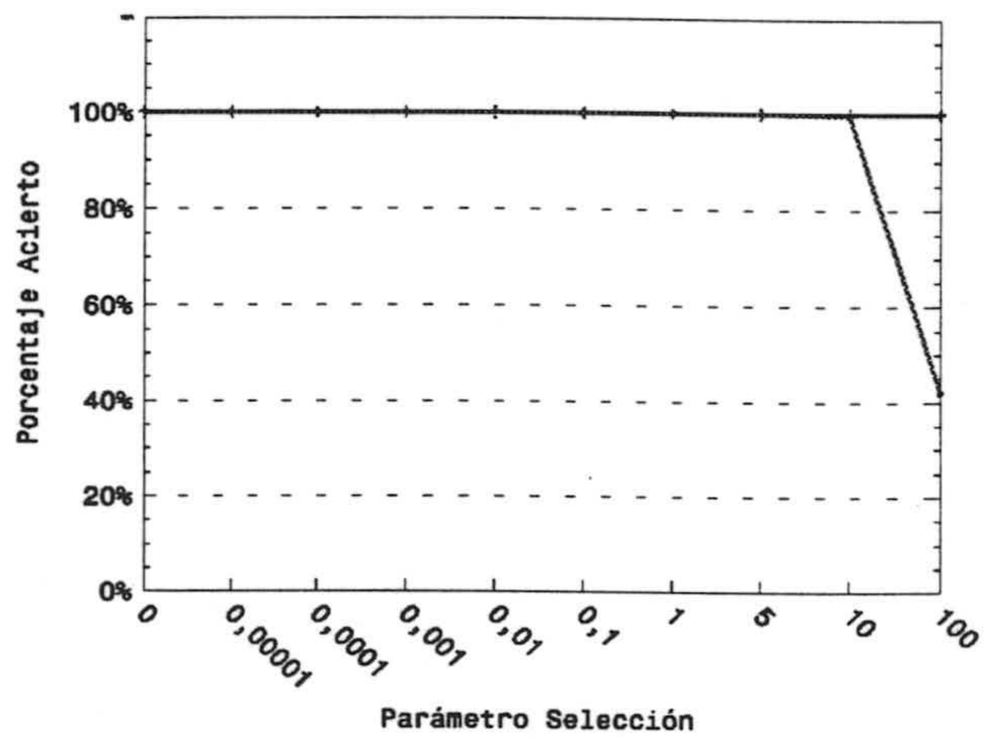


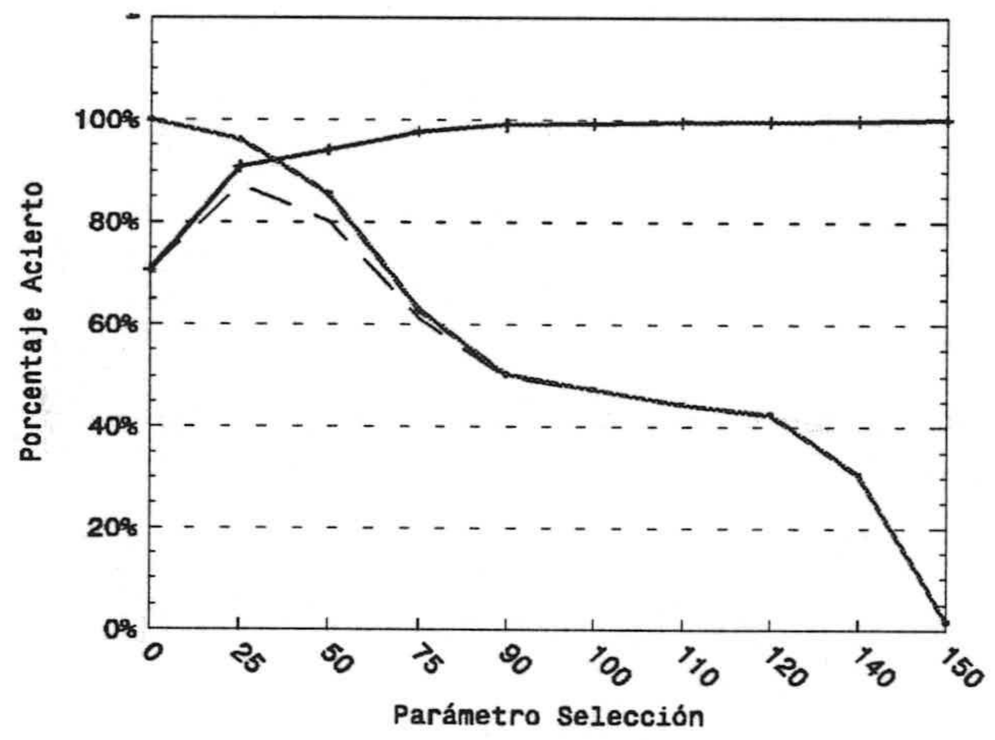
Figura 27: Ilustración de la dependencia entre la calidad de la máscara y el umbral de selección determinado por el parámetro ρ, para el caso de una imagen de la clase de radiografías, empleando un tamaño 3×3 de vantana. Para ruido impulsivo *sal y pimienta*, las graficas superiores muestran que un valor pequeño de ρ da resutados excelentes. En cuanto al ruido impulsivo uniformemente distribuído, las mácaras de mayor calidad se obtienen para valores de ρ en torno a 5 y 25. En todos los casos la máscara empleada ha sido binaria; y los mejores resultados de calidad de máscara superan el 85%.



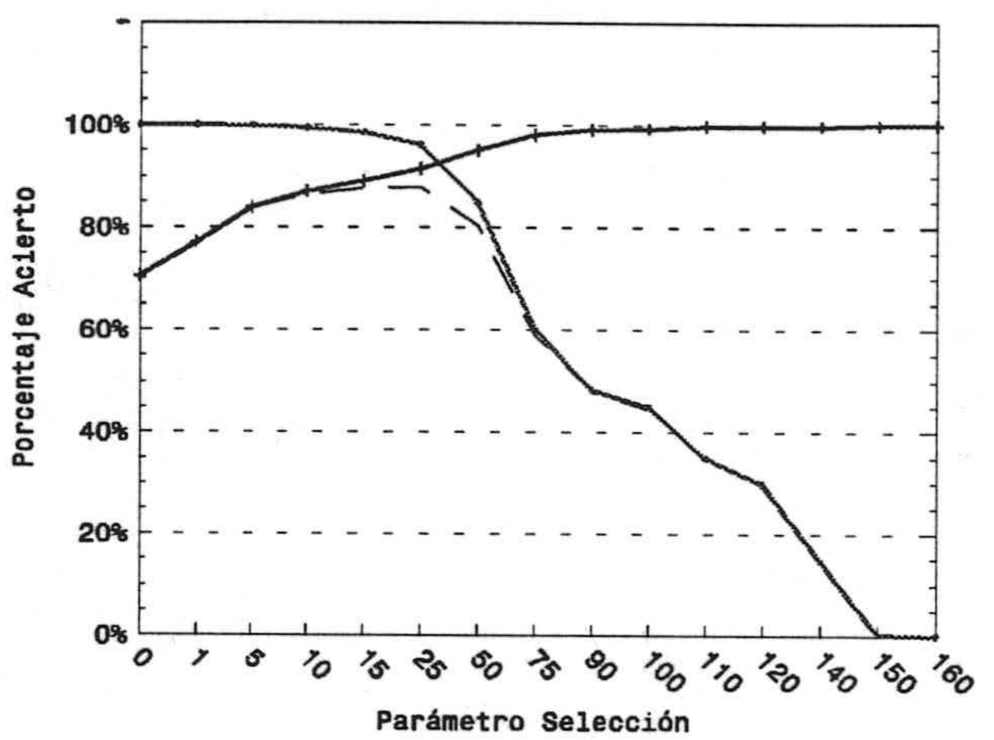
5 % Ruido Impulsivo *sal y pimienta*
[MAE(S-O) = 1.59400]



30 % Ruido Impulsivo *sal y pimienta*
[MAE(S-O) = 9.558975]



5 % Ruido Impulsivo *uniforme*
[MAE(S-O) = 1.14965]



20 % Ruido Impulsivo *uniforme*
[MAE(S-O) = 4.686675]

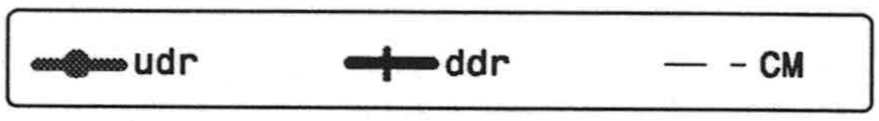
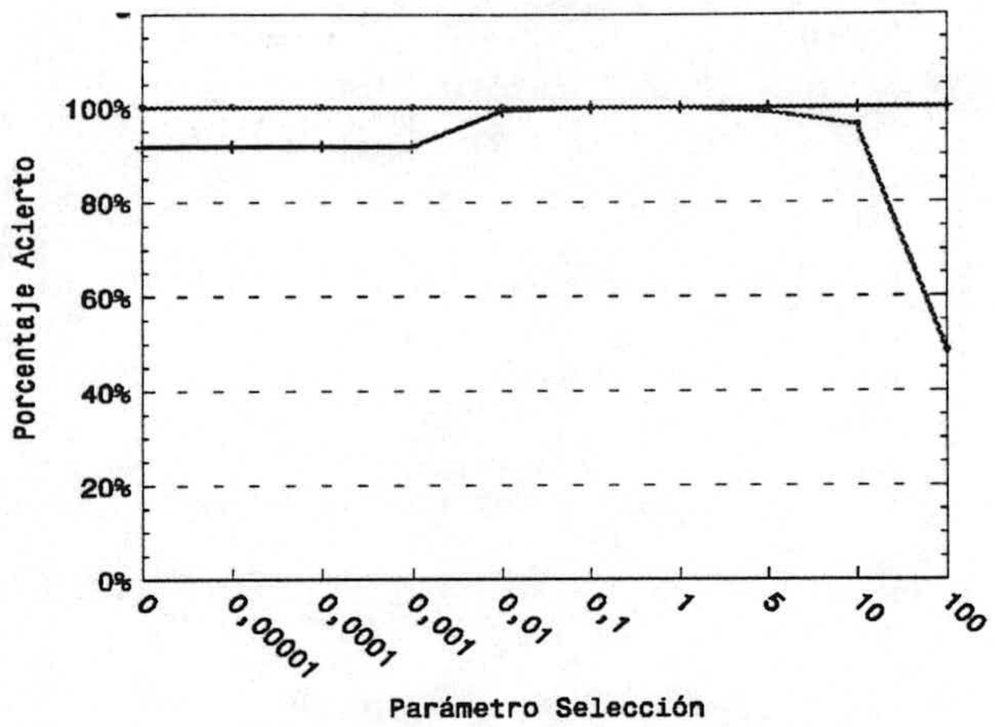
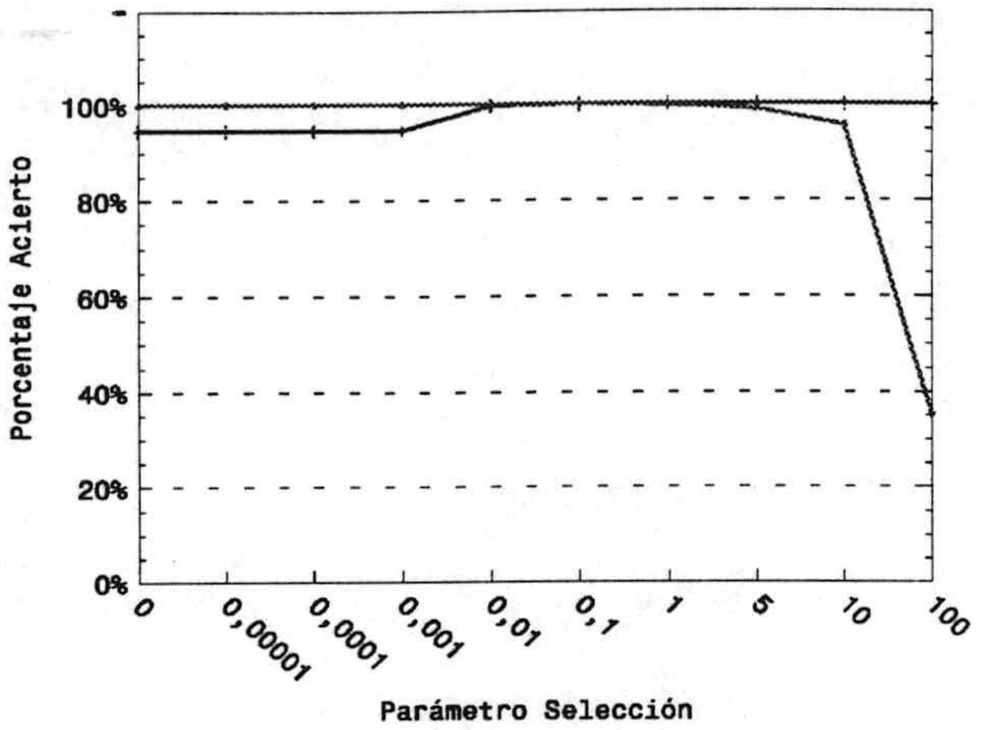


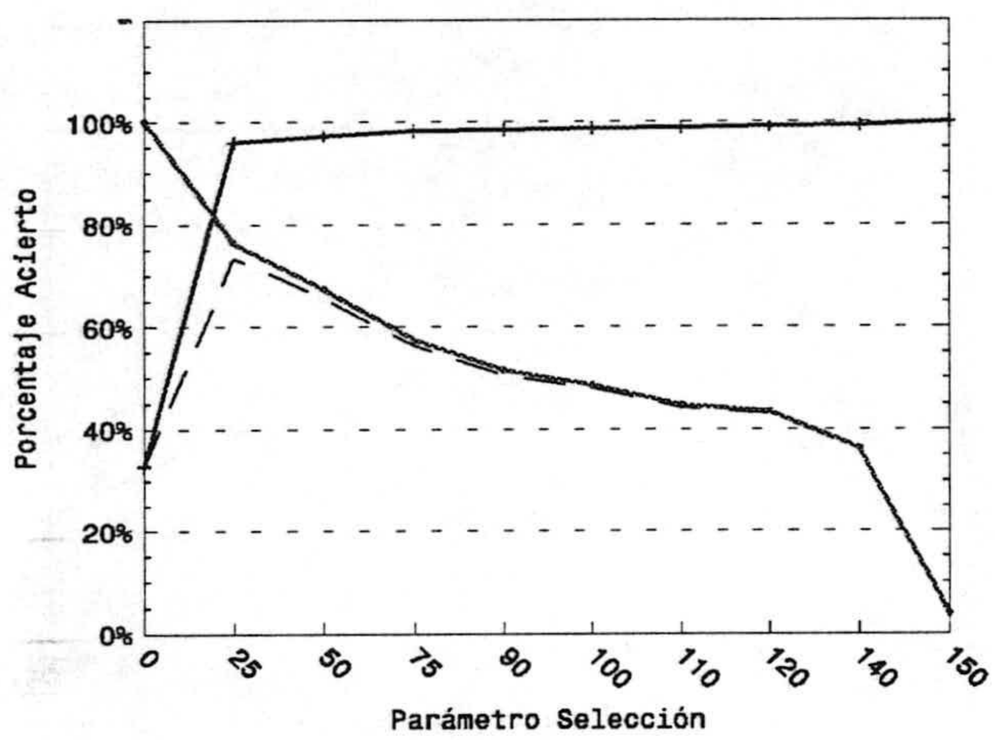
Figura 28: Comportamiento de la calidad de la máscara cuando varía el umbral de selección al cambiar el parámetro de selección ρ, en el caso de una imagen de la clase de ecografías. El tamaño de la ventana empleada es de 5×5. Para ruido imulsivo *sal y pimienta* el acierto de la máscara es total cuando el parámetro ρ es pequeño (de 0 a 10, independientemente de la cantidad de pulsos). Para ruido impulsivo uniformemente distribuído, la calidad de la máscara es mayor cuando ρ alcanza valores en el entorno de 25. Obsérvese el parecido con el comportamiento para ventana de tamaño menor, aunque en este caso la medida *CM* se ha desplazado hacia la derecha para ruido impulsivo uniforme. En todos los casos se ha utilizado máscara binaria. los mejores resultados se acercan al 85% de acierto.



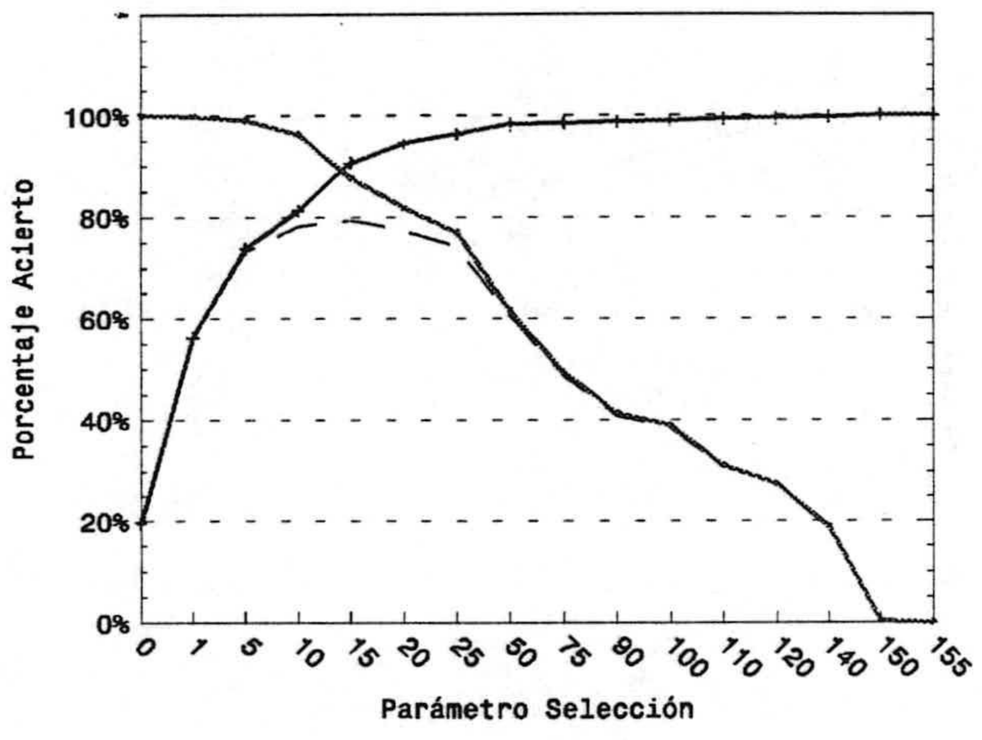
5 % Ruido Impulsivo *sal y pimienta*
[MAE(S-O) = 1.55865]



30 % Ruido Impulsivo *sal y pimienta*
[MAE(S-O) = 9.27985]



5 % Ruido Impulsivo *uniforme*
[MAE(S-O) = 1.294675]



20 % Ruido Impulsivo *uniforme*
[MAE(S-O) = 5.1038]

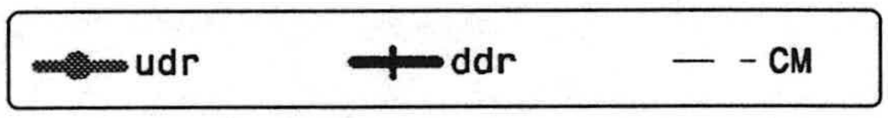
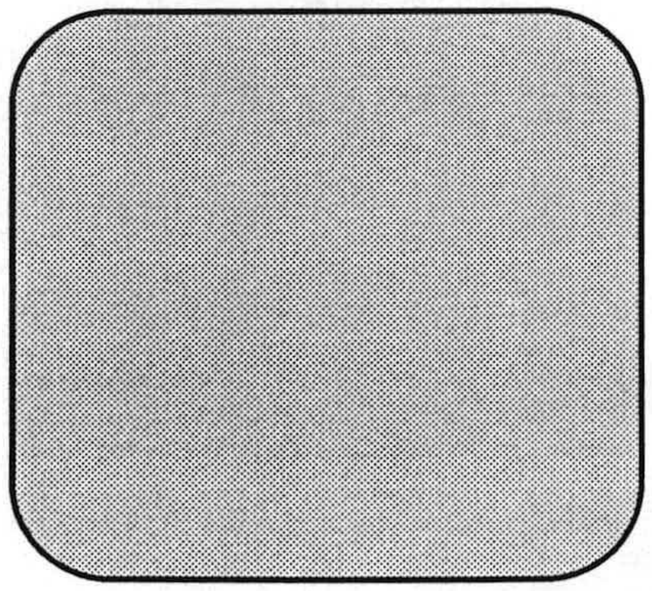
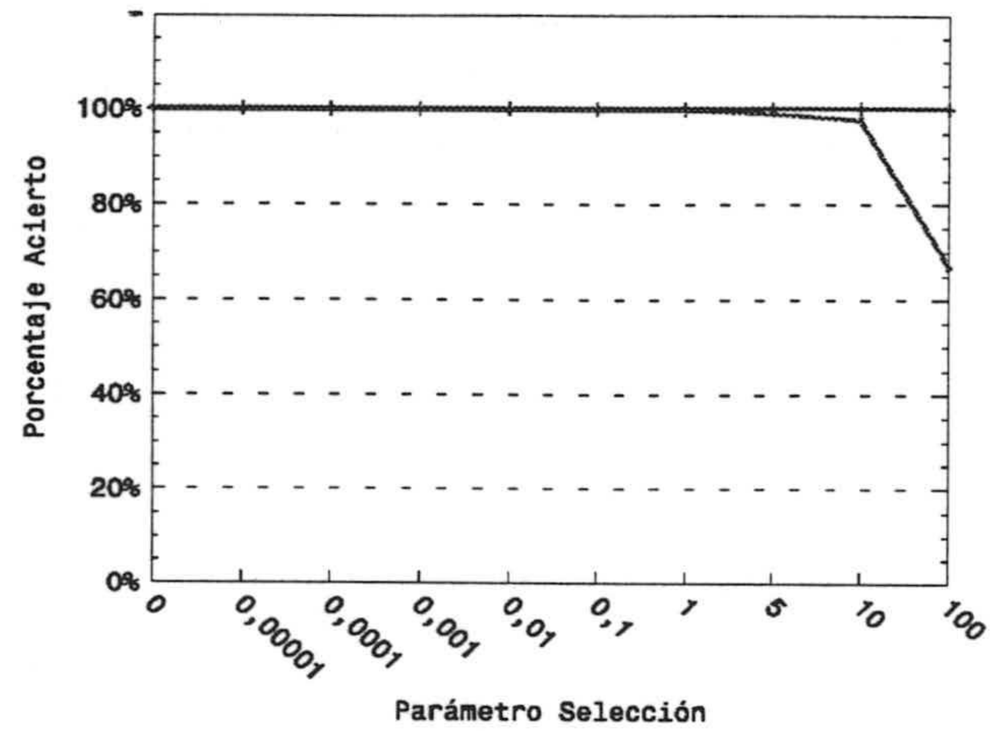
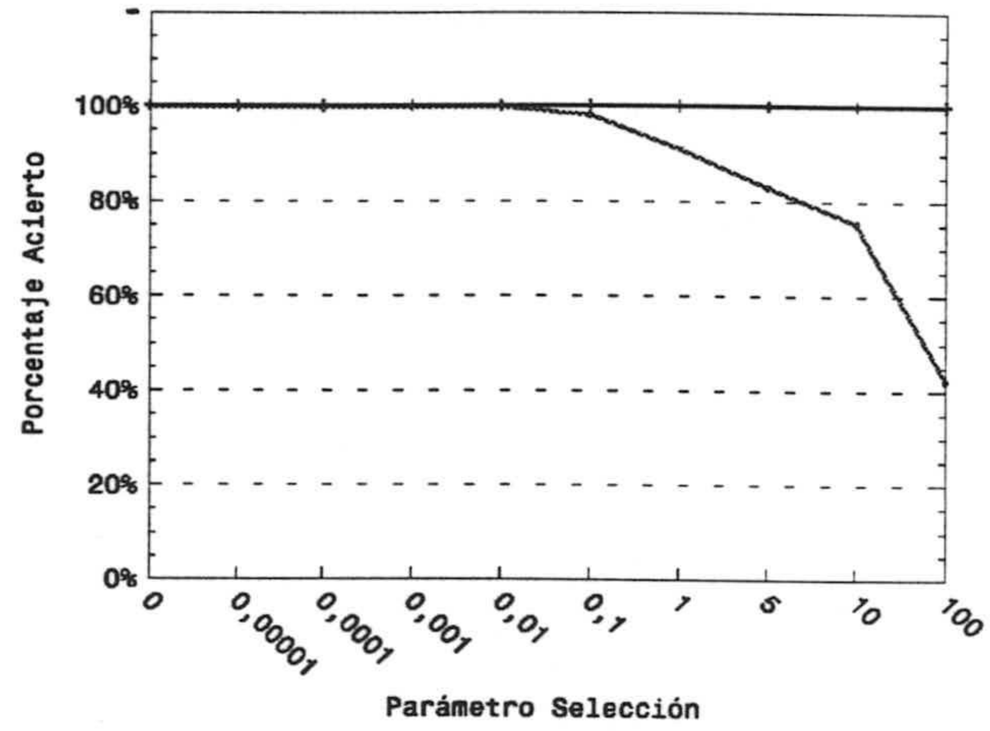


Figura 29: Se muestra la calidad de la máscara para diferentes valores de parámetros ρ (que determinan el umbral de selección) y para distintos tios y cantidades de ruido impulsivo. La imagen es de clase texto mecanografiado, y el tamaño de ventana utilizado es 5×5. Los resultados para ruido *sal y pimienta* son de un 100% de acierto para valores $0.01 \leq \rho \leq 1$, independientemente de la cantidad de distorsión. Para ruido impulsivo uniformemente distribuído, las mejores máscaras se obtienen para ρ en el entorno de 10, 20 ó 25. En todos los casos, la máscara empleada ha sido binaria. Los mejores resultados se acercan, en el peor de los casos, al 75% de acierto de máscara.

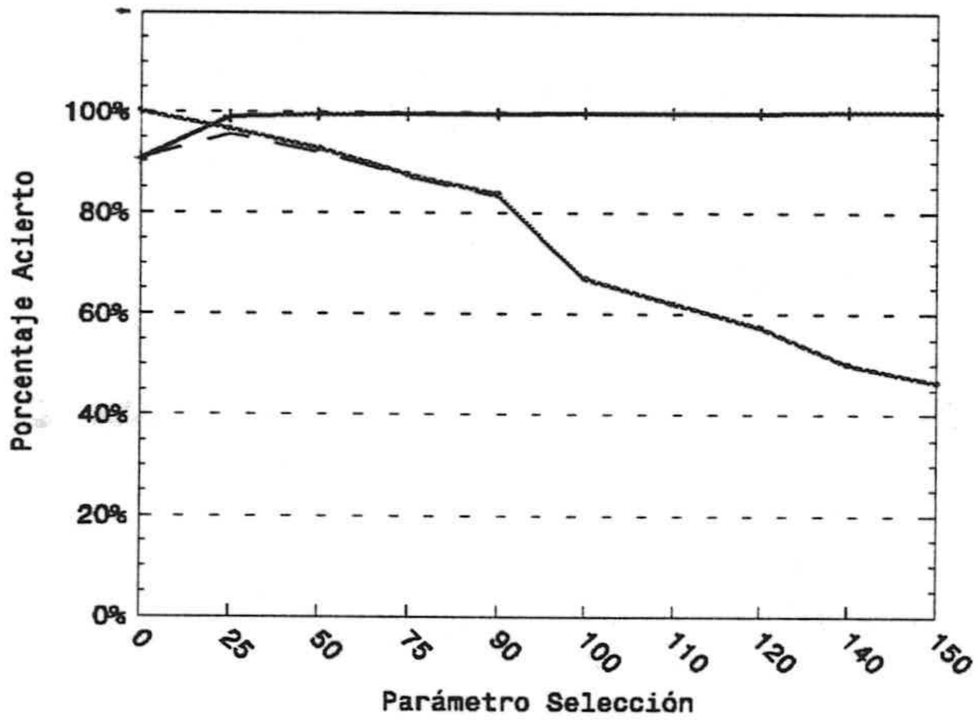




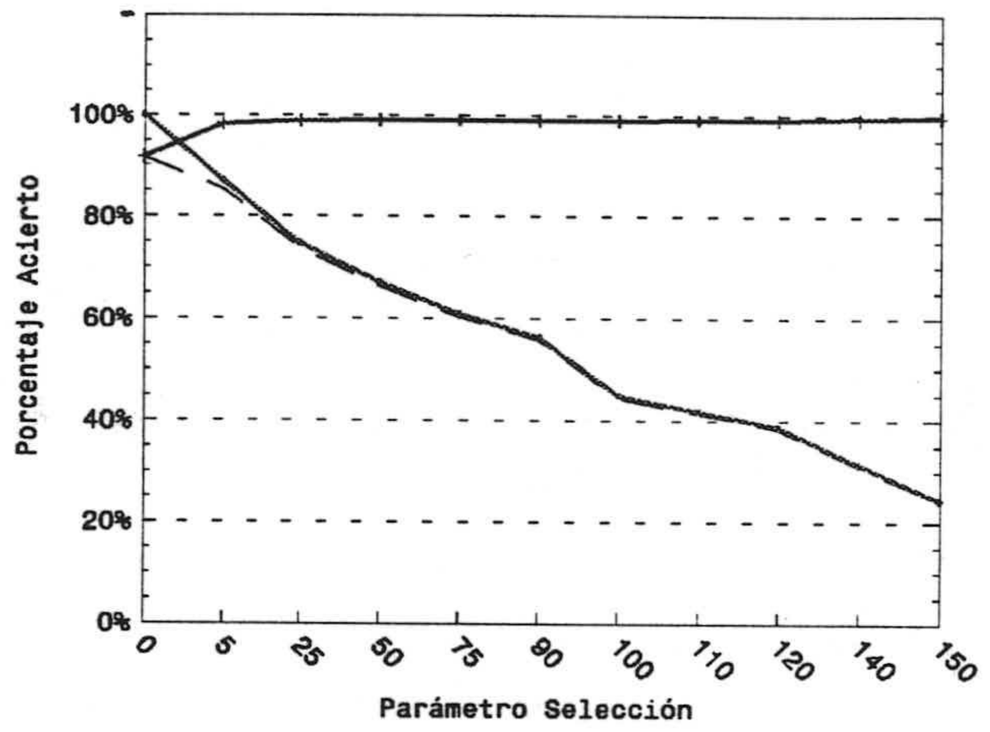
5 % Ruido Impulsivo *sal y pimienta*
[MAE(S-O) = 1.582125]



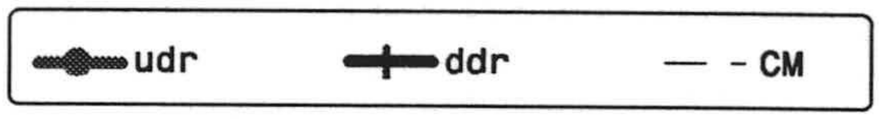
30 % Ruido Impulsivo *sal y pimienta*
[MAE(S-O) = 9.56045]



5 % Ruido Impulsivo *uniforme*
[MAE(S-O) = 1.30025]



20 % Ruido Impulsivo *uniforme*
[MAE(S-O) = 5.474025]



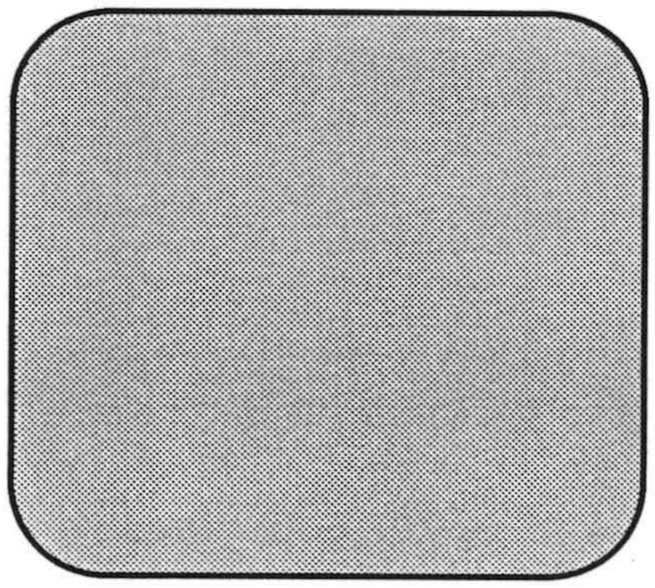


Figura 30: Aparecen gráficas que muestran la influencia del umbral de selección, determinado por el parámetro ρ, para el caso de una imagen de la clase radiografías, en la que se ha empleado una ventana de tamaño 5×5. Las gráficas superiores ilustran el problema para ruido impulsivo *sal y pimienta*, de las cuales se desprende que un valor muy pequeño de ρ es lo más indicado para obtener una máscara de calidad. En las gráficas inferiores aparece el comportamiento frente a ruido imulsivo uniformemente distribuído. Cuando la imagen está poco afectada por la distorsión, la mejor máscara se obtiene para ρ en torno a 25. Pero si el ruido es más abundante, lo mejor es ρ = 0. En todos los casos los mejores resultados superan el 90% de acierto. Todas las máscaras empleadas son binarias.

INFLUENCIA DE LA CALIDAD DE LA MATRIZ DE COMPOSICIÓN EN LA MÁSCARA DE ESTIMACIÓN DE RUIDO

La matriz de composición $Q(C)$ juega un papel muy importante en la construcción de la máscara de preprocesamiento. Es de esperar que cuanto mayor sea la calidad de la matriz de composición, mejor sea la calidad de la máscara, en el sentido de que mejor detectará el ruido, según indican los valores de *ddr, udr* y *CM*. Este comportamiento esperado es lo que se aprecia numéricamente en la tabla 2.

Q(O)	JS	udr	ddr	Tamaño Ventana
texto	1,65	99,28	49,76	3x3
ecografía	0,75	100	79,82	3x3
radiografía	0,03	96,91	96,68	3x3
texto	0,8	99,93	46,59	5x5
ecografía	0,48	99,97	77,54	5x5
rediografía	0,05	96,72	96,58	5x5

Tabla 2: Se muestra la dependencia de la matriz de composición en la calidad de la máscara de forma numérica. Para ello se ha construído una máscara binaria empleando el modelo indicado. Los resultados de *udr* y *ddr* son elocuentes.

Se construyó la máscara $\underline{M}$ mediante el método ya explicado, a partir de una imagen fuente con 20% de ruido impulsivo uniforme y un parámetro de selección de 0% y diversos tamaños de ventana. La razón por la que se eligió ese valor de parámetro de selección se expondrá más adelante. La calidad de $Q(C)$ se evaluó empleando la divergencia de Jensen-Shannon. De los valores de la tabla se puede inferir el hecho de que cuanto menor es el valor de JS, mejor es el ajuste de la máscara de estimación de ruido.

La influencia del modelo sobre la calidad de la máscara $\underline{M}$ puede ser analizada desde otro punto de vista. Empleando la misma imagen fuente S, perteneciente a una clase C, es posible construir una nueva máscara de preprocesamiento $\underline{M'}$, pero utilizando una matriz de composición representativa de una clase diferente C'. Es de esperar que la máscara $\underline{M}$ obtenida a partir de $Q(C)$ posea mayor calidad que la $\underline{M'}$ construida empleando $Q(C')$, por razones de compatibilidad modelo-máscara. Este comportamiento se ilustra en el ejemplo cuyos resultados aparrecen en la tabla 3, en el que la imagen fuente S es una radiografía contaminada por un 30% de ruido impulsivo uniforme.

Tabla 3: Efecto de emplear un modelo de una clase diferente a la de la imagen que se procesa.

Modelo Empleado	*udr*	*ddr*
Radiografía Q(O)	90,62	95,77
Ecografía Q(O')	51,78	75,23

El tamaño de la ventana es de 5 × 5 y se usaron modelos de radiografías, de textos y de ecografías para fabricar la máscara. Como se aprecia, la calidad de la máscara está directamente emparentada con la clase de imágenes cuyo modelo se emplea al construírla.

CONCLUSIONES

En este capítulo se ha presentado un nuevo método de preprocesamiento aplicable al ruido impulsivo, consistente en la localización previa de los pulsos, para así filtrar tan sólo las distorsiones de la imagen fuente. Este objetivo se logra mediante una máscara binaria, denominada máscara deestimación de ruido, $\underline{M}$, en la cual se determina el carácter de cada píxel (ruido o no ruido). La construcción de la citada máscara se fundamenta en el criterio de preselección de píxeles, que emplea información apriorística acerca de la probabilidad de aparición de un cierto nivel de gris k dentro de una región de nivel de gris k', la cual viene proporcionada por la matriz de composición, o modelo.

La obtención de la máscara $\underline{M}$ ha de realizarse mediante un proceso iterativo, dado que la presencia de ruido falsea la información proporcionada por el modelo. La convergencia se asegura mediante la introducción de los conceptos de descartes y persistencia, que se han mostrado ciertamente útiles. Como

medida de calidad de máscara se ha escogido el acierto en la detección de píxeles limpios y ruidosos (*udr* y *ddr* respectivamente); y una medida conjunta, definida como el producto de las anteriores cantidades, denominada *CM* (calidad de máscara).

Además de la capacidad de localizar adecuadamente los pulsos en la imagen fuente, una buena máscara habrá de ser capaz de corregir el sesgo introducido en la LUT por la presencia de ruido. Esta corrección se lleva a cabo mediante una operación de promedio simple, que da lugar al nivel de gris estimado de ventana, el cual constituye la entrada de la LUT. Dado que cuando las distorsiones son pulsos de nivel de gris uniformemente distribuído la detección de las mismas es más difícil, y por lo tanto la corrección de la entrada de la LUT es menos fiable, se diseñó una máscara de niveles de gris —ya no binaria— en la cual cada píxel tenía un nivel de gris proporcional a la probabilidad de aparición en la región centrada en él. De esta manera el sesgo en la LUT es menor, pero las mejoras que proporciona este tipo de máscara no compensan la complejidad introducida, por lo que ha sido abandonado su uso práctico.

Las máscaras binarias tienen una calidad mejor para ciertos valores del parámetro ρ, el cual determina el umbral de selección a partir del cual se considera que un píxel es limpio. Se ha calculado el valor de este parámetro empíricamente, mostrando a su vez la inflencia del mismo tanto en la capacidad correctora como en la localizadora de la máscara. Asímismo se ha mostrado la dependencia de la calidad de la máscara con la calidad del modelo empleado.

En conclusión, tanto la filosofía como el procedimiento de obtención de máscaras se muestra ciertamente eficaz a la hora de marcar los pulsos en una imagen, cumpliendo además la tarea de permitir una corrección del nivel de gris de ventana de manera excelente. Los resultados obtenidos hablan por sí mismos, mostrando una elevada calidad.

UNIVERSIDAD DE GRANADA
15 JUL. 1998
283

CAPITULO IV

FILTROS BASADOS EN MATRICES DE COMPOSICIÓN

IV.I INTRODUCCIÓN.....287

IV.II DESCRIPCIÓN DE LOS FILTROS.....290

IV.II.1 Prefiltrado 290

IV.II.2 Filtro Máscara-mediana.....292

IV.II.3 Filtro Máscara-Modelo.....294

IV.II.4 Filtro Máscara-Modelo Ponderado.....297

IV.II.5 Filtro de Mediana Combinado.....300

IV.II.6 Filtro Restaura-Modelo.....302

IV.III RESULTADOS Y DISCUSIÓN..... 306

IV.III.1 Discusión..... 329

IV.III.2 Comentarios y propiedades.....333

IV.III.4 Influencia de los parámetros.....335

IV.IV CONCLUSIÓN.....337

INTRODUCCIÓN

Como ya se dijo, los filtros adaptativos, para eliminar ruido impulsivo, deben en primer lugar detectarlos y en segundo lugar, procesarlos. La primera de estas fases fue abordada, para el caso que nos ocupa, en el anterior capítulo. Es necesario ahora tratar el problema de la estimación del nivel de gris de los píxeles que se han marcado mediante el uso de la máscara.

Para llevar a cabo esa labor, es preciso emplear cierto conocimiento acerca de la señal original. Lo mejor es utilizar la información local proveniente de los píxeles vecinos al que se procesa, y que no estén marcados como ruidosos por la máscara. En efecto, se sabe que los píxeles de la imagen original son altamante redundantes debido a la fuerte correlación espacial existente. Así, dada una resolución m', es posible realizar una estimación del nuevo nivel de gris del píxel —a resolución mayor m— en función del de sus vecinos, merced a la filosofía de la multirresolución.

Por lo tanto, todos los filtros que aquí se diseñen habrá de trabajar con esta idea. Harán uso, por lo tanto, de ventanas deslizantes sobre la imagen fuente S, de manera que vayan observando las vecindades adecuadas. Además, emplearán la información contenida en las matrices de composición, no sólo para la construcción de la máscara, sino como fuente de información apriorística en forma de regularidades encontradas sobre imágenes originales de la misma clase que la fuente, S, que se procesa. Esas matrices de composición, o modelos, habrán sido construídos empleando las resoluciones adecuadas al problema; es decir, el tamaño de la ventana usado para su elaboración, y que determina la resolución mínima con que se opera, ha de ser el mismo que el utilizado para definir las vecindades en la fase de filtrado. Por otra parte, las características principales que deben presentar los filtros propuestos son, por un lado, la total supresión del ruido; y por otro, la preservación de los contornos de los detalles de la imagen, en especial los de bajo contraste, que son los que suelen ser suavizados durante el procesamiento.

Una vez aclarados los objetivos, a continuación se describirán los cuatro filtros diseñados basándose en el esquema de las matrices de composición. Los tres primeros calculan la mediana de una determinada distribución, pero teniendo en cuenta la información proporcionada por la máscara de estimación de ruido.. Así, el primero realiza dicha operación sobre los píxeles encerrados por una ventana deslizante; el segundo sobre una columna del modelo; mientras el tercero lo hace sobre una tabla de probabilidades construída a partir de la matriz de composición. Para cada filtro se ofrecerán ejemplos del modo de operación empleando imágenes que pertenezcan a una determinada clase C. Los resultados son visiblemente buenos, y comparables con los obtenidos a partir de otros filtros ampliamente utilizados en la literatura. Aplicando criterios expuestos en el primer capítulo, se realizarán aquí comparaciones con los filtros de mediana tradicional, mediana pesada en el centro y el algoritmo adaptativo de Sun & Venetsanopoulos. También aparecerán los resultados del filtro de mediana multietapa y el FMH Junto con los ejemplos visuales se proporcionan resultados numéricos.

Algunas de las razones esgrimidas para escoger los filtros de comparación son los siguientes. Por un lado, se introducen filtros no adaptativos, que no requieren información previa acerca de la imagen ni una computación de estadísticas locales dentro de la ventana —que además no es modificable— dando lugar a algoritmos rápidos y simples. De otra parte, se ha escogido un filtro adaptativo con ventana fija que precisa el cálculo de la esperanza y varianza dentro de la ventana para estimar el ruido y así poder proporcionar un resultado que minimice el error cuadrático medio. De esta manera, los filtros aquí propuestos pueden ser comparados con filtros bien conocidos que mejoran los resultados de la mediana tradicional, uno de los cuale4s emplea información previa como los nuestros. Otros tipos de filtros, como los de mediana híbrida (*FIR-median hybrid*) y multietapa (*multistage*) se han incluido en nuestra comparación a pesar de que preservan estructuras de ruido en lugar de eliminarlas cuando la probabilidad de aparición de ruido impulsivo es alta (20%, 30%), y de que el filtro de mediana multietapa tiene probabilidades de colapso más altas que las de la mediana tradicional, por lo que se comporta peor que ésta cuando el ruido es abundante. Por esta razón nuestros filtros trabajan mucho mejor que el filtro de mediana multietapa, como se demostrará, porque son mejores que el filtro de mediana clásico en cuanto a eliminación de ruido y preservación de bordes.

En cuanto al tipo de ruido empleado en nuestros experimentos, debe advertirse de nuevo que tan sólo se ha tenido en cuenta el caso del ruido impulsivo, tanto *sal y pimienta* como uniformemente distribuido. Esto es debido a la carencia de métodos de procesamiento que sean realmente efectivos frente a la presencia de tales distorsiones, y que mantengan a su vez unas propiedades adecuadas de preservación de bordes. Por esa razón se diseñó la máscara de estimación de ruido estudiada en el anterior capítulo, que sólo puede definirse si el ruido no afecta más que a una fracción de los píxeles de la imagen, lo cual sólo sucede en el ruido impulsivo.

DESCRIPCIÓN DE LOS FILTROS

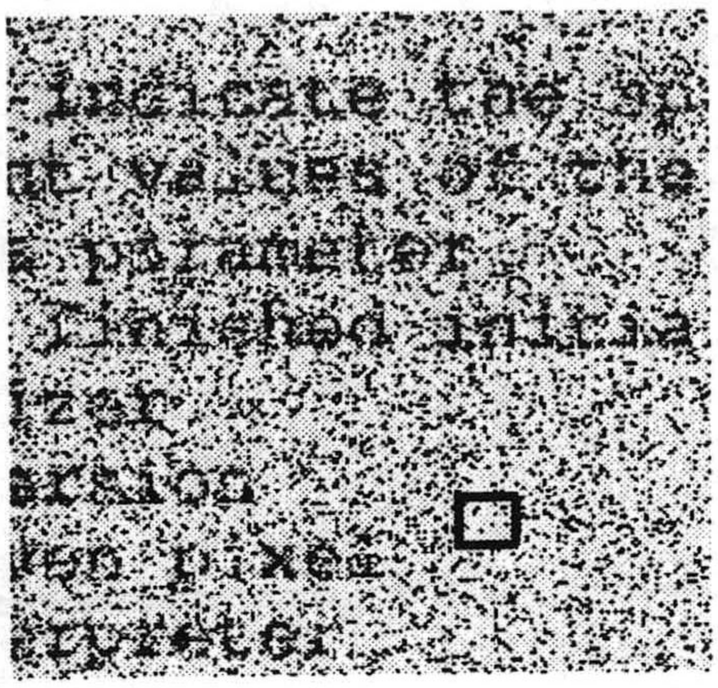
PREFILTRADO

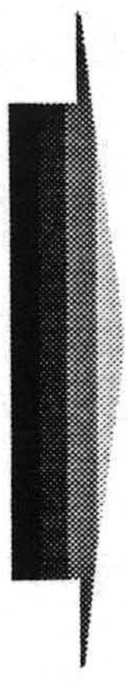
Todos los filtros propuestos se fundamentan en el empleo de la máscara de estimación de ruido. Por lo tanto, existe un paso previo al procesamiento que es común a todos los algoritmos. Este prefiltrado consiste, básicamente, en la construcción de $\underline{M}$. En el esquema del lateral de esta página se ilustra este proceso.

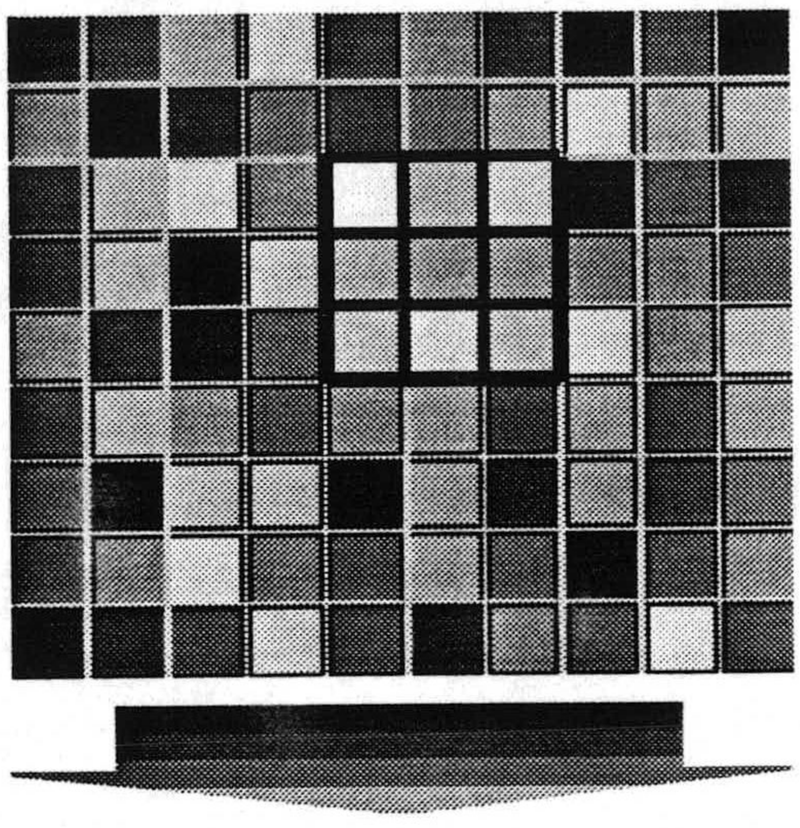
Se realiza un desplazamiento de la ventana deslizante sobre la imagen fuente, calculándose el nivel de gris de ventana k' en cada punto. A partir de dicho dato se consulta la tabla de acceso directo (LUT) que ha sido construída a partir de las probabilidades que conforman la matriz de composición que caracteriza la clase a la que pertenece la imagen que se procesa. Mediante la elección de un umbral adecuado, como se explicó en el anterior capítulo, se consigue seleccionar aquéllos píxeles de la imagen que poseen menos probabilidades de pertenecer a la original, de acuerdo con el modelo. De este modo se obtiene una imagen binaria: la máscara de estimación de ruido, $\underline{M}_0$.

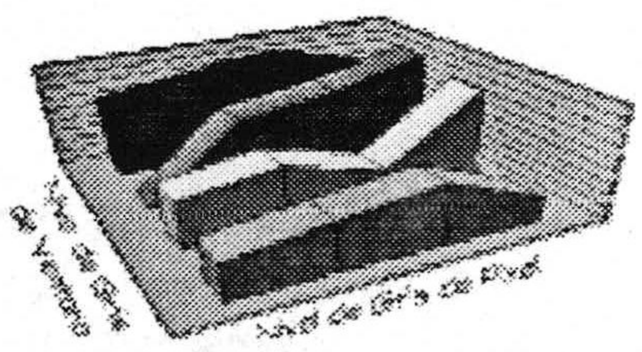
Sin embargo, este resultado no es satisfactorio. Casi todos los píxeles resultan ser sospechosos de estar afectados por el ruido impulsivo. Ello es debido al error cometido al acceder a la LUT, puesto que como dato de entrada se le ha proporcionado un nivel de gris k' que no corresponde a la columna adecuada del modelo. En efecto, ese nivel k' fue computado a partir de la imagen fuente, contaminada de ruido, que entra a formar parte de la región cubierta por la ventana, y por tanto contribuye al nivel de gris de ventana. Como la LUT se construye a partir de la matriz de composición, que a su vez se obtiene empleando imágenes originales sin ruido, el conflicto es evidente.

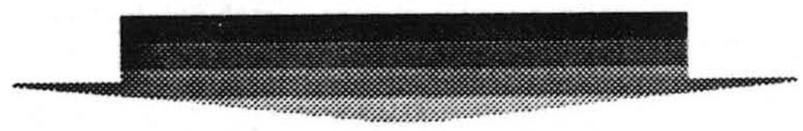
Siguiendo el método iterativo propuesto en el anterior capítulo, se puede realizar la corrección del nivel de gris k' de manera que la entrada de la LUT sea más acertada. De este modo se consiguen resultados mucho más cercanos a la realidad, y la máscara inicial $\underline{M}_0$ se mejora hasta obtener la máscara definitiva tras un número definido de iteraciones.

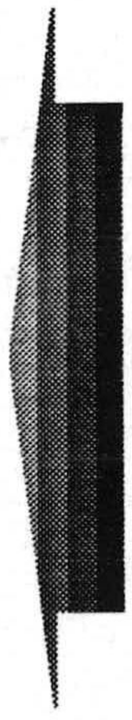












PREFILTRADO
LUT
TABLA DE ACCESO DIRECTO

Una vez calculada esta máscara $\underline{M}$, como resultado convergente del proceso iterativo, se está en condiciones de iniciar el procesamiento propiamente dicho, que consiste en la estimación del nivel de gris de los píxeles que, según esta máscara, están contaminados por ruido impulsivo. A este objetivo dedicaremos el resto del capítulo.

FILTRO MÁSCARA-MEDIANA

Este sencillo filtro no es más que uno de mediana tradicional operando sobre los píxeles que le indica la máscara de estimación de ruido. El modelo $Q(C)$ se emplea aquí tan sólo para construir $\underline{M}$, pero no interviene en el proceso de filtrado. El método que sigue el filtro Máscara-Mediana es simple: para cada posición (i,j) marcada como ruido por la máscara ($\underline{M}_{ij} = 0$) se computa la mediana de la ventana centrada $W_{ij}(S)$. De este modo, el filtro puede ser descrito como:

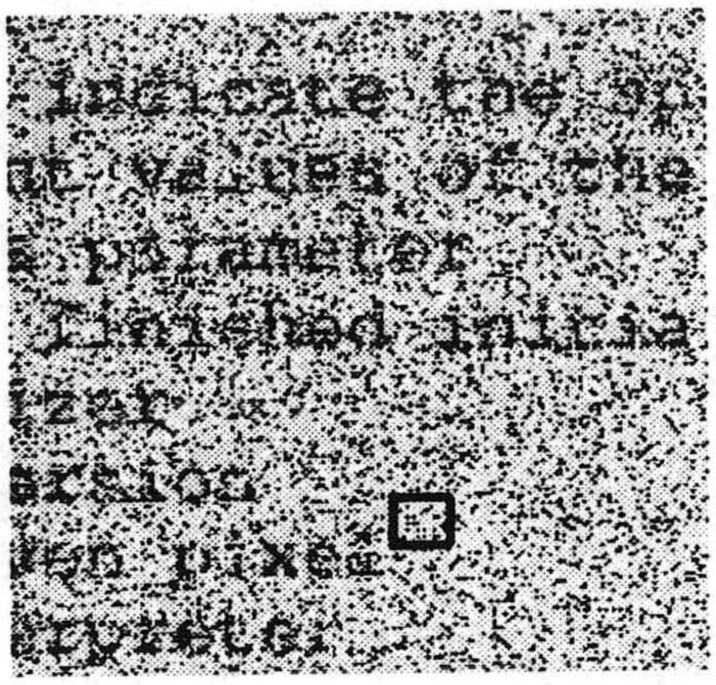
$$R_{ij} = S_{ij} \cdot \underline{M}_{ij} + \text{MED}[W_{ij}(S)] \cdot (1 - \underline{M}_{ij})$$

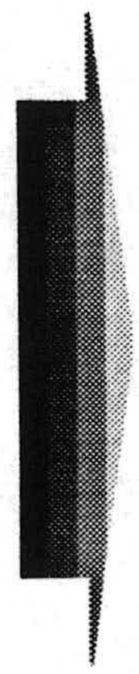
Este esquema de funcionamiento puede apreciarse en el esquema de la página siguiente. Se recorre la imagen fuente mediante una ventana deslizante, pero centrando ésta solamente en aquellos píxeles que fueron dados por malos en la máscara de estimación de ruido, calculada en el prefiltrado. Una vez posicionada la ventana, se realiza un filtrado de mediana tradicional. De esta manera, solventamos la mayor parte de los inconvenientes de la mediana, puesto que al procesar tan sólo aquellos píxeles que realmente lo necesitan, no se producen alteraciones notables en los contornos y detalles de la imagen, pero se mantienen las características de eliminación de ruido.

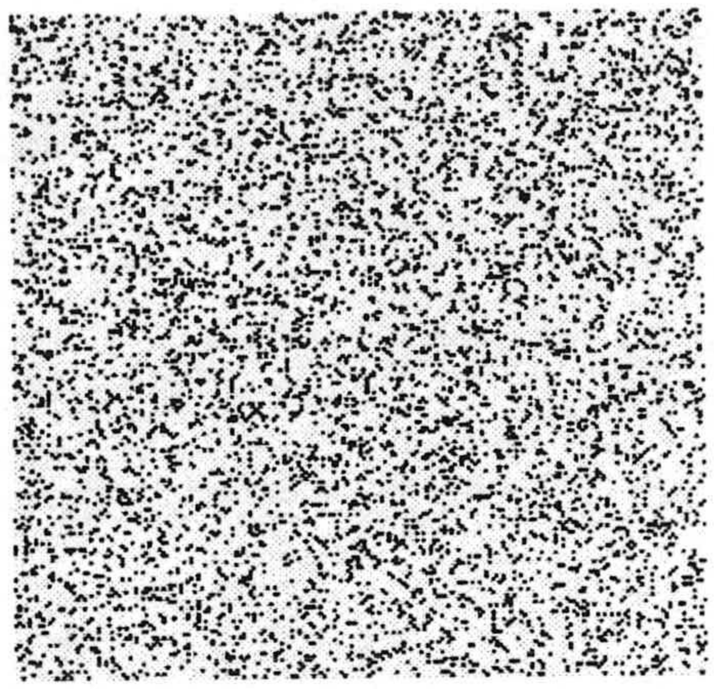
PSEUDOCÓDIGO DEL ALGORITMO DEL FILTRO MÁSCARA-MEDIANA.

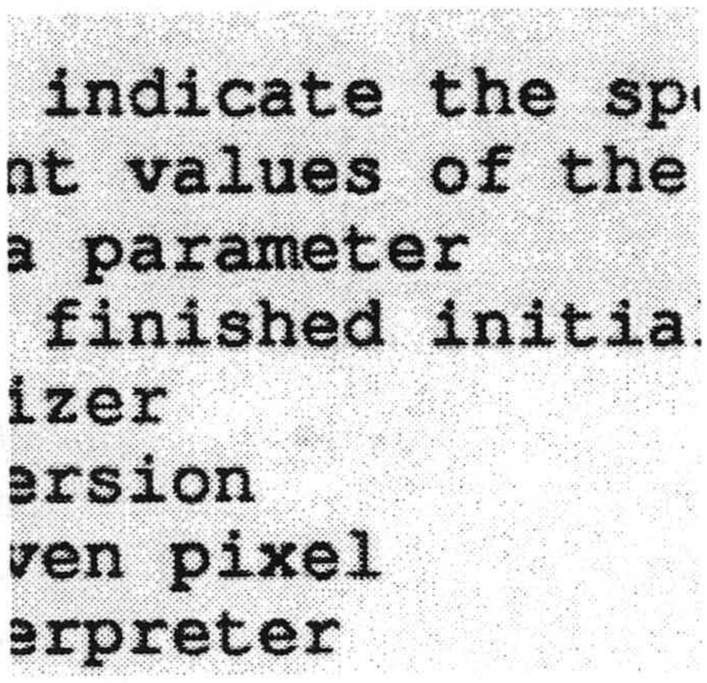
Algorithm *Filtro_Máscara_Mediana*(*Imagen_Fuente*, *Máscara*):Imagen_Resultado;

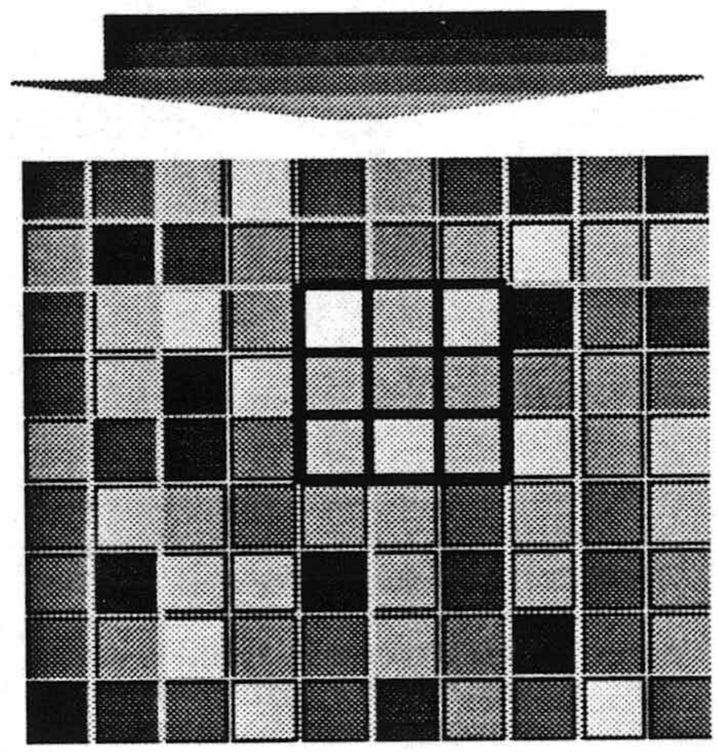
FILTRO MÁSCARA-MEDIANA











```
Begin

     For todas posiciones (i,j) en Imagen_Fuente do

     If Máscara(i,j) = 1              //no ruido

     Then Imagen_Resultado(i,j):= Imagen_Fuente(i,j);

     Else Imagen_Resultado(i,j) :=
          := Mediana[ventana centrada en (i,j)];

End;
```

FILTRO MÁSCARA-MODELO

Este filtro sustituye el nivel de gris contaminado por la mediana, pero no de los valores de niveles de gris de los píxeles cubiertos por la ventana deslizante, sino de la columna del modelo $Q_{k'}$ correspondiente al nivel de gris de ventana computado. Como ésta es una distribución de probabilidad, su mediana es el nivel de gris correspondiente a la probabilidad acumulada de 0.5.

Los pasos del filtro aparecen esquemáticamente en la página siguiente, y pueden especificarse como sigue:

❶ ***Estimación del nivel de gris de la ventana.*** Para cada posición de la ventana (i,j) tal que $\underline{M}_{ij} = 0$, se calcula el *nivel de gris corregido de ventana*, $\underline{k'}$, definido como:

$$\underline{k'} = \underline{\Sigma}\left[W_{ij}(S), \underline{M}\right] = \Sigma\left[W_{ij}(S \cdot \underline{M})\right] \frac{(2v_w + 1) \times (2h_w + 1)}{\Sigma\left[W_{ij}(\underline{M})\right]}$$

Es decir, se suman los niveles de gris de todos los píxeles dentro de la ventana dados por buenos por la máscara, para a continuación estimar el nivel de gris de la ventana total mediante un promedio.

❷ ***Mediana de la columna de la matriz de composición.*** El modelo $Q(C)$ proporciona la disribución de probabilidad para cada nivel de gris de píxel, k. Una vez que se conoce la estimación del nivel de gris de ventana $\underline{k'}$, se busca la correspondiente columna del modelo, suponiendo que la estimación es correcta $(\underline{k'} = k')$. Entonces el filtro devuelve una estimación del nivel de gris del píxel central de la ventana:

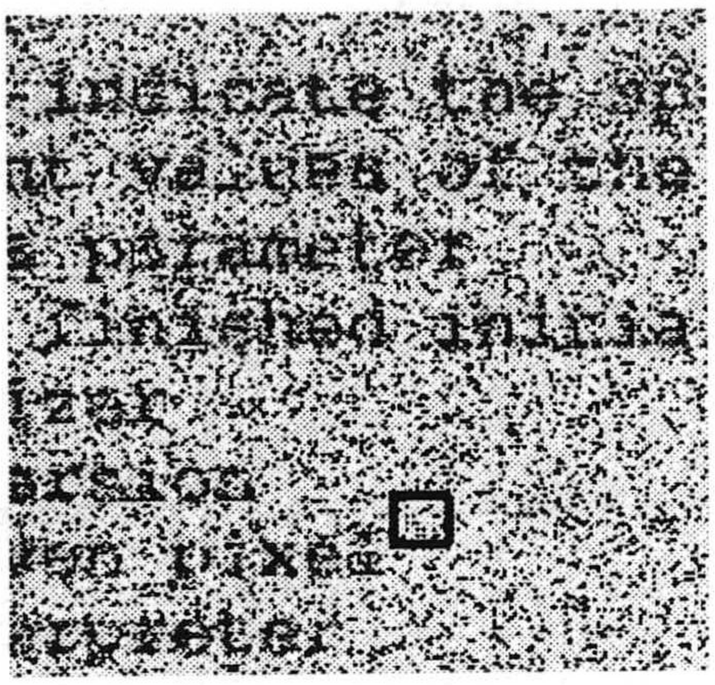
$$\underline{k} = \text{MED}[Q_{k'}]$$

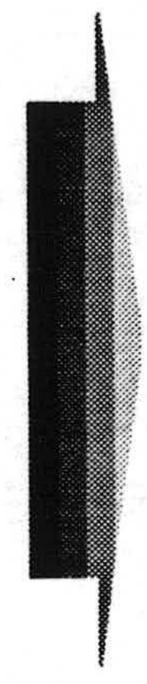
En resumen, la operación de filtrado puede expresarse como:

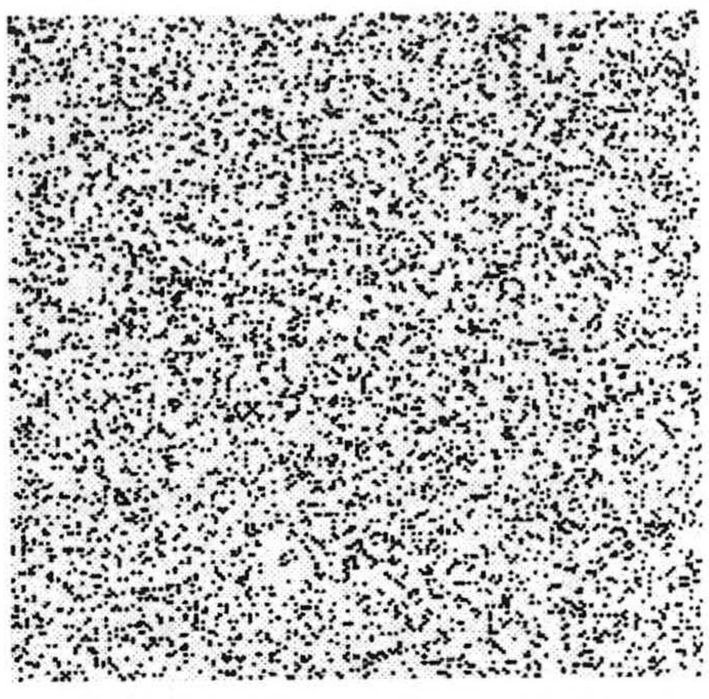
$$R_{ij} = S_{ij} \cdot \underline{M}_{ij} + \text{MED}[Q_{k'}] \cdot (1 - \underline{M}_{ij})$$

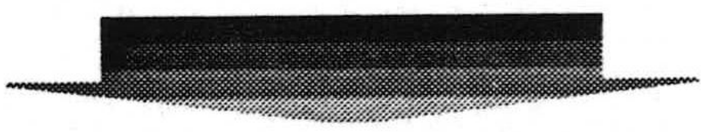
donde $k' = \underline{k'}$, el cual se calcula como se indica en ❶.

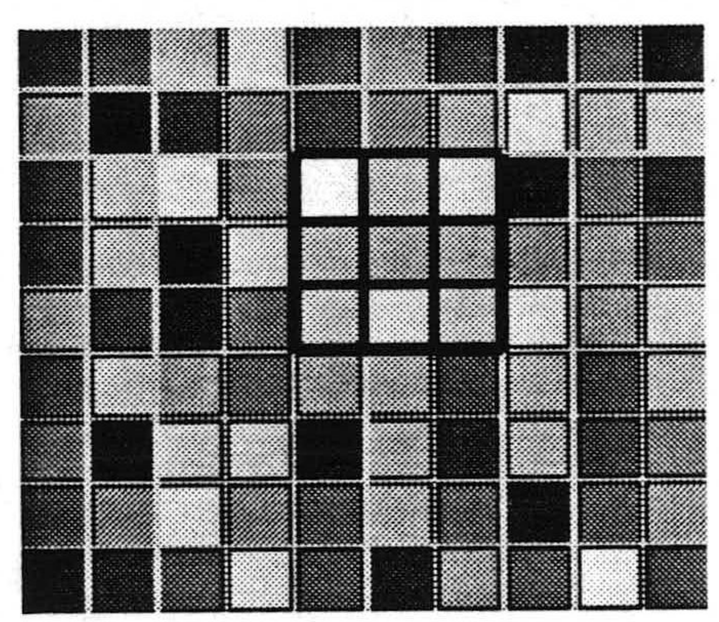
FILTRO MÁSCARA-MODELO

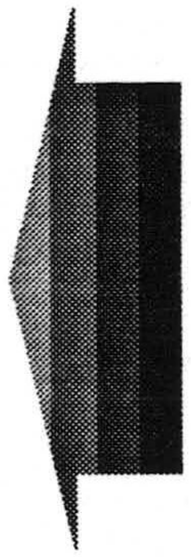


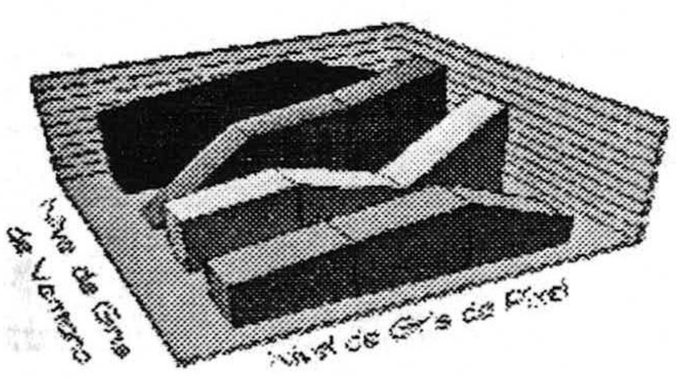


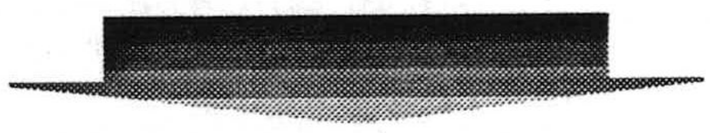




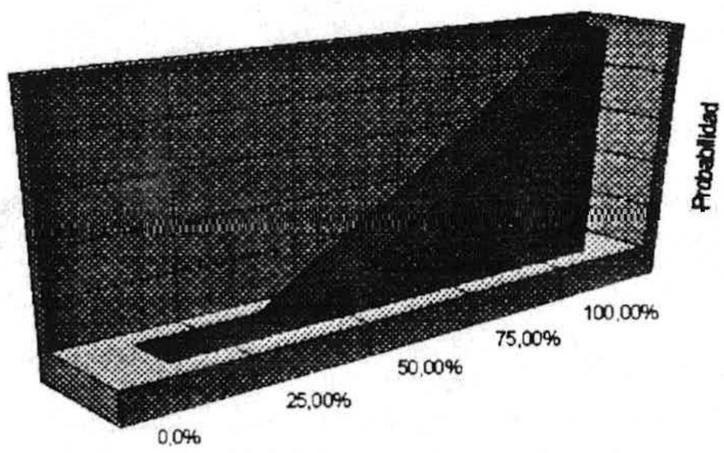


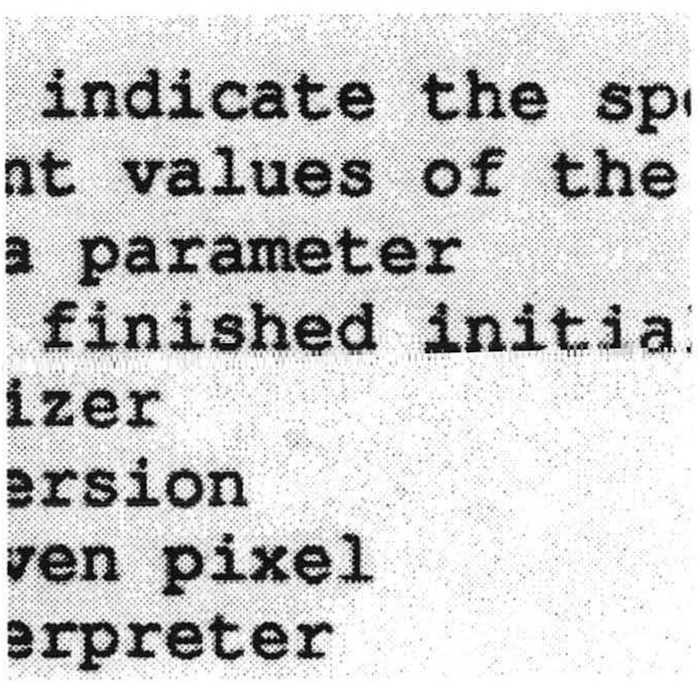






Ejemplo de columna de modelo





PSEUDOCÓDIGO DEL ALGORITMO DEL FILTRO MÁSCARA-MODELO

```
Algorith Filtro_Máscara-Modelo(Imagen_Fuente, Más-
cara): Imagen:Resultado;

Begin

     For toda posición (i,j) en Imagen_Fuente do

     If Máscara(i,j) := 1          //no ruido

     Then Imagen_Resultado(i,j):=Imagen_Fuente(i,j);

     Else begin

          k':=Σ[ventana centrada en (i,j),Máscara];

          if Columna_Matriz[k']<>Null

          then Imagen_Resultado(i,j):=
               := Mediana[Columna_Matrix[k']];

          else Imagen_Resultado(i,j):=
               := Mediana[ventana centrada en (i,j)];

     end;

End;
```

Como veremos en los ejemplos, los detalles de la imagen original se preservan mejor, y se aprecia la ausencia de blotches en el fondo cuando se emplea este filtro.

FILTRO MÁSCARA-MODELO PONDERADO

Este algoritmo surge como refinamiento del anterior, al tener en cuenta también el histograma de la ventana, pesado con las probabilidades de la correspondiente columna de la matriz. La operación del filtro aparece en la página siguiente, de forma resumida.

Puede describirse como sigue: el histograma de la ventana $\mathcal{P}[W_{ij}(S \cdot \underline{M})]$ es ponderado por la distribución de probabilidad que constituye la columna $Q_{k'}$ de la matriz de composición correspondiente. El nivel de gris de la ventana k' se obtiene como se indicó en ❶ del anterior apartado. Como resultado se obtiene un vector de $(G + 1)$ componentes:

$$\mathcal{P}^*(\mathcal{P},Q_{k'}) = \mathcal{P}[W_{ij}(S \cdot \underline{M})] \cdot Q_{k'}$$

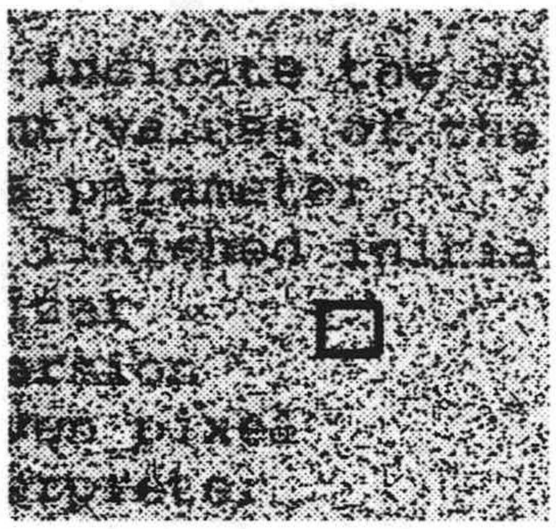
Realizando la mediana del histograma de la ventana ponderado se obtiene la salida del filtro. En esencia:

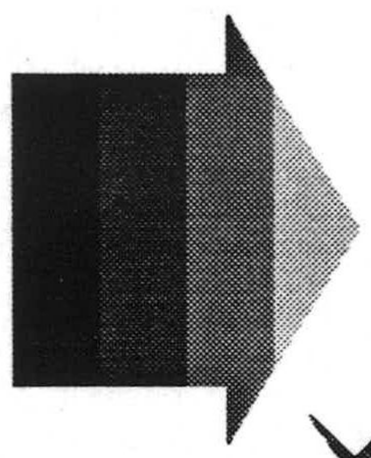
$$R_{ij} = S_{ij} \cdot \underline{M}_{ij} + \text{MED}[\mathcal{P}^*] \cdot (1 - \underline{M}_{ij})$$

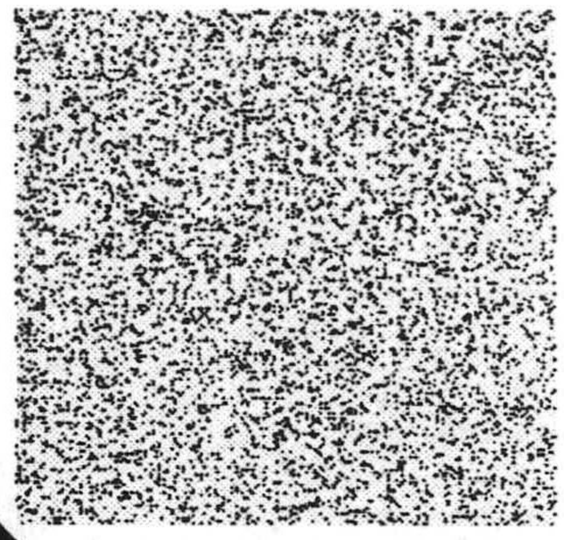
PSEUDOCÓDIGO DEL ALGORITMO DEL FILTRO MÁSCARA-MODELO PONDERADO.

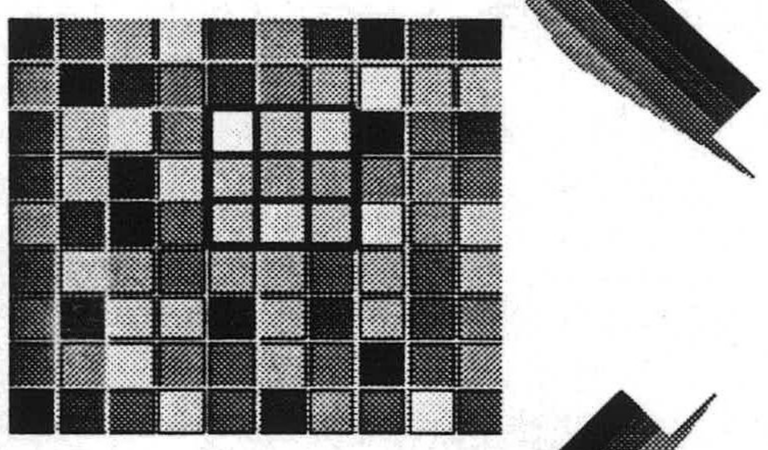
```
Algorithm Filtro_Máscara-Modelo_Ponderado
(Imagen_Fuente, Máscara): Imagen_Resultado;

Begin

    For toda posición (i,j) en Imagen_Fuente do

    If Máscara(i,j) := 1          //no ruido

    Then Imagen_Resultado(i,j):=Imagen_Fuente(i,j);

    Else begin

         k':=Σ[ventana centrada en (i,j),Máscara];

    if Columna_Matriz[k']<>Null

    then begin

         Histograma_Ponderado :=
              := Histograma[ventana  centrada
              en (i,j), Máscara];
```

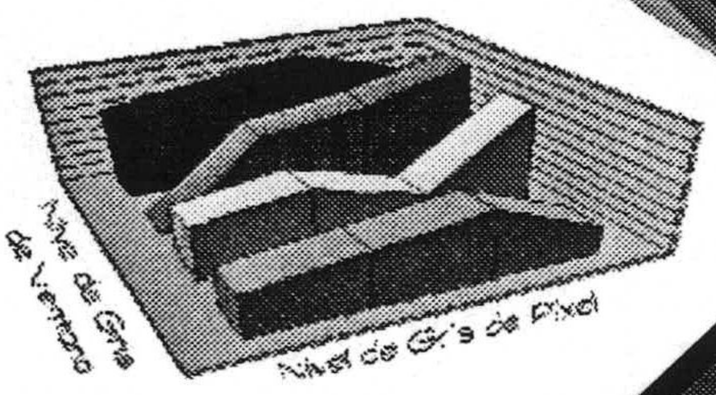
FILTRO MÁSCARA-MODELO PONDERADO



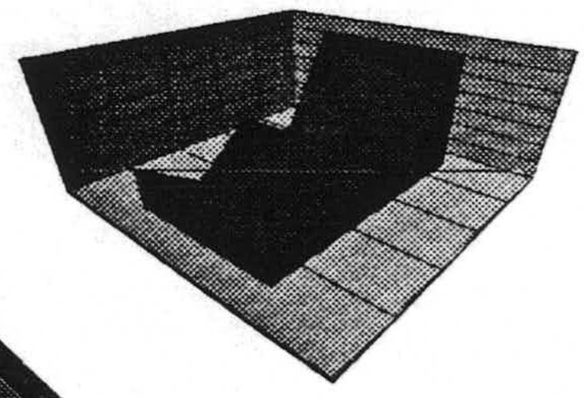








HISTOGRAMA DE VENTANA



PRODUCTO

MEDIANA

```
                Imagen_Resultado(i,j):=
                     := Mediana[Columna_Matrix[k']];

          end;

          else Imagen_Resultado(i,j):=
               := Mediana[ventana centrada en (i,j)];

     end;

End;
```

Obsérvese el gran parecido con el anterior algoritmo. Es por ello por lo que hemos considerado que el filtro Máscara-Modelo Ponderado es un tipo especial de filtro de Máscara-modelo. Pero la idea de ponderar el histograma de la ventana con las probabilidades proporcionadas por la matriz de composición puede interpretarse también como un filtro WM, donde los pesos en este caso tienen una interpretación física, esto es, no son *ad hoc*.

En efecto, procediendo de la manera expuesta, se da mayor importancia a aquellos niveles de gris que tienen más probabilidad de aparecer en la imagen original, de acuerdo con el modelo, lo cual tiene plena justificación.

Sin embargo, es posible que se produzca un efecto inconveniente, conocido por lo sque emplean métodos numéricos de resolución de problemas. Se trata del problema de realizar una ponderación excesivamente violenta, con lo que los resultados serán desvirtuados. Este riesgo no es despreciable, desgraciadamente. Obsérvese que las columnas de las matrices de composición, al menos de las empleadas en los experimentos realizados en el presente trabajos, suelen ser bastante degeneradas, es decir, tienen pocos valores muy elevados, y la mayoría son despreciables e incluso nulos. Esta característica de los modelos provocará que al realizar el producto del vector del histograma de ventana con el vector de columna de la matriz de composición, la mayor parte de los niveles de gris se desprecien, dando lugar a resultados espúreos. En un intento de solucionar este problema, se diseñó otro filtro, que a continuación describiremos.

FILTRO DE MEDIANA COMBINADO

Aunque prevalece la idea básica del anterior filtro, el modo de llevarla a cabo es sutilmente diferente. En efecto, en lugar de realizar el producto de los vectores histograma de ventana y columna del modelo, como antes, ahora se lleva a cabo la siguiente operación:

$$\mathcal{P}^*(\mathcal{P},Q_{k'}) = \mathcal{P}[W_{ij}(S \cdot \underline{M})] \cdot (G^{-1} - Q_{k'})$$

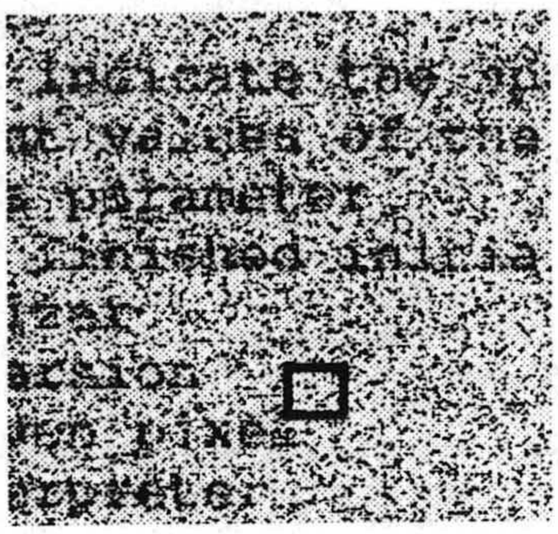
De este modo, no se despreciará ningún nivel de gris, puesto que la lista a la que se realiza la mediana tiene ahora un aspecto más uniforme. La elección del elemento que se suma a la probabilidad proporcionada por la matriz de composición es, en principio, arbitraria, siempre y cuando su valor no sea excesivo (en cuyo caso tenderíamos al filtro de mediana tradicional) ni nulo (pues entonces se obtendría el filtro Máscara-Modelo). En nuestro caso se ha escogido el sumando $1/G$, puesto que de ese modo el resultado tiene cierta independencia del tamaño de la escala de niveles de gris empleada. Entonces el filtro opera de esta manera:

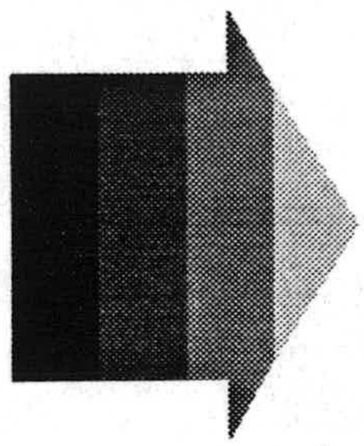
$$R_{ij} = S_{ij} \cdot \underline{M}_{ij} + \text{MED}[\mathcal{P}^*] \cdot (1 - \underline{M}_{ij})$$

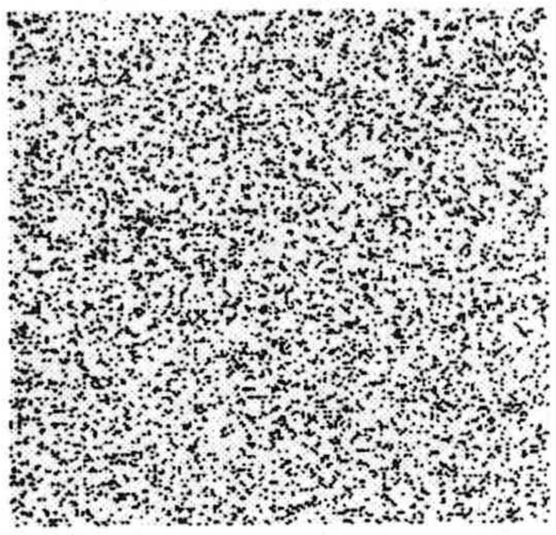
PSEUDOCÓDIGO DEL ALGORITMO DEL FILTRO DE MEDIANA COMBINADO.

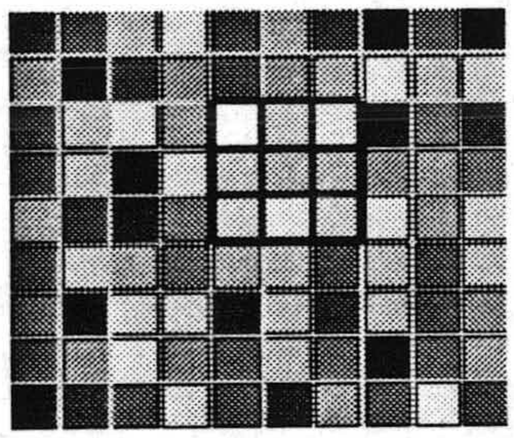
```
Algorithm Filtro_Mediana_Combinado (Imagen_Fuente,
Máscara): Imagen_Resultado;

Begin
     For toda posición (i,j) en Imagen_Fuente do
     If Máscara(i,j) := 1          //no ruido
     Then Imagen_Resultado(i,j):=Imagen_Fuente(i,j);
     Else begin
          k':=Σ[ventana centrada en (i,j),Máscara];
          if Columna_Matriz[k']<>Null
          then begin
               Histograma_Ponderado :=
               := Histograma[ventana centrada en
               (i,j), Máscara];
               Imagen_Resultado(i,j):=
                    := Mediana[Columna_Matrix[k']];
               end;
               else Imagen_Resultado(i,j):=
               := Mediana[ventana centrada en (i,j)];
          end;

End;
```

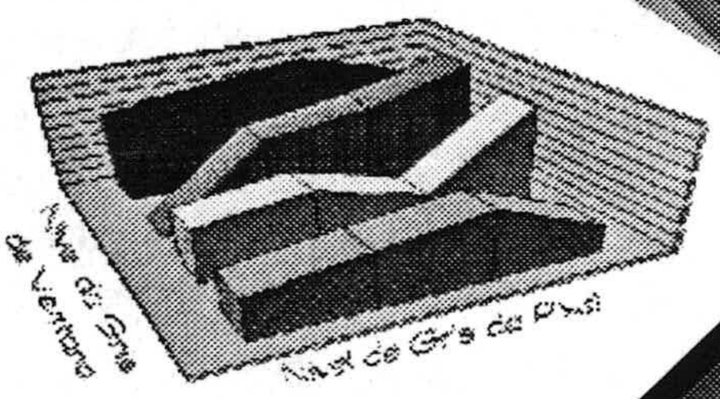
FILTRO MEDIANA COMBINADO



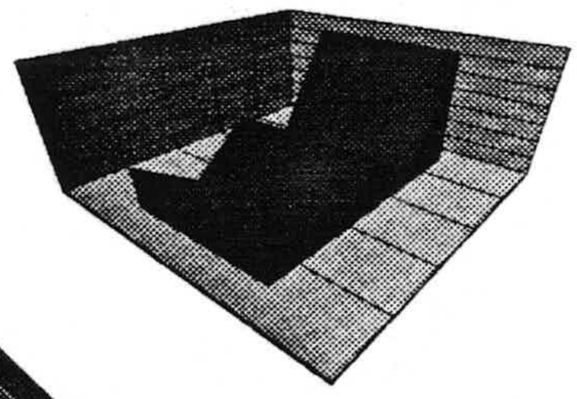






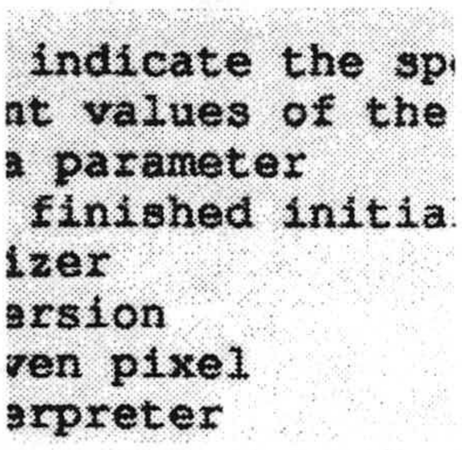


HISTOGRAMA DE VENTANA



COMBINAR

MEDIANA



FILTRO RESTAURA MODELO

El procedimiento que se presenta a continuación resulta más elaborado que los anteriores. Se ha implementado un criterio de proximidad entre la imagen filtrada y el modelo que se emplea en el procesamiento. El nivel de gris final del píxel filtrado es aquel que maximiza la probabilidad de que la región de imagen cubierta por la ventana pertenezca al modelo con mayor proximidad. Los pasos de la operación, que se ilustran en la página siguiente, se detallan a continuación:

❶ ***Estimación del nivel de gris de la ventana.*** Este paso es totalmente análogo al primero dado en el proceso de filtrado empleando el algoritmo de máscara-mediana, excepto en un detalle. En este caso, no se calcula el nivel de gris corregido de la ventana completa, sino el de la ventana reducida de tamaño $(2v_w + 1)\times(2h_w + 1)$-1:

$$\underline{k'}^* = \underline{\Sigma}^*\left[W_{ij}(S), \underline{M}\right] = \Sigma\left[W_{ij}(S \cdot \underline{M})\right]\frac{(2v_w + 1)\times(2h_w + 1) - 1}{\Sigma\left[W_{ij}(\underline{M})\right]}$$

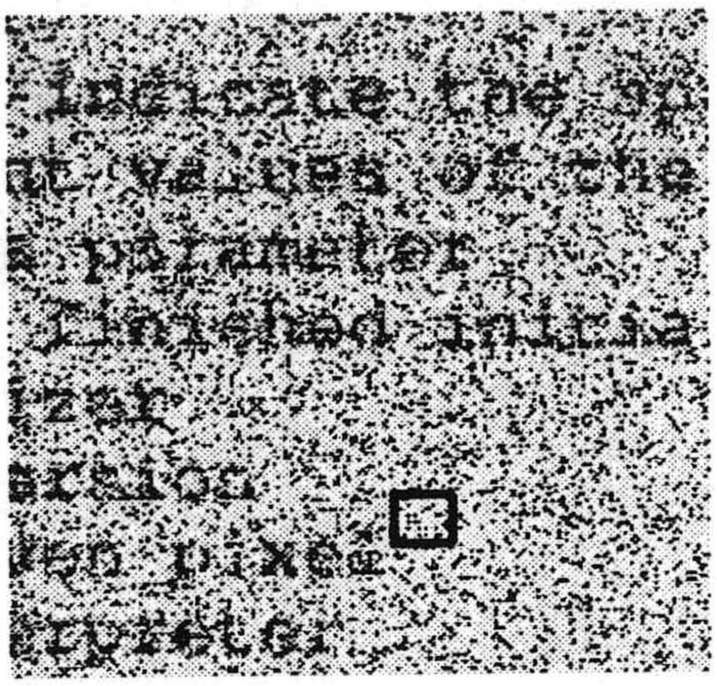
❷ ***Prueba de*** k**.** En este estadio del procesamiento se va añadiendo cada posible valor del nivel de gris del píxel central de la ventana, k =0,1,2,...,G, al nivel de gris estimado $\underline{k}^*$, de modo que el nivel de gris de la ventana completa sería:

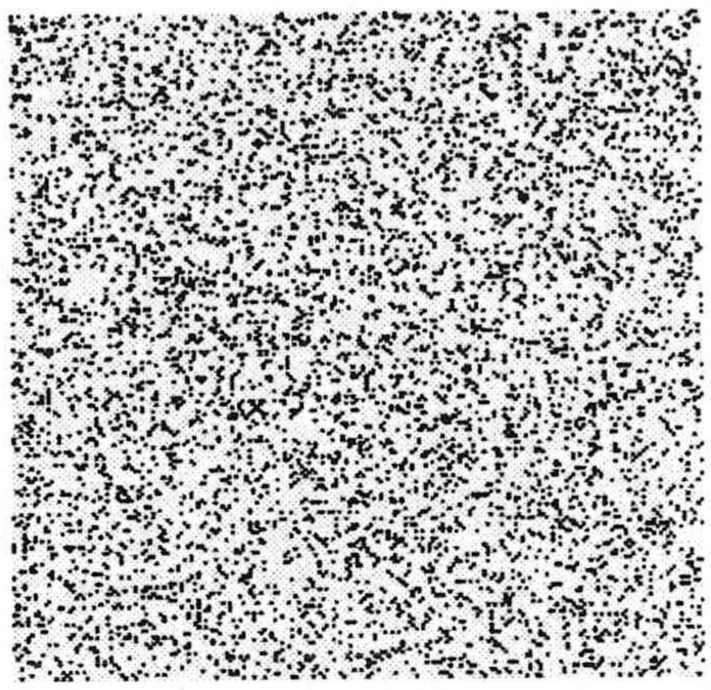
$$k' = \underline{k}^* + k$$

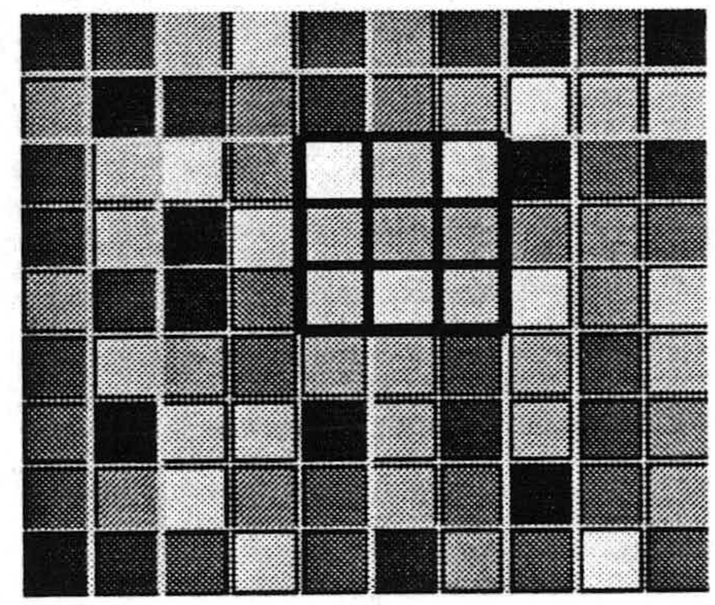
Obsérvese que se obtendrá un conjunto de posibles valores de k', según el nivel de gris de píxel que se le añada.

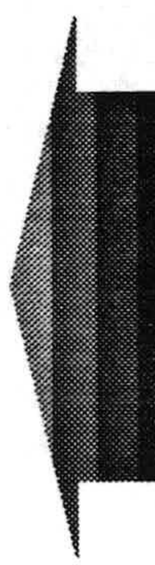
Para cada uno de los valores de nivel de gris de ventana k' obtenidos, la probabilidad de que contengan el nivel de gris de píxel que se está probando, k, viene dado por el correspondiente elemento de la columna de la matriz, $Q_{\underline{k}^* + k'}$, que recordemos era una distribución de probabilidad. Por lo tanto, la probabilidad de que una ventana de nivel de gris $\underline{k}^* + k$ contenga un píxel central de nivel de gris k está situada en un lugar diferente de la matriz de composición para cada k. El conjunto de todas estas probabilidades se sitúa sobre una diagonal del modelo, que empieza en la columna $Q_{\underline{k}'^*}$. Debe notarse que este conjunto de probabilidades no constituye una distribución, ya que la suma de todas ellas no resulta la unidad, al pertenecer cada una de ellas a una columna diferente.

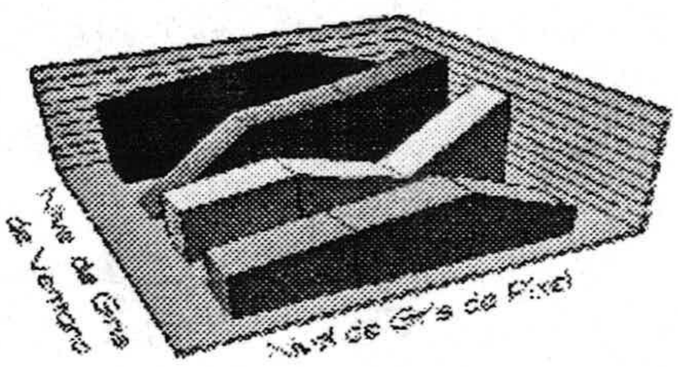
FILTRO RESTAURA-MEDIANA

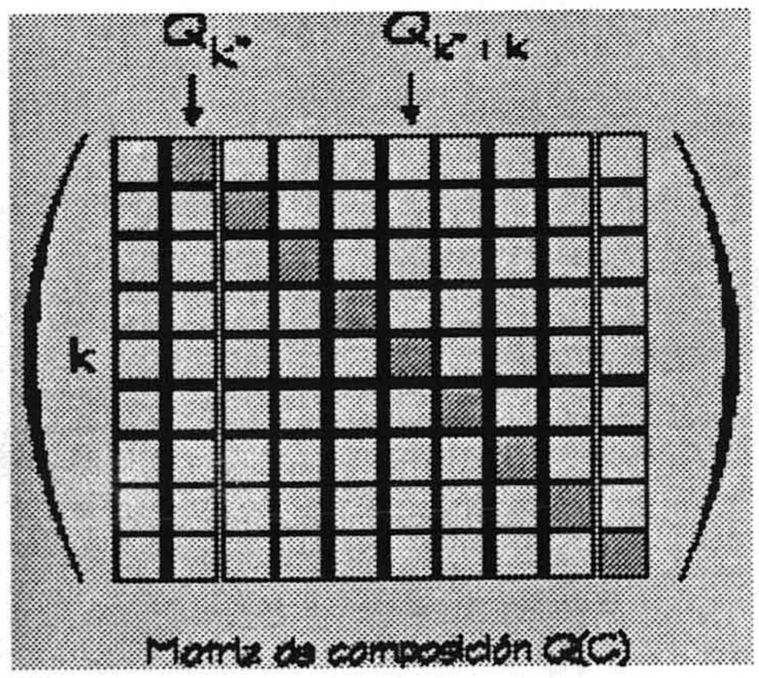


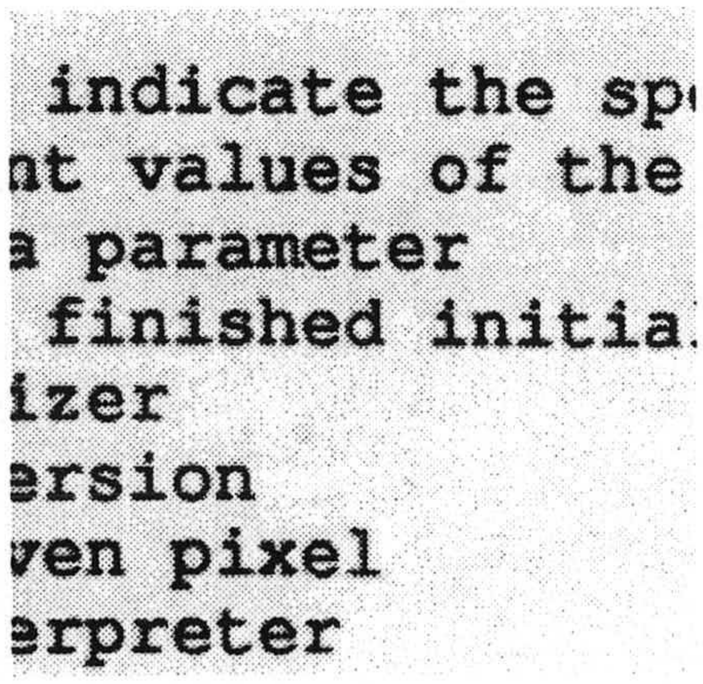












❸ ***Búsqueda del máximo.*** En este momento se determina el máximo de todas las probabilidades de la diagonal. El nivel de gris al que corresponda dicho máximo es la salida del filtro, $\underline{k}$, que se asigna al píxel central de la ventana.

A la vista de lo expuesto, podemos resumir el modo de operación del algoritmo como:

$$R_{ij} = S_{ij} \cdot \underline{M}_{ij} + \underline{k} \cdot (1 - \underline{M}_{ij})$$

donde

$$(Q_{\underline{k}'^* + \underline{k}})_{\underline{k}} \geq (Q_{\underline{k}'^* + k})_k \quad \forall k = 0,1,2,....,G$$

Este filtro puede ser más fácilmente implementado si se construye una Tabla de Consulta de Diagonales (DLUT) a partir de la matriz de composición $Q(C)$. Para cada $\underline{k}^*$ se toma la diagonal hacia la derecha en la matriz, y se determina su elemento de máximo valor. EL valor de nivel de gris de píxel k correspondiente a dicho máximo se asigna a la entrada $\underline{k}^*$ de la DLUT.. El algoritmo completo se expone en el siguiente parágrafo.

PSEUDOCÓDIGO DEL ALGORITMO DEL FILTRO RESTAURA-MODELO.

Algorithm *DLUT* (Matriz_Composición): DLUT;

Begin

For $\underline{k}'^*$ = 0 **to** G*(tamaño_ventana - 1) **do**

DLUT(k') := $\underline{k}$; //definido en el texto

End;

```
Algorithm  Filtro_Restaura-Modelo(Imagen_Fuente,
Máscara): Imagen_Resultado;

Begin

     For toda posición (i,j) en Imagen_Fuente do

     If Máscara(i,j) := 1          //no ruido

     Then Imagen_Resultado(i,j):=Imagen_Fuente(i,j);

     Else begin

          k'* := {como en el texto};

          if DLUT[k'*] <> Null

          then Imagen_Filtrada(i,j) := DLUT[k'*];

          else Imagen_Resultado(i,j):=
               := Mediana[ventana centrada en (i,j)];

     end;

End;
```

RESULTADOS Y DISCUSIÓN

Una vez descritos los algoritmos de filtrado diseñados basándose en la información proporcionada por los modelos, llega el turno de verificar su validez y estudiar sus propiedades. Para ellos se han preparado una serie de experimentos, cuyos resultados son el núcleo del presente apartado.

Clases de imágenes utilizadas en experimentos. Formato y escala de grises

Se ha trabajado con tres clases de imágenes fundamentales: radiografías, ecografías y textos mecanografiados, todas ellas escaneadas a 8 bits, en formato TIFF, y luego la escala de grises se redujo a 6 bits (0-63 niveles) para un manejo más sencillo. De tales clases se han construido las matrices de composición adecuadas, empleando dos tipos diferentes de resolución mínima —recordemos que la máxima siempre es la del nivel de píxel— una con ventana de tamaño 3×3, y otra con 5×5. Además, se ha incluido una tercera imagen, que podría constituir una clase: una fotografía en blanco y negro. De esta última no se ha realizado un modelo empírico, sino que se ha empleado en su procesamiento un modelo teórico. La razón para ello es la siguiente: recordemos que el modelo hipergeométrico se construía cuando no existía ninguna restricción a la composición de las regiones.

Tamaños de ventana empleados

Aparte de por razones estrictamente científicas —e incluso éticas— que obligan al uso del modelo hipergeométrico cuando no existe información previa acerca de la composición de las regiones de una imagen, existe una razón teórica: en una fotografía de un objeto real (no en una imagen artificial, como las de las clases antes aludidas) las regiones se deben componer, en promedio, del mayor número posible de niveles de gris. No hay, en principio, razones para que unos niveles de gris tengan más oportunidades que otros de aparecer en la foto. Es por ello por lo que el modelo empírico para una fotografía como la que se trata sería muy parecido al hipergeométrico.

Por otro lado, el hecho de realizar experimentos con modelos teóricos permite una comprobación de la bondad de los modelos empíricos, que demostrarán ser adecuados para la tarea impuesta.

Tipo de ruido utilizado

Todas las imágenes antedichas han sido contaminadas por diferentes cantidades de ruido impulsivo, tanto *sal y pimienta* como uniformemente distribuido. Se obtienen así las imágenes fuente que constituirán la entrada de los filtros.

Algoritmos empleados en los experimentos

En cuanto a los algoritmos empleados, además de los diseñados en el presente trabajo, y que empleen modelos durante el filtrado o prefiltrado —Restdiver con máscara, Máscara-Mediana, Máscara-Modelo, Máscara-Modelo Ponderado, Mediana Combinada, Restaura-Modelo—, se utilizarán también algunos procesamientos de amplia extensión en la literatura, con objeto de ralizar una comparación de comportamientos, en igualdad de condiciones. Se han escogido el filtro de mediana tradicional [TUK77]; el de mediana pesado en el centro (CWM) [KO91][ZEN94a,b]; el de mediana multietapa [ARC89]; el FIR-*median-hybrid* [NIE87]; y un algoritmo adaptativo [SUN88].

Iteraciones

Por otra parte, es bien sabido que, con frecuencia, la imagen resultado puede mejorarse si se itera el filtrado. Por esta razón se han realizado aquí 5 iteraciones del algoritmo de procesamiento, a no ser que se indique lo contrario. Debe tenerse en cuenta que la aplicación sucesiva del filtro implica realizar algunas modificaciones en la máscara de estimación de ruido que se emplea.

Medida de calidad

Se presentarán resultados mediante impresión de algunas imágenes significativas filtradas con cada algoritmo. Además, se incluyen gráficas y tablas que resumen los resultados numéricos. Como medida objetiva que indique el grado de mejora que ha sufrido la imagen gracias al procesamiento, utilizaremos el CMAE, como se dijo en el primer capítulo.

Se presentará un primer conjunto de gráficas, una para cada algoritmo, en la que se muestra el comportamiento de cada filtro frente a las diferentes clases de imágenes, variando el tipo y cantidad de ruido y el tamaño de la ventana. En estas representaciones aparecerán en trazo discontinuo los resultados correspondientes al filtro *sin* usar la máscara de estimación de ruido. De esta manera se pareciará diáfanamente la importancia de $\underline{M}$ en el procesamiento.

Otro grupo de gráficas se exhiben a continuación, en las que se presenta un análisis comparativo entre los distintos filtros, tanto los de nuevo diseño como los de referencia, cuando se enfrentan a la misma imagen fuente. Así, aparecerá una representación por cada clase de imágenes empleada, en la que se variará el tipo y cantidad de ruido impulsivo, y el tamaño de la ventana utilizada.El trazo discontinuo se reserva ahora para los filtros de comparación.

Abreviaturas de las tablas y gráficas

Debe hacerse un pequeño inciso en las abreviaturas y símbolos que se han utilizado en la construcción de las gráficas y tablas que acompañan a este capítulo. Los algoritmos tienen un nombre demasiado largo en muchas ocasiones, por lo que se han empleado los siguientes apócopes:

FILTRO	ABREVIATURA	FILTRO	ABREVIATURA
Mediana	MED	Máscara-Mediana	Mas-Med
Mediana Ponderada en el Centro	CWM	Máscara-Modelo	Mas-Mod
Mediana Multietapa	MM	Máscara-Modelo Ponderada	MMP
Híbrido FIR-Mediana	FM	Mediana Combinada	Med-Com
Adaptativo de Sun & Venetsanopoulos	Asa-Sun	Restaura-Modelo	Res-Mod

Además de ello, aparece en las tablas una primera columna que identifica el experimento llevado a cabo, mediante el uso de la siguiente notación:

v/r/%/m

donde *v* indica el tamaño de la ventana empleada; *r* el tipo de ruido impulsivo (teniendo en cuenta que *SP* = *sal y pimienta*; y *G* = uniformemente distribuido); % el porcentaje de ruido que tiene la imagen fuente; y *m* indica si se ha utilizado o no la máscara de estimación de ruido. Esto último se ha incluido en los resultados para que exista una comparación entre el comportamiento de los algoritmos desarrollados cuando trabajan con la máscara de estimación de ruido y cuando no lo hacen.

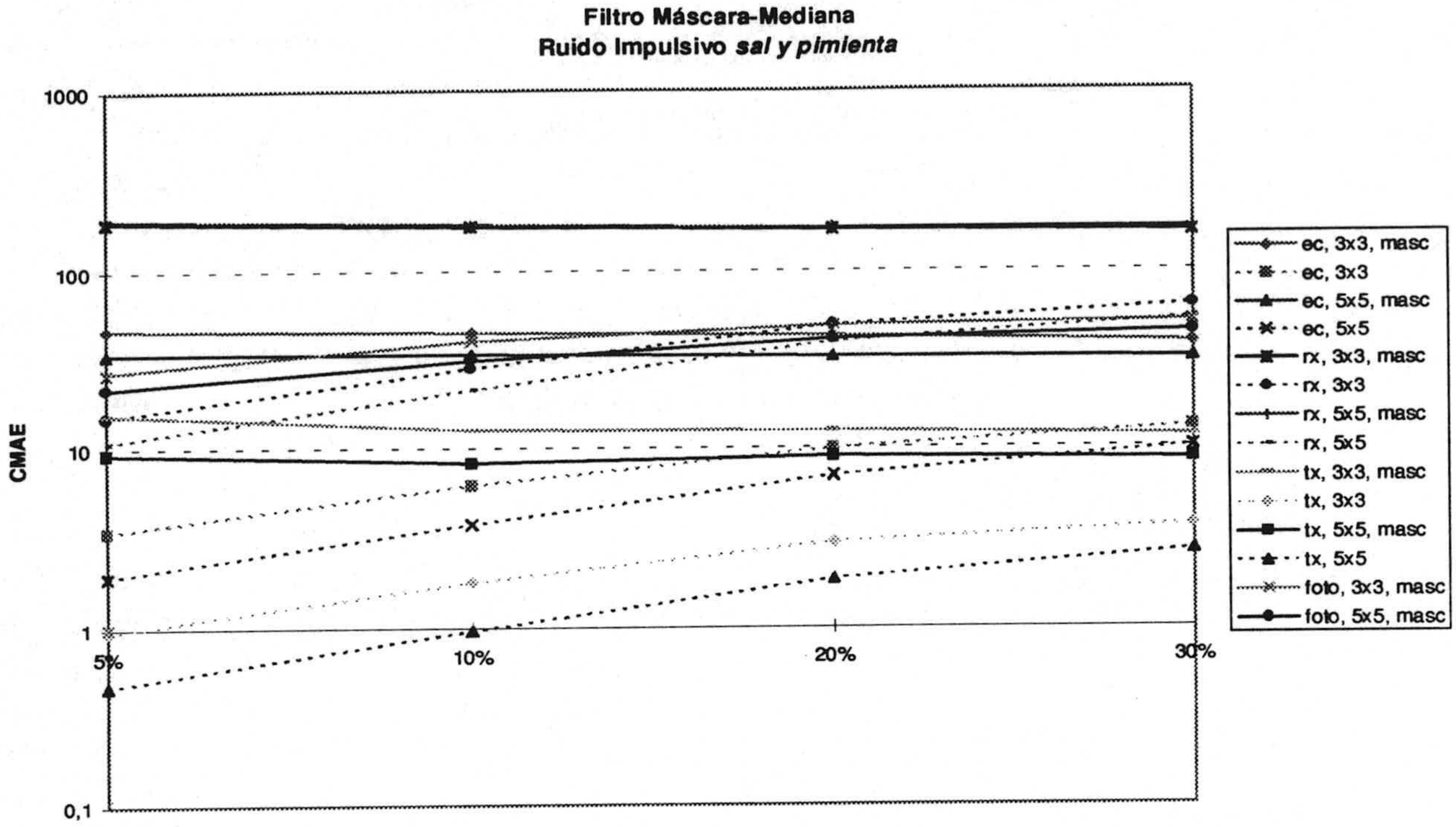
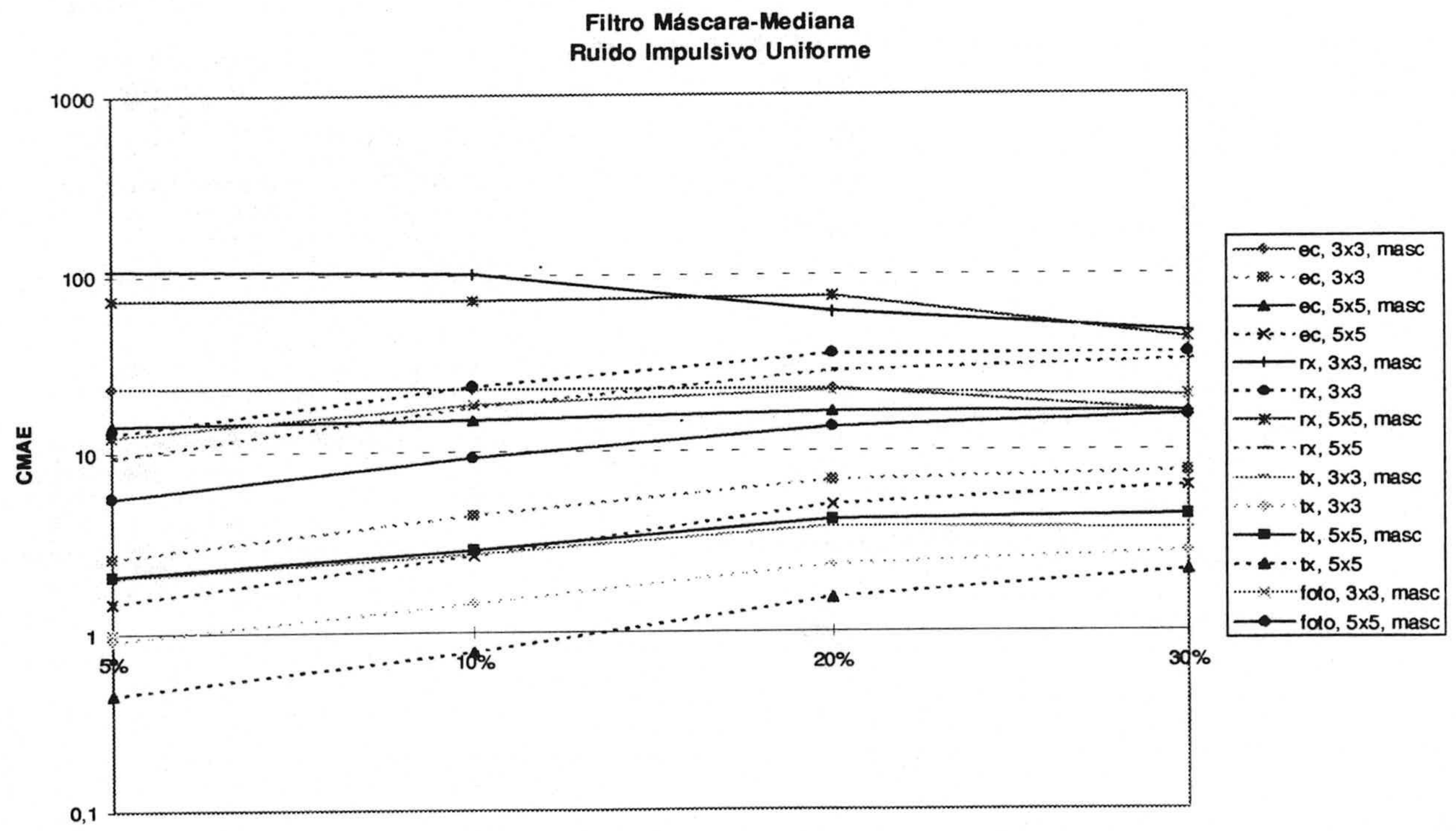


Figura 1: Resultado de los experimentos empleando el filtro **Máscara-Mediana**. En la gráfica superior se estudia el caso del ruido impulsivo *sal y pimienta*; mientras que en la inferior, se trata el caso del ruido impulsivo uniformemente distribuido. En trazo discontínuo, los algoritmos sin emplear máscara.



En las tablas, se han empleado tipo de letra negrita cuando el mejor valor de CMAE se ha obtenido en una iteración de filtrado anterior a la quinta, y después el resultado ha empeorado. El tipo cursiva y subrayado indica que el mejor valor de CMAE se produce en la quinta y última iteración, por lo que es posible que posteriores iteraciones mejoren aún más el resultado. El tipo de letra normal significa que se ha alcanzado un valor convergente a lo largo de las 5 iteraciones programadas. Por otro lado, las casillas con fondo oscuro indican el mejor valor de CMAE en cada fila, de manera que se destaca rápidamente el mejor resultado.

Umbral de selección

Obsérvese que no se ha hecho mención al umbral de selección como parámetro que haya de tenerse en cuenta en los experimentos. En efecto, si bien es una variable de extraordinaria importancia, se trabajará con el valor, obtenido en el anterior capítulo, que mejor máscara proporciona.

Se han de hacer dos advertencias. En primer lugar, se apreciará que sólo se presentan resultados para *restdiver* con máscara en el caso de texto. Ello es así porque ninguna de las otras imágenes propuestas tienen bordes definidos, lo cual es requisito imprescindible para el uso de este algoritmo. Además, no se muestran gráficamente los resultados para *restdiver* sin máscara, puesto que ya se hizo en el capítulo II. Por otra parte, no se han realizado experimentos para la fotografía sin emplear máscara, puesto que esta imagen se añadió a la batería de experimentos tan sólo para mostrar el comportamiento del modelo hipergeométrico.

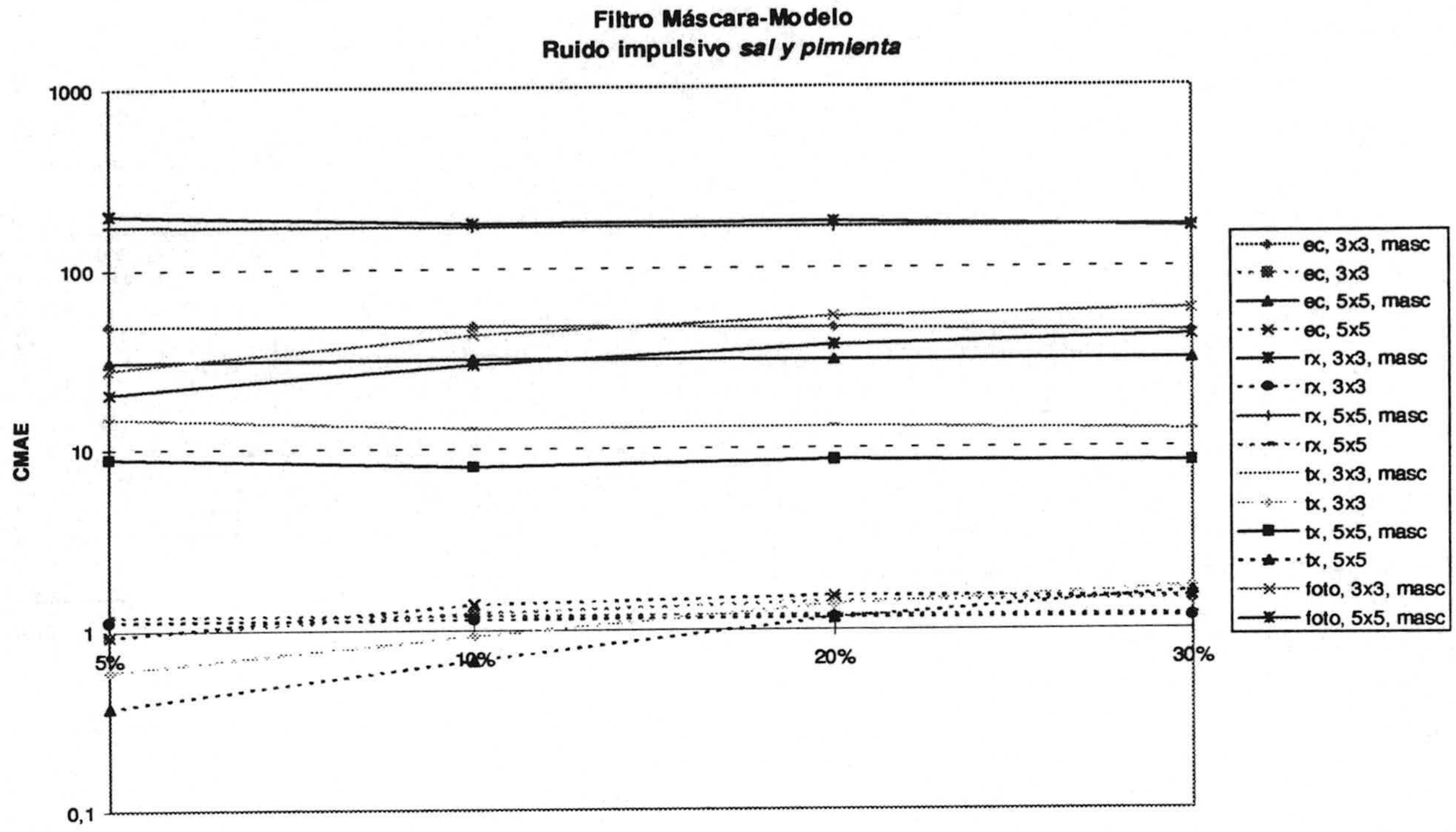
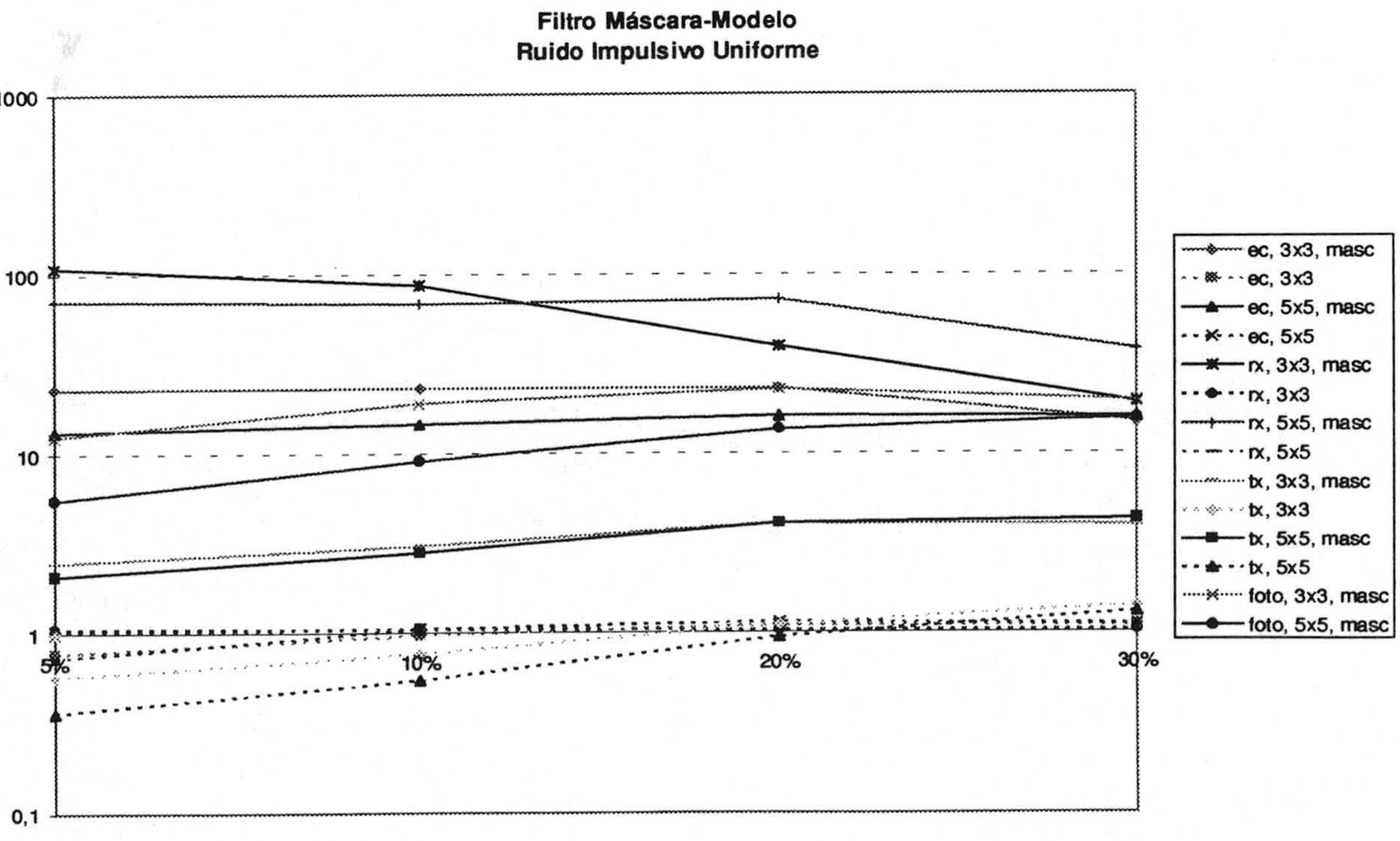


Figura 2: Resultados del filtrado con el algoritmo **Máscara-Modelo**. En trazo discontinuo aparecen los valores de CMAE obtenidos al no emplear la máscara de estimación de ruido. En la gráfica superior se muestra el caso del ruido impulsivo *sal y pimienta*; mientras que abajo aparece el correspondiente al ruido impulsivo uniformemente distribuido.



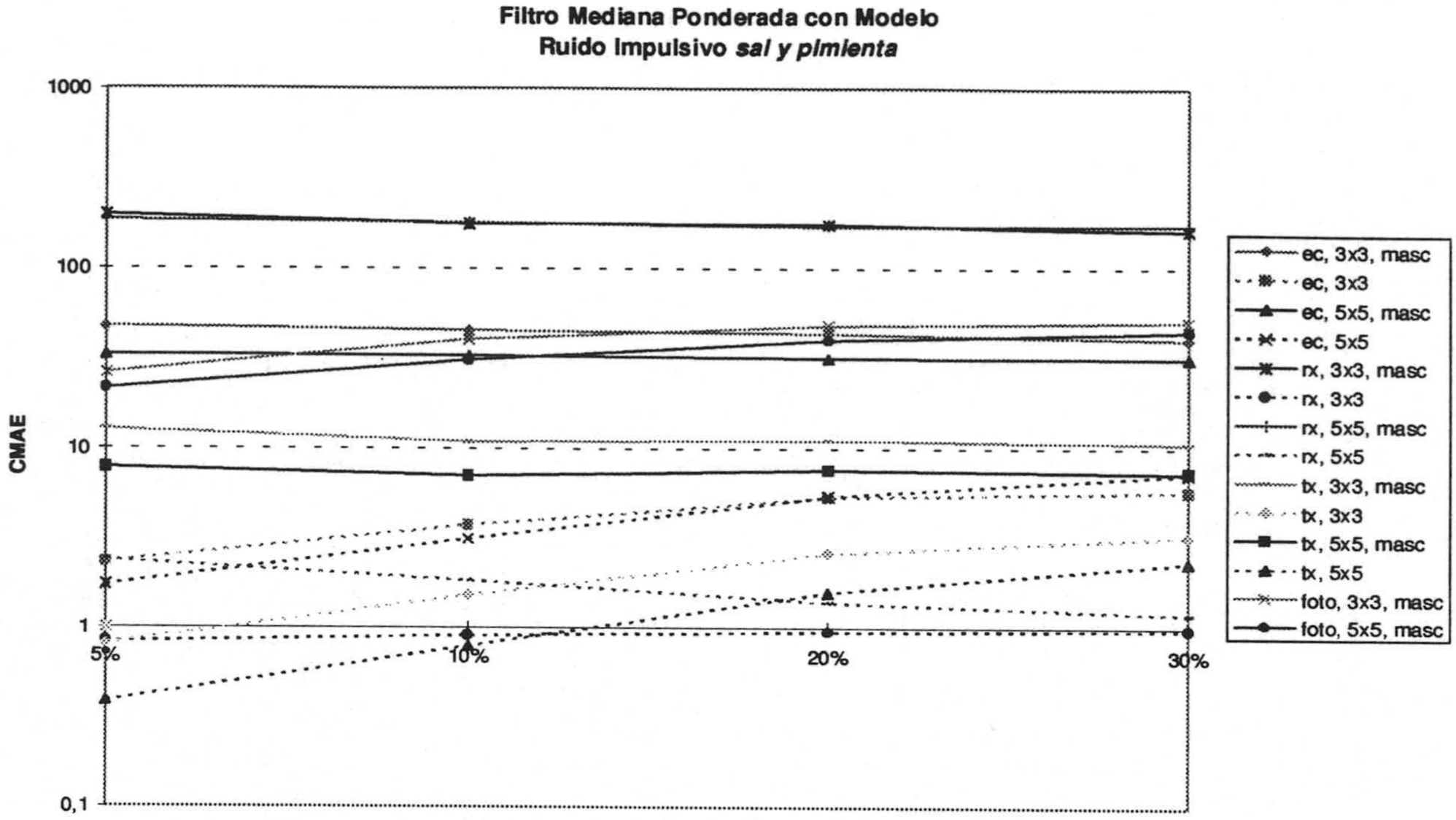
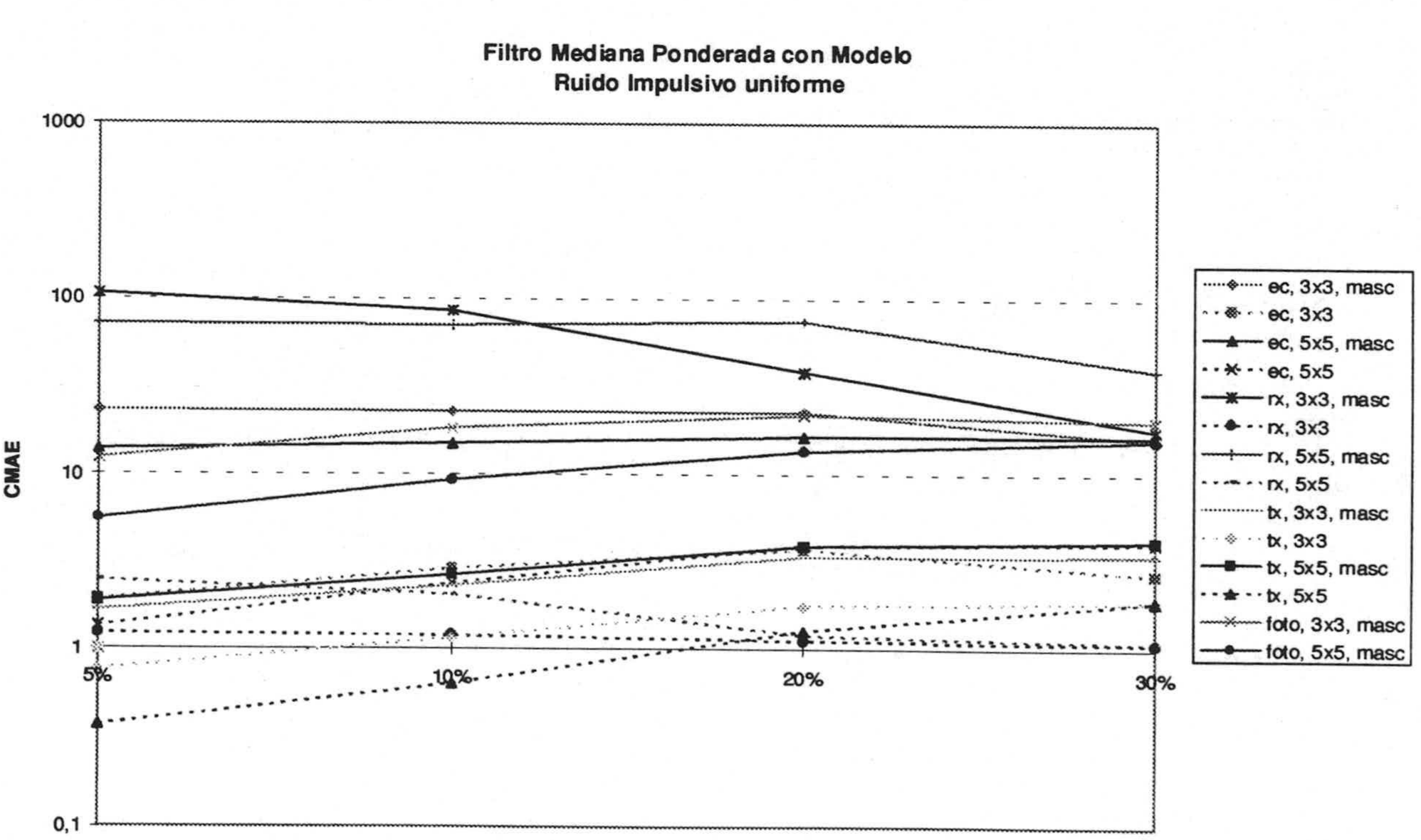


Figura 3: Resultados del filtrado con el algoritmo **Máscara-Modelo Pesado**. En trazo discontinuo aparecen los valores de CMAE obtenidos al no emplear la máscara de estimación de ruido. En ambas gráficas la imagen fuente está afectada de ruido impulsivo. En la superior es del tipo *sal y pimienta*; mientras abajo es uniformemente distribuido.



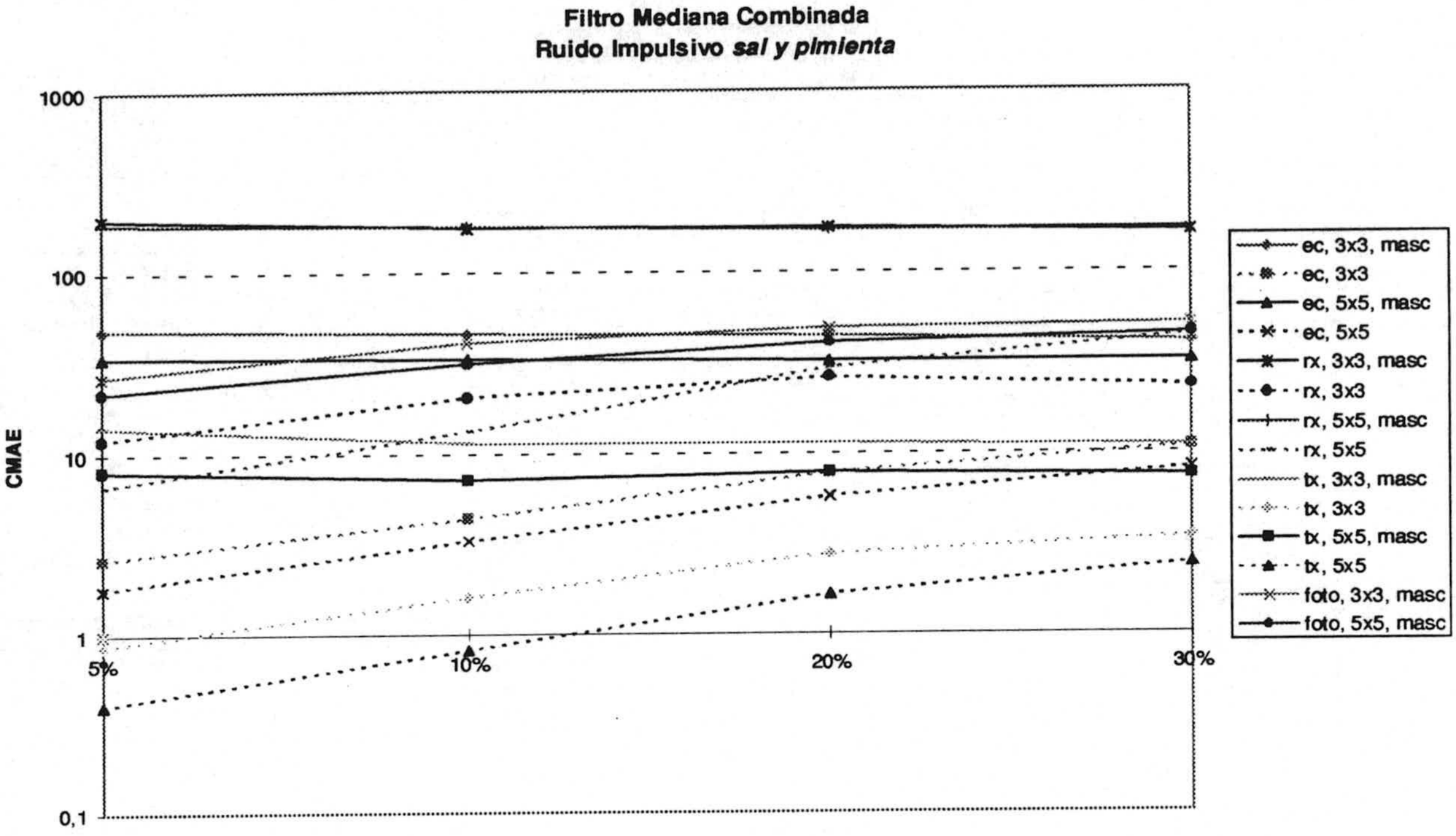
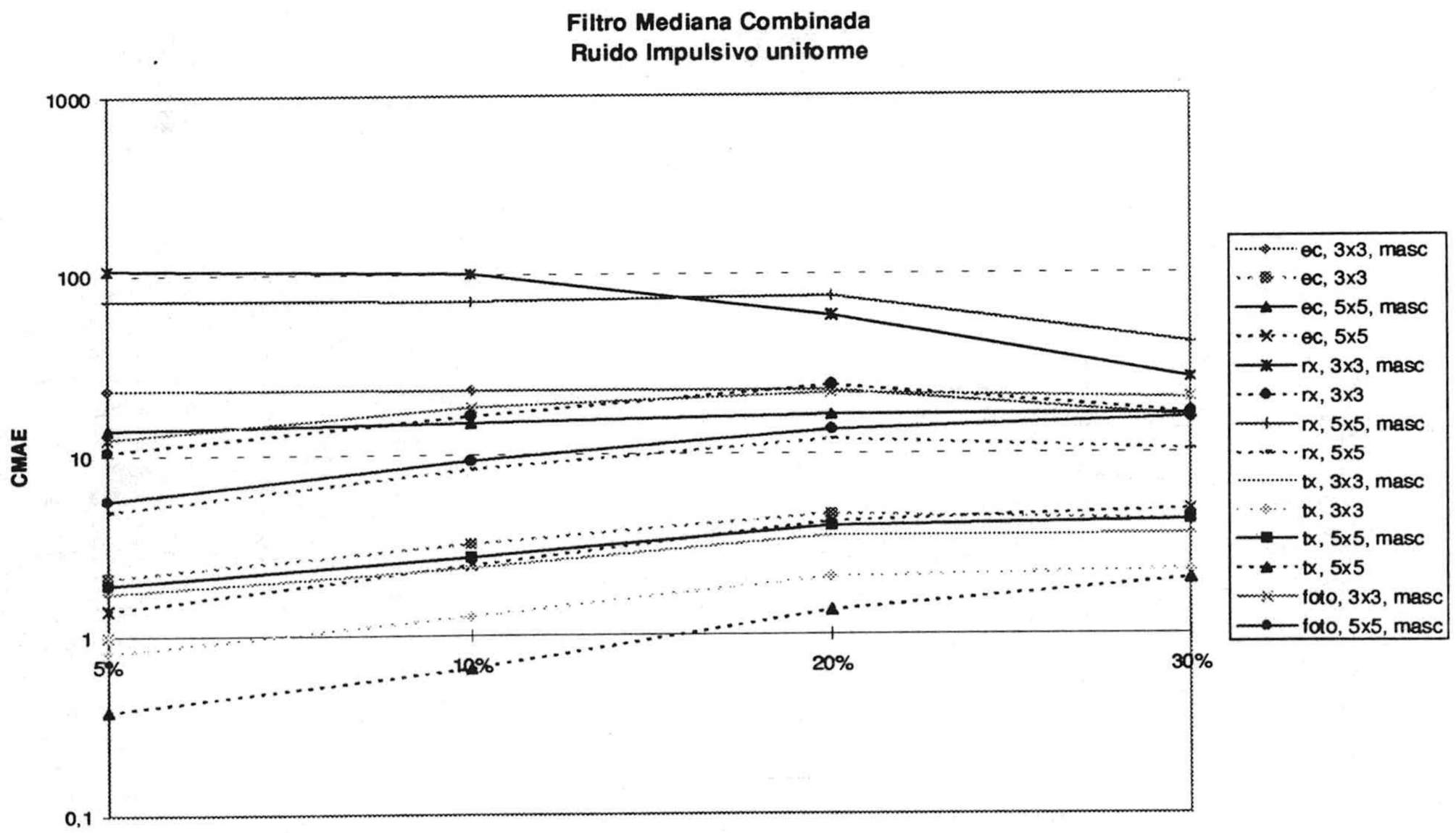


Figura 4: Valores de CMAE correspondientes al procesado de imágenes fuente afectadas de ruido impulsivo *sal y pimienta* (arriba), y uniformemente distribuido (abajo), utilizando el filtro **Mediana-Combinada**. El trazo discontinuo indica que no se utiliza la máscara de estimación de ruido.



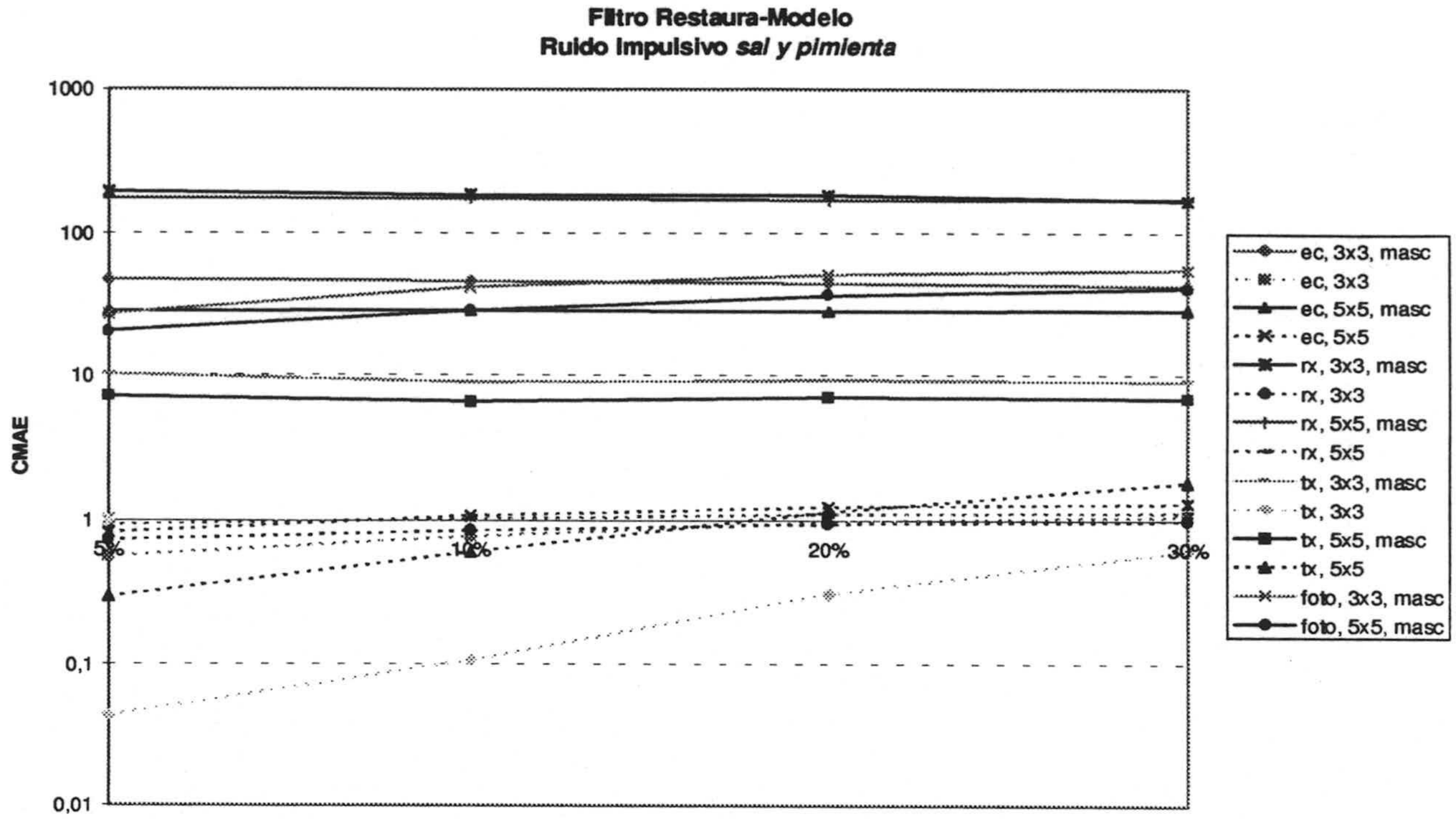
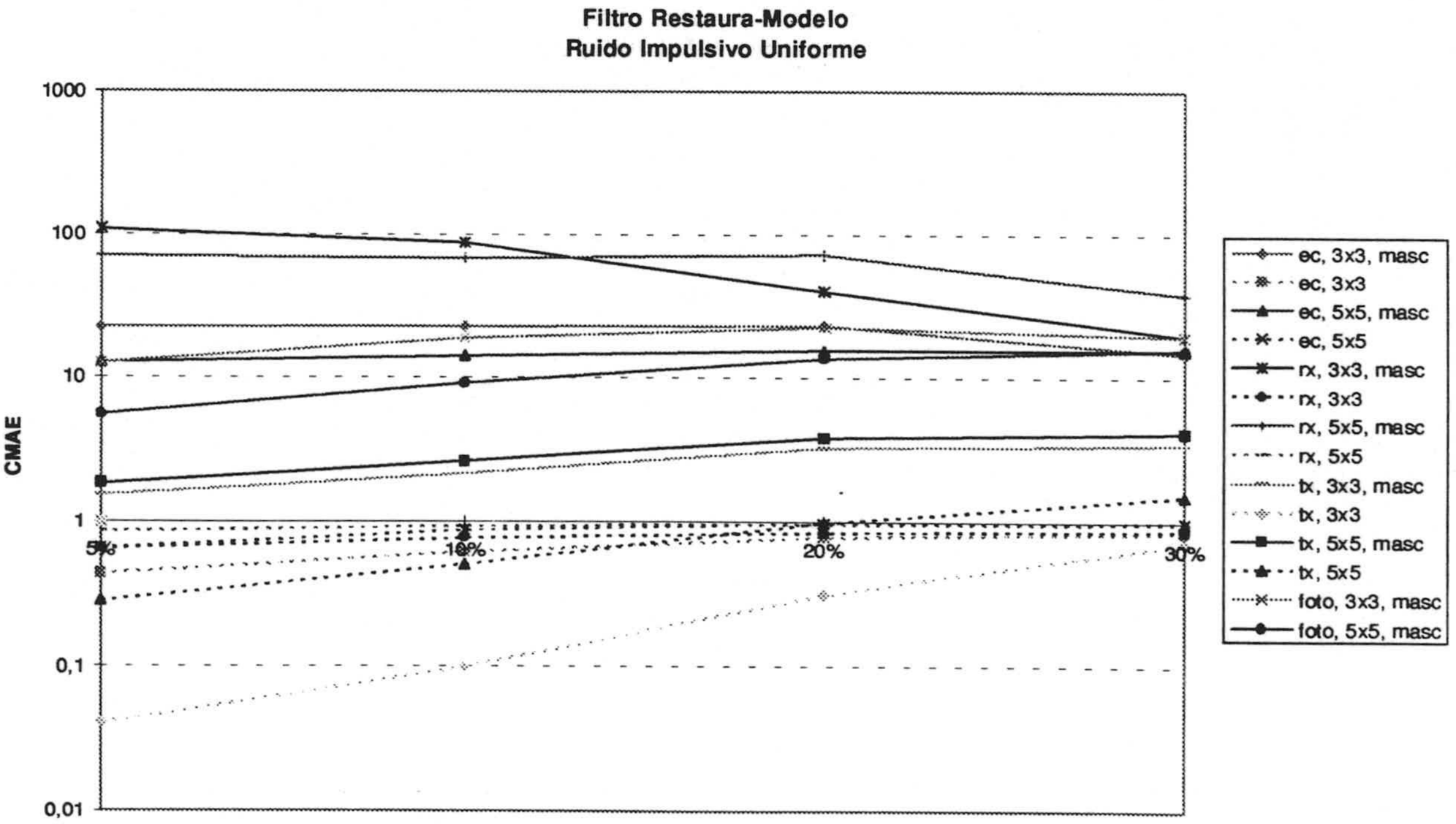


Figura 5: Valores de CMAE obtenidos empleando el filtro **Restaura-Modelo**. El trazo discontinuo significa que no se ha empleado la máscara en el procesado. EL ruido impulsivo *sal y pimienta* se muestra arriba; mientras que bajo estas líneas se representa el caso del ruido impulsivo uniformemente distribuido.



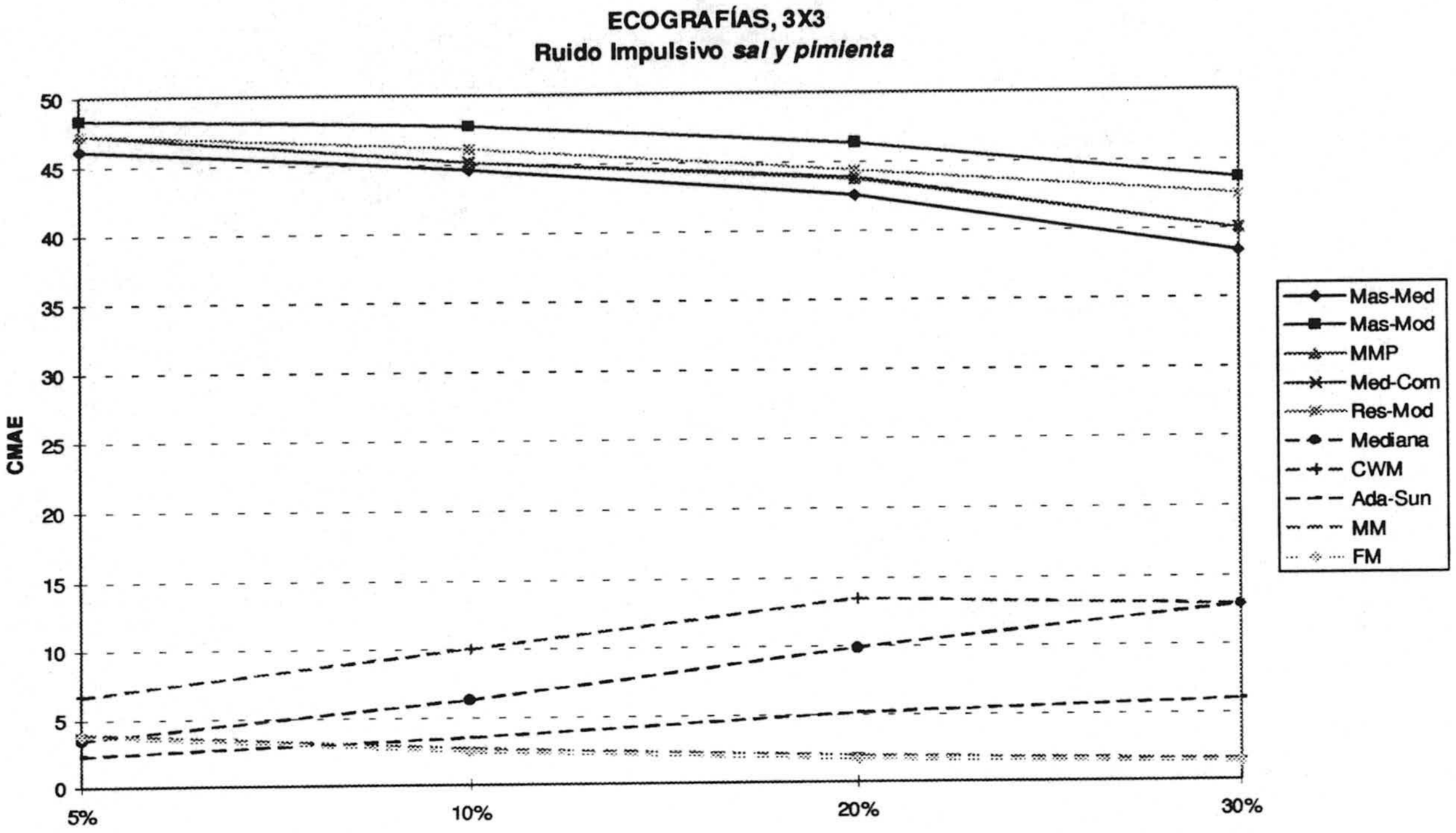
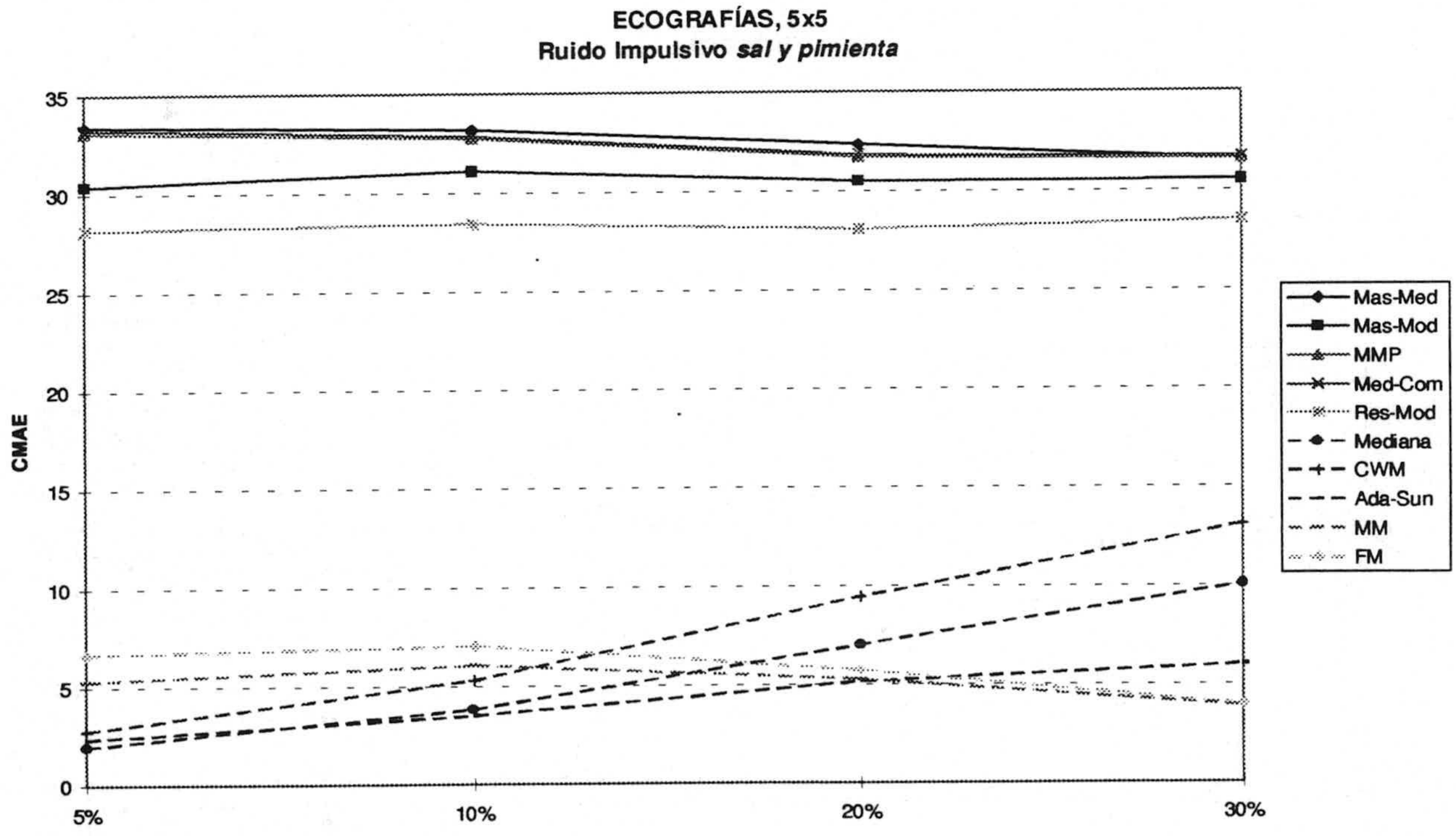


Figura 6a: Resultados comparativos entre todos los filtros al procesar imágenes fuente de **ecografías**, afectadas por ruido impulsivo *sal y pimienta* (arriba), y uniforme (abajo). En ambos casos se ha empleado ventana de tamaño 3×3. Los filtros de comparación se representan con trazo discontinuo. En todos los casos se emplea la máscara de estimación de ruido.



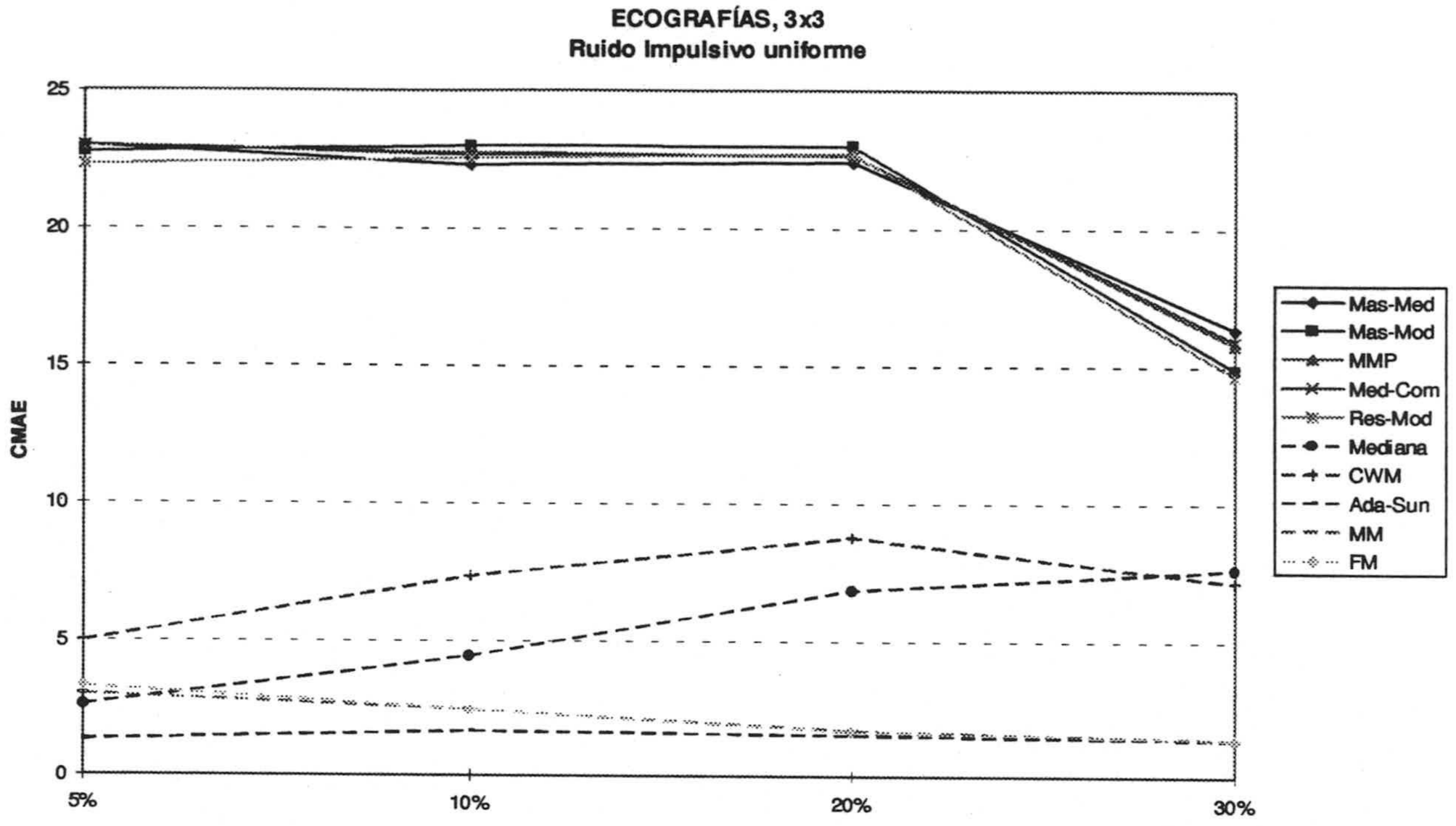
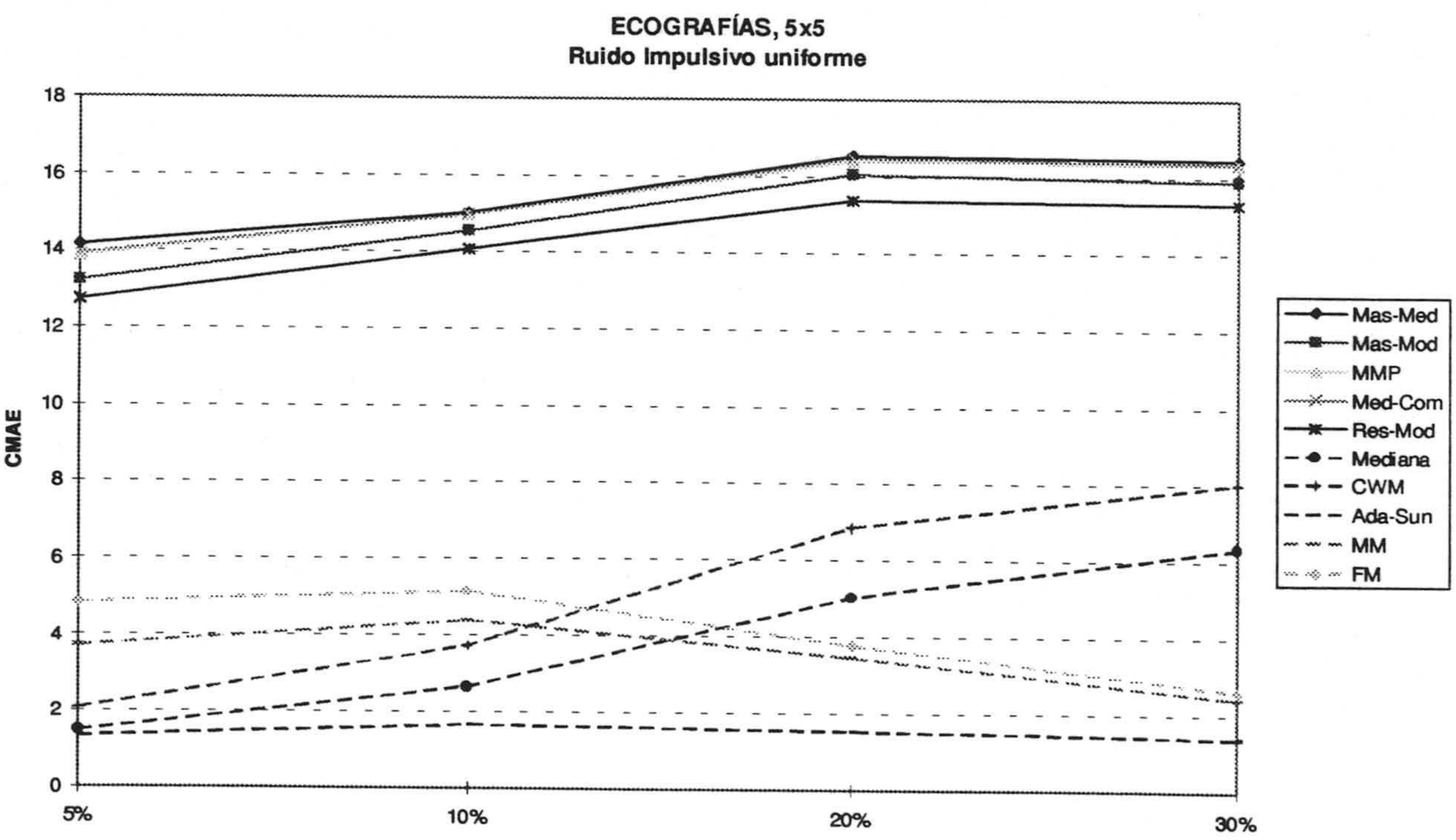


Figura 6b: Se representan los valores de CMAE obtenidos por los distintos algoritmos de filtrado al procesar imágenes fuente de **ecografías**, pero utilizando ventana de tamaño 5×5. Arriba se ilustra el caso de ruido *sal y pimienta*; y abajo el ruido impulsivo uniforme. Los resultados con los filtros utilizados para comparar aparecen con trazo discontinuo.



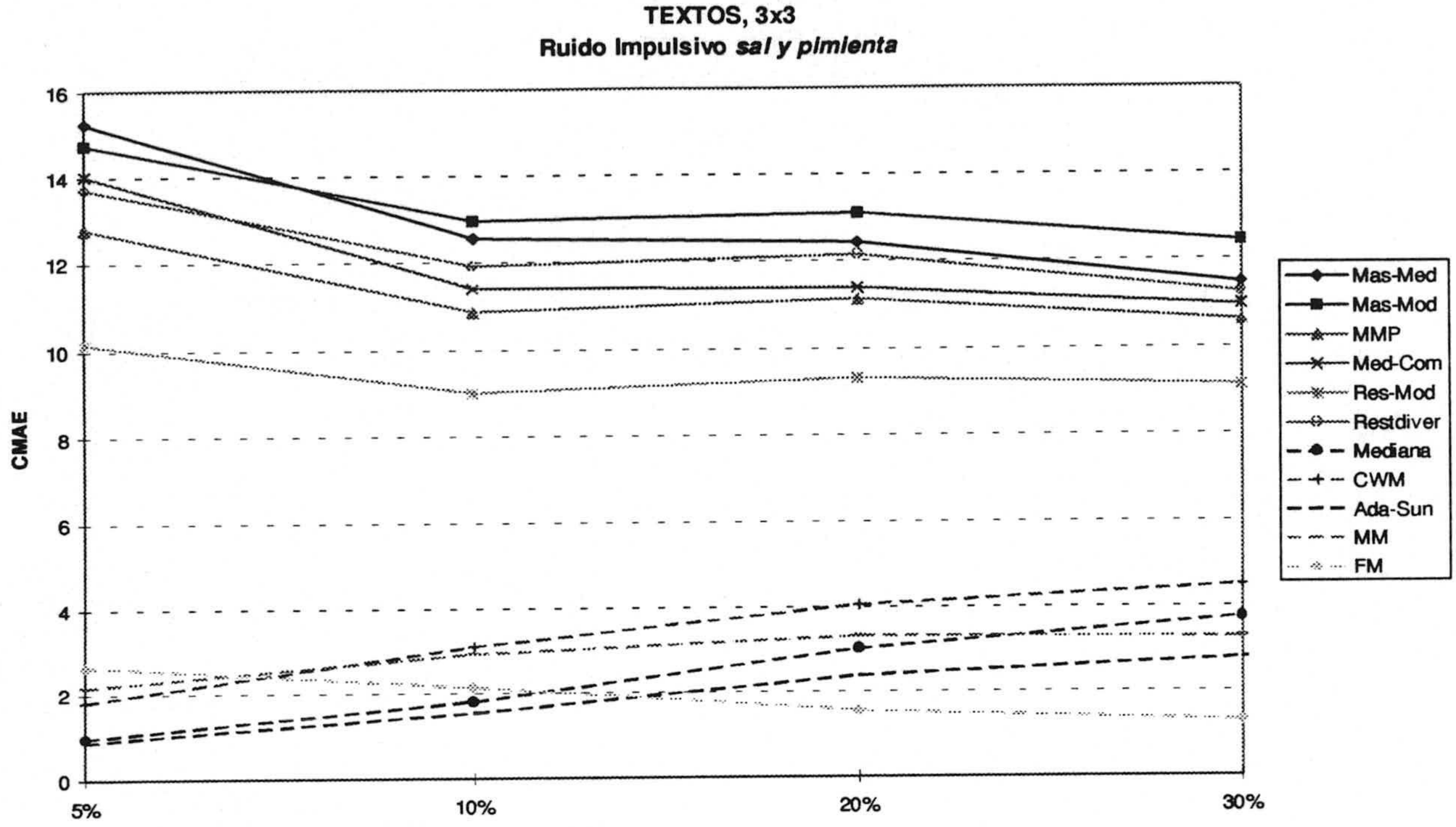
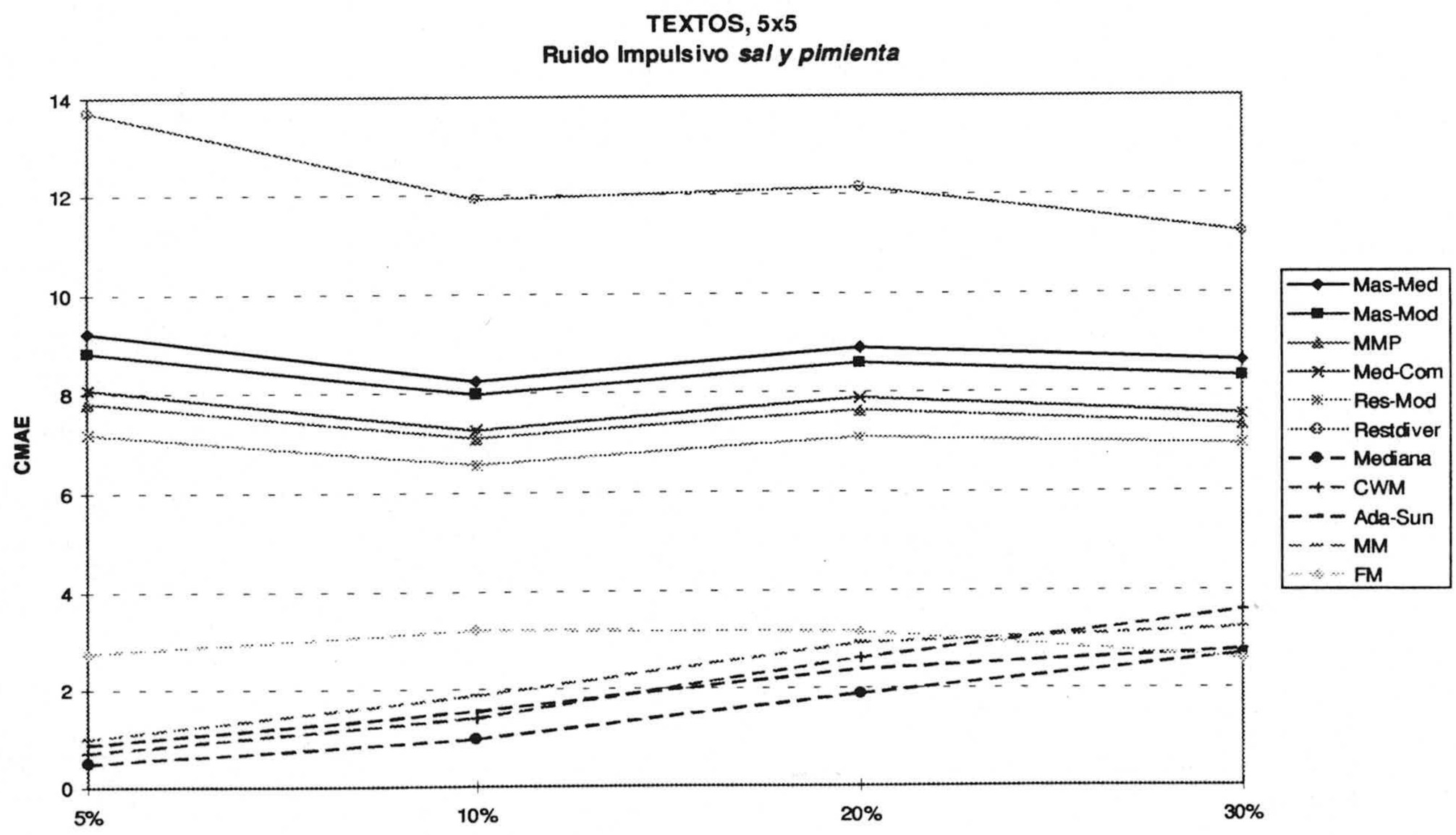


Figura 7a: Valores de CMAE obtenidos por los diferentes filtros al procesar una imagen de **texto mecanografiado**, corrompida con diferentes cantidades de ruido impulsivo *sal y pimienta* (arriba) , y uniformemente distribuido (abajo). El tamaño de ventana utilizado es 3×3. Siempre se usa la máscara de estimación de ruido. En trazo discontinuo aparecen los resultados correspondientes a los filtros de comparación.



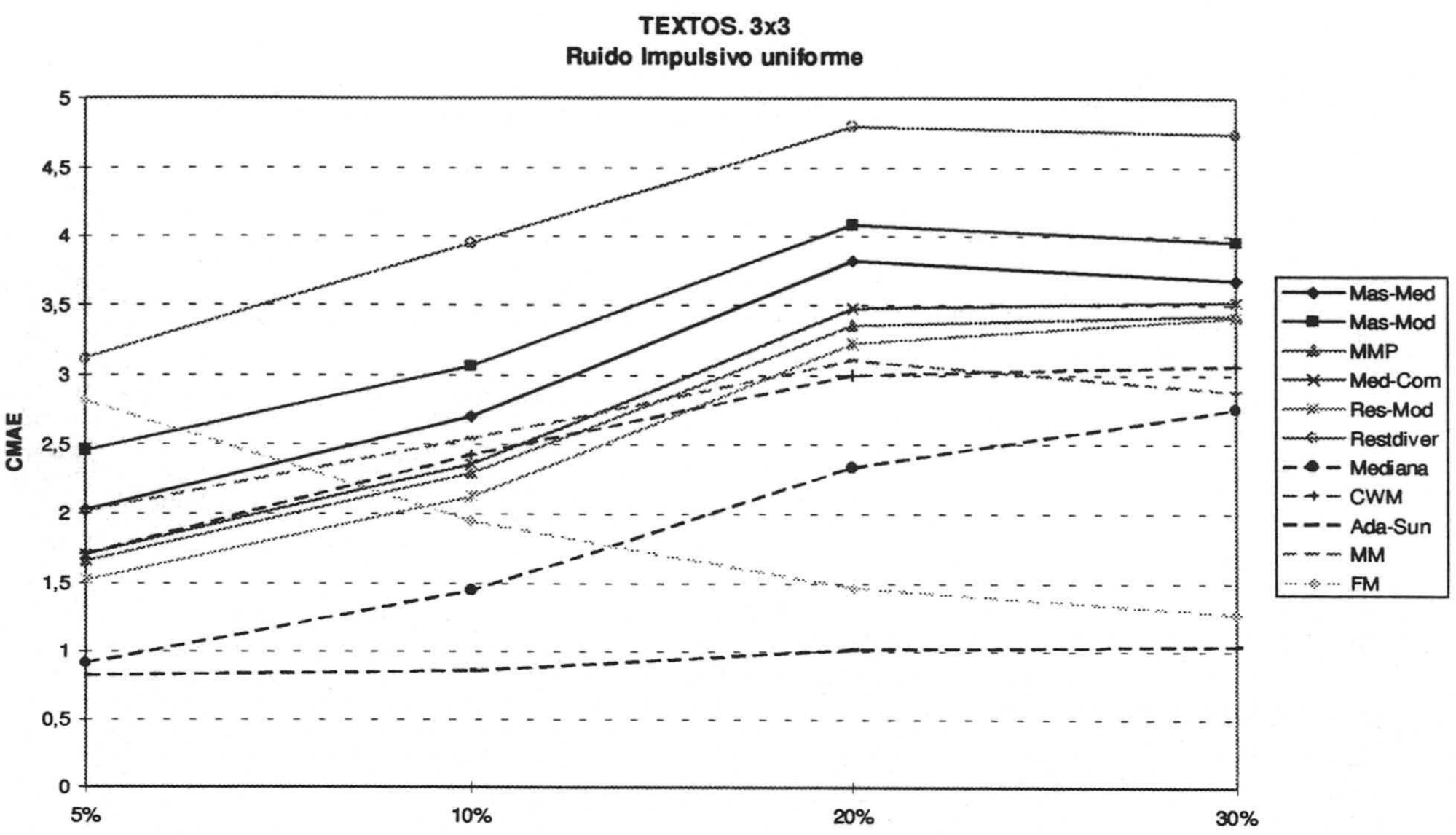
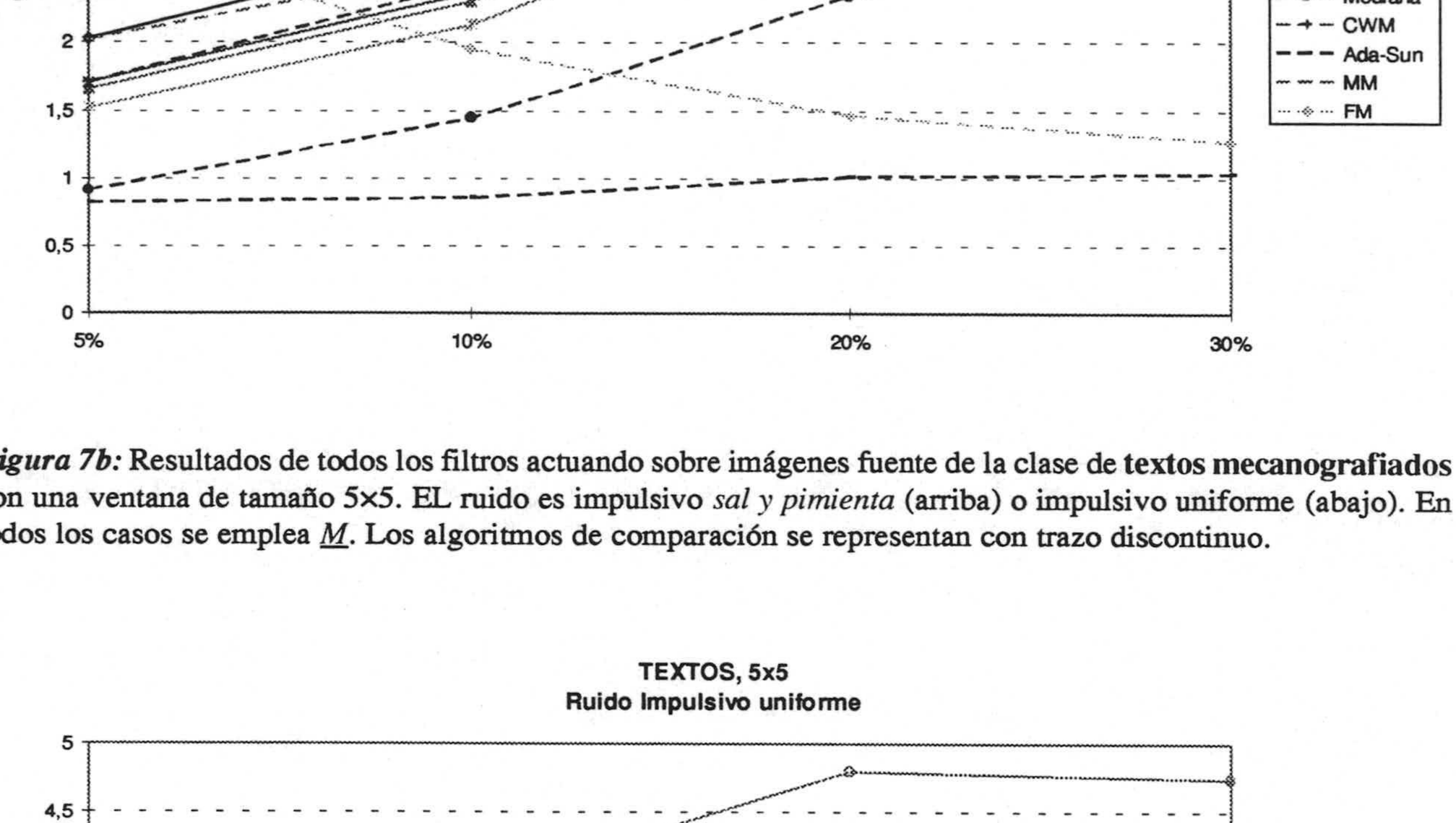


Figura 7b: Resultados de todos los filtros actuando sobre imágenes fuente de la clase de **textos mecanografiados** con una ventana de tamaño 5×5. EL ruido es impulsivo *sal y pimienta* (arriba) o impulsivo uniforme (abajo). En todos los casos se emplea $\underline{M}$. Los algoritmos de comparación se representan con trazo discontinuo.



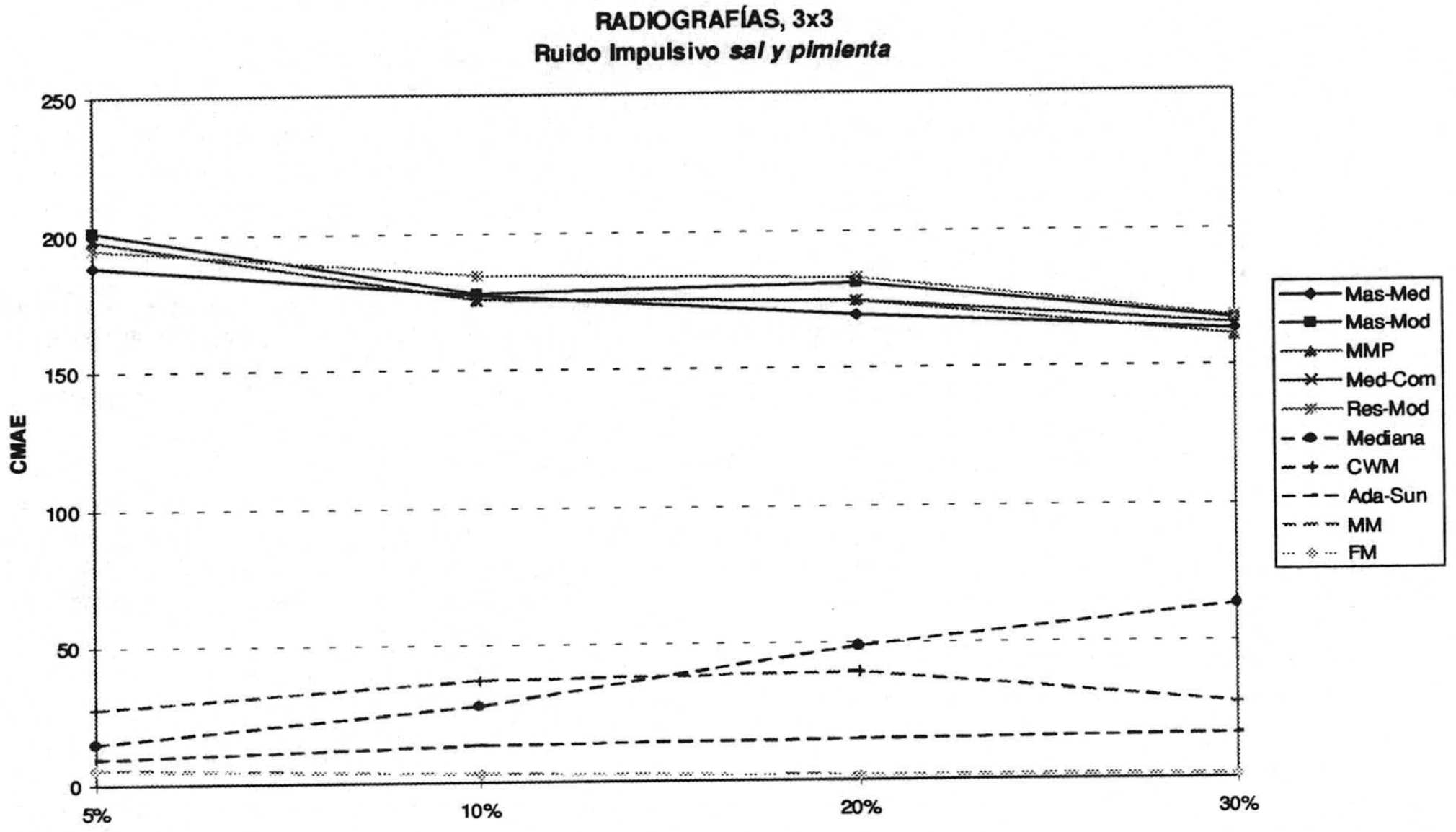
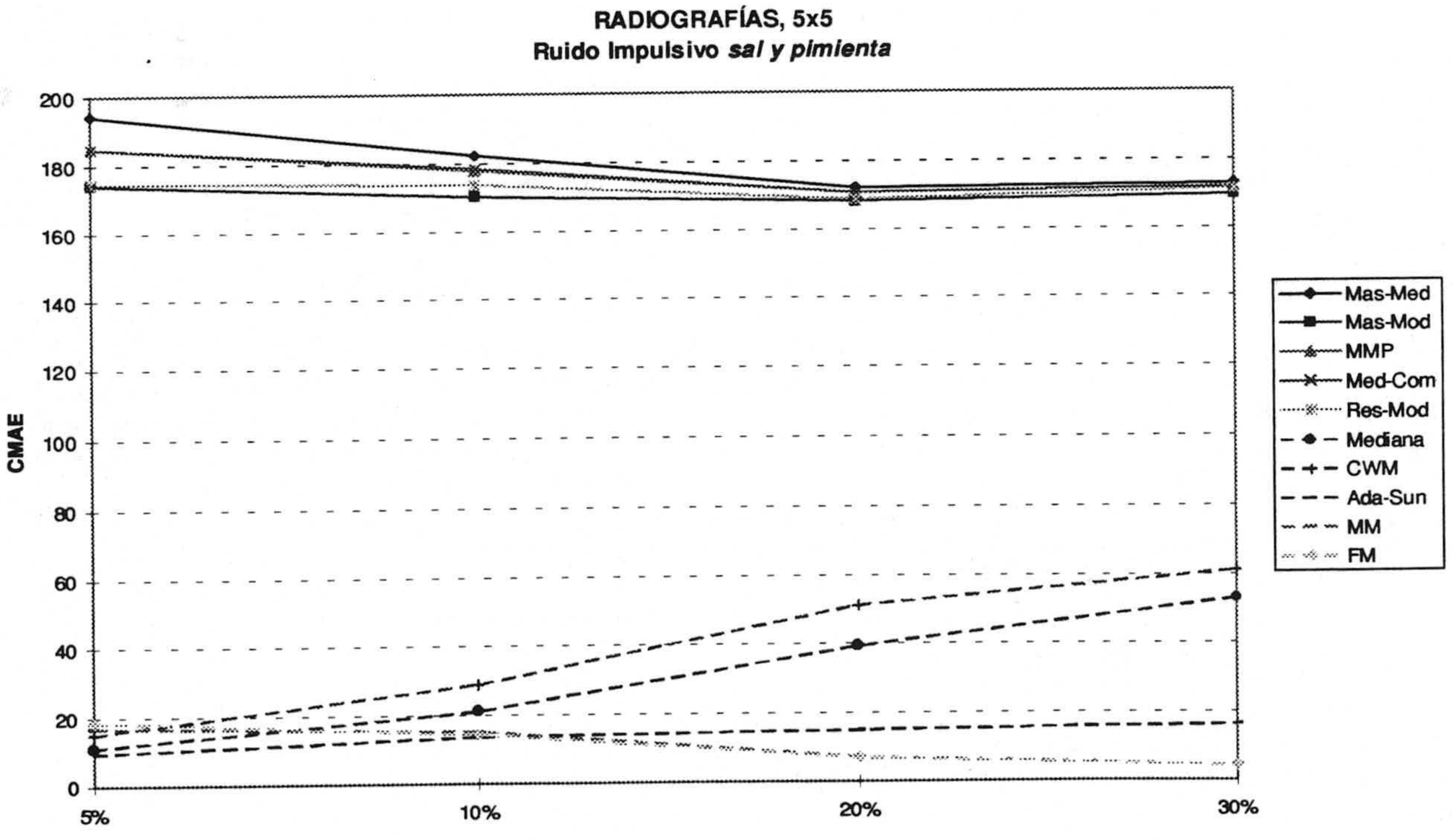


Figura 8a: Resultados procedentes del procesado con todos los filtros de imágenes fuente de la clase de **radiografías**, empleando una ventana de 3×3. EL ruido es impulsivo *sal y pimienta* (arriba) y uniforme (abajo). Los filtros de comparación aparecen en la gráfica con trazo discontinuo.



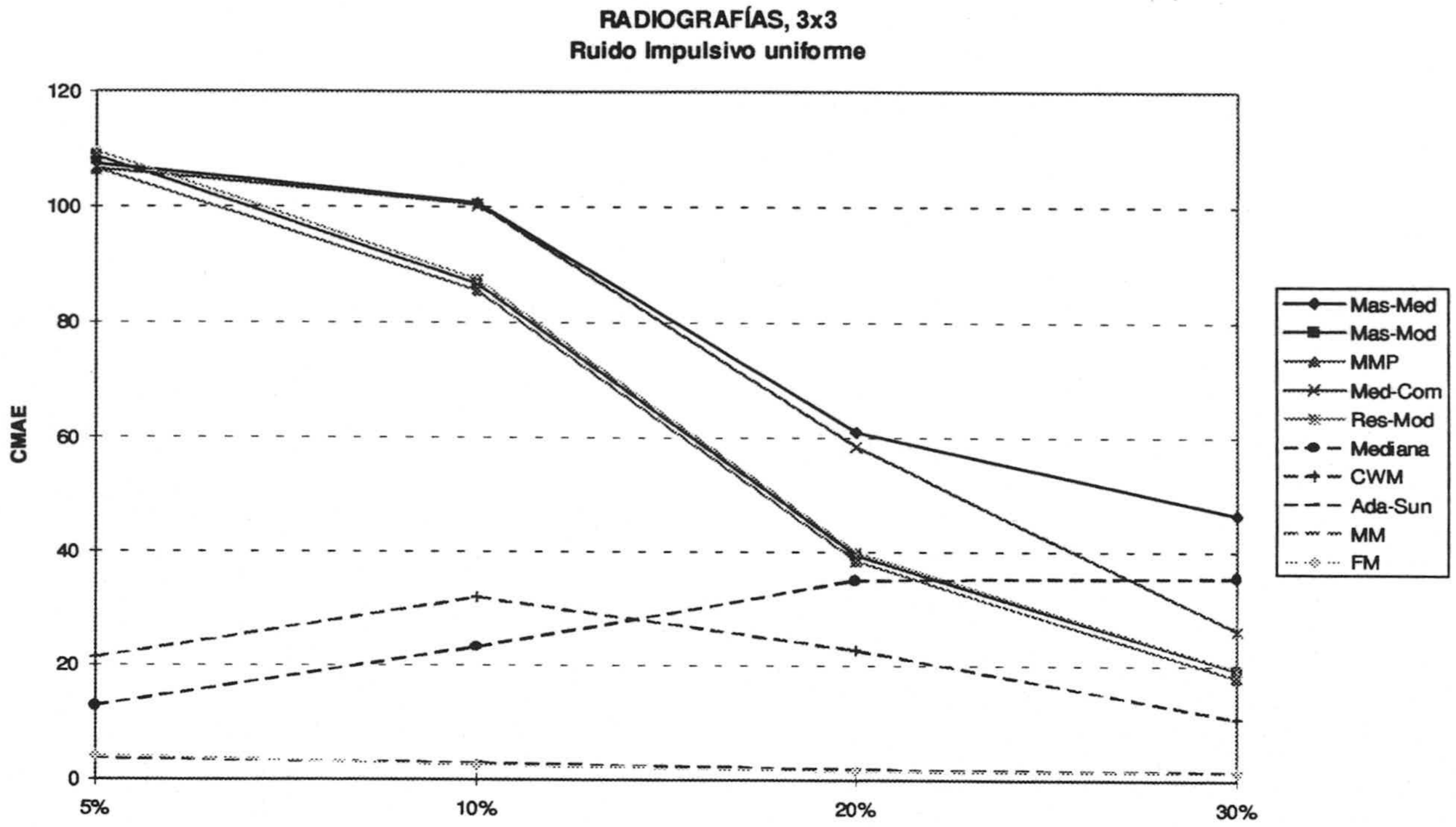
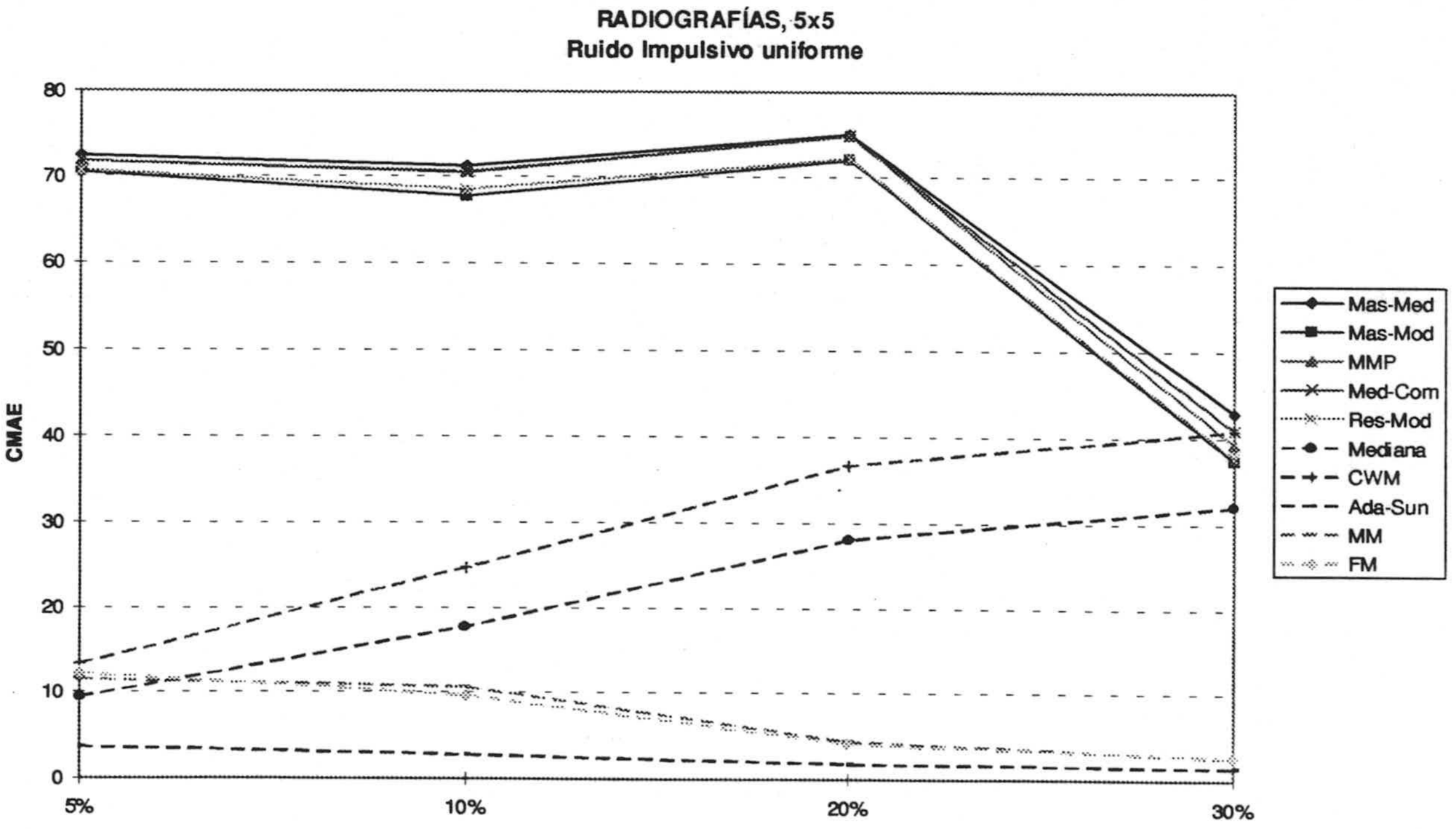


Figura 8b: CMAE obtenido al filtrar imágenes fuente de la clase **radiografías** corrompidas por ruido impulsivo uniforme y *sal y pimienta* (abajo y arriba, respectivamente) con todos los algoritmos diseñados. El tamaño de ventana empleado es 5×5. Las líneas de trazo discontinuo representan los resultados obtenidos con los filtros de comparación.



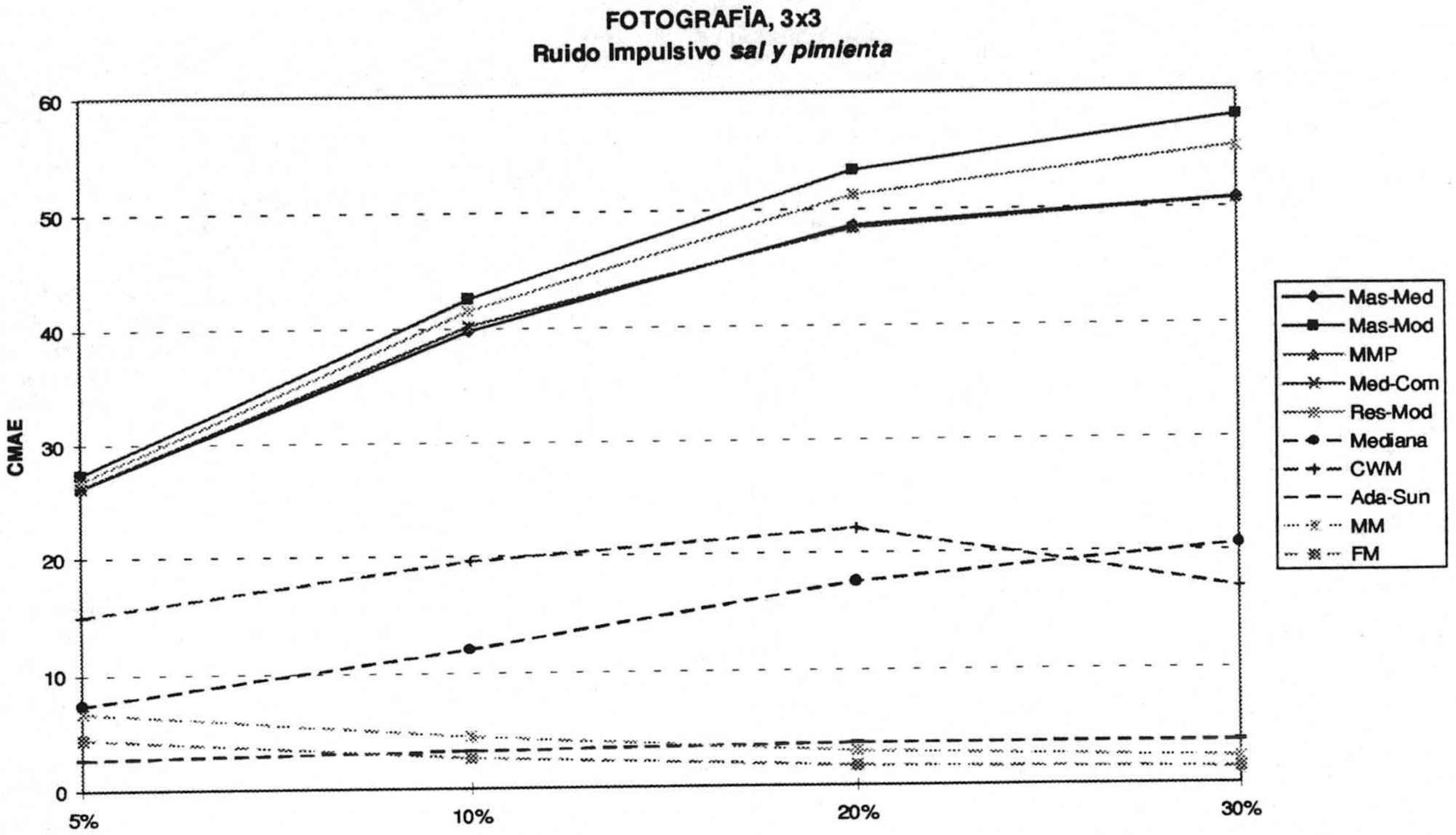
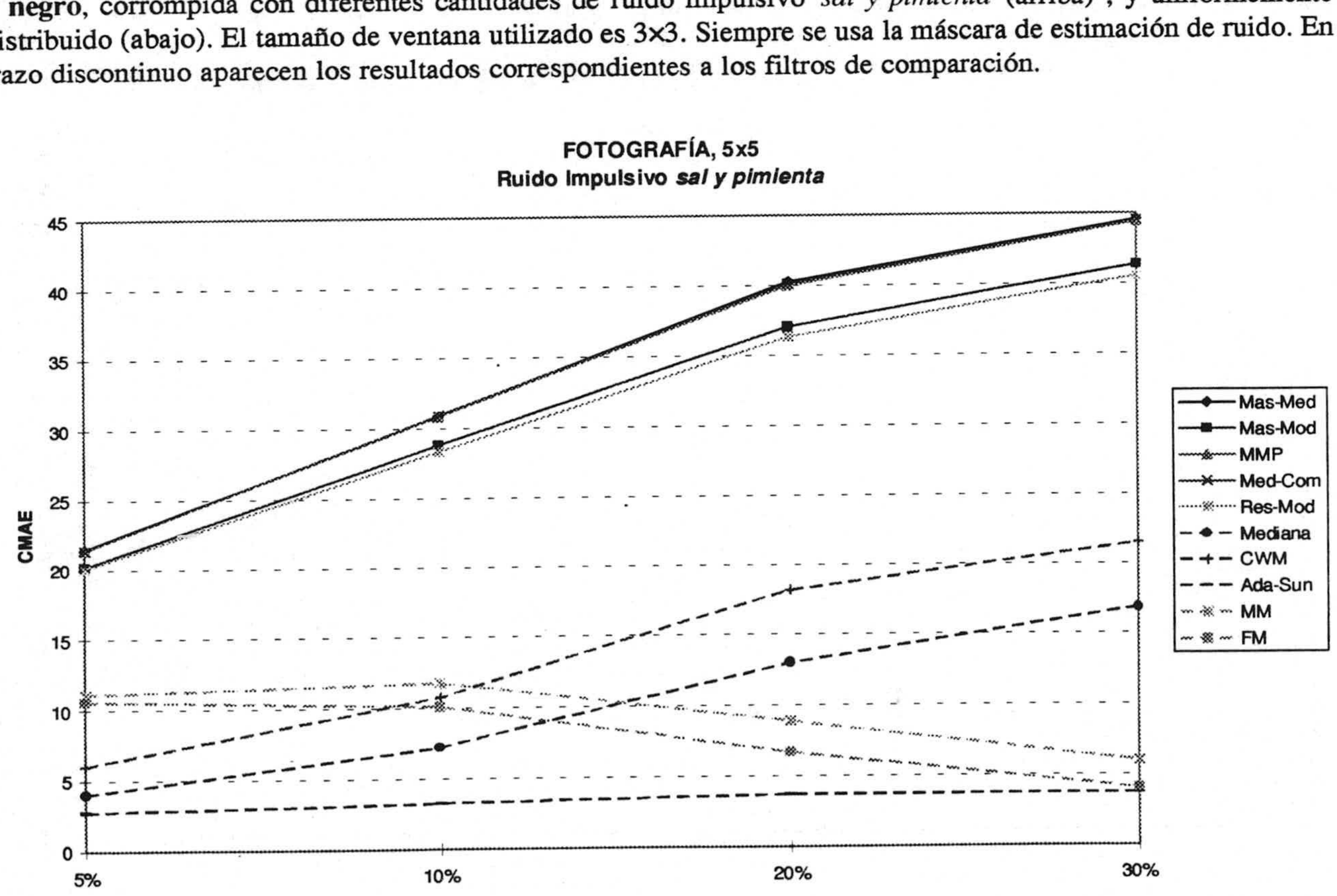


Figura 9a: Valores de CMAE obtenidos por los diferentes filtros al procesar una imagen de **fotografía en blanco y negro**, corrompida con diferentes cantidades de ruido impulsivo *sal y pimienta* (arriba) , y uniformemente distribuido (abajo). El tamaño de ventana utilizado es 3×3. Siempre se usa la máscara de estimación de ruido. En trazo discontinuo aparecen los resultados correspondientes a los filtros de comparación.



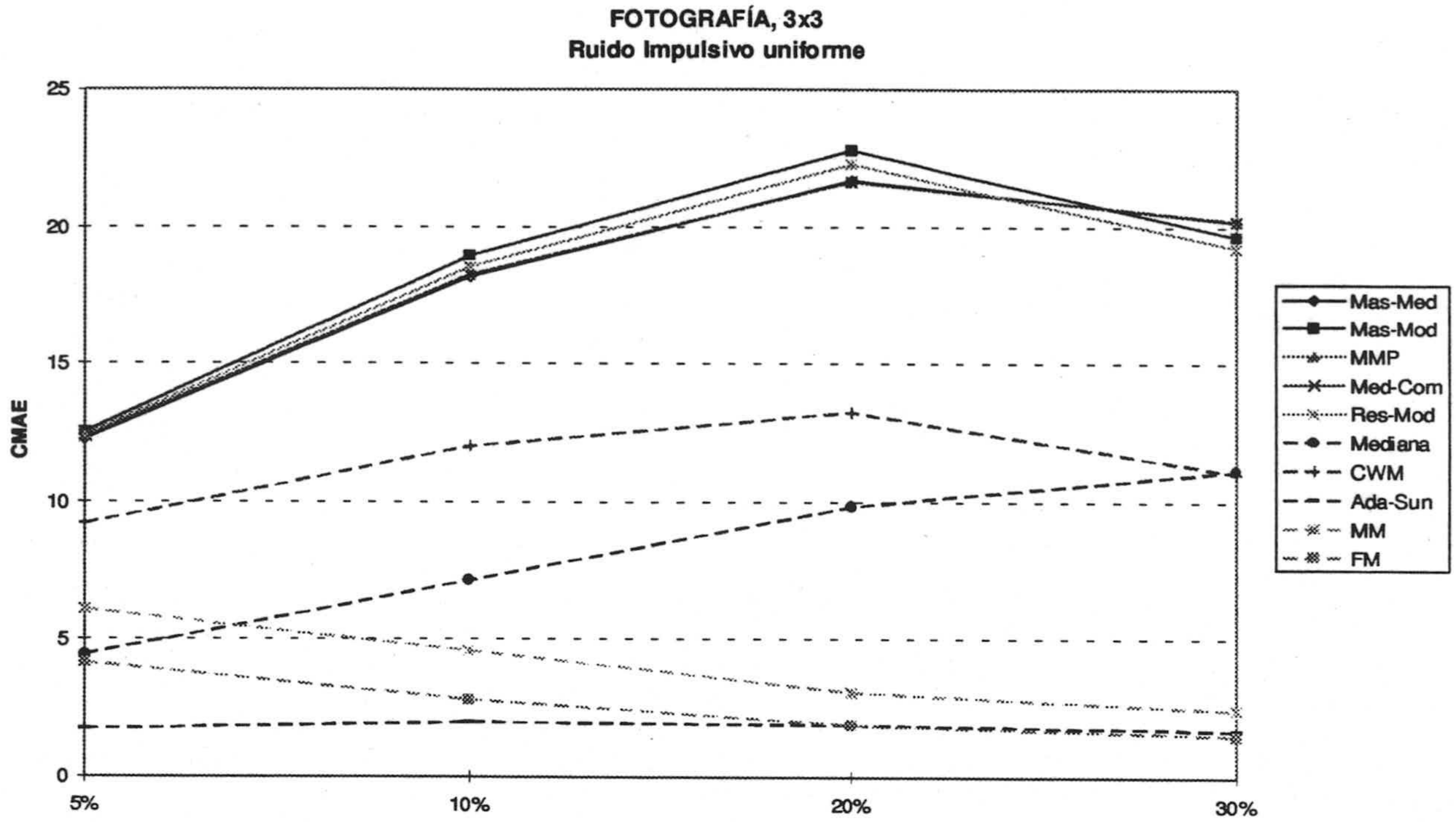
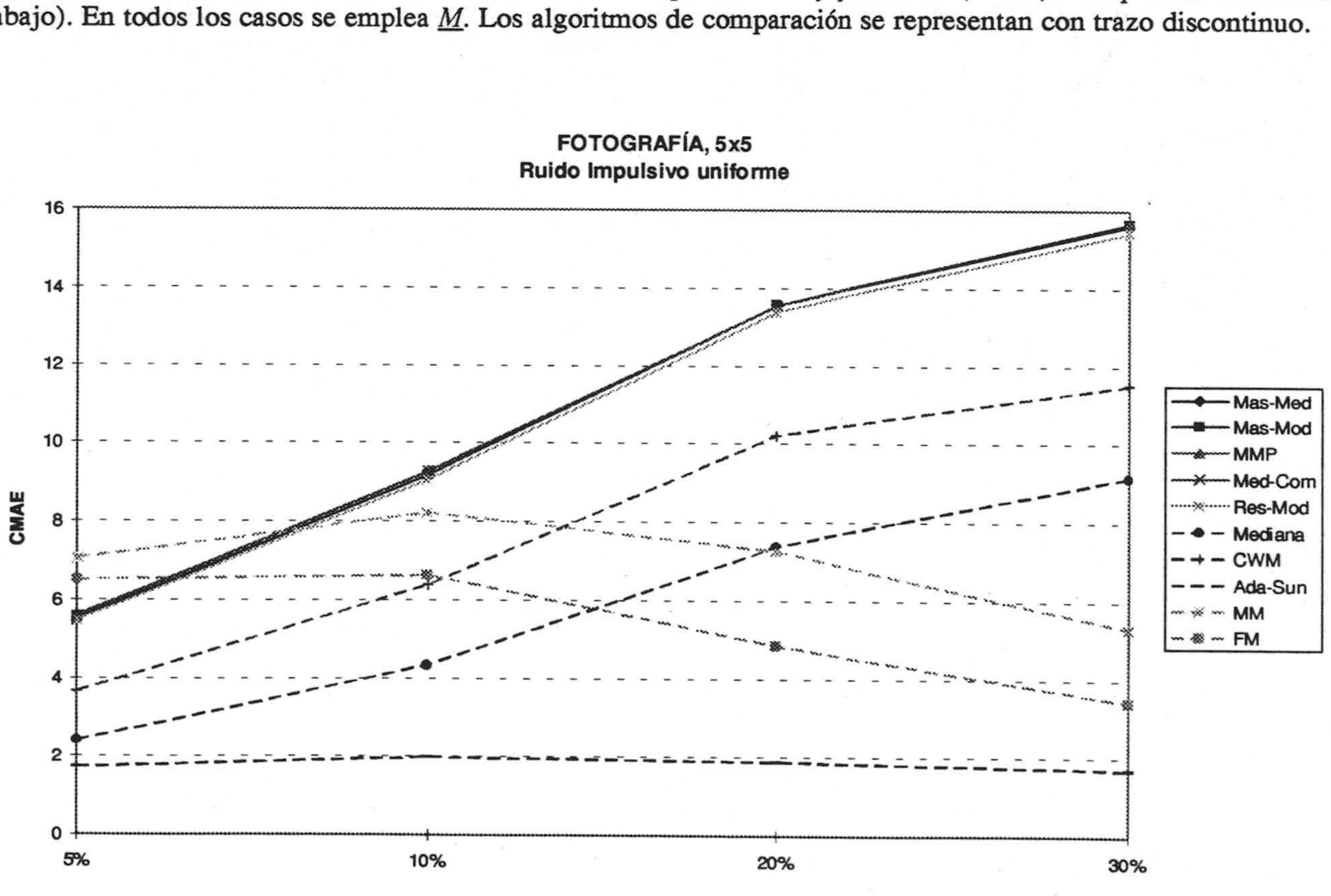


Figura 9b: Resultados de todos los filtros actuando sobre imágenes fuente de la clase de **fotografía en blanco y negro** con una ventana de tamaño 5×5. EL ruido es impulsivo *sal y pimienta* (arriba) o impulsivo uniforme (abajo). En todos los casos se emplea $\underline{M}$. Los algoritmos de comparación se representan con trazo discontinuo.



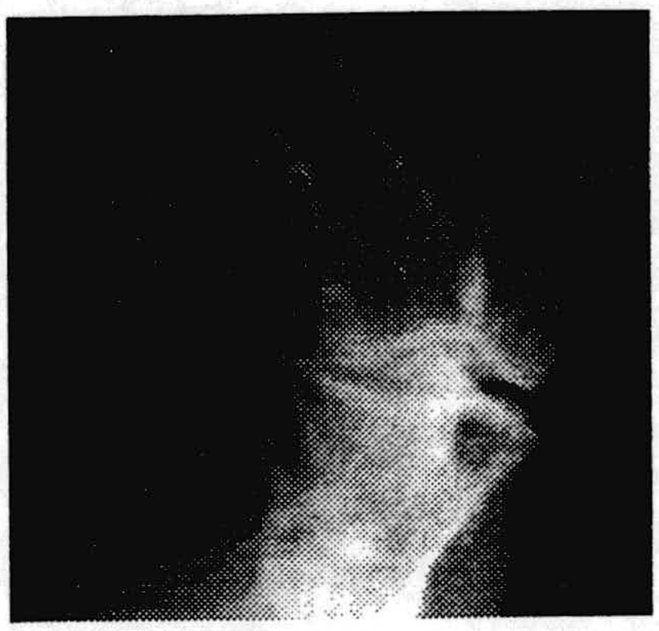
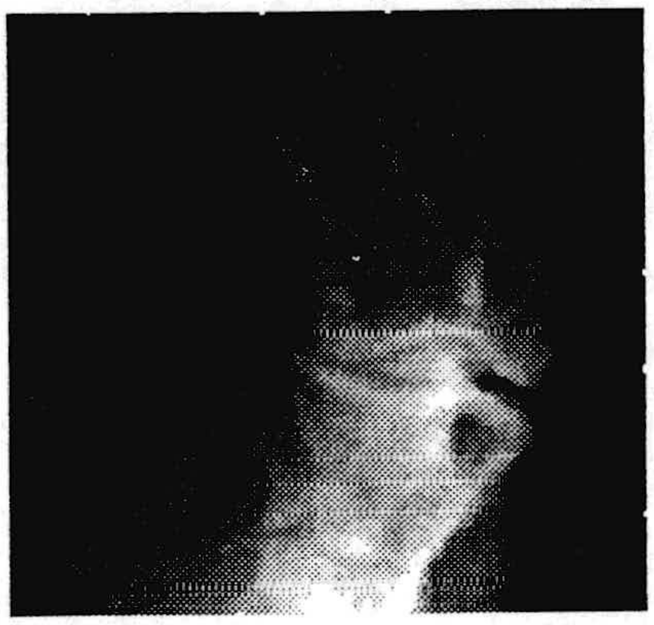
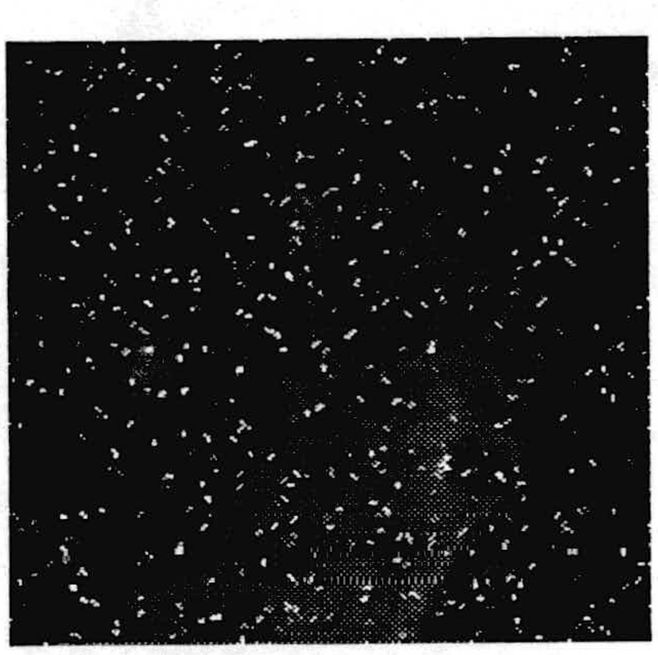
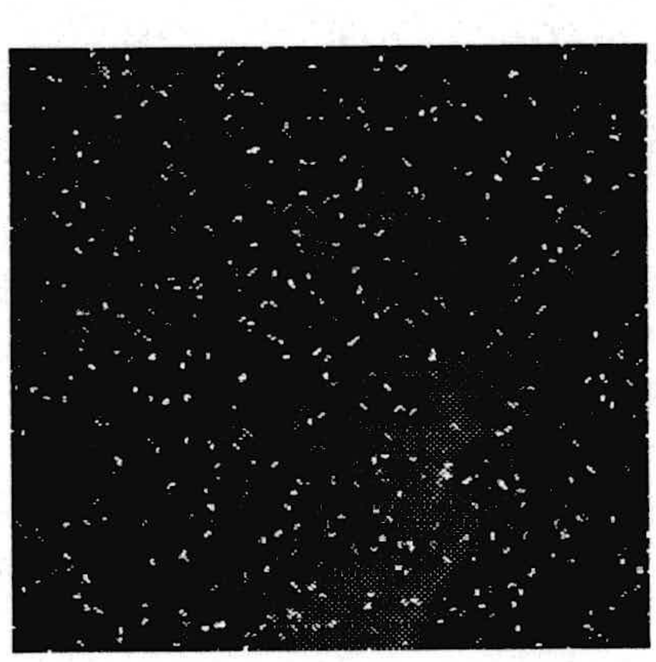


Figura 10: Se muestran visualmente los resultados del filtrado de una imagen fuente de radiografías contaminada por 10% de ruido impulsivo uniforme, con los distintos algoritmos, empleando una ventana de tamaño 3×3. Junto a estas líneas, la imagen **original** y la imagen **fuente**.

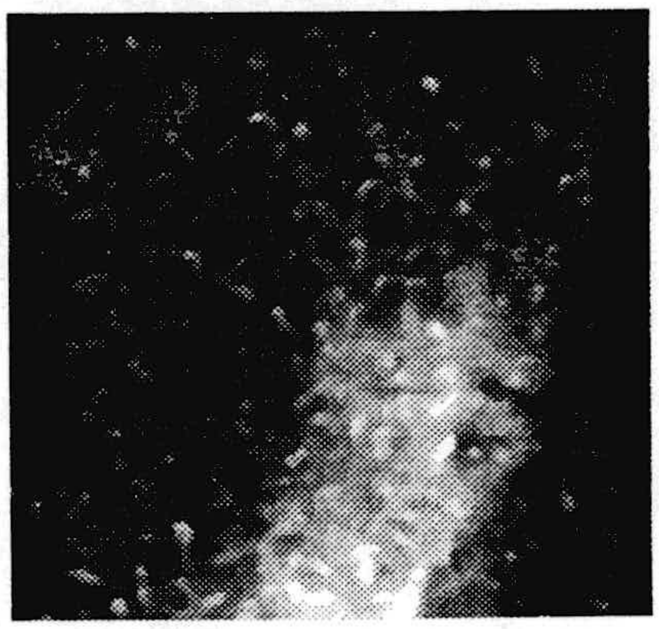


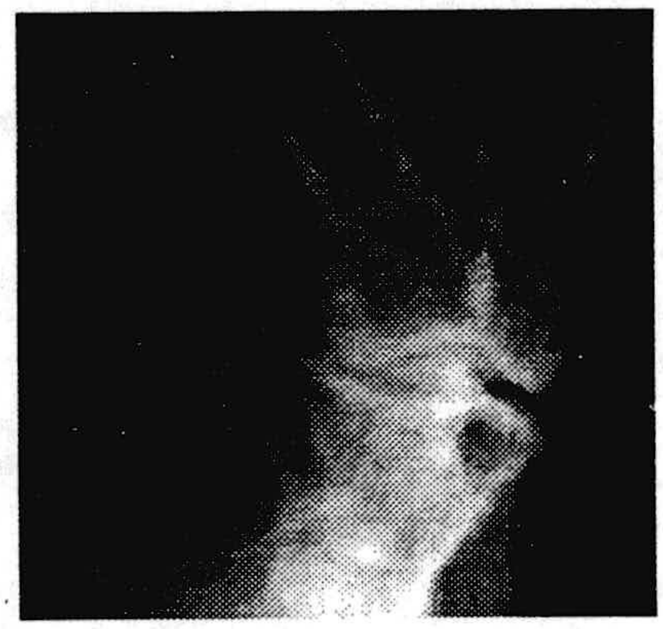
A la izquierda se muestra la imagen resultado aplicando el filtro de **mediana** tradicional. A la derecha, el correspondiente al uso del filtro de **mediana pesado en el centro**. Los resultados son muy buenos, pero se advierte cierta pérdida de detalle y aparición de *blotches*.





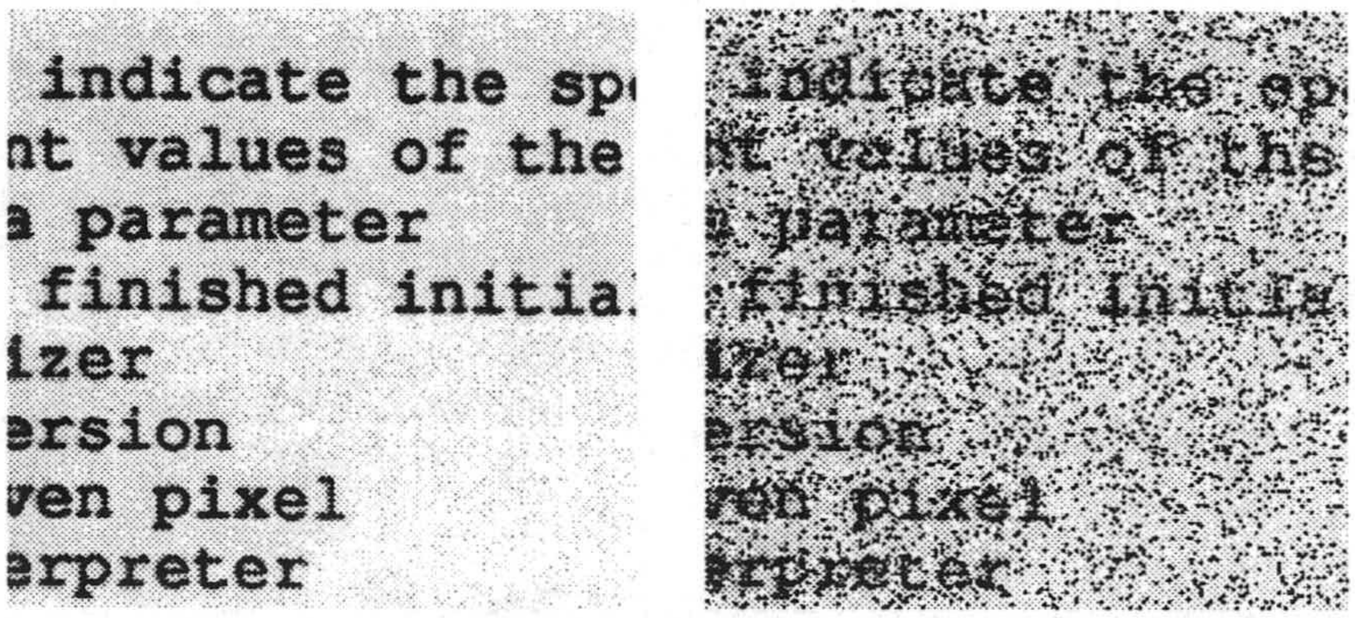
Resultado de filtrar con el híbrido **FIR-mediana** (izquierda) y **mediana multietapa** (derecha). Obsérvese cómo respetan estructuras de ruido, impidiendo una total eliminación de las distorsiones.



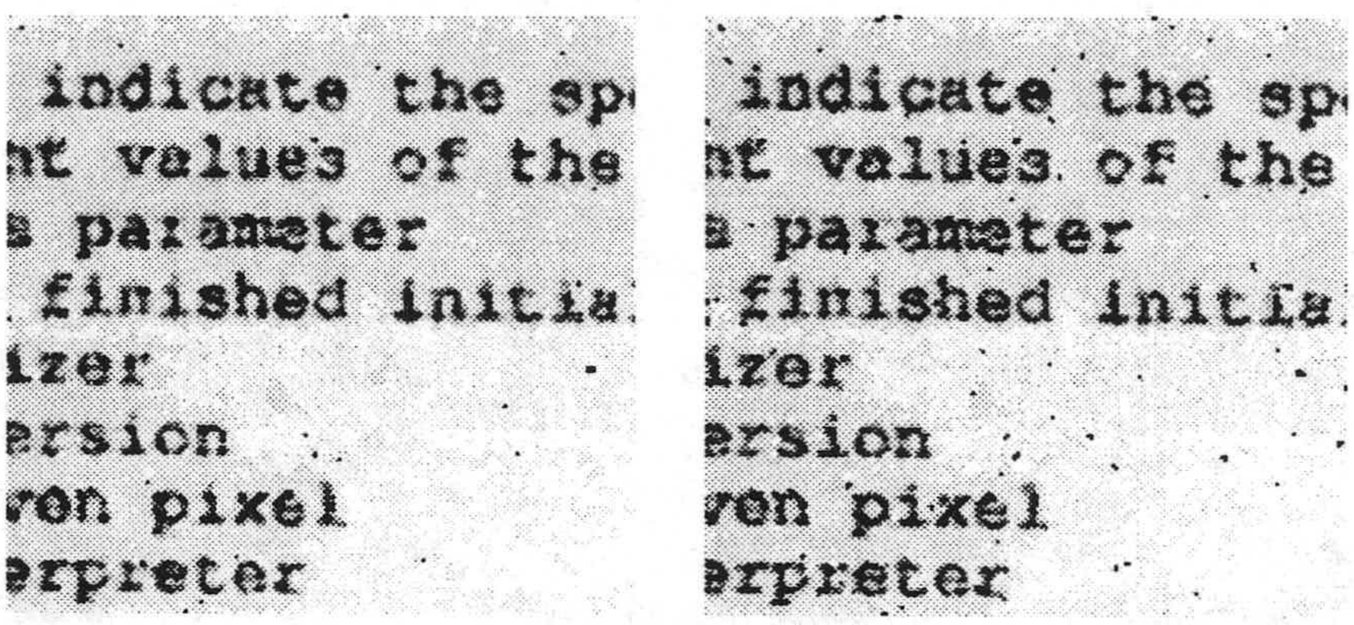


El filtro **adaptativo de Sun&Venetsanopoulos** produce una salida como la de la izquierda. El comportamiento no es bueno. Sin embargo, el filtro **Máscara-Mediana** proporciona un resultado excelente. Si bien deja algunos pulsos inalterados, el respeto por los detalles es máximo. No se aprecia presencia de *blotches*.

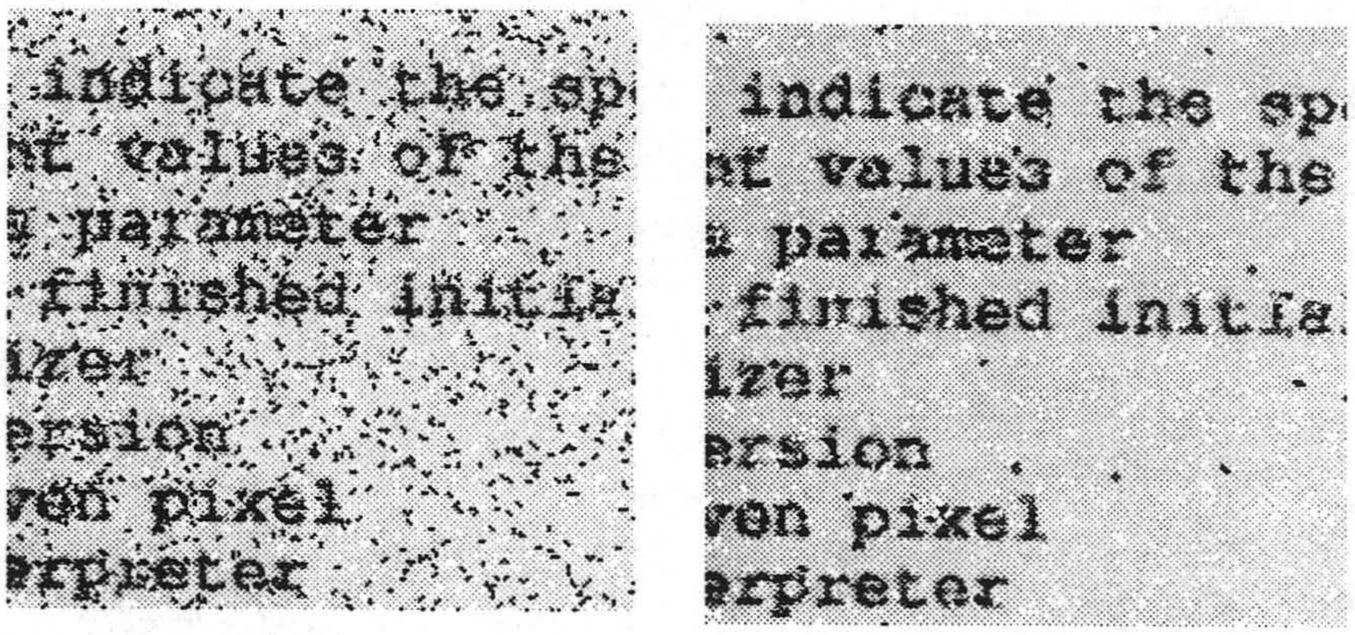
Figura 11: Resultado de filtrar con los diferentes algoritmos una imagen de texto, con ventana 3×3. Junto a estas líneas, imagen **original** de textos mecanografiado; e imagen **fuente** contaminada por 20% de ruido impulsiv *sal y pimienta*.



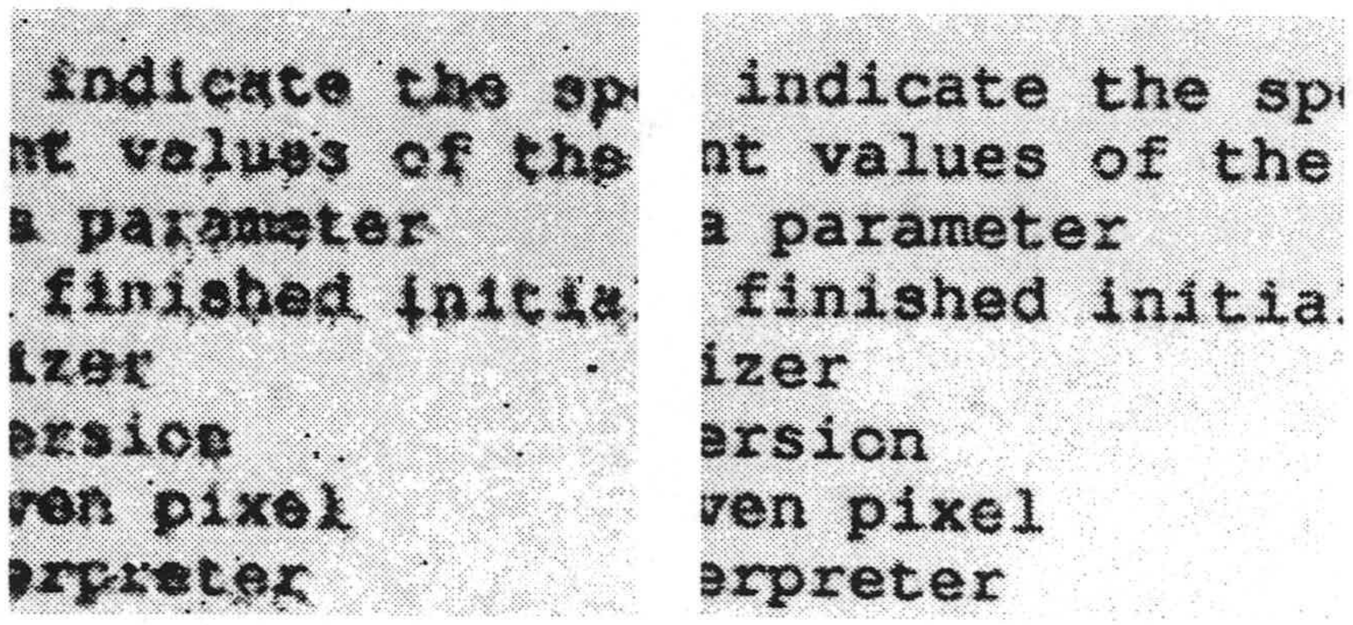
El filtro de **mediana** estándar produce una salida como la de la izquierda, junto a estas líneas. A la derecha, resultado de aplicar el filtro de **mediana ponderado en el centro.** Este último caso, si bien proporciona unas letras más legibles, elimina menios ruido que la mediana clásica. Los *blotches* son evidentes.



A la izquierda, el resultado de filtrar la imagen fuente anterior con el algoritmo híbrido **FIR-mediana**. A la derecha, salida del filtro de **mediana multietapa**. Aunque este último modifica más las letras, elimina mejor el ruido. La presencia de *blotches* también es notable. Obsérvese cómo se respetan las agrupaciones de ruido en las direcciones de los subfiltros empleados.



Se aprecia el resultado de aplicar el algoritmo **adaptativo de Sun&Venetsanopoulos** en la imagen de la izquierda. Aparecen *blotches*, se estropean las letras y no se elimina todo el ruido. En cambio, a la derecha se observa cómo el filtro **Máscara-Modelo** produce una salida de excelente calidad. No hay *blotches* notables, y la eliminación de ruido es total, preservando al máximo la legibilidad.



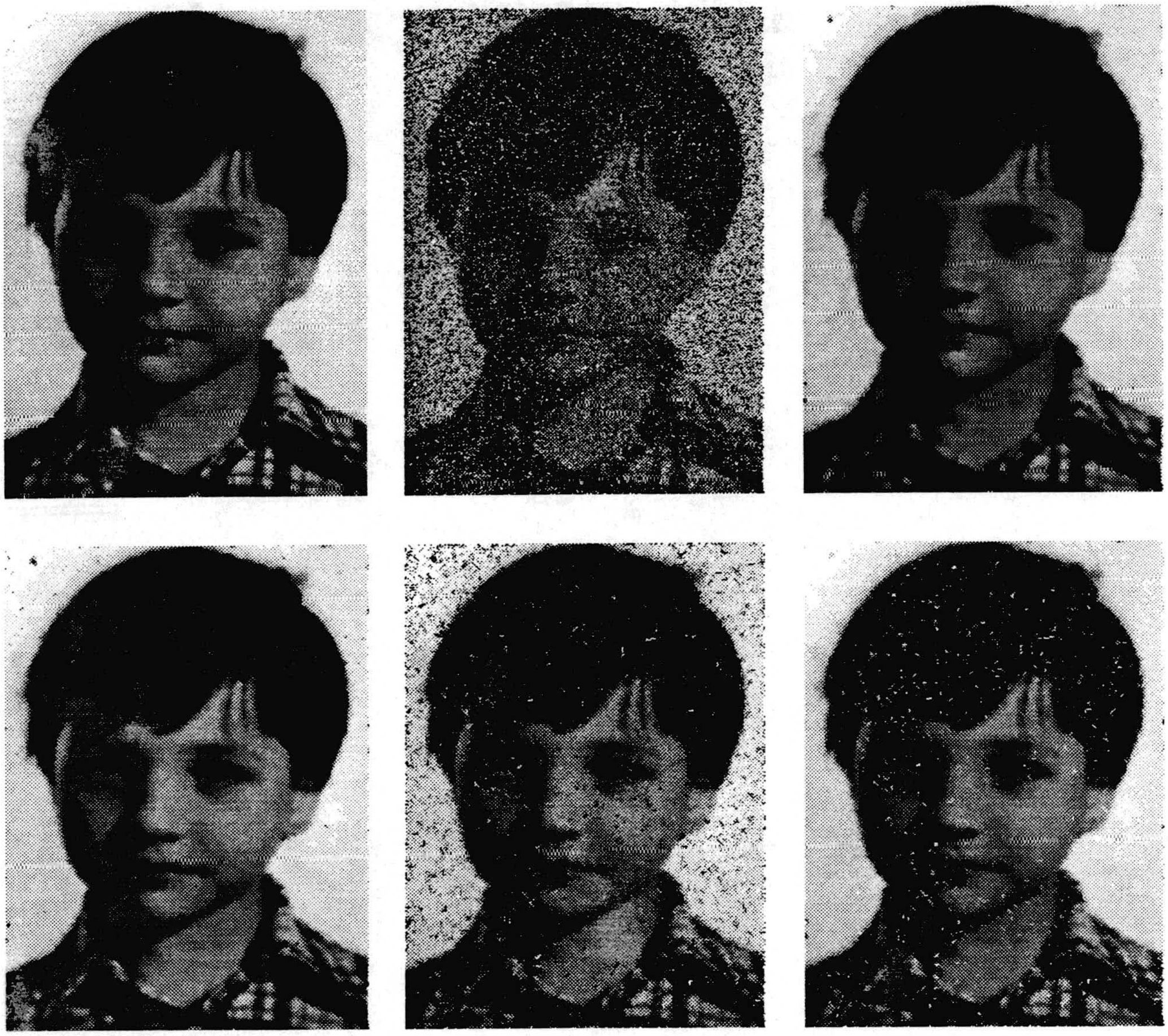
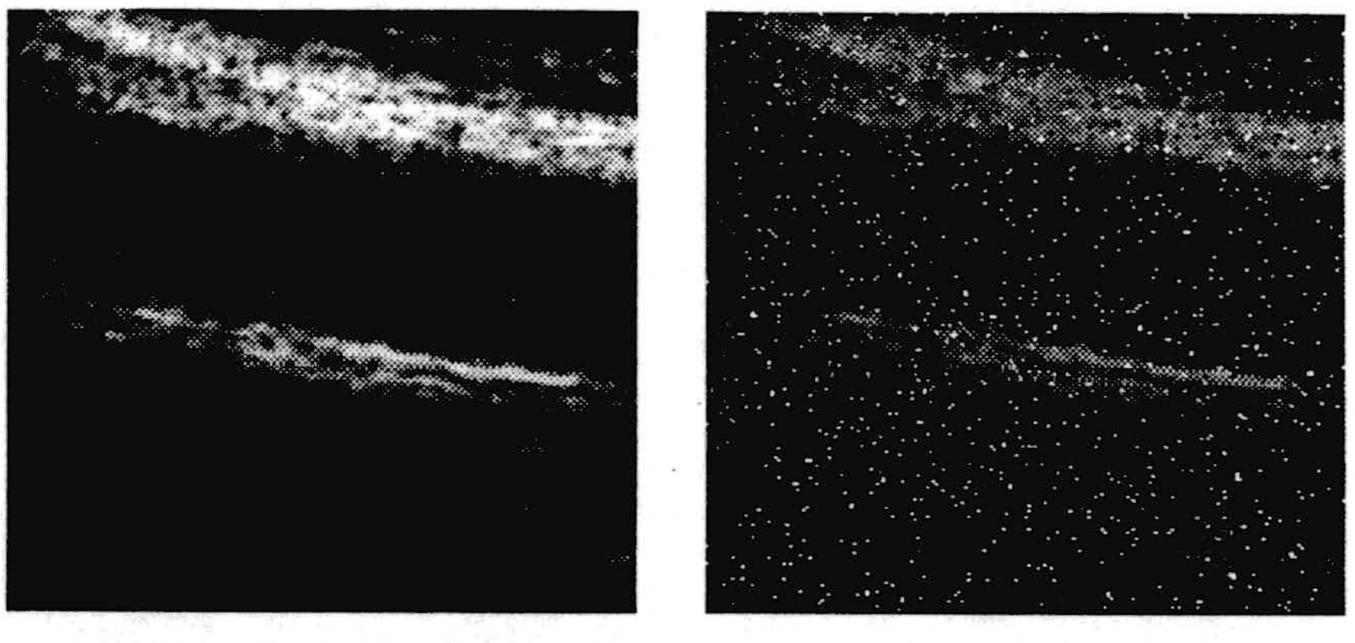
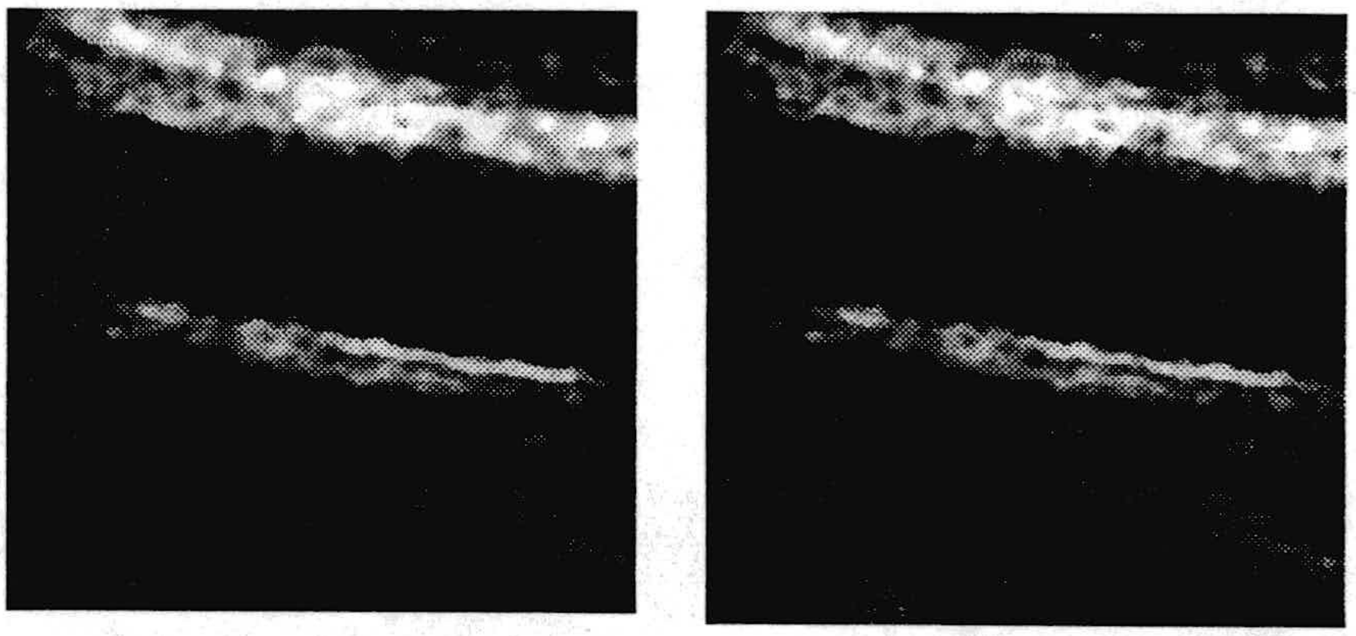


Figura 12: De izquierda a derecha, y de arriba a abajo: imagen original; imagen fuente (30% ruido uniforme); resultado con mediana, mediana ponderada, híbrido FIR-mediana, mediana multietapa, adaptativo de Sun & Venetsanopoulos y Máscara-Modelo Ponderado. En todos los casos se ha empleado una ventana de 5×5. Mmientras FM y MM preservan estructuras de ruido, y el adaptativo de Sun no realiza un buen filtrado, el Máscara-Modelo se comporta muy bien, mejorando a la mediana estándar y ponderada en preservación de detalles.

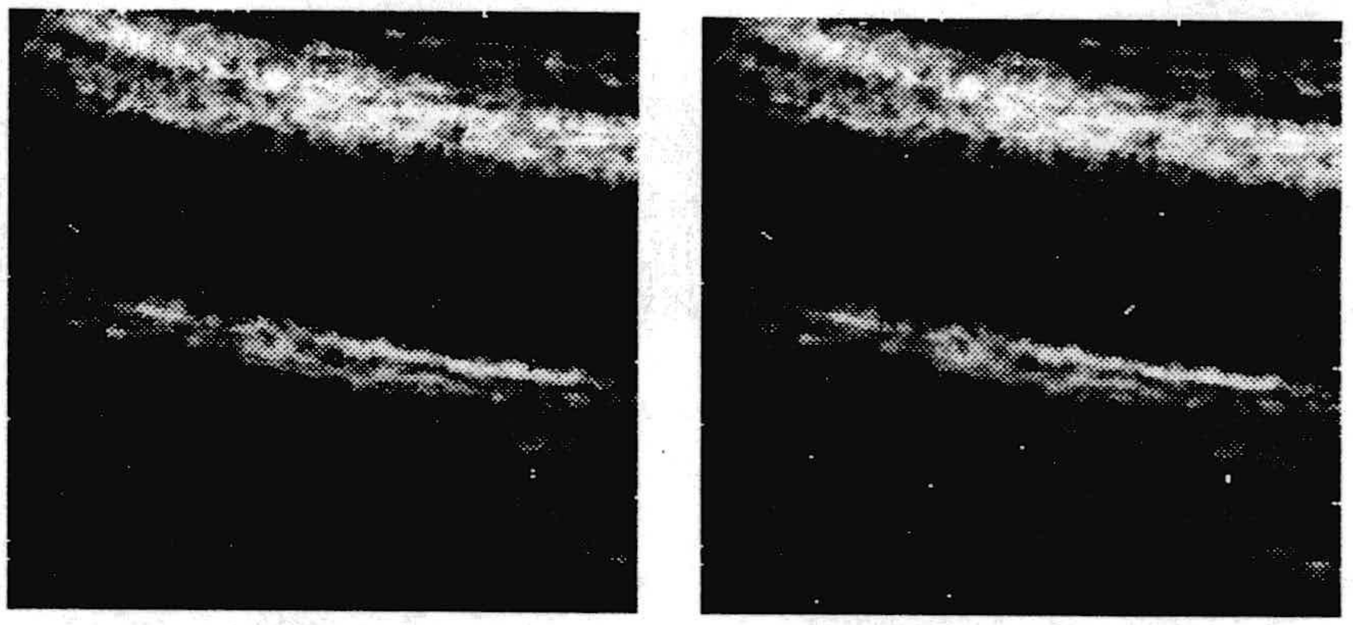
Figura 13: Imágenes de resultado obtenidas al aplicar los distintos filtros a una imagen fuente de ecografías con ventana de tamaño 5×5.. A la izquierda de estas líneas, imagen original y fuente (5% ruido impulsivo *sal y pimienta*).



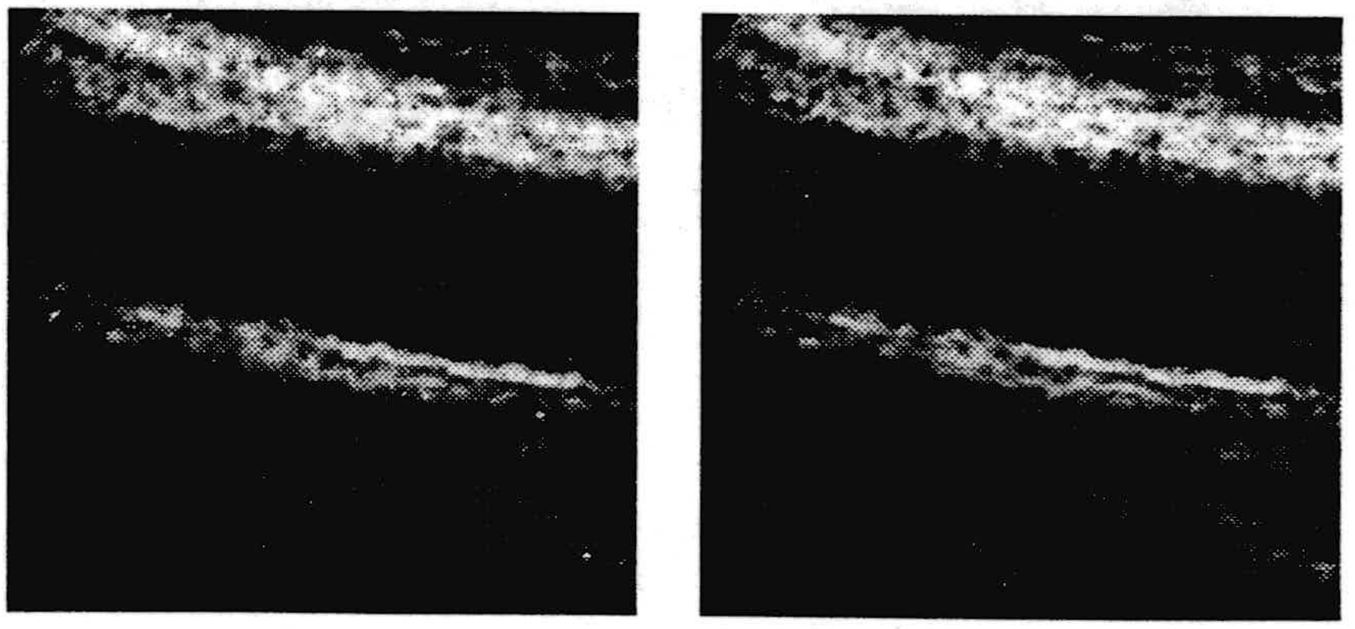
El filtro de **mediana** elimina todo el ruido, pero difumina los detalles (izquierda). El filtro de **mediana pesado en el centro** adolece del mismo problema, si bien en menor medida (derecha).

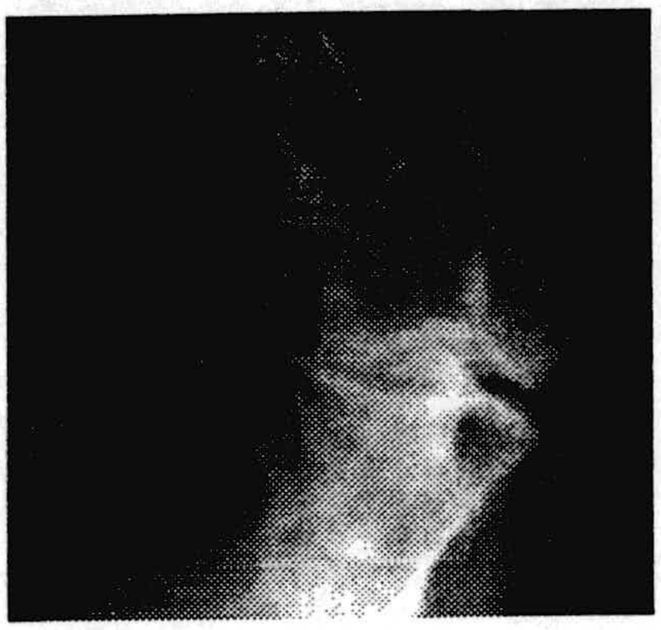


Los filtros de **mediana multietapa** (derecha) e híbrido **FIR-mediana** (izquierda) preservan ruido que se disponga espacialmente según la dirección de los subfiltros que utilizan. En el primero de los casos mencionado el efecto es más acentuado.



A la izquierda, el resultado de utilizar el algoritmo **adaptativo de Sun y Venetsanopoulos**. El procesamiento es visualmente mejor que los que produjo este mismo filtro en casos anteirores. Pero sigue manteniendo ruido sin eliminar. Mientras tanto, el filtro de **Mediana-Combinada** proporciona un resultado muy bueno, con supresión de pulsos y preservación de detalles.





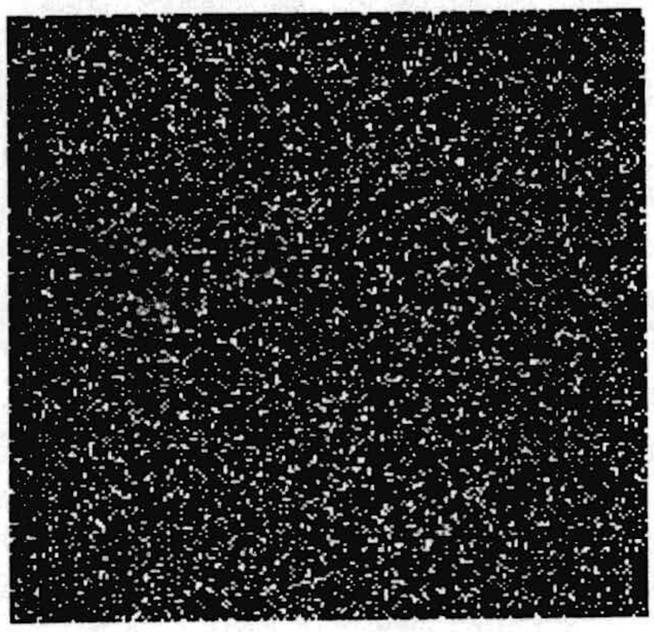
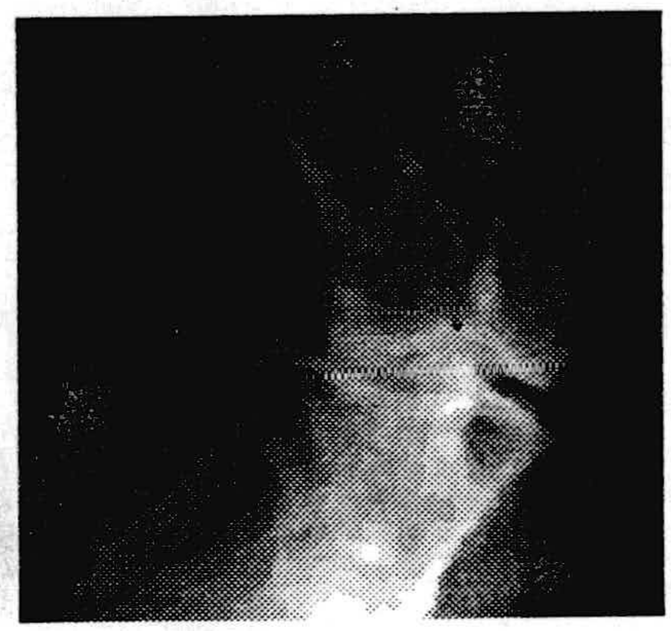
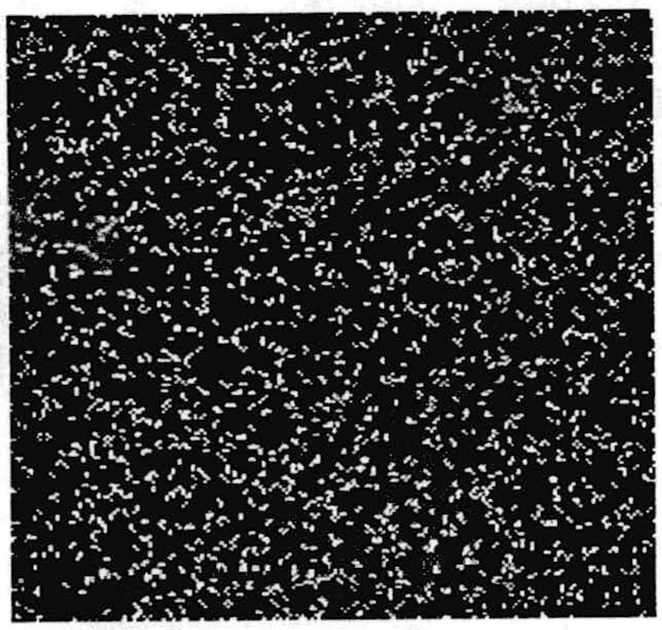
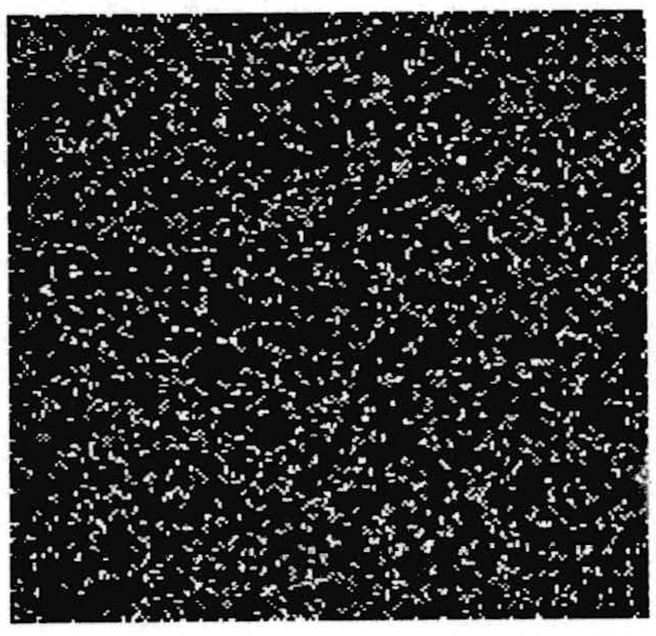


Figura 14: Imágenes resultado obtenidas al procesar una imagen fuente de radiografías (afectada por 30% de ruido impulsivo *sal y pimienta*) con todos los filtros, empleando ventanas de tamaño 3×3. A la izquierda, la imagen **oroginal** y la imagen **fuente**.

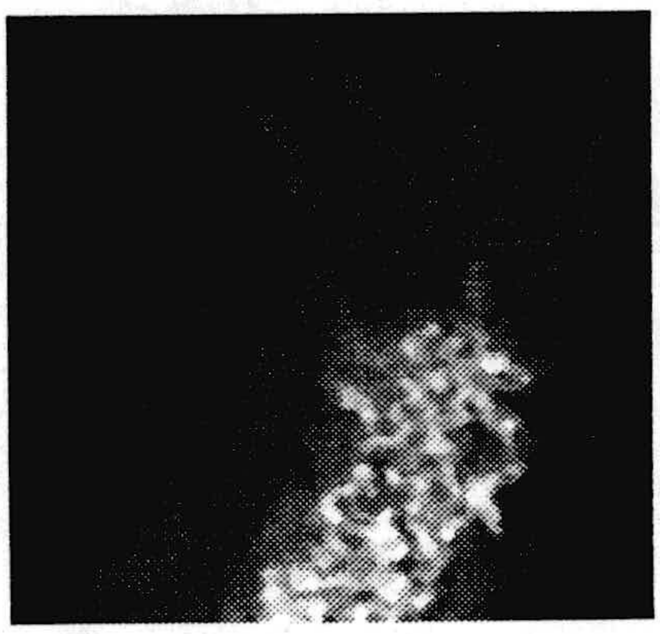


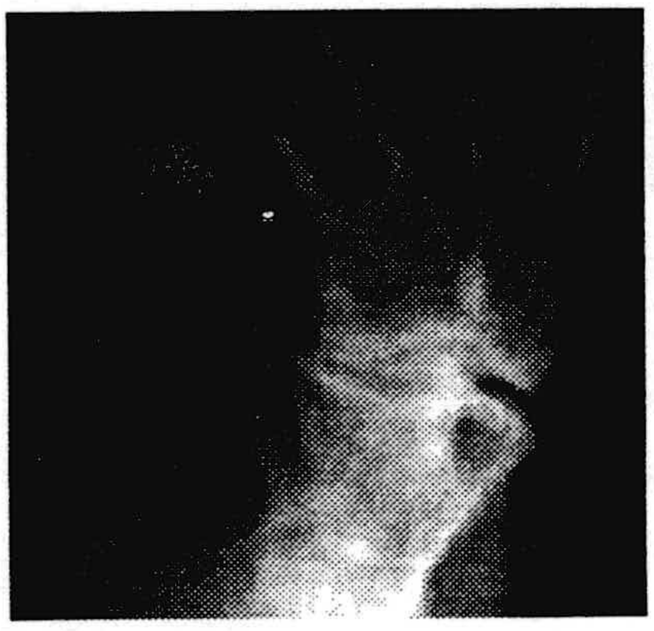
A la izquierda, el filtro de **mediana** estándar hace un procesamiento muy bueno, aunque no elimina por completo el ruido, y ha difuminado los ocntornos. A la derecha, el filtro de **mediana ponderado en el centro** elimina menos ruido, aunque preserva mejor algunos detalles. En ambos casos se originan *blotches*.





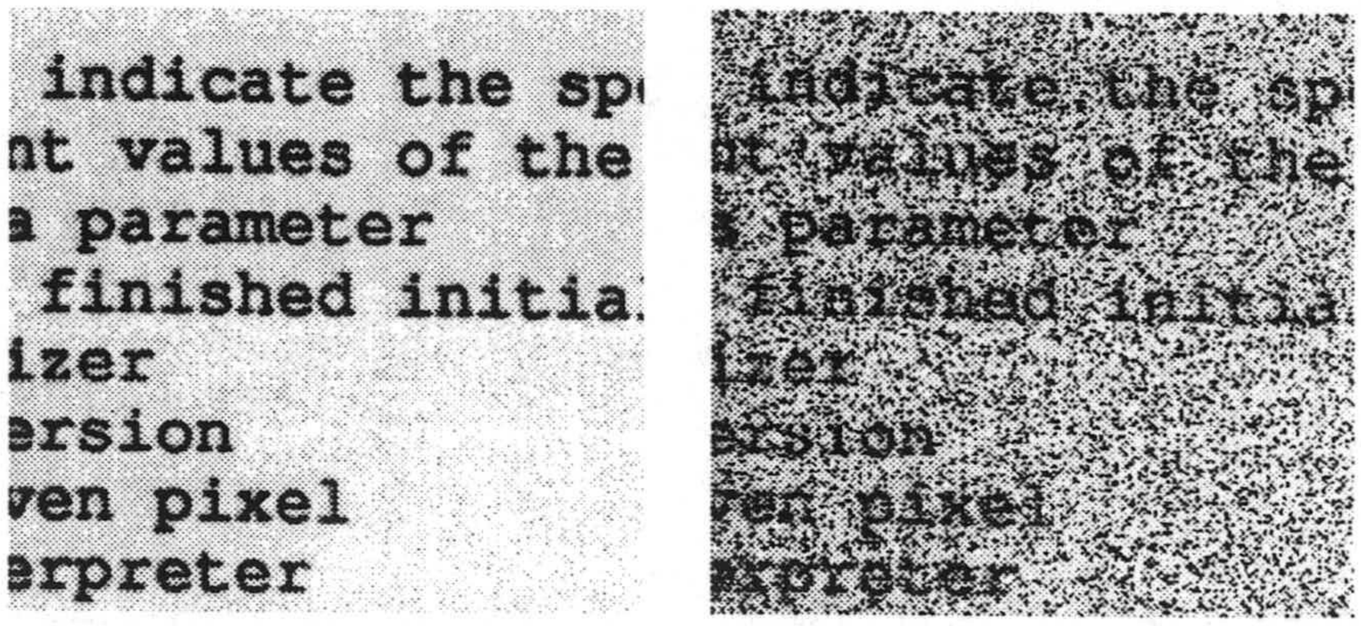
El filtro híbrido **FIR-mediana** (izquierda) no se muestra muy indicado para este procesamiento. Tampoco lo es el filtro de **mediana multietapa** (derecha). Apenas eliminan ruido, pues los subfiltros que utilizan tienen pequeño tamaño, y alcanzan en seguida la probabilidad de colapso, provocando que la salida del filtrado sea un píxel ruidoso.



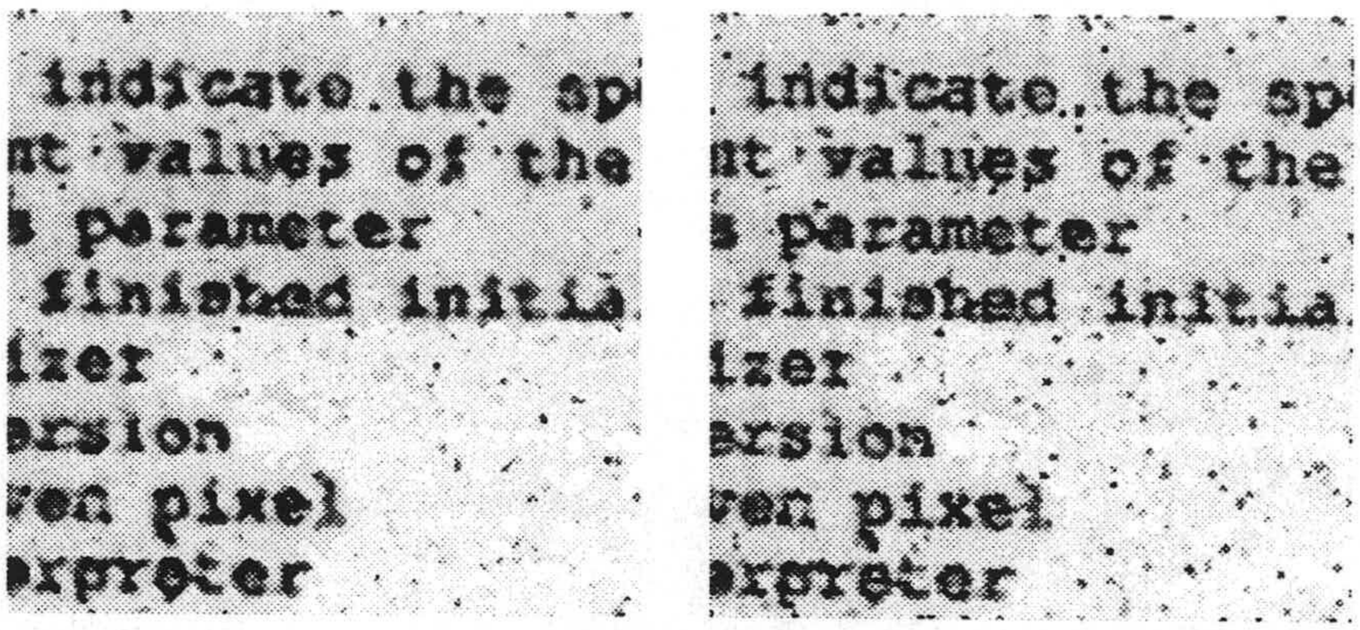


El algoritmo **adaptativo de Sun y Venetsanopoulos** (izquierda) demustra ser efectivo para eliminar el ruido,pero destruye los detalles del hueso. A la derecha, el resultado del filtro **Restaura-Modelo** es excelente. Aunque deja tres píxeles de ruido, no afecta a la visibilidad de la imagen. No se producen alteraciones notables de los contornos y detalles. No se originan *blotches*.

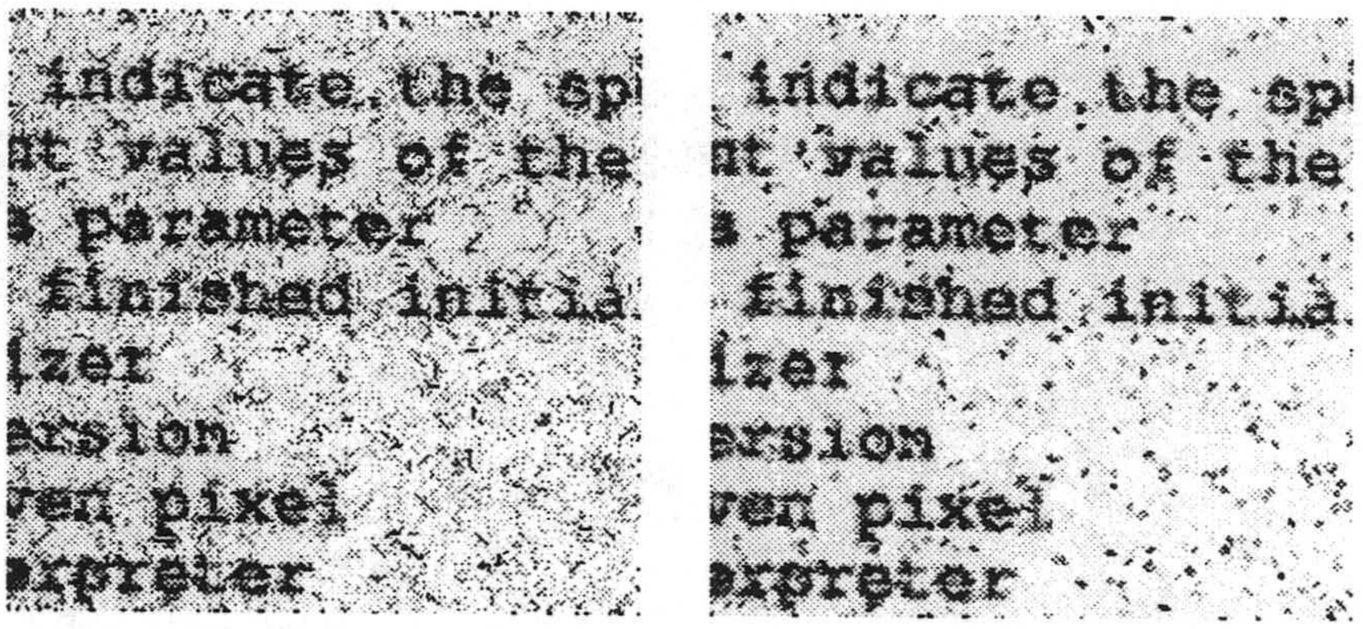
Figura 15: Resultados de filtrar una imagen de la clase de textos con diferentes filtros, empleando ventana de tamaño 5×5. Junto a estas líneas, la imagen **original** y la **fuente** (30% de ruido impulsivo uniformemente distribuido).



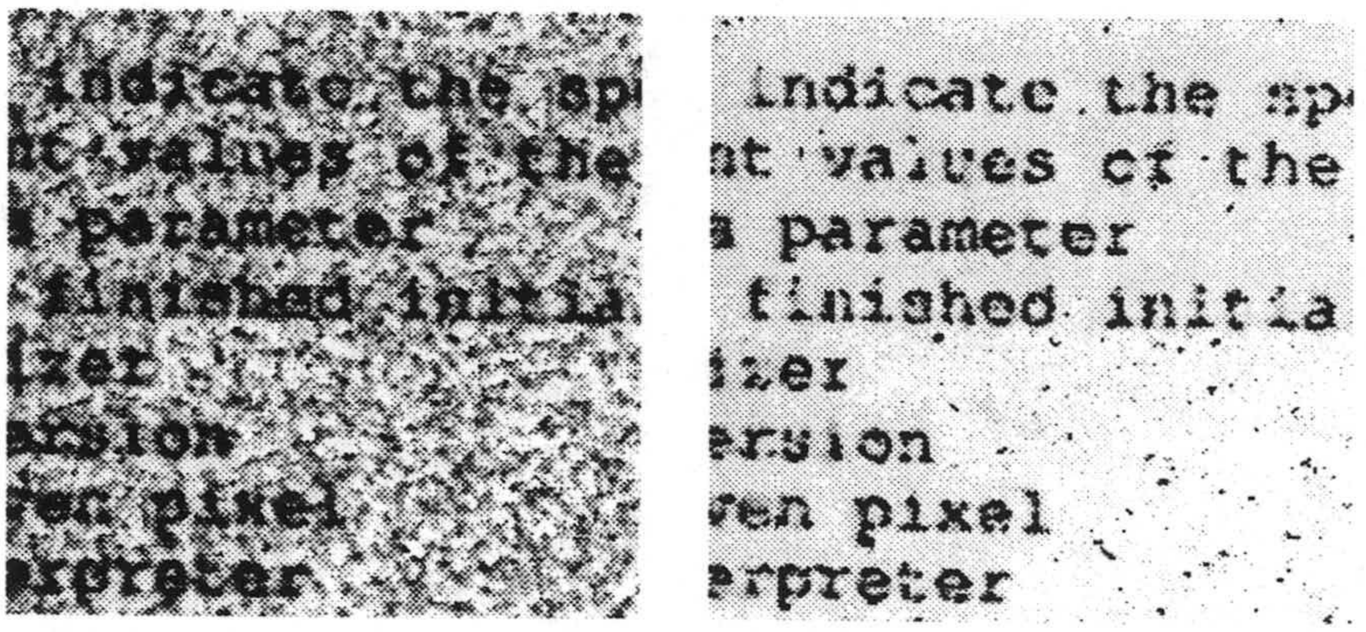
El filtro de **mediana** tradicional no es capaz de eliminar todo el ruido, ni de mantener un texto legible, aunque el resultado no es malo. El filtro de **mediana ponderado en el centro** consigue mayor respeto por los contornos de las letras a cambio de una menor supresión de ruido. En ambos casos se aprecia la aparición de *blotches*.



A la izquierda, imagen resultado de aplicar el filtro híbrido **FIR-mediana**. El ruido no se elimina satisfactoriamente debido a la direccionabilidad de los subfiltros empleados. Aunque ello ocurre también en el resultado de utilizar el filtro de **mediana multietapa**, no lo hace de forma tan acusada. La legibilidad es buena en este último caso.



El esquema **adaptativo de Sun & Venetsanopoulos** (izquierda) no demuestra ser muy adecuado en este caso. En cambio, el filtro **restdiver**, empleando máscara de estimación de ruido, elimina ruido aceptablemente y preserva la legibilidad al menos tan bien como el filtro de mediana A la vista de los demás resultados, puede considerarse que proporciona el texto más parecido al original.



DISCUSIÓN

Los filtros diseñados propuestos demuestran ser, en general, mucho mejores que los utilizados para comparar. El CMAE llega a valer hasta 9 veces más. En nigún caso se ven nuestros algoritmos superados por los de comparación. Tan sólo en el caso de radiografías afectadas de ruido impulsivo gris se produce poca diferencia. En cualquier caso, los resultados visuales son mejores en el caso de los filtros propuestos. A la vista de las gráficas y tablas, podemos realizar los siguientes comentarios:

• Los resultados con nuestros algoritmos, empleando $\underline{M}$, son muy similares, aun cuando varíen parámetros como porcentaje y tipo de ruido. Ello pone de manifiesto la importancia de la máscara de estimación de ruido, que es la característica común.

• Los valores de CMAE son más elevados para ruido impulsivo *sal y pimienta* que en el caso del ruido uniformemente distribuido. Además, en el primer caso, los resultados apenas varían al crecer el porcentaje de ruido, cosa que no sucede con ruido uniforme, donde las radiografías crecen a medida que lo hace la cantidad de ruido introducida; mientras que el texto y la fotografía hacen lo contrario. Las ecografías siempre son estables frente a ese parámetro. Esta última característica de los modelos se explica por la cohesión de los mismos. En efecto, un modelo poco cohesionado significa que una región de nivel k' puede componerse de muchos valores de k diferentes, siempre que mantengan la consistencia. Por lo tanto, en esos modelos hay un cierto número de niveles de gris k que poseen probabilidad elevada. La detección de ruido impulsivo uniforme será, por lo tanto, poco exacta con este tipo de matrices de composición. Pero además, el nivel de gris estimado pertenecerá a un intervalo ciertamente amplio. Como resultado, el filtro actuará de una manera poco adecuada, aunque siempre eliminando parte del ruido, a costa de estropear algunos detalles. Si el nivel de ruido es pequeño, el efecto neto es malo, pues distorsiona más que limpia, pero cuando el ruido es muy abundante, se compensan sus características y el resultado es más que aceptable. Lo contrario sucede si el modelo es muy cohesionado. Detectará mejor los pulsos y preservará mejor los detalles. Sólo tendrá dificultades cuando el ruido esté presente en gran porcentaje: entonces, aunque realizará un filtrado bueno, no podrá superar a los resultados obtenidos cuando el ruido era menos abundante.

• Se obtienen mejores salidas de los filtros cuando se emplea

ventana pequeña.

• Todos los algoritmos, incluso los convencionales, proporcionan mejores resultados cuando se trabaja con radiografías; y los peores cuando lo hacen con textos. Ello evidencia que la cohesión de la clase de imágenes es un factor primordial.

• Nuestros filtros suelen mejorar al repetir el proceso de filtro, convergiendo a un valor óptimo antes de las cinco iteraciones programadas. Ello no sucede con los filtros de referencia, que, por lo general, empeoran su comportamiento con la repetición del procesamiento.

• Los filtros propuestos no funcionan de manera aceptable cuando no emplean la máscara, aunque en muchos casos son comparables con filtros de referencia. Ello no indica necesariamente que el proceso de filtrado en sí mismo tenga una comportamiento desastroso; tan sólo subraya la enorme importancia que tiene la máscara de estimación de ruido en el proceso. En efecto, recuérdese que al no emplear $\underline{M}$, los algoritmos trabajan con la columna equivocada del modelo. Por esta razón, los resultados que produzcan no serán los adecuados. No podemos deducir que el diseño de filtro sea erróneo, pero sí que es poco robusto frente al ruido cuando no se emplea la máscara.

• Puede decirse que el filtro Máscara-Modelo se ha reveleado como sorprendentemente bueno, teniendo en cuenta que una vez que la máscara se ha obtenido, el filtro tan solo utiliza la escasa información proveniente del nivel de gris de ventana k', sin considerar en nigún caso la que proporciona el histograma de la ventana. Este hecho pone de manifiesto el preponderante papel que juega la máscara de estimación de ruido.

• Ningún caso es favorable al filtro Máscara-Modelo Ponderado. Este filtro ha sido diseñado para, de alguna manera, reforzar mutuamente la buena calidad de los algoritmos de Máscara-Mediana y Máscara-Modelo. Sin embargo, aunque ciertamente comparables, los resultados no son mejores en ningún caso. Podemos apreciar cómo el filtro Máscara-modelo Ponderado se comporta, como era de esperar, como un híbrido entre Máscara-Mediana y Máscara-Modelo, con sólo observar las figuras 1, 2 y 3, sobre todo al no emplear la máscara, que es cuando se aprecia más fácilemente el comportamiento del filtro por sí solo. En efecto, los resultados para radiografías, para todos los tama-

ños de ventana, mejoran a medida que aumenta el porcentaje de ruido, y son relativamente buenos, en el caso del Máscara-Mediana y Máscara-Modelo. Pero en el caso del Máscara-Modelo Ponderado, el comportamiento de las radiografías es muy distinto: los valores de CMAE son realmente malos, y además decrecen a medida que aumenta la cantidad de ruido. Esta anomalía es achacable a la ponderación excesiva: como el modelo de radiografías es muy cohesionado, presenta poca dispersión y todas las columnas de la matriz sólo presentan poquísimos niveles de gris con alta probabilidad, presentando el resto probabilidad prácticamente nula. Al realizar el producto entre los vectores histograma de ventana y columna de modelo, la mayoría de los niveles alcanzan valores despreciables, sesgando mucho el resultado, y proporcionando una salida del filtro inexacta. Este comportamiento extraño e indeseable es el que se pretendió corregir con el diseño del Mediana Combinado. Y, a la vista de la figura 4, se logra ese objetivo. Obsérvese que en ese caso, las radiografías de nuevo presentan las mismas características que en los filtros Máscara-Mediana y Máscara-Modelo.

• Se obtienen buenos resultados con el filtro Restaura-Modelo, el único que tiene un diseño basado en un criterio diferente. Su mejor comportamiento tiene lugar con radiografías afectadas de ruido *sal y pimienta* y con la fotografía en blanco y negro. Este último caso tiene particular interés, puesto que se ha empleado un modelo teórico. Se observa además en la figura 5 una mayor diferencia entre el resultado obtenido al usar la máscara de estimación de ruido, y los conseguidos sin su utilización, que en los casos de los filtros Máscara-Modelo, Máscara-modelo Ponderado y Mediana Combinado. Ello evidencia la necesidad de una buena corrección del acceso a la LUT (columna adecuada de la matriz de composición) que, por construcción, tiene el Restaura-Modelo.

• Se observa en las figuras 10-15 que nuestros filtros eliminan la casi totalidad de los pulsos del ruido, pero que a veces permanecen pequeños cúmulos de ellos sin procesar. Esto se prouce sólo en casos de ruido muy abundante, y con más facilidad cuando el tamaño de la ventana es pequeño. La explicación de este fenómeno es la siguiente: cuando el ruido está presente en un porcentaje elevado, es posible que se formen agrupaciones de píxeles afectados por el mismo. La máscara de estimación de ruido detecta como ruidosos ese cúmulo de píxeles, de manera que cubren la totalidad de la ventana deslizante. Entonces el

filtro no tiene píxeles limpios en esa vecindad para poder realizar la estimación del nivel de gris central. Ante esta indeterminación, el algoritmo no realiza operación alguna, dejando el píxel central con el valor original, que en este caso es un impulso. Este resultado indeseable se corregirá en las sucesivas iteraciones del filtro, puesto que el mismo irá eliminando el ruido empezando por las orillas del cúmulo inicial, reduciéndolo paulatinamente hasta su total eliminación. Además, la probabilidad de que el ruido forme esos grumos de tamaño mayor o igual que el de la ventana es poco probable.

• Otro comportamiento anómalo se observa en las radiografías contaminadas por ruido impulsivo uniforme. Cuando se las filtra con nuestros algoritmos, con ventana pequeña, se observa una fuerte caída de los valores de CMAE, desde 110 con 5% de ruido a 20 con 30%. Este efecto se aprecia visualmente como una serie de "*islas*" de nivel de gris medio, que emborronan la imagen. La causa que las origina está relacionada con la forma del modelo de radiografías, que recordemos era muy cohesionada, y de bajo contraste. Esto indica que los niveles de gris medios son muy probables. Como el ruido es uniformemente distribuído, algunos de sus pulsos tienen niveles de gris considerados probables por el modelo. Es decir, algunos píxeles contaminados se consideran píxeles limpios. En consecuencia, los píxeles con nivel intermedio sobrevivirán al filtrado, siempre que no estén aislados en una región de nivel de gris muy distinta. Pero ello sucederá pocas veces si el ruido es abundante, pues entonces incluso en esas regiones homogéneas, como puede ser el fondo de la radiografías, puede haber pulsos suficientemente cercanos con niveles de gris parecidos. Al realizar las primeras iteraciones de filtrado, estos impulsos entran a formar parte de la información local que provocará la estimación de los niveles filtrados. Ello se traducirá en la aparición de píxeles con nivel de gris intermedio en los alrededores de esos pulsos, que también tenían nivel de gris medio. De este modo, la aparición de esas "*islas grises*" queda explicada por la utilización de una máscara de estimación de ruido con un modelo de bajo contraste. Cuanto menor sea el tamaño de la ventana, más acusado será este molesto efecto, puesto que menos elementos formarán parte de la vecindad que da lugar a la estimación.

• El tiempo de cómputo no es muy superior a la de los filtros de referencia (ecepto la mediana y mediana ponderada).

COMENTARIOS Y PROPIEDADES

A la vista de las gráficas y las tablas adjuntas, pueden deducirse las siguientes propiedades.

• ***Comportamiento frente a ruido poco abundante.*** Para niveles bajos de ruido, los filtros propuestos presentan más ventajas que el resto. Ello se debe a que el principal objetivo es dejar inalterados los píxeles no contaminados. Por lo tanto, en una imagen casi libre de ruido, apenas interviene el algoritmo y no estropea los contornos y detalles. Obsérvese que en tal caso, la imagen fuente es casi una raíz del filtro.

• ***Raíces.*** El problema de calcular raíces [COY89] en los filtros aquí diseñados es extremadamente complejo y depende de la matriz de composición. Sin embargo, el cálculo se simplifica enormemente si se supone una máscara de estimación de ruido ideal , $\underline{M} = M$, sea cual sea el modelo. Entonces la subimagen no filtrada ($\underline{M}_{ij} = M_{ij} = 1$) es claramente una subraíz, mientras que la imagen filtrada ($\underline{M}_{ij} = M_{ij} = 0$) es procesada. En el caso particular del algoritmo Máscara-Mediana, se realiza un filtrado de mediana a la subimagen con $\underline{M}_{ij} = M_{ij} = 0$, por lo que en ese caso se presenta el mismo comportamiento, en cuanto a raíces, que el filtro de mediana estándar (*cf.* capítulo I y II). Por tanto, podemos decir, *grosso modo*, que el filtro de Máscara-Mediana tiene las mismas propiedades de generación de raíces que el filtro de mediana, aunque se espera una convergencia más rápida. Este aserto se comprueba por completo a la vista de las tablas. En cuanto al estudio de las raíces del resto de los filtros propuestos, es de una complejidad extraordinaria, incluso bajo la hipótesis simplificadora aquí empleada.

• ***Apilamiento.*** Debido al empleo de la matriz de composición, nuestros filtros no poseen la propiedad de descomposición por umbral —*threshold decomposition*— por lo que no pueden ser filtros apilados —*stack filters*—. Si así fuese, tendría que realizarse una descomposición en un conjunto de modelos binarios, esto es, con dos filas, una para cada nivel de apilamiento. Sin embargo, ese conjunto representaría, o bien un modelo inconsistente, o bien uno universal, independiente de la clase de imágenes.

• ***"Blotches".*** Los filtros propuestos atenúan la conocida e indeseable propiedad de los filtros tipo mediana consistente en generar parches, o regiones de nivel de gris constante, como resultado colateral de su procesamiento. Pitas [PIT90] discute

este inconveniente y propone una medida del efecto mediante la probabilidad de obtener la misma salida para dos píxeles, bajo ciertas suposiciones, como función de la distancia entre ellos. Este y otros análisis son aplicables a los filtros diseñados, pero hay que tener en cuenta que en nuestro caso no todos los píxeles son procesados. Para que sea aplicable al caso que nos ocupa, cualquier medida válida para dos píxeles en el filtro de mediana tradicional debe ser multiplicada por la probabilidad de que esos píxeles sean filtrados. Dicha probabilidad se calcula directamente bajo la hipótesis de máscara perfecta, $\underline{M} = M$. En efecto, dado que el ruido impulsivo es espacialemente independiente, si p es la probabilidad de filtrar un pñixel dado, la probabilidad de procesar dos es p^2. Este es, por lo tanto, el factor de reducción aproximado de la aparición de *"blotches"*. Puede comprobarse visualmente la atenuación de este efecto con más claridad en las figuras 11 y 15, en el fondo del texto; y en la figura 12.

• ***Robustez.*** La probabilidad de colapso (*cf.* capítulo I) constituye una medida de la robustez de l filtro frente a la presencia de ruido impulsivo. Se podía definir como el número de píxeles aberrantes que debe haber dentro de la región cubierta por la ventana para que la salida del filtro sea también aberrante [LON92]. Esta medida es muy difícil de calcular teóricamente en el caso de nuestros filtros, debido a la complejidad de la estadística inherente, que incluye la interrelación entre la matriz de composición con la imagen original y resultado. Sin embargo, de nuevo utilizando la hipótesis de máscara perfecta $\underline{M} = M$, la probabilidad de colapso puede estimarse fácilmente, ya que una máscara ideal detecta todos los píxeles aberrantes, por lo que ninguno de ellos entra a formar parte de la vecindad que se emplea en el filtrado. De este modo, ningún píxel aberrante podrá aparecer en la salida del algoritmo. Sólo existe un caso en el que es posible que el resultado del procesamiento sea un píxel aberrante: cuando todos los píxeles en la región cubierta por la ventana son aberrantes. En ese caso, el filtro no puede actuar, al no disponer de información local fiable, por lo que deja los valores de los píxeles inalterados. En este caso, la probabildad de colapso es del 100%. Pero la probabilidad de que toda la ventana sea ruidosa es pequeña (tanto menor cuanto mayor sea el tamaño de la ventana),por lo que podemos asegurar que la robustez de nuestros algoritmos supera con creces la de los empleados con fines comparativos. La robustez puede observarse visualmente en las figuras 10-15.

INFLUENCIA DE LOS PARÁMETROS DE FILTRADO

Los algoritmos de filtrado que se han desarrollado poseen una serie de parámetros cuyo valor condiciona el resultado final. En esta sección estudiaremos cómo afecta cada uno de ellos al procesamiento [GAR??].

• ***La ventana.*** Los filtros propuestos trabajan mejor con ventanas de tamaño 3×3 que con 5×5, respecto a los resultados de los otros algoritmos. Sin embargo, esta diferencia decrece a medida que el ruido aumenta. La razón para ello estriba en que cuanto mayor sea la ventana, más elementos entran a formar parte de la vecindad que determina el nivel de gris estimado. Por lo tanto, aunque la estadística es más rica, la localidad se pierde. Además, los detalles pequeños se pierden tanto más fácilamente cuanto más grande es la ventana.

• ***El ruido.*** El comportamiento general con respecto al ruido se ilustra en la siguiente tabla, la cual muestra los resultados de un experimento llevado a cabo sobre una imagen fuente libre de ruido ($S = O$). En este caso se expresa el MAE entre el original y el resultado (pues el CMAE es nulo al serlo el MAE entre fuente y original).

Mas-Med	Mas-Mod	MMP	MED-Gom	Res-Mod	Mediana	Ada-Sun	CWM	MM	FM
0.0028	0.0029	0.0029	0.0029	0.0029	3.31	1.36	2.19	2.25	2.26

Se aprecia una mejora importante en comparación con los otros filtros. En los algoritmos aquí presentados, sólo unos pocos píxeles (16) han sido procesados, frente al gran número de píxeles en la imagen (40000). Estas cifras corroboran el éxito de los filtros basados en modelos multirresolución. Por otro lado, los resultados son mejores cuanto menor es el ruido, evidentemente, puesto que es más fácil de detectar, y más sencillo de eliminar, al ser la mayor parte de los píxeles limpios. Además, el ruido *sal y pimienta* proporciona mayores valores de CMAE, pues al estar constituidos por pulsos de niveles de gris extremos, resulta más sencillo de detectar.

• ***El umbral de selección.*** Cuanto más cohesión, medida según su divergencia de Jensen-Shannon (*cf.* capítulo I), tiene una matriz de composición, mayor es la confianza que puede depositarse en su uso, en el siguiente sentido: si el modelo dice que un determinado nivel de gris es probable, podemos asegurar que realmente lo es. En el extremo opuesto, un modelo poco

cohesionado podrá asignar probabilidades elevadas a niveles de gris que realmente no aparecen en la región. Por tanto, cuanto mayor sea la cohesión de la matriz de composición, menor umbral de selección habrá de elegirse, dado que el objetivo principal es eliminar el ruido. Esto concuerda con las conclusiones obtenidas en el anterior capítulo.

• ***La matriz de composición.*** Pueden apreciarse sutilmente diferencias locales con respecto al modelo empleado. Por ejemplo, una radiografía se filtra mejor que otra clase de imágenes, con respecto a los filtros de referencia. Podemos apreciarlo por los elevados valores del CMAE. Sin embargo, un análisis teórico que proporcione resultados concluyentes acerca de la influencia del modelo en el resultado precisa de un estudio más detallado y complejo de la calidad de la matriz de composición en relación con las operaciones involucradas en la fabricación de la máscara y en el filtrado propiamente dicho. En cualquier caso podemos decir que los modelos más cohesionados proporcionan los mejores resultados, como cabía esperar.

• ***La máscara de estimación de ruido.*** Queda patente en todas las gráficas y tablas que $\underline{M}$ es un elemento fundamental del mecanismo de procesamiento que se ha desarrollado en el presente trabajo. Su doble función de localización de ruido y de corrección de la entrada a la LUT la hacen imprescindible. En las figuras 1-5 queda patente su importancia, al presentar una enorme mejora cuando el algoritmo la emplea. Además, cuando se usa $\underline{M}$ todos los filtros producen resultados muy similares, lo cual subraya el papel de la máscara de estimación de ruido. En otro orden de cosas, la máscara provoca que los algoritmos alcancen una raíz antes que si no la usaran, debido a que entonces sólo se filtra una parte de la imagen, y no la totalidad. Esta característica se aprecia en las tablas, en las que se observa que todos los filtros propuestos, cuando usan máscara, alcanzan valores convergentes antes de la quinta iteración.

• ***Iteraciones.*** Tan solo los resultados obtenidos con la fotografía en blanco y negro presentan un diferente comportamiento en cuanto a convergencia, según sea el filtro. Ello es achacable, sin duda, a que el modelo hipergeométrico, si bien es adecuado, no es tan acertado como el correspondiente empírico. Los algoritmos Restaura-modelo y Máscara-modelo alcanzan raíz porque su modo de operar está más determinado por el modelo.

CONCLUSIONES

En este último capítulo se han presentado nuevos algoritmos de filtrado. Todos ellos se diseñaron para reducir los daños colaterales que usualmente producen los filtros tipo mediana sobre los píxeles que no han sido afectados por el ruido tipo impulsivo. Este objetivo se consigue mediante el empleo de unamáscara de estimación de ruido, la cual preselecciona aquellos píxeles que van a ser procesados. Dicha máscara se construye, como se especificó en el anterior capítulo, a partir de la información proporcionada *a priori* por una adecuada matriz de composición, o modelo, el cual se obtiene basándose en la idea de la multirresolución. El comportamiento de estos filtros ha sido examinado respecto a diferentes configuraciones de operación.

Los resultados han mostrado que estos algoritmos son más eficientes en eliminar ruido impulsivo que otros filtros tipo mediana conocidos propuestos en la literatura. Además, presentan un buen comportamiento en cuanto a obtención de raíces, robustez, reducción del efecto creado por la aparición de *blotches*, a pesar de no ser filtros apilados.

Tablas 1, 2 y 3: En las páginas siguentes se muestran los resultados numéricos para ecografías, texto, radiografías y fotografías. En negrita se indica que el mejor resultado se obtuvo antes de la 5ª iteración, y q ue luego empeoró. Cursiva y subrayado signifoca que el mejor CMAE se produce en la última iteración, por lo que es posible una mejora si se continúa la repetición del filtrado. El estilo normal quiere decir que se alcanza raíz antes de la 5ª iteración.

ECOGRAFÍA

Config	Mas-Med	Mas-Mod	MMP	Med-Com	Res-Mod	Mediana	Ada-Sun	CWM	MM	I
3/e/05/m	46,1381295	46,3650075	47,2601326	47,2601326	47,2601326	3,46042195	*2,3021035*	6,66444976	3,95193493	
3/e/10/m	44,6219848	47,8010336	45,1667248	45,1824843	46,1157407	6,25932908	*3,4948047*	9,98943146	2,70996568	
3/e/20/m	42,5670068	46,3593258	43,7806542	43,8573681	44,4256301	9,87430961	*5,1289754*	13,4762289	1,93693024	
3/e/30/m	38,3547999	43,6981714	39,9873457	39,9998954	42,5126751	12,899666	5,98118166	12,9350135	*1,5446352*	
3/g/05/m	23,0142045	22,7769447	22,9924314	23,0142045	22,3168044	2,60901771	1,32138973	4,97553486	3,02031939	
3/g/10/m	22,3109696	23,009596	22,7567433	22,688745	22,5595444	4,43612463	*1,647346*	7,37797571	2,44238347	
3/g/20/m	22,4001673	23,0020859	22,6491482	22,6655785	22,7067587	6,87472955	*1,5211785*	8,77120666	1,67939048	
3/g/30/m	16,3255014	14,8762402	15,7662295	15,9079741	14,6648821	7,61833449	*1,3660361*	*7,1667925*	1,38618968	
5/e/05/m	33,4020833	30,3655303	33,0918473	33,2634855	28,1157387	1,93483377	*2,3021035*	2,73111319	5,27791951	
5/e/10/m	33,2203694	31,0534772	32,7664474	32,8662437	28,4287596	3,80705004	*3,4948047*	5,30100704	6,05900243	
5/e/20/m	32,3126775	30,4642162	31,6988349	31,8197305	28,0819028	7,00260191	*5,1289754*	9,43935737	5,24846401	
5/e/30/m	31,5920846	30,4911483	31,5868649	31,6705873	28,4811173	10,0567859	5,98118166	13,051577	*3,8893195*	
5/g/05/m	14,1419843	13,2261224	13,8557583	13,9033181	12,7274156	1,46659827	1,32138973	2,07336945	3,72003674	
5/g/10/m	15,00873	14,530059	14,9422762	14,96682	14,065761	2,65526285	*1,647346*	3,73312029	4,39249904	
5/g/20/m	16,5592262	16,0777873	16,4156743	16,4820644	15,3812767	5,03159053	*1,5211785*	6,83313286	3,46621921	
5/g/30/m	16,4499365	15,9177516	16,3367359	16,3724138	15,3054317	6,32771373	*1,3660361*	8,00404585	2,41507642	
3/e/05/sm	3,46042195	*0,9341879*	2,31866662	2,60773391	0,56316903	3,46042195	*2,3021035*	6,66444976	3,95193493	
3/e/10/sm	6,25932908	*1,2404613*	3,77024981	4,40902281	0,77462329	6,25932908	*3,4948047*	9,98943146	2,70996568	
3/e/20/sm	9,87430961	*1,4228364*	5,38602569	7,83368283	0,9646133	9,87430961	*5,1289754*	13,4762289	1,93693024	
3/e/30/sm	12,899666	*1,4543618*	*5,7960406*	10,8038484	1,05503666	12,899666	5,98118166	12,9350135	*1,5446352*	
3/g/05/sm	2,60901771	0,77305765	1,91113907	2,09211036	0,4343428	2,60901771	1,32138973	4,97553486	3,02031939	
3/g/10/sm	4,43612463	*0,9759541*	2,88503309	3,18449656	0,60777343	4,43612463	*1,647346*	7,37797571	2,44238347	
3/g/20/sm	6,87472955	*1,0998744*	*3,7095734*	4,58578767	0,77144691	6,87472955	*1,5211785*	8,77120666	1,67939048	
3/g/30/sm	7,61833449	1,09788808	*2,6718438*	*4,1909526*	0,83908175	7,61833449	*1,3660361*	*7,1667925*	1,38618968	
5/e/05/sm	1,93483377	0,92443855	1,74006946	1,7804553	0,82378934	1,93483377	*2,3021035*	2,73111319	5,27791951	
5/e/10/sm	3,80705004	*1,3804047*	3,14051852	3,30676711	1,07388211	3,80705004	*3,4948047*	5,30100704	6,05900243	
5/e/20/sm	7,00260191	*1,523285*	5,456	5,79719745	1,24159313	7,00260191	*5,1289754*	9,43935737	5,24846401	
5/e/30/sm	10,0567859	*1,4954124*	7,19396049	8,23144819	1,30430734	10,0567859	5,98118166	13,051577	*3,8893195*	
5/g/05/sm	1,46659827	0,72701437	1,35494662	1,37772109	0,65451167	1,46659827	1,32138973	2,07336945	3,72003674	
5/g/10/sm	2,65526285	*1,0513696*	2,36399958	2,42916556	0,86803849	2,65526285	*1,647346*	3,73312029	4,39249904	
5/g/20/sm	5,03159053	*1,1463277*	3,91960776	4,16852708	0,9659961	5,03159053	*1,5211785*	6,83313286	3,46621921	
5/g/30/sm	6,32771373	*1,109899*	*4,0394759*	4,83191424	0,99096279	6,32771373	*1,3660361*	8,00404585	2,41507642	

Config	Mas-Med	Mas-Mod	MMP	Med-Com	Res-Mod	Restdiver	Mediana	Ada-Sun	CWM	MM
3/e/05/m	15,2468477	14,7427902	12,7696994	14,019621	10,1432489	13,7078701	0,95150039	0,86109475	1,81387567	2,1547582
3/e/10/m	12,543474	12,9428339	10,8456359	11,3923359	8,99166137	11,9061477	1,79417363	1,5341886	3,09340432	2,9185723
3/e/20/m	12,408432	13,0839183	11,1125193	11,3649625	9,31390184	12,1361466	3,0072779	2,37399057	4,04206823	3,3040926
3/e/30/m	11,4590807	12,4165914	10,5985781	10,9499985	9,1245053	11,2309461	3,73724112	2,77864778	4,4968139	*3,275250*
3/g/05/m	2,02193352	2,45181881	1,65543154	1,70687936	1,52054204	3,11402141	0,91217708	0,82004681	1,71180437	2,018323
3/g/10/m	2,69973211	3,06265642	2,29105376	2,33991942	2,12610395	3,944097	1,44543499	0,85583335	2,42269621	2,5406574
3/g/20/m	3,81963778	4,0809995	3,35748705	3,47309505	3,2236732	4,7979E-06	2,33524742	1,01235743	3,00042621	*3,106059*
3/g/30/m	3,67731939	3,95186616	3,42506652	3,52155062	3,40601875	4,7422E-06	2,75099616	1,03769544	3,06259767	*2,878329*
5/e/05/m	9,23183086	8,8361439	7,80124069	8,08304409	7,16068785	13,7078701	0,4739643	0,86109475	0,70003674	0,9829602
5/e/10/m	8,24099741	7,97799235	7,06737392	7,25267898	6,54599239	11,9061477	0,96064294	1,5341886	1,39638068	1,8534201
5/e/20/m	8,89732348	8,58439653	7,62023438	7,87220797	7,0791628	12,1361466	1,86667705	2,37399057	2,61520634	2,9193939
5/e/30/m	8,61338902	8,3037448	7,31762804	7,53525101	6,9310802	11,2309461	2,69837601	2,77864778	3,58599969	3,2290093
5/g/05/m	2,06449835	2,07968902	1,87285116	1,89836523	1,84148253	3,11402141	0,45226055	0,82004681	0,66377171	0,9438896
5/g/10/m	2,83830798	2,81861804	2,65970659	2,68506426	2,60894394	3,944097	0,78286002	0,85583335	1,11102104	1,5281411
5/g/20/m	4,13112631	4,06256467	3,87414605	3,9162846	3,78992704	4,7979E-06	1,52973265	1,01235743	2,04682127	2,4587147
5/g/30/m	4,35252819	4,28209357	4,16902491	4,22339316	4,10258627	4,7422E-06	2,14485632	1,03769544	2,56958214	2,8768573
3/e/05/sm	0,95150039	0,59657679	0,81029395	0,85760659	0,04325044	0,86966985	0,95150039	0,86109475	1,81387567	2,1547582
3/e/10/sm	1,79417363	0,92189594	1,52616555	1,60952501	0,10677859	1,66492147	1,79417363	1,5341886	3,09340432	2,9185723
3/e/20/sm	3,0072779	1,3725333	2,60144523	2,76728432	0,30594784	2,88710547	3,0072779	2,37399057	4,04206823	3,3040926
3/e/30/sm	3,73724112	1,69499621	3,24566742	3,42199441	0,60893502	3,57425952	3,73724112	2,77864778	4,4968139	*3,275250*
3/g/05/sm	0,91217708	0,57448012	0,76366232	0,80837516	0,04080129	0,84355677	0,91217708	0,82004681	1,71180437	2,018323
3/g/10/sm	1,44543499	0,76876767	1,15401455	1,27230982	0,09902121	3,65027981	1,44543499	0,85583335	2,42269621	2,5406574
3/g/20/sm	2,33524742	1,1008644	1,76125197	2,05800462	0,30690549	2,26498213	2,33524742	1,01235743	3,00042621	*3,106059*
3/g/30/sm	2,75099616	1,41743423	1,87349516	2,23995517	0,7011321	2,59487316	2,75099616	1,03769544	3,06259767	*2,878329*
5/e/05/sm	0,4739643	0,37775228	0,3844315	0,40246299	0,29400653	0,86966985	0,4739643	0,86109475	0,70003674	0,9829602
5/e/10/sm	0,96064294	0,67922	0,78402026	0,82018474	0,60038881	3,9504278	0,96064294	1,5341886	1,39638068	1,8534201
5/e/20/sm	1,86667705	1,15745826	1,57705637	1,64809709	1,1431719	2,88710547	1,86667705	2,37399057	2,61520634	2,9193939
5/e/30/sm	2,45610762	1,60085047	2,32589353	2,43302199	1,82791156	3,57425952	2,69837601	2,77864778	3,58599969	3,2290093
5/g/05/sm	0,45226055	0,36252972	0,36578963	0,38295126	0,28169014	0,84355677	0,45226055	0,82004681	0,66377171	0,9438896
5/g/10/sm	0,78286002	0,54834236	0,63042372	0,66257372	0,49804986	3,65027981	0,78286002	0,85583335	1,11102104	1,5281411
5/g/20/sm	1,52973265	0,94433934	1,27131764	1,34720893	0,97709368	2,26498213	1,52973265	1,01235743	2,04682127	2,4587147
5/g/30/sm	2,14485632	1,30164239	1,87221349	1,9970548	1,51430679	2,59487316	2,14485632	1,03769544	2,56958214	2,8768573

TEXTO MECANOGRAFIADO

RADIOGRAFÍA

Config	Mas-Med	Mas-Mod	MMP	Med-Com	Res-Mod	Mediana	Ada-Sun	CWM	MM	I
3/e/05/m	188,176471	201,261589	197,983713	197,983713	194,810897	14,8608802	9,2470714	27,3911672	5,57521556	
3/e/10/m	177,189729	177,69671	175,437853	175,437853	184,561664	27,7936899	13,51872	37,0554893	3,03313716	
3/e/20/m	169,074967	180,814603	174,205517	174,205517	182,644975	48,726466	14,786513	39,2354769	1,8328109	
3/e/30/m	163,008525	166,776276	160,612348	164,906425	167,727193	63,2095868	16,215833	27,415442	1,432642	
3/g/05/m	107,363095	108,656627	106,517717	106,517717	109,536437	12,7770956	3,6194649	21,3119338	3,97495042	
3/g/10/m	100,612431	86,6011015	85,391777	100,245902	87,3571429	23,1823926	2,7256519	32,0343423	2,51066353	
3/g/20/m	60,9749373	39,4524324	38,4344392	58,4519487	39,9855734	35,0674247	1,7719592	22,678509	1,56744432	
3/g/30/m	46,4672628	19,175385	17,862512	26,180548	19,343459	35,445985	1,3984913	10,606142	1,31243457	
5/e/05/m	194,188498	174,157593	184,744681	184,744681	174,658046	10,7557954	9,2470714	14,7526699	16,7302505	
5/e/10/m	182,393539	170,384088	177,951289	178,462644	173,72028	21,0703986	13,51872	28,8457966	15,0648878	
5/e/20/m	172,304229	168,174434	170,674324	178,462644	168,736139	39,1624806	14,786513	51,029899	6,95536526	
5/e/30/m	172,883363	169,737239	171,796047	171,873258	171,334229	52,9371539	16,215833	60,9528212	4,1566722	
5/g/05/m	72,3409091	70,4570313	71,765252	71,765252	70,5488918	9,4401605	3,6194649	13,3409763	11,5031888	
5/g/10/m	71,1966365	67,6937269	70,512492	70,4673496	68,4089497	17,7303479	2,7256519	24,6517357	10,5918014	
5/g/20/m	75,0380398	72,0029596	74,8073112	74,8840629	72,1928783	28,0611303	1,7719592	36,7137827	4,37065352	
5/g/30/m	42,8747838	37,4612254	39,2692261	40,9147937	37,7199719	32,0537207	1,3984913	40,7595193	2,5588602	
3/e/05/sm	14,8608802	1,1227672	0,82736888	11,9389118	0,7308571	14,8608802	9,2470714	27,3911672	5,57521556	
3/e/10/sm	27,7936899	1,1381421	0,90011812	20,4292763	0,84779775	27,7936899	13,51872	37,0554893	3,03313716	
3/e/20/sm	48,726466	1,16093243	0,94764999	25,9821024	0,9410761	48,726466	14,786513	39,2354769	1,8328109	
3/e/30/sm	63,2095868	1,16425444	0,96727751	22,9887586	0,98385109	63,2095868	16,215833	27,415442	1,432642	
3/g/05/sm	12,7770956	1,053358	1,22739645	10,4663443	0,64948268	12,7770956	3,6194649	21,3119338	3,97495042	
3/g/10/sm	23,1823926	1,0442182	1,19368832	16,4161074	0,76280172	23,1823926	2,7256519	32,0343423	2,51066353	
3/g/20/sm	35,0674247	1,0329324	1,11381787	23,953725	0,83113556	35,0674247	1,7719592	22,678509	1,56744432	
3/g/30/sm	35,445985	1,03382842	1,07342653	16,689319	0,86630746	35,445985	1,3984913	10,606142	1,31243457	
5/e/05/sm	10,7557954	1,1960997	2,3890963	6,61453912	0,93225252	10,7557954	9,2470714	14,7526699	16,7302505	
5/e/10/sm	21,0703986	1,1943499	1,8408845	13,2901776	1,02129584	21,0703986	13,51872	28,8457966	15,0648878	
5/e/20/sm	39,1624806	1,2074012	1,3876038	29,2867246	1,09883981	39,1624806	14,786513	51,029899	6,95536526	
5/e/30/sm	52,9371539	1,20640399	1,186414	43,7249028	1,1213782	52,9371539	16,215833	60,9528212	4,1566722	
5/g/05/sm	9,4401605	1,0220036	2,4644077	4,86172507	0,8513909	9,4401605	3,6194649	13,3409763	11,5031888	
5/g/10/sm	17,7303479	1,0452793	2,041244	8,301531	0,92405722	17,7303479	2,7256519	24,6517357	10,5918014	
5/g/20/sm	28,0611303	1,0367472	1,2014123	12,032808	0,95441944	28,0611303	1,7719592	36,7137827	4,37065352	
5/g/30/sm	32,0537207	1,03981721	1,0843362	10,421917	0,97046767	32,0537207	1,3984913	40,7595193	2,5588602	

FOTOGRAFÍA

Config	Mas-Med	Mas-Mod	MMP	Med-Com	Res-Mod	Mediana	Ada-Sun	CWM	MM	F
3/e/05/m	26,2010319	27,3464004	26,2121952	26,2121952	26,7849931	**7,34918885**	*2,682947*	**14,9848519**	**6,69871602**	
3/e/10/m	39,7751372	42,562681	40,2047967	40,2047967	41,4951347	**11,9993013**	*3,2231339*	**19,5702819**	**4,46566553**	
3/e/20/m	48,6084458	53,3294901	48,3593904	48,3593904	51,1951639	**17,5058957**	*3,6683682*	**22,146405**	*2,9573495*	
3/e/30/m	50,7864822	57,9523324	50,6985774	50,6890997	55,2598493	**20,6078773**	*3,7032272*	*16,875035*	*2,3111755*	
3/g/05/m	12,2395128	12,5004484	12,3099293	12,3030995	12,4072967	**4,42786736**	*1,7243064*	**9,21892569**	**6,08134556**	
3/g/10/m	18,2001658	18,9622875	18,2017092	18,2017092	18,558602	**7,14262685**	*1,9814736*	**12,0066752**	*4,5549336*	
3/g/20/m	*21,702085*	22,7985795	*21,669083*	*21,662641*	22,3072155	**9,84983956**	*1,8774876*	**13,2399929**	*3,0664083*	
3/g/30/m	*20,19685*	19,6598558	*20,247126*	*20,241589*	19,2566803	**11,1189002**	*1,6905855*	*11,027633*	*2,4098057*	
5/e/05/m	21,4193147	20,1744674	21,3866672	21,3941044	20,0284605	**3,97795871**	*2,682947*	**6,02130983**	**11,1285482**	
5/e/10/m	30,8921477	28,8261706	30,8472134	30,9293398	28,3772866	**7,24924695**	*3,2231339*	**10,7879417**	**11,7818204**	
5/e/20/m	40,2813631	37,0439274	39,971938	40,0833238	36,3129141	**12,9861523**	*3,6683682*	**18,1367597**	**8,90765773**	
5/e/30/m	44,5711173	41,3715332	44,4485785	44,5380677	40,5509332	**16,8411959**	*3,7032272*	**21,4831308**	**6,03673355**	
5/g/05/m	**5,60181544**	5,50056658	**5,5804769**	**5,59282443**	5,4735351	**2,40276259**	*1,7243064*	**3,66132676**	7,05862139	
5/g/10/m	**9,24855895**	9,16019176	**9,25201391**	**9,25201391**	9,03597923	**4,32884196**	*1,9814736*	**6,38886794**	**8,21323436**	
5/g/20/m	**13,5492404**	13,5767741	**13,5509045**	*13,565947*	13,4070832	**7,37972332**	*1,8774876*	**10,194453**	**7,26240752**	
5/g/30/m	*15,603722*	15,6600882	**15,6476017**	**15,6487322**	15,4277879	**9,13775733**	*1,6905855*	**11,5069487**	**5,30953081**	

CONCLUSIONES

CONCLUSIONES DEL TRABAJO

Dedicaremos este apartado a resumir de forma sucinta el trabajo realizado, destacando las ideas y resultados más importantes, y extrayendo las principales consecuencias que de ellos se derivan.

Multirresolución.

Este concepto es uno de los pilares sobre el que se fundamentea esta obra. Con este término se da a entender que una imagen es susceptible de ser observada —y por tanto procesada— bajo distintos niveles de resolución. Empleando el concepto de multirresolución es posible observar una imagen con poco detalle (ventana grande), sin perder información acerca de los elementos de mayor resolución (píxeles). Esta filosofía será la base de posteriores herramientas desarrolladas en este trabajo, como las matrices de composición.

Información. Entropía.

Para cuantificar el conocimiento adquirido mediante la observación de una imagen a una determinada resolución se ha escogido la entropía discreta de Shannon, introducida en el capítulo I, aplicada a la distribución de probabilidad representada por el histograma de niveles de gris, a la resolución correspondiente.Esta función se ha empleado en la medida de divergencia entre distribuciones de Jensen-Shannon, al cual ha sido muy útil en dos aspectos de la investigación. Por una parte ha permitido el diseño de un método de detector de bordes muy robusto frente al ruido impulsivo; mientras que por otro lado forma parte de la medida empleada como base del umbral de selección utilizado en la fabricación de la máscara de estimación de ruido.

Matrices de composición.

Las matrices de composición, o modelo, proporcionan información acerca de la probabilidad de hallar un píxel con nivel de gris *k* dentro de una región de nivel de gris de ventana k'. En este trabajo, hacemos la hipótesis de que una matriz de composición determina una clase de imágenes. Así, existirán modelos de radiografías, textos, etc. Se construyen mediante entrenamiento de la muestra de imágenes de la misma clase, buscando las regularidades que se presentan, como un sistema complejo adaptativo.

Esta herramienta permite llevar a cabo la idea de la multirresolución: proporciona información acerca del nivel más alto de resolución a partir de cierto conocimiento a un nivel inferior.

Máscara de estimación de ruido.

Es la herramienta más potente introducida en la presente obra. A partir de la información contenida en el modelo, y mediante un proceso de umbralización basado en la información (medida entrópicamente) se permite la construcción de una máscara inicial, que señala la posición de los píxeles afectados por ruido impulsivo. Esta máscara inicial se refina mediante un proceso iterativo, cuya convergencia se fuerza mediante la introducción de los conceptos de descartes y persistencia. Finalmente, la máscara definitiva, binaria, localiza de forma muy aproximada los pulsos en la imagen.

La calidad de la máscara se cuantifica mediante el porcentaje de acierto en detección de píxeles ruidosos y limpios. Los resultados son elocuentes, mostrando estas máscaras una capacidad de localización casi perfecta cuando el ruido es *sal y piomienta*, y más que aceptable cuando el ruido es impulsivo uniforme.

Hay que señalar que la máscara tiene dos objetivos. Por un lado, localizar los pulsos; por otra, corregir el nivel de gris de la ventana, que se ve alterado por la presencia de ruido impulsivo. Esto último se realiza con dificultad cuando el ruido impulsivo tiene niveles de gris uniformemente distribuidos. Por ello se probó la idea de construir máscaras con niveles de gris, propor-

cionales a la probabilidad de aparición del nivel del píxel en la región. Se han probadoen la práctica, demostrándo que el uso de la máscara gris no se justifica, debido a que las mejoras introducidas no compensan la complejidad que su empleo conlleva.En cualquier caso, las máscaras binarias muestran un comportamiento corrector excelente en la mayor parte de los casos, siempre que se elija el umbral adecuado —lo cual no presenta un grado de arbitrariedad excesivo, puesto que el valor del umbral puede predecirse con gran aproximación—.

Nuevos algoritmos de filtrado.

Se han diseñado e implementado novedosos métodos de procesamiento de imágenes contaminadas con ruido impulsivo. Describámoslos brevemente.

FILTROS NO BASADOS EN MODELOS. RESTDIVER; MEDIANA RELAJADA.

El filtro de mediana relajada es una variante muy simple del filtro de mediana tradicional, que emplea un intervalo central dentro del cual el nivel de gris se considera limpio y no sufre procesado. Se ha realizado un estudio estadístico acerca de sus capacidades, tanto en supresión de ruido como en respeto por los contornos de la imagen. Se aprecia cómo mejora al filtro de mediana tradicional en cuanto a preservación de bordes, mientras que mantiene un aceptable nivel de eliminación de ruido.

En cuanto al filtro *restdiver*, que utiliza una ventana que se adapta a la presencia de bordes en la imagen —los cuales se detectan mediante un innovador método basado en la divergencia de Jensen-Shannon, y que es muy robusto frente al ruido impulsivo—, puede decirse que presenta un comportamiento muy adecuado en cuanto a filtrado, y preservación de bordes. Frente a la mediana tradicional, se muestra más efectivo en este último aspecto.

FILTROS BASADOS EN MODELOS.

Se han realizado diseños de filtros que emplean la información procedente de la matriz de composición. Algunos son

muy simples, como el que estima el nivel de gris resultante usando la mediana de la columna adecuada del modelo. Otros son más sofisticados, como el que realiza una ponderación del histograma de la ventana empleando los valores de la correspondiente columna del modelo; o el que estima el nivel de gris resultante como el de mayor probabilidad sobre la diagonal del modelo que comienza en el nivel de gris de la ventana.

También se han implementado algoritmos que sólo emplean el modelo para construir la máscara de estimación de ruido, procesando tan sólo los píxeles ruidosos marcados por ésta. Es el caso del filtro de máscara-mediana —que realiza la mediana sobre los puntos indicados por la máscara—, o el *restdiver* con máscara.

Todos ellos proporcionan resultados excelentes. Se han realizado comparaciones con el filtro de mediana tradicional, el filtro de mediana de centro ponderado, y el filtro adaptativo propuesto por Sun y Venetsanopoulos.

Debe reseñarse un hecho importante: la matriz de composición se emplea tanto para construir la máscara de estimación de ruido como para permitir la estimación del nivel de gris en el procesamiento.

PROBLEMAS ABIERTOS

La terminación de una Tesis Doctoral no supone la conclusión de la investigación. Antes al contrario, sugiere nuevos caminos, tanto como continuación, como innovación paralela al campo aquí estudiado.

Como problemas abiertos podemos indicar algunos que pueden tener un futuro desarrollo, puespueden originar aplicaciones prácticas o estudios teóricos de interés. Pasemos a numerarlos.

Univocidad de la matriz de composición.

La hipótesis más fuerte realizada en el presente trabajo es la que aduce que cada clase de imágenes está determinada por una matriz de composición. Desgraciadamente, este enunciado no es unívoco. En efecto, es posible construir modelos a partir de una muestra muy representativa de imágenes de una determinada clase; radiografías, por ejemplo. Sin embargo, este modelo se habrá construido bajo ciertas condiciones de brillo y contraste con que fueron adquiridas las imágenes. De este modo, una radiografía que se salga de esos parámetros no pertenecerá a la clase de radiografías del modelo en cuestión. Bien es cierto que este problema se soluciona realizando una nueva matriz de composición para ese otro grupo de radiografías, pero no deja de ser una limitación al método. Sería interesante realizar un esfuerzo investigador orientado a buscar matrices de composición más robustas.

Ventaneo múltiple.

A lo largo del presente trabajo se han empleado algoritmos que utilizan ventanas deslizantes, las cuales determinan una región de la imagen centrada en el píxel que se procesa. De esta manera se consigue información local. Pero es posible obtener información local de otra manera: a partir de los píxeles encerrados en todas las ventanas que contengan al píxel que se filtra. Evidentemente existirá una cierta redundancia, pero se proporcionará mayor conocimiento acerca de la vecindad, qunque tal vez de no tanta calidad. En cualquier caso, la implementación no parece complicada.

Mejora de la idea de la máscara gris.

La máscara gris que aquí se ha presentado de forma somera emplea niveles proporcionales a la probabilidad de aparición de los niveles de gris k de los píxeles dentro de una región de nivel k'. Pero esa proporcionalidad en nuestro caso estaba determinada por la diferencia con el umbral de selección previamente determinado. Es posible definir otro tipo de proporcionalidad, tal vez incluso directamente tomado del modelo, sin necesidad de umbral. De este modo podrian obtenerse máscaras grises mejoradas. Su uso podría estar indicado para proce-

sar imágenes cuyo modelo es muy disperso (poco cohesionado); o para tipos de ruido distintos al impulsivo.

Ampliación del campo de aplicación de los filtros.

En la presente obra se han realizado diseños de filtros con una orientación restringida: imágenes de niveles de gris, contaminadas por ruido impulsivo. Una extensión natural de este trabajo sería la construcción de algoritmos, basados en la misma filosofía, que pudiesen obtener buenos resultados con ruidos de otra clase (gaussiano, laplaciano, mezcla de impulsivo y gaussiano...). En estos casos no puede emplearse la máscara binaria, puesto que el ruido afecta a la totalidad de la imagen. Tal vez el uso de máscaras grises proporcione la información precisa para abordar el problema.

Por otro lado, sería interesante intentar realizar el procesado de imágenes en color. Obsérvese que sería una superposición de tres imágenes monocromáticas —rojo, verde, azul—, o bien una imagen cuyas regiones (definidas por la ventana deslizante) ofrecen un vector de tres intensidades o niveles de color. Podría trabajarse con tres histogramas, aunque esto presentaría problemas a la hora de definir la cantided de información contenida, y por tanto los modelos realizadso carecerían de validez. Es necesario, por lo tanto, cambiar las bases de la teoría de modo que se defina una distribución de probabilidad adecuada para realizar máscaras de composición realistas.

BIBLIOGRAFÍA

[ARC82] G.R. Arce, N.C. Gallagher, *Jr.* "State description for the root-signal set of median filters". *IEEE Transactions on Accoustics, Speech and Signal Processing*, vol. ASSP-30, pp. 849-902. Dic. 1982.

[ARC89] G.R. Arce, R.E. Foster. "Detail-preserving ranked-order based filters for image processing".*IEEE Transactions on Accoustics, Speech and Signal Processing*, vol. ASSP-37, pp. 83-98. Ene. 1989.

[BAR95] V. Barranco López, P.L. Luque Escamilla, J. Martínez Aroza, R. Román Roldán. "Entropic texture-edge detection for image segmentation". *Electronic Letters,* vol 31, pp 867-869. May. 1995.

[BEN??] A. Ben Hamza, P.L. Luque Escamilla, J. Martínez Aroza, R. Román Roldán. "Removing noise and preserving details with relaxed median filters". En preparación.

[BER87] R. Bernstein. "Adaptive nonlinear filters for simultaneous removal of different kinds of noise in images". *IEEE Transactions on Circuits and Systems*, vol CAS-34, pp. 1275-1291. Nov. 1987.

[BOV83] A. Bovik, T.S. Huang, D.C. Munson. "A generalization of median filtering using linear combinations of order statistics". *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol ASSP-31, pp. 1342-1350. Dic. 1983.

[CHE92] R. Chellappa. *Digital image processing*. IEEE Computer Society Press, 1992.

[COY89] E.J. Coyle, J.H. Lin, M. Galdbouj. "Optimal stack filtering and the estimation and structural approaches to image processing", *IEEE Transactions on Accoustics, Speech and Signal Processing*, vol. ASSP-37, pp. 2037-2066. Dic. 1989.

[DAV88] E.R. Davies. "On the noise suppression and iamge enhancement characteristics of the median, truncated median and mode filters". *Pattern Recognition Letters*, vol 7, pp. 87-97. Feb. 1988.

[DAV89] E.R. Davies. "Edge location shifts produced by median filters: theoretical bounds and experimental results". *Signal Processing*, vol 16, pp. 83-96. Feb. 1989.

H.A. Davis. *Order Statistics*. John Wiley, 1980. [DAV80]

J.P. Fitch, E.J. Coyle, N.C. Gallagher, *Jr.* "Root properties and convergence rates of median filters". *IEEE Transactions on Accoustics, Speech and Signal Processing*, vol. ASSP-33, pp. 230-240. Feb. 1985. [FIT85]

N.C. Gallagher, *Jr.*, G.L. Wise. "A theoretical analysis of the properties of median filters". *IEEE Transactions on Accoustics, Speech and Signal Processing*, vol. ASSP-29, pp. 1136-1141. Jul. 1981. [GAL81]

L. García Cabrera, M.J. García Salinas, P.L. Luque Escamilla, J. Martínez Aroza, R. Román Roldán. "Median-type filters with model -based preselection masks". *Image and Vision Computing*. Aceptado y en prensa. [GAR96]

L. García Cabrera, P.L. Luque Escamilla, J. Martínez Aroza, A.M. Robles Pérez, R. Román Roldán. "Some pixel-preselection methods for median-type filtering". Enviado a *IEE Proceedings Vision, Image and Signal Processing*. [GAR??]

M.J. García Salinas. *Modelos empíricos multirresolución de imágenes digitales. Aplicación a filtrado.* Tesis de Licenciatura. Departamento de Física Aplicada. Universidad de Granada, 1994. [GAR94]

R.C. González "Image Enhancement and restoration". En *Handbook of pattern recognition imagen processing*, Academic Press, San Diego (CA), 1986. [GON86]

R. Gordon, R.M. Rangayyan. "Feature enhancement of film mammograms using fixed and adaptive neighborhoods". *Applied Optics*, vol. 23, pp 560-564. Feb. 1984. [GOR84]

M.S. Gupta, ed. *Selected papers on noise in circuits and systems*. IEEE Press, 1988. [GUP88]

P.S. Huber. *Robust Statistic*. John Wiley, 1981. [HUB81]

D.P. Huttenlocher, G.A. Klanderman, W.J. Rucklidge. "Comparing images using the Haussdorf distance". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 15, pp 850-863. Sep. 1993. [HUT93]

[HWA95] H. Hwang, R.A. Haddad. "Adaptive median filters: new algorithms and results". *IEEE Transactions on Image Processing*, vol 4, pp 499-502. Abr. 1995.

[JAI89] A.K. Jain. *Fundamentals of digital imagen processing*. Prentice-Hall Int. Ed. New Jersey, 1989.

[JUS81] B.I. Justunsson. "Median filtering: statistical properties". En *Two dimensional digital signal processing*, vol II. T.S. Huang ed. Springer Verlag, 1981.

[KO91] S.J. Ko, Y.H. Lee. "Center weighted median filters and their applications to imagen enhancement". *IEEE Transactions on Circuits ans Systems*, vol. 38, pp 984-993. Sep. 1991.

[LEE85] Y.H. Lee, S.A. Kassam. "Generalizad median filtering and related nonlinear filtering techniques". *IEEE Transactions on Accoustic, Speech, and Signal Processing*, vol 33, pp 672-683. Jun. 1985.

[LEE90] Y.H. Lee, S. Tantaratana. "Decision-based order statistic filters". *IEEE Transactions on Accoustic, Speech, and Signal Processing*, vol 38, pp 406-420. Mar. 1990.

[LIN91] J. Lin. "Divergence measures based on the Shannon entropy". *IEE Transactions on Information Theory*, vol 37, pp. 145-150. Ene. 1991.

[LON92] H. Longbotham, D. Eberly. "Statistical properties, fixed points and decompositions with WMMR filters". *Journal of Mathematical Imaging and Vision*, vol 2, pp. 99-116. Feb. 1992.

[MAR90] J. Martínez Aroza. *Contribución al análisis y estimación en imágenes digitales mediante la Teoría de la Información*. Tesis Doctoral. Departamento de Matemática Aplicada. Univerisdad de Granada, 1990.

[MAR95] J. Martínez Aroza, R. Román Roldán. "Probabilistic linear models for grey-level image estimation". *Multidimensional System and Signal Processing*, vol MSSP-6, pp. 7-35. Ene. 1995.

[NET89] A. Netravali, B.G. Haskell. *Digital pictures. Representation and compression*. Plenum Press, 1989.

[NIE87] P. Nieminen, Y. Heinonen, Y. Neuvo. "A new class of detail-

preserving filters for image processing". *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol PAMI-9, pp. 74-90. Ene. 1987.

T.A. Nodes, N.C. Gallagher, *Jr.* "Median filters: some modifications and their properties". *IEEE Transactions on Accoustic, Speech and Signal Processing,* vol. ASSP-30, pp. 739-746. Oct. 1982. [NOD82]

I. Pitas, A.N. Venetsanopoulos, *Nonlinear digital filters. Principles and applications*. Kluwer Academic Publishers, 1990. [PIT90]

I. Pitas, A.N. Venetsanopoulos. "Nonlinear mean filters in image processing". *IEEE Transactions on Accoustic, Speech and Signal Processing,* vol. ASSP-34, pp. 573-584. Jun. 1986. [PIT86]

W.K. Pratt. *Digital Image Processing*. Wiley, 1978. [PRA78]

G. Ramponi. "Detail-preserving filter for noisy images"*Electronic Letters,* vol 21, pp. 865-866. May. 1995. [RAM95a]

G.Ramponi."A rational edge-preserving smoother". En *Proceedings of International Conference on Image Processing*. IEEE Company Society Press, 1995. [RAM95b]

A. Restrepo Palacios, L. Chacón. "A smoothing property of the median filter". *IEEE Transactions on Signal Processing,* vol 42, pp. 1553-1555. Jun. 1994. [RES94]

X. Roca y otros. *Curso de tratamiento y análisis de imágenes*. Facultad de Ciencias. Departamento de Informática. Universidad Autónoma de Barcelona. 1991. [ROC91]

R. Román Roldán, J. Martínez Aroza, M.J. García Salinas, P.L. Luque Escamilla, J.C. Segura Luna. "Multirresolution empirical model for grey-level images". Enviado a*Image Communications*. [ROM??]

R. Román Roldán, J.J. Quesada Molina, J. Martínez Aroza. "Multiresolution-information analysis for images" *Signal Processing (Elsevier)*, vol 24, pp. 77-91. Jul. 1991. [ROM91]

R.J. Schalkoff. *Digital image processing and computer vision*. John Wiley & Sons, 1989. [SCH89]

[SON93] M. Sonka, V. Hlavac, R. Boyle. *Image processing, analysis and machine vision.* Chapman & Hall, 1993.

[SON88] W.J. Song, W.A. Pearlman. "Edge-preserving noise filtering based on adaptive windowing". *IEEE Transactions on Circuits and Systems,* vol 35, pp 1048-1054. Ago. 1988.

[STA95] V.V. Staroitov. "Towards a measure of diversity between grey-scale images". En *CAIP'95 Proceedings.* Hlavàc, Sàra, Eds. Springer-Verlag, 1995

[SUC95] R. Sucher. "A recursive nonlinear filter for removal of impulse noise". En *Proceedings of International Conference on Image Processing*. IEEE Company Society Press, 1995.

[SUN88] X.Z. Sun, A.N. Venetsanopoulos. "Adaptive schemes for noise filtering and edge detection by use of local statistics". *IEEE Transactions on Circuits and Systems,* vol 35, pp. 57-70. Ene. 1988.

[TUK77] J.W. Tukey. *Exploratory data analysis.* Addison-Wesley, Memlo Park (CA), 1977.

[WEN86] P.D. Wendt, E.J. Coyle, N.C, Gallagher, *Jr.* "Stack filters". *IEEE Transactions on Accoustics, Speech and Signal Processing*, vol. ASSP-34, pp. 898-911. Ago. 1986.

[ZEN94a] B. Zeng. "Optimal median-type filtering under structural constraints". *IEEE Transactions on Signal Processing.* vol 42, pp 921-931. Jul. 1994.

[ZEN94b] B. Zeng. "Convergence properties of median and weighted median filters". *IEEE Transactions on Signal Processing.* vol 42, pp 3515-3519. Dic. 1994.