

APLICACIÓN DE TÉCNICAS DE INTERPOLACIÓN BASADAS EN LÓGICA DIFUSA AL PROCESADO DE IMÁGENES DE VIDEO

P. Brox, I. Baturone, S. Sánchez-Solano

Instituto de Microelectrónica de Sevilla (CNM-CSIC)
Avda. Reina Mercedes S/N. Edificio CICA. 41012 Sevilla (España)
e-mail: brox@imse.cnm.es

ABSTRACT

Muchas tareas básicas de procesamiento de imágenes requieren la manipulación de grandes volúmenes de información que, en ocasiones, puede resultar ambigua y/o imprecisa como consecuencia de las características propias de las imágenes (gran cantidad de detalles con grandes contrastes de valores de luminancia y secuencias con un elevado grado de movimiento) o de los defectos de las mismas (presencia de ruido, falta de nitidez, etc.). En esta comunicación se analizan nuevas técnicas de interpolación basadas en lógica difusa que proporcionan soluciones eficaces para dos aplicaciones típicas de procesamiento de imágenes: el desentrelazado de señales de vídeo y el incremento de resolución de imágenes.

1. INTRODUCCIÓN

La eficacia de la lógica difusa para manejar la ambigüedad e imprecisión que aparece en numerosos problemas ha motivado en los últimos años una creciente aplicación de dichas técnicas de inferencia al procesamiento de imágenes [1]. Muchos de los trabajos realizados se han centrado en dos aplicaciones que son objeto de una gran demanda en la actualidad: el desentrelazado de señales de vídeo [2] y el incremento de resolución de imágenes [3]. La conversión de señales de vídeo entrelazadas a progresivas es requerida por numerosos dispositivos como proyectores, DVDs, televisores de alta definición o monitores LCD, para adecuar su formato de presentación al formato de transmisión entrelazado empleado por los sistemas convencionales de televisión. Ello ha fomentado el desarrollo de numerosos algoritmos de desentrelazado que implementan desde simples esquemas de interpolación espacio-temporal hasta complejas técnicas de compensación de movimiento. Por otro lado, la ampliación de tamaño de imágenes es especialmente necesaria en aplicaciones de análisis e interpretación donde un aumento de la resolución en ciertas áreas de la imagen puede resultar crucial. Para ambas aplicaciones se han propuesto técnicas basadas en lógica difusa que mejoran notablemente la calidad de las imágenes interpoladas pero, habitualmente, a costa de un elevado coste computacional [4]-[5].

En esta comunicación se proponen nuevas técnicas de interpolación basadas en lógica difusa que proporcionan soluciones eficaces para las dos aplicaciones anteriores sin implicar un incremento excesivo del coste computacional.

2. DESENTRELAZADO DE SEÑALES DE VÍDEO

El formato entrelazado empleado por los actuales sistemas de transmisión de señales de televisión permite reducir a la mitad el ancho de banda de la señal transmitida sin afectar significativamente a la calidad de las imágenes. Por este motivo, dicho formato sigue empleándose a pesar de que muchos de los dispositivos actuales realizan un barrido de presentación progresivo que requiere disponer del número total de líneas de la imagen. Los algoritmos de desentrelazado aplican diferentes técnicas de interpolación para generar la imagen completa a partir de la información transmitida en los sucesivos campos. De entre ellos, los más eficientes son aquellos que adaptan la estrategia de interpolación a las características específicas de la imagen como el grado de movimiento o la presencia de bordes.

La detección del grado de movimiento permite dirigir y optimizar el proceso de desentrelazado. Si no existe movimiento, las líneas no transmitidas en un campo pueden obtenerse a partir de la información del campo anterior mediante lo que se denomina una técnica de interpolación temporal. Por el contrario, cuando existe movimiento, la información de las líneas del campo anterior no es fiable, siendo preferible aplicar una interpolación espacial entre los píxeles del campo actual. La idea básica de los algoritmos adaptativos al grado de movimiento fue propuesta en [6] y puede expresarse matemáticamente mediante la expresión:

$$I_o(x, y, t) = (1 - \gamma(x, y, t)) \cdot I_T(x, y, t) + \gamma(x, y, t) \cdot I_S(x, y, t) \quad (1)$$

donde $I_o(x, y, t)$ representa la luminancia de un determinado píxel, las variables x e y son las coordenadas espaciales del píxel en el fotograma, t indica el número de orden del fotograma en la secuencia,

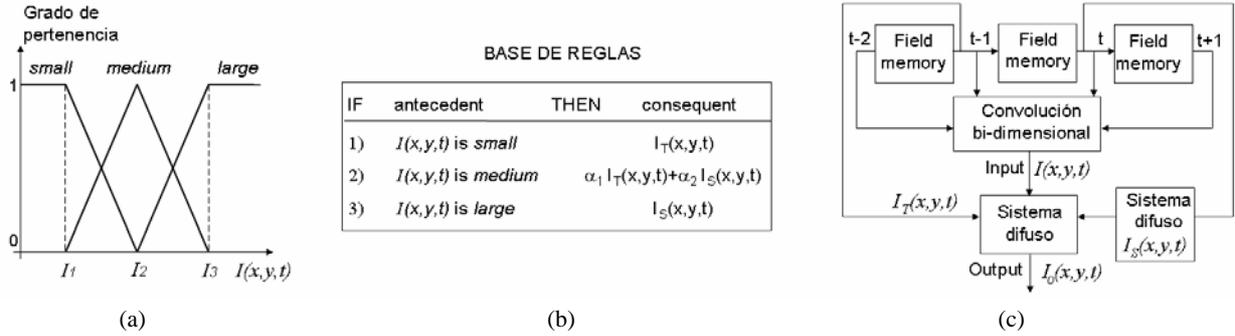


Fig. 1. Funciones de pertenencia utilizadas por el sistema difuso para realizar el desentrelazado de la señal de vídeo adaptativo con el grado de movimiento. (b) Base de reglas de dicho sistema. (c) Diagrama de bloques del sistema.

I_S e I_T representan los valores de luminancia obtenidos al utilizar técnicas de interpolación espacial y temporal, respectivamente, y γ indica el grado de movimiento mediante un valor comprendido entre 0 y 1.

Los trabajos más recientes se centran en determinar el valor de γ para que sea lo suficientemente robusto y, por tanto, no se produzcan detecciones erróneas de movimiento. Nuestra propuesta consiste en utilizar un sistema basado en lógica difusa para estimar el grado de movimiento. La entrada al sistema (I), es la convolución bi-dimensional de la diferencia de valores de luminancia entre dos campos con el mismo orden de paridad [7]. Dicha variable de entrada se fuzzifica utilizando el conjunto de funciones de pertenencia triangulares que se muestra en la Fig. 1(a). Aunque se ha analizado el efecto de emplear un número mayor de funciones de pertenencia, la propuesta que emplea tres ofrece el mejor compromiso entre complejidad y eficacia [8]. La Fig. 1(b) recoge la base de reglas del sistema difuso. Su interpretación es simple: cuando el grado de movimiento es pequeño se realizará una interpolación temporal, cuando es grande una interpolación espacial, en cualquier otro caso se llevará a cabo una combinación de ambas. Los parámetros (I_1 , I_2 , I_3) que definen las funciones de pertenencia de la Fig. 1(a), así como los factores (α_1 , α_2) que determinan la combinación entre I_S e I_T , son ajustados mediante técnicas de aprendizaje supervisado. Los patrones de aprendizaje utilizados han

sido generados a partir de distintos fotogramas de secuencias progresivas de vídeo estándares.

Para la implementación del algoritmo se selecciona como método de interpolación temporal (I_T) la inserción del píxel del campo anterior con las mismas coordenadas espaciales. Para realizar la interpolación espacial (I_S) se propone un algoritmo que utiliza un sistema basado en lógica difusa para detectar de forma robusta la presencia de bordes en la imagen. Las entradas del sistema son las diferencias en valor absoluto de la luminancia en las tres direcciones (a, b, c) que se muestran en la Fig. 2(a). La Fig. 2(b) contiene la base de reglas del sistema. Las dos primeras reglas se activan cuando la correlación es grande en una dirección y al mismo tiempo es pequeña en las direcciones contrarias. En ambos casos el resultado interpolado se obtiene como valor medio de la luminancia en la dirección donde se encuentra el borde. La tercera regla describe una situación en la que no existe un borde porque la correlación es muy grande en dos direcciones al mismo tiempo. En este caso el resultado se obtiene interpolando los cuatro píxeles de las direcciones a y c . En cualquier otro caso el valor más adecuado se obtiene interpolando en la dirección vertical. Las etiquetas de la base de reglas (*small*, *strongly small* and *large*) se describen mediante las funciones de pertenencia de la Fig. 2(c). Los parámetros que las definen (a_1 , a_2 , a_3) también han sido ajustados mediante técnicas de aprendizaje supervisado.

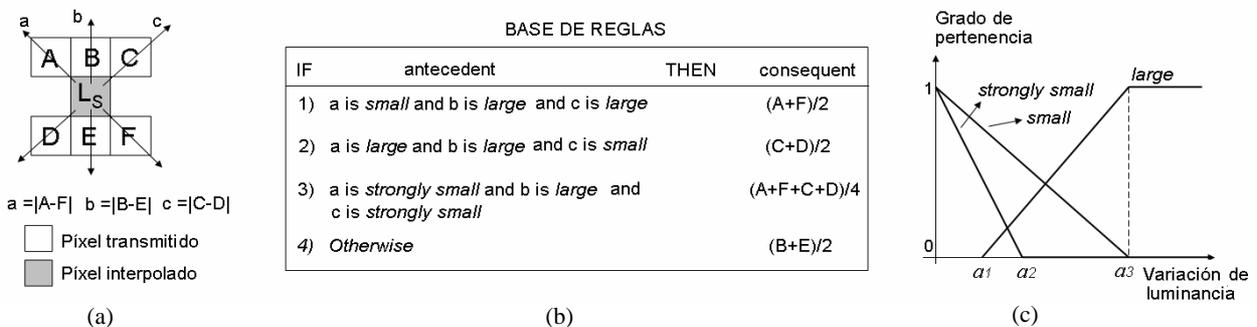


Fig. 2. (a) Píxeles utilizados por el sistema difuso para implementar la interpolación espacial adaptativa con los bordes de la imagen. (b) Base de reglas del sistema. (c) Funciones de pertenencia utilizadas.

Tabla 1. Valor medio de MSE obtenido al procesar distintos fotogramas de varia secuencias de vídeo estándares.

Secuencia	Missa	París	Trevor	Salesman	News	Mother	Carphone
Formato	CIF	CIF	CIF	CIF	QCIF	QCIF	QCIF
Repetición línea	14.75	283.19	51.05	68.87	197.27	42.86	97.29
V. medio líneas	5.83	139.98	20.37	28.84	77.28	16.56	35.65
Pixel anterior	9.48	67.15	23.82	15.706	31.62	15.81	60.12
VT 2 fields [2]	6.13	54.96	14.19	14.42	18.49	7.11	25.41
VT 3 fields [2]	5.76	47.43	12.504	13.12	17.62	2.301	22.86
ELA [2] 3+3	7.31	182.01	25.23	40.01	141.27	18.79	45.92
ELA [2] 5+5	9.05	223.39	30.34	62.52	166.37	24.72	35.32
Técnica [4]	6.48	31.701	18.83	11.24	21.88	7.31	38.55
Técnica [7]	6.23	19.27	13.93	9.64	11.53	4.22	21.63
Propuesta	5.42	16.907	11.35	8.87	8.51	3.92	20.75

3. INCREMENTO DE RESOLUCIÓN

La técnica de interpolación espacial propuesta para el algoritmo de desentrelazado descrito previamente puede también emplearse para incrementar la resolución de imágenes. Con este objetivo se ha diseñado una estrategia para interpolar mediante esta técnica no solo los píxeles de las nuevas líneas de la imagen (como requiere la aplicación de desentrelazado) sino también los píxeles pertenecientes a las nuevas columnas. Para describir el método propuesto consideremos un factor de amplificación igual a 2 como se ilustra en la Fig.3. El proceso consta de dos fases. Durante la primera se interpolan los píxeles de las nuevas columnas y filas, que se muestran con los símbolos 'c' y 'f' respectivamente. Por ejemplo, para interpolar el píxel 'f' sombreado en gris se utilizan los píxeles originales {A, B, C, D, E, F} y {A,B,D,E,G,H} para el píxel 'c' en negro. Los píxeles mostrados con el símbolo 'fc' son interpolados en una segunda fase utilizando como entrada los cuatro píxeles originales y los cuatro píxeles interpolados en la fase previa más cercanos. El resultado final se obtiene como el valor medio de los resultados obtenidos al aplicar el algoritmo en los 3+3 píxeles de las líneas superior e inferior, y los 3+3 píxeles de las columnas más próximas. Esta estrategia puede ser fácilmente extendida a otros factores de amplificación [9].

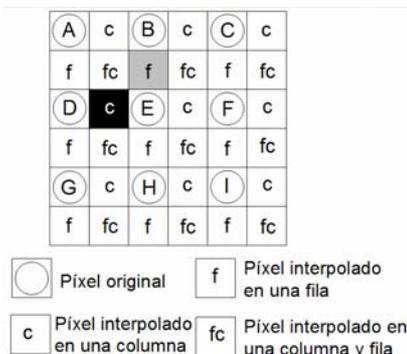


Fig. 3. Píxeles que participan en el proceso de aumento de resolución de la imagen para un factor de amplificación igual a 2.

4. RESULTADOS

Las dos aplicaciones descritas han sido testadas utilizando una amplia batería de imágenes y secuencias de vídeo estándares. Para obtener los ficheros de entrenamiento se ha partido de imágenes en formato progresivo y se han eliminado filas (para la aplicación de desentrelazado) y también columnas (para la aplicación de aumento de resolución). Los píxeles eliminados son interpolados aplicando no solo las técnicas descritas en los apartados anteriores sino también otras técnicas convencionales descritas en la literatura [2]-[3]. Las Tablas 1 y 2 muestran los errores obtenidos al comparar las imágenes interpoladas con las originales. Para la aplicación de desentrelazado se muestran los valores medios obtenidos tras procesar varios fotogramas de las secuencias. Se observa que los errores más pequeños corresponden en todos los casos al método propuesto. Algo que además se corrobora en las Fig.4 y Fig.5, donde aparecen las imágenes interpoladas aplicando distintos métodos para la aplicación de desentrelazado (Fig.4) y la de aumento de resolución (Fig.5).

El estudio realizado se ha extendido también al análisis de distintas opciones de implementación de los algoritmos. Entre estas cabe destacar el empleo de distintos tamaños de la ventana de convolución [10] y de los píxeles usados en la técnica de interpolación espacial [11]. Asimismo, se ha realizado un estudio de la eficacia de las técnicas basada en lógica difusa cuando las imágenes contienen ruido [12].

Tabla 2. Resultados de MSE obtenidos al aumentar la resolución de distintas imágenes estándares. Se compara la técnica propuesta con otras técnicas polinómicas convencionales.

Imagen	Elaine	Nematode	X-ray	Lena	Boat
Resolución inicial	256x256	256x256	256x256	256x256	256x256
Resolución final	512x512	512x512	512x512	512x512	512x512
NNI [3]	88.2	19.5	382.8	76.9	189.4
BLI [3]	43.6	80.2	132.3	21.1	80.7
BCI [3]	46.4	84.6	119.5	18.6	78.2
CSI [3]	49.7	91.1	123.9	18.5	81.6
Propuesta	42.7	79.9	115.6	20.01	79.4

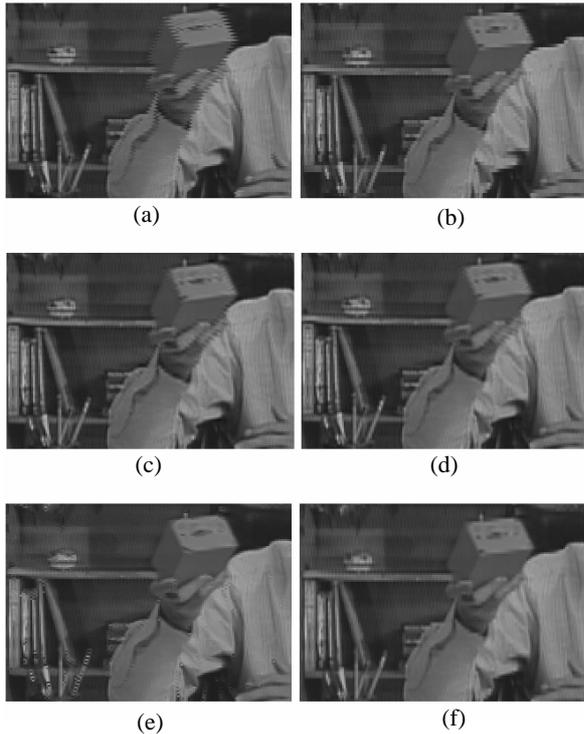


Fig.4. Zoom de un fotograma desentrelazado de la secuencia. 'Salesman' aplicando distintas técnicas: (a) píxel anterior, (b) v.medio líneas, (c) VT 2 fields, (d) VT 3 fields, (e) ELA 3+3, (f) propuesta.

5. CONCLUSIONES

En esta memoria se presentan dos sistemas basados en lógica difusa, que adaptan la interpolación de píxeles a la presencia de movimiento y/o bordes en la imagen. Una combinación de ambos sistemas se presenta para el desentrelazado de la señal de vídeo, mientras que el sistema adaptativo con la presencia de bordes en la imagen es utilizado como técnica para aumentar la resolución de la imagen. Ambos sistemas alcanzan unos resultados que superan los obtenidos por otras técnicas convencionales de interpolación, poniendo de manifiesto la eficacia de las técnicas basadas en lógica difusa para el procesamiento de imágenes.

REFERENCIAS

[1] E. E. Kerre, M. Nachtegael, *Fuzzy Techniques in Image Processing*, Springer, 2000.

[2] G. De Haan, E. B. Bellers, "De-interlacing-An overview", *Proc. of the IEEE*, vol. 86, pp.1839-1857, 1988.

[3] S. Baker, T. Kanade, "Limits on super-resolution and how to break them", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.24, no.9, pp.1167-1183, 2002.

[4] D. Van de Ville, B. Rogge, W. Philips, I. Lemahieu, "De-interlacing using fuzzy-based motion detection", *Int. Conf. on Knowledge-Based Intelligent Information Engineering Systems*, pp.263-267, 1999.

[5] T. Aso, N. Suetake, T. Yamakawa, "A code-reduction technique for an image enlargement by using a som-based fuzzy interpolation", *Int. Conf. on Neural Information Processing*, vol.3, pp.1281-1286, 2002.

[6] A. M. Bock, "Motion adaptive standards conversion between formats of similar field rates" *Signal Processing: Image Communication*, vol.6, no.3, pp.275-280, 1994.

[7] J. Gutiérrez-Ríos, F. Fernández-Hernández, J. C. Crespo, G. Triviño, "Motion adaptive fuzzy video de-interlacing method based on convolution techniques", *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, 2004.

[8] P. Brox, I. Baturone, S. Sánchez-Solano, J. Gutiérrez-Ríos, F. Fernández-Hernández, "Fuzzy motion adaptive algorithm for video de-interlacing", *Lectures Notes in Computer Science*, vol.4252/2006, no.363-370, 2006.

[9] P. Brox, I. Baturone, S. Sánchez-Solano, "Image enlargement using the fuzzy-ELA algorithm", *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, 2006.

[10] P. Brox, I. Baturone, S. Sánchez-Solano, "A fuzzy motion adaptive algorithm for interlaced-to-progressive conversion", *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, 2006.

[11] P. Brox, I. Baturone, S. Sánchez-Solano, "Progressive scan conversion based on edge-dependent interpolation using fuzzy logic", *Int. Conference of the European Society for Fuzzy Logic and Technology*, 2006.

[12] P. Brox, I. Baturone, S. Sánchez-Solano, "A fuzzy edge-dependent interpolation algorithm", *Soft Computing in Image Processing - Recent Advance*, Springer (to appear in Nov. 2006).

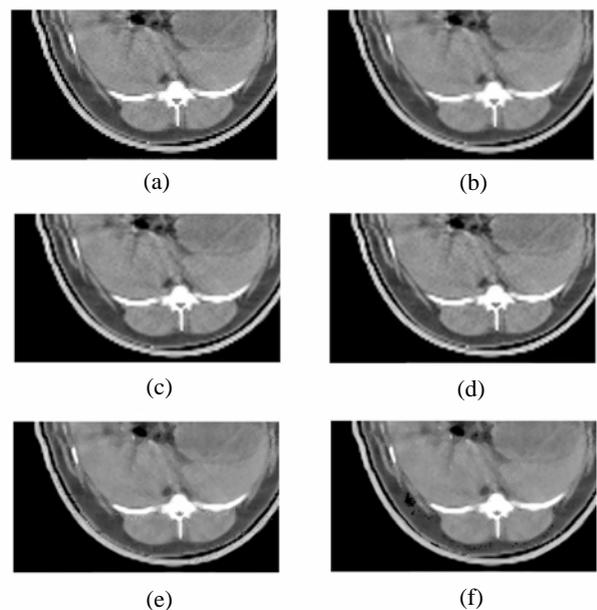


Fig.5. Zoom de un fotograma ampliado de la imagen 'Xray' aplicando distintas técnicas: (a) NNI, (b) BLI, (c) BCI, (d) CSI, (e) ELA 3+3, (f) propuesta.