

SEGMENTACION DE IMAGENES MEDIANTE MODELOS DE GIBBS-MARKOV

Ferran MARQUÉS ¹ Antoni GASULL ¹ Christian ROUX ²

¹ Dept. Teoría de la Señal y Comunicaciones
E.T.S.E.T.B. - U.P.C.

² Dept. mathématiques et systèmes de communication
E.N.S.Telecom. Bretagne

ABSTRACT

A coding oriented image segmentation algorithm is presented. This new method is based on Gibbs-Markov Random Fields (GMRF) for contour information, a stationary gaussian model for region internal information, and a preprocessing task for human visual system model. This is an extension of a previous work [4], which has been improved because the number of regions to be found is not fixed a priori. This fact gives better results, and is quite independent of the initial segmentation approach.

1.-INTRODUCCION

La segmentación de imágenes es una técnica que se aplica actualmente en un gran número de disciplinas que hacen uso de imágenes. Su utilización puede ser tanto como paso previo para obtener un resultado, como un fin en sí misma. De este modo, se pueden hallar técnicas de segmentación aplicadas a materias tan dispares como la Medicina (análisis de radiografías, Citología, etc...), la Robótica (análisis de imágenes, estereovisión, etc...), la Agricultura (estudio de las cosechas mediante imágenes aéreas) o la compactación de información (codificación de imágenes, compensación de movimiento, etc...). Para cada una de estas aplicaciones, el método de segmentación a utilizar deberá cumplir una serie de requisitos o condicionantes que impone la propia aplicación.

En este artículo se presenta un método de segmentación de imágenes fijas en blanco y negro cuya finalidad es servir como paso previo a la codificación de imágenes para su posterior análisis visual. Esta finalidad impone dos condiciones al método. En primer lugar, la segmentación deseada debe aproximarse lo máximo posible a la que realiza el ojo humano para, de este modo, codificar posteriormente sólo la información útil para él. Por tanto la técnica que se propone se engloba en los denominados *Métodos de Segunda Generación* [1].

Este trabajo ha sido subvencionado parcialmente por la Acción Integrada hispano-francesa n.172 A

Además, se debe hallar una segmentación fácilmente codificable, para así ahorrar en lo posible a la hora de transmitir o almacenar las imágenes. Esta segunda condición lleva a buscar en la imagen original regiones homogéneas y con contornos suaves.

Ambos requisitos han sido satisfechos principalmente mediante el modelo de imagen que se ha tomado. Dicho modelo incluye tanto información referente al interior de la región (nivel de gris de los puntos), como a los contornos (forma de la región). De este modo se puede controlar ambas características y, para discernir si un conjunto de puntos dado forma una región, se utiliza conjuntamente ambas informaciones, al igual que hace el ojo humano.

2.- MODELO DE IMAGEN

El modelo de imagen que se utiliza a lo largo de este artículo presupone que la imagen está compuesta de un conjunto de regiones inconexas $R = \{R_n\}$. Estas regiones forman una partición (distribución espacial de las regiones representada por un conjunto de etiquetas $q = \{q_{ij}\}$) que viene regida por un campo aleatorio de Gibbs-Markov (GMRF) $Q = \{Q_{ij}\}$ [2]. Por su parte, el conjunto de niveles de gris dentro de cada región R_n ($y_n = \{y_{ij} : (i, j) \in R_n\}$) se supone una realización de un proceso gaussiano estacionario. Este tipo de modelo estocástico jerárquico de imagen, que se ilustra en la figura 1, permite introducir el concepto de distribución espacial de los puntos en el interior de una región, aún utilizando una descripción de su textura tan simple como un proceso gaussiano.

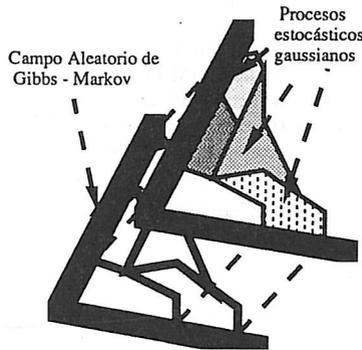


Fig 1.- Esquema de un modelo estocástico jerárquico (o compuesto).

Cabe destacar que un modelo de imagen únicamente basado en procesos gaussianos sería independiente de la forma y distribución de las regiones en la imagen (p. e.: dado el número de regiones que componen la imagen y los parámetros característicos de cada uno de los procesos gaussianos, el modelo no haría distinción entre una imagen con todas las regiones de forma circular u otra con todas de forma cuadrangular).

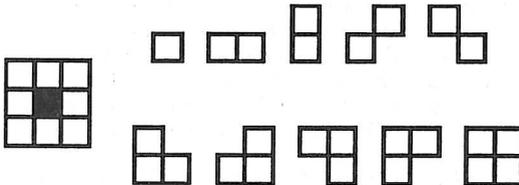


Fig. 2.- Un vecindario de segundo orden y sus 10 cliques asociados.

El GMRF utilizado ha sido definido sobre un vecindario isotrópico de segundo orden, lo cual da lugar a un conjunto de 10 *cliques* diferentes [2], como se muestra en la figura 2. De este conjunto de *cliques*, sólo se utilizan los compuestos por dos puntos con distinta etiqueta (pertenezcan a regiones distintas), ya que el resto no aporta información para la descripción de los contornos de las regiones.

Además, los *cliques* que relacionan pares de puntos en diagonal se consideran análogos (se les asocia el mismo potencial V_2), así como los horizontales y verticales (potencial V_1). De este modo, la expresión general de una distribución de Gibbs:

$$P(Q) = \frac{1}{Z} \exp \left(-\frac{1}{T} U(Q) \right) \quad \text{with } U(Q) = \sum_{c \in Q} V_c, \quad (1)$$

se reduce a

$$P(Q) = \frac{1}{Z} \exp \left(-\frac{1}{T} (n_1 V_1 + n_2 V_2) \right), \quad (2)$$

donde N_1 y N_2 son respectivamente, el número de *cliques* de tipo 1 y 2 que hay en la imagen, mientras que Z y T son constantes.

3.- ALGORITMO DE SEGMENTACION

El modelo de imagen descrito en el apartado 2 ha sido utilizado anteriormente como base de un método de relajación de contornos [3], así como en un algoritmo de segmentación, para el cual se fijaba a priori el número máximo de regiones que debía tener la segmentación final [4]. El método que se describe en este apartado es una extensión del anterior, en el cual el algoritmo no está condicionado en este sentido, sino que él mismo halla el número de regiones final que ha de tener la segmentación.

La segmentación se obtiene mediante la búsqueda de la realización de máxima verosimilitud del proceso Q dada una imagen inicial Y . Por tanto se trata de un criterio de Máximo a Posteriori (MAP), en el que se pretende maximizar $P(Q/Y)$. Mediante la regla de Bayes se puede comprobar que, en este caso, maximizar la probabilidad conjunta $P(Q, Y)$ lleva a la misma solución que el problema anterior, dado que $P(Q, Y) = P(Q) P(Y/Q)$. Para poder calcular el término $P(Y/Q)$ se estiman los parámetros que definen cada uno de los campos gaussianos, y de este modo la expresión a maximizar es [4]:

$$P(Y, Q) \propto \exp \left(-\frac{1}{T} (n_1 V_1 + n_2 V_2) \right) \prod_n P(y_n / \hat{m}_n, \hat{\sigma}_n^2) \quad (3)$$

Esta técnica de segmentación podría ser utilizada tomando como punto de partida cualquier segmentación inicial, y en concreto suponiendo que cada punto en la imagen original es una región aislada. Para acelerar el proceso, se toma una estimación inicial de la imagen segmentada que, aunque burda, aproxime el punto inicial de búsqueda al máximo deseado. A partir de esta estimación, el algoritmo debe buscar aquellos puntos en la partición que al variar su etiqueta lleven a una realización del proceso conjunto de mayor verosimilitud (maximicen $P(Q, Y)$), y cambiar su etiqueta, hasta que no haya ningún cambio posible de etiqueta que aumente la probabilidad conjunta (se esté en un máximo).

Para realizar esta búsqueda se ha optado por un algoritmo subóptimo que no alcanza el máximo por el camino del máximo gradiente, y por tanto se deben realizar más pasos de los estrictamente necesarios. Este algoritmo efectúa un barrido normal de la imagen (desde la esquina izquierda superior hasta la derecha inferior) realizando los cambios de etiqueta que

lleven a una realización más probable del proceso conjunto. Tras cada cambio de etiqueta, el conjunto de parámetros del modelo es actualizado. Al finalizar el barrido, éste se repite en sentido contrario y el proceso se itera hasta que no haya más cambios de etiqueta posibles que lleven a una realización de mayor verosimilitud.

Contrariamente a lo propuesto en [4], aquí se permite al algoritmo crear nuevas regiones mediante la separación de una región original, formando otras dos disjuntas. Esto se hace así para dotar al algoritmo de más flexibilidad y de este modo poder enmendar errores que se hayan producido al hallar la estimación previa, así como para obtener resultados finales con probabilidades mayores. De este modo se consiguen resultados de mayor calidad, sin que el número final de regiones aumente apreciablemente (el aumento medio del número de regiones es del orden del 5%). Esto es así ya que el algoritmo también puede anular regiones y, además, tras obtener la segmentación, se eliminan todas las regiones que contengan un número de puntos menor que un umbral dado, ya que se debe recordar que la finalidad de esta segmentación es ser utilizada para la codificación, y dichas regiones casi no aportan información.

4.- RESULTADOS

En esta sección se presentan dos resultados distintos que pretenden ilustrar las características del método de segmentación anteriormente propuesto. En las figuras 3 a) y b) se muestra respectivamente el cuadro #6 original de la secuencia *Miss America*, y la segmentación mediante el algoritmo anterior, con cada una de sus regiones rellena con la media del nivel de gris de sus puntos. Cabe destacar la alta calidad de la reconstrucción de la imagen, aún mediante 128 regiones. Para mostrar la ductilidad que ofrece el permitir al algoritmo crear nuevas regiones, se ha tomado como segmentación inicial de la imagen *Cameraman.*, la imagen mostrada en la figura 4 a), que como puede verse corresponde a una escena completamente diferente. El resultado obtenido es el mostrado en la figura 4 b), donde, aunque la calidad de la reproducción no es tan buena como en el caso anterior, las componentes más relevantes de la imagen han sido recuperadas.

5.- CONCLUSIONES

El método de segmentación que se ha presentado permite obtener resultados de gran calidad aún con un número bajo de regiones. Además se ha comprobado la robustez del método frente a posibles errores en la estimación inicial. Pese a estas características positivas, el algoritmo sigue adoleciendo de una visión de las características globales de la imagen. Esto es debido a que el método utiliza básicamente información local de la imagen. Para subsanar este problema, actualmente se está trabajando en un algoritmo de segmentación multinivel, con lo cual se puede tener una visión más global de la imagen.

REFERENCIAS

- [1] *M. Kunt, A. Ikonomopoulos, M. Kocher*, "Second-Generation image coding techniques", Proc. IEEE vol. 73 num. 4 April 1985.
- [2] *Geman, Geman*, "Stochastic Relaxation, Gibbs Distribution and the Bayesian restoration", IEEE Trans. PAMI vol 5 1984.
- [3] *Mester, Franke*, "Statistical model based image segmentation using region growing, contour relaxation and classification" SPIE Conf. on Visual Comm. and Image Proc. '88.
- [4] *F. Marqués, A. Gasull, T. Reed, M. Kunt*, "Coding-oriented segmentation based on Gibbs-Markov Random Fields and human visual system knowledge", Proc. ICASSP-91, pp. 2749-2752, Toronto, mayo 1991.



Fig. 3.- Imágenes original a) y segmentada b) (nivel de gris medio de cada región).



Fig. 4.- a) Segmentación inicial incorrecta de la imagen Cameraman. b) Resultado final