

Capítulo 2

Segmentación de Imágenes de Ultrasonido para la Detección Asistida de Tumores de Mama

Lilia García Rosales¹
Miguel Angel López Díaz²

Abstract: The segmentation of images is one of the most important problems in areas such as pattern recognition and artificial intelligence. In the field of medicine, and particularly in oncology, the obtaining of a diagnosis by non-invasive media is being done more frequently through the segmentation of images. The analysis of the image serves, from the clinical point of view, to give a quantitative estimate of the size of the lesion to verify the effectiveness of the treatment. When images are obtained in grayscale, many segmentation techniques use the intensities histogram to find one or more thresholds that identify the components of the image. In this work the subject of grayscale image segmentation is addressed, considering that the histogram of intensities is a sample of a mixture of normal distributions. It is known that one of the disadvantages when using a mixture of models to analyze an image is that the number of components must be proposed beforehand; to solve this problem we use a Bayesian method of reversible transdimensional jumps alternative to the RJMCMC, with which the number of the components of the mixture is inferred.

Resumen: La segmentación de imágenes es uno de los problemas más importantes dentro de áreas como el reconocimiento de patrones y la inteligencia artificial. En el campo de la medicina, y de manera particular en la oncología, la obtención de un diagnóstico a través de medios no invasivos se realiza cada vez con mayor frecuencia mediante la segmentación de imágenes. El análisis de la imagen sirve, desde el punto de vista clínico, para dar una estimación cuantitativa acerca del tamaño de la lesión y monitorear su crecimiento para comprobar la eficacia del tratamiento. Cuando las imágenes se obtienen en escala de grises, muchas técnicas de segmentación emplean el histograma de las intensidades para hallar uno o varios umbrales

¹jarolil@hotmail.com. Facultad de Ciencias UAEMéx

²mald@uaemex.mx. Facultad de Ciencias UAEMéx

que identifiquen los componentes de la imagen. En este trabajo abordamos el tema de la segmentación de imágenes en escala de grises, considerando que el histograma de intensidades es una muestra de una mezcla de distribuciones normales. Se sabe que una de las desventajas cuando se usa una mezcla de modelos para analizar una imagen es que el número de componentes debe ser propuesto de antemano, para resolver este problema se emplea un método bayesiano de saltos transdimensionales reversibles alternativo a los saltos reversibles Monte Carlo vía cadenas de Markov (RJMCMC) con el que se infiere el número de componentes en la mezcla.

Palabras clave: (Ecografía, atenuación, campo de distorsión)

2.1 Introducción

En la medicina actual, hacer diagnósticos utilizando imágenes es invaluable, el procesamiento de imágenes provee un medio no invasivo y efectivo de delinear o delimitar la anatomía de un sujeto. Con el incremento en tamaño y número de imágenes médicas, se ha hecho necesario el uso de cómputo intensivo para facilitar el procesamiento y análisis de las mismas. En particular, los algoritmos para el delineamiento de estructuras anatómicas y otras regiones son un componente clave para asistir y automatizar ciertas tareas radiológicas (ver [1]).

Las imágenes de ultrasonido o ecografía es una técnica de adquisición de imágenes, no invasivas y accesible que permite la evaluación del sistema musculoesquelético en tiempo real, con la ventaja de no someter al paciente a radiación ionizante. Su principio se basa en la emisión y recepción de ondas sonoras, cuya frecuencia está por encima de la capacidad del oído humano para percibirlos (ver [2]).

Además, en los últimos años se han utilizado con gran éxito para operaciones e intervenciones guiadas por imágenes y para terapias especiales. Por estas razones es de suma importancia desarrollar métodos automáticos de segmentación e interpretación de este tipo de imágenes que además sean eficientes, muy precisos y funcionen en tiempo real (ver [3]).

Sin embargo, existen diversos parámetros que afectan la calidad de la imagen de ultrasonido, y que es importante tomar en cuenta. Dos de ellos son: La interacción con los tejidos: cuando una onda de ultrasonido atraviesa un tejido sucede una serie de hechos; entre ellos, la reflexión o rebote de los haces ultrasónicos hacia el transductor que es llamado eco. Una reflexión ocurre en el límite o interface entre dos materiales y provee la evidencia de que un material es diferente a otro. Esta propiedad es conocida como la impedancia acústica y es el producto de la densidad y velocidad de propagación. Cuando se emplea la escala de grises, las reflexiones más intensas se observan en tono blanco y las más débiles en diversos tonos de gris y cuando no hay reflexiones son en un tono negro. Otro factor a considerar es la atenuación; mientras las ondas se propagan a través de las diferentes interfases reticulares, la energía ultrasónica pierde potencia y su intensidad disminuye progresivamente a medida que inciden estructuras más profundas, circunstancia conocida como atenuación (ver [4]).

2.2 Adquisición de la imagen

Las imágenes de ultrasonido están formadas por una matriz de elementos fotográficos, las imágenes en escala de grises están generadas por la visualización de los ecos, regresando al transductor como elementos fotográficos (píxeles). Su brillo dependerá de la intensidad del eco que es captado por el transductor en su viaje de regreso (ver [5]).

El modo B es la representación gráfica de la suma de los ecos en diferentes direcciones favoreciendo que el equipo reconozca la posición espacial y la dirección del haz. Esta es la modalidad empleada en todos los equipos de ecografía en tiempo real y se trata de una imagen bidimensional estática, en el modo B dinámico se obtienen varias imágenes por segundo y es el modo ultrasonográfico más utilizado en medicina (ver [6]).

2.2.1 Segmentación

El ultrasonido modo B es la herramienta de diagnóstico médico más empleada debido a que la imagen se obtiene en tiempo real, existe poco riesgo para la salud del paciente así como el bajo costo comparado con otro tipo de adquisición de imágenes como las resonancias.

La segmentación es a menudo una etapa importante en el análisis de imágenes de ultrasonido modo B. Por ejemplo, la segmentación puede ser un paso preliminar para medir de manera cuantitativa el tamaño de la lesión, desde el punto de vista clínico esto es muy valioso pues permite monitorear el crecimiento de la lesión y a la vez ser un indicador que permita planear un tratamiento o una cirugía. En el modo B de imágenes de ultrasonido, se han estudiado el efecto de varios artefactos, por ejemplo el ruido de moteado, la atenuación y la interacción con los tejidos que han sido estudiados ampliamente.

Existen diversos enfoques para la segmentación de imágenes ecográficas, basados en métodos clásicos de detección de bordes y aplicados a un problema clínico concreto, estudios relacionados con ginecología y obstetricia (ver [7] y [8]).

Debido al moteado (speckle, en inglés) característico que este tipo de imágenes posee, es muy útil utilizar distribuciones estadísticas para modelar los datos.

En este trabajo nos enfocaremos en resolver el problema de corregir la atenuación relacionada con las inhomogeneidades en la intensidad, es decir aquella que es causada por pequeños cambios en la intensidad y contraste que no se deben al moteado.

2.3 Modelo estadístico

El modelo estadístico para la corrección de las inhomogeneidades en la intensidad y la segmentación se propone como sigue:

Sea S un subconjunto de N números naturales que nos permita identificar los píxeles que componen una imagen. Sean además $I = \{I_1, \dots, I_N\}$ e $I^* = \{I_1^*, I_2^*, \dots, I_N^*\}$ las intensidades observada e ideal, es decir sin la distorsión debida a la falta de homogeneidad en la intensidad.

Supongamos que la intensidad ideal para el i -ésimo pixel, puede ser expresada a través de una función de valor real f que depende tanto de la intensidad observada I_i y

de la intensidad adicional d_i debida a la inhomogeneidad, de modo que,

$$I_i^* = f(I_i, d_i).$$

Típicamente f se define como:

$$f(I_i, d_i) = \frac{I_i}{d_i},$$

por lo que la distorsión en la intensidad del i -ésimo pixel puede ser expresada a través de un modelo multiplicativo de la forma:

$$I_i = I_i^* \times d_i. \quad (2.3.1)$$

Esta definición facilita el análisis debido a que con una transformación logarítmica, la intensidad observada es simplemente la suma del logaritmo de la intensidad ideal, denotada por y_i^* , y del logaritmo de la intensidad debida a la inhomogeneidad, denotado por d_i . Esto es, si, \mathbf{y} y \mathbf{y}^* los vectores de logaritmos de las intensidades observada e ideal, entonces:

$$\mathbf{y} = \mathbf{y}^* + \mathbf{d}, \quad (2.3.2)$$

donde \mathbf{d} es el logaritmo del campo de distorsión.

La segmentación, por otro lado, puede considerarse como un problema estadístico de clasificación, el cual consiste en asignar a cada pixel una y sólo una etiqueta que lo identifique. Sea \mathcal{L} el conjunto de etiquetas disponible para la imagen. Una etiquetación \mathbf{x} , es una colección de variables x_i con $x_i \in \mathcal{L}$ que denota la etiqueta asignada por algún método al i -ésimo pixel. De este modo suponemos que, dada la etiqueta x_i , la intensidad ideal y_i^* para el pixel i sigue una distribución gaussiana con parámetros $\theta(x_i) = (\mu_{x_i}, \sigma_{x_i}^2)$, es decir $p(y_i^* | x_i) = \mathcal{N}(y_i^* | \mu_{x_i}, \sigma_{x_i}^2)$.

Entonces, tomando en cuenta el campo de distorsión, la distribución anterior puede escribirse como:

$$p(y_i^* | x_i, d_i) = \mathcal{N}(y_i - d_i | \mu_{x_i}, \sigma_{x_i}^2).$$

Como el valor de las etiquetas no es conocido de antemano, es claro que se debe calcular la probabilidad siguiente:

$$p(y_i^*, x_i | d_i) = p(y_i^* | x_i, d_i) \cdot p(x_i | d_i).$$

La cual se simplifica marginalizando con respecto a los valores de las etiquetas

$$p(y_i^* | d_i) = \sum_{j \in \mathcal{L}} p(y_i^*, x_i = j | d_i) = \sum_{j \in \mathcal{L}} p(y_i^* | x_i = j, d_i) \cdot p(x_i = j). \quad (2.3.3)$$

Suponiendo que las intensidades en los píxeles son estadísticamente independientes, la distribución de probabilidad del logaritmo de las intensidades, dado el campo de distorsión es:

$$p(\mathbf{y}^* | \mathbf{d}) = \prod_{j \in \mathcal{L}} \mathcal{N}(y_j^* | \mu_j, \sigma_j^2) p(x_i = j). \quad (2.3.4)$$

De este modo, la distribución de la intensidad del i -ésimo pixel se modela como una mezcla de distribuciones gaussianas. Así, dado el campo de distorsión, la distribución de probabilidad para la imagen completa es:

$$p(\mathbf{y} | \mathbf{d}) = \prod_{i=1}^N p(y_i | d_i). \quad (2.3.5)$$

A través del teorema de Bayes podemos calcular la distribución final del campo de distorsión, dadas las intensidades observadas:

$$p(\mathbf{d} | \mathbf{y}) \propto p(\mathbf{y} | \mathbf{d}) \cdot p(\mathbf{d}), \quad (2.3.6)$$

y donde la distribución inicial para el campo de distorsión $p(\mathbf{d})$ se modela como una distribución normal con media cero para reflejar la suavidad del campo.

Sin embargo, pretendemos ser más generales en el sentido en que la variable aleatoria \mathbf{d} pueda ser modelada como la suma de dos variables aleatorias independientes v_i y ψ_i , donde v_i refleja la heterogeneidad en la distorsión y puede ser modelada en efecto como una normal con media cero, mientras que ψ_i refleja la dependencia espacial en torno a una vecindad del pixel i , cuya distribución puede ser modelada a través de un Campo Aleatorio de Markov (RMF, por sus siglas en inglés). Más aún, si de algún modo se puede extraer información adicional con respecto del tejido u órganos que se estén analizando, esta podría ser incorporada a través de un modelo de regresión, por ejemplo.

Empleamos ahora el principio de máxima distribución a posteriori (MAP) para estimar el campo de distorsión $\hat{\mathbf{d}}$, dado los valores del logaritmo de las intensidades observadas, esto es:

$$\hat{\mathbf{d}} = \arg_{\mathbf{d}} \text{máx } p(\mathbf{d} | \mathbf{y}),$$

cuya solución debe satisfacer la condición de primer orden siguiente:

$$\frac{\partial}{\partial d_i} \ln p(\mathbf{d} | \mathbf{y}) \Big|_{d=\hat{\mathbf{d}}} = 0, \quad \forall i = 1, 2, \dots, N.$$

$$\begin{aligned} \ln p(\mathbf{d} | \mathbf{y}) &= \ln p(\mathbf{y} | \mathbf{d}) + \ln p(\mathbf{d}) \\ &= \ln \prod_{i=1}^N p(y_i | d_i) + \ln p(\mathbf{d}). \end{aligned}$$

En este punto, el empleo de las herramientas variacionales Bayesianas jugarán un papel predominante. Por lo general, se emplean algoritmos genéticos o del tipo templado simulado en la solución de la ecuación anterior, los cuales tienen una tasa de convergencia muy lenta. Los métodos variacionales bayesianos tratan de dar una aproximación a través de una distribución de probabilidad conocida, la cual bien puede ser una mezcla de distribuciones, de $p(\mathbf{d} | \mathbf{y})$ en la ecuación anterior.

2.4 Modelo propuesto

Si se tiene una imagen con N pixeles y para cada pixel i se define la intensidad observada I_i y la intensidad real I_i^* , puede considerarse el factor de distorsión D_i tal que:

$$I_i = I_i^* \times D_i.$$

Tomando en consideración los logaritmos de tal manera que:

$$y_i = y_i^* + d_i,$$

se supone que la imagen está compuesta por un conjunto de \mathcal{L} clases y que para cada pixel i , la etiqueta $x_i \in \mathcal{L}$ denota al componente al que pertenece el pixel.

Si se considera para el pixel i el valor de x_i , entonces se supone que:

$$y_i^* \sim \mathcal{N}(\mu_{x_i}, \sigma_{x_i}^2),$$

es decir, tomando en cuenta el campo de distorsión

$$y_i - d_i \sim \mathcal{N}(\mu_{x_i}, \sigma_{x_i}^2), \quad (2.4.1)$$

de modo que si suponemos que $x_i = j$ y el valor de la distorsión d_i para cada pixel,

$$p(y_i | x_i = j, d_i) = \mathcal{N}(y_i - d_i | \mu_j, \sigma_j^2). \quad (2.4.2)$$

Asumiendo que la etiquetación es independiente de la distorsión

$$p(y_i | d_i) = \sum_{j \in \mathcal{L}} p(y_i | x_i = j, d_i) \cdot p(x_i = j);$$

la cual podemos reescribir, por (2.4.2), como:

$$p(y_i | d_i) = \sum_{j \in \mathcal{L}} \mathcal{N}(y_i - d_i | \mu_j, \sigma_j^2) \cdot p(x_i = j). \quad (2.4.3)$$

De este modo, dada la distorsión, la distribución de la intensidad del pixel se modela a través de una mezcla de distribuciones normales, que se puede estimar vía MCMC, por ejemplo.

Si suponemos la independencia entre las intensidades de los pixeles, (es decir, que el ruido generado por el campo de distorsión es esencialmente un ruido blanco), entonces la distribución de probabilidad para la imagen completa es:

$$p(\mathbf{y} | \mathbf{d}) = \prod_{i=1}^N p(y_i | d_i). \quad (2.4.4)$$

Además, podemos usar el teorema de Bayes para calcular la distribución a posterior del campo de distorsión dada la intensidad observada:

$$p(\mathbf{d} | \mathbf{y}) \propto p(\mathbf{y} | \mathbf{d}) \cdot p(\mathbf{d})$$

Una vez obtenida la distribución a posterior del campo de distorsión, a través del MAP podemos formular un estimador $\hat{\mathbf{d}}$ del campo de distorsión como sigue:

$$\hat{\mathbf{d}} = \arg_{\mathbf{d}} \text{máx } p(\mathbf{d} | \mathbf{y}).$$

Las condiciones de primer orden para la solución optima $\hat{\mathbf{d}}$ deben satisfacer

$$\left[\frac{\partial}{\partial d_i} \ln p(\mathbf{d} | \mathbf{y}) \right]_{d=\hat{\mathbf{d}}} = 0, \quad \forall i = 1, 2, \dots, N.$$

Tomando logaritmos

$$\begin{aligned}\ln p(\mathbf{d} | \mathbf{y}) &= \ln p(\mathbf{y} | \mathbf{d}) + \ln p(\mathbf{d}), & \text{por (2.4)} \\ &= \sum_{i=1}^N \ln p(y_i | d_i) + \ln p(\mathbf{d}). & \text{por (2.4.4)}\end{aligned}$$

Ahora, se deriva con respecto a \mathbf{d}

$$\frac{\partial}{\partial \mathbf{d}} \ln p(\mathbf{d} | \mathbf{y}) = \frac{\frac{\partial}{\partial d_i} p(y_i | d_i)}{p(y_i | d_i)} + \frac{\frac{\partial}{\partial \mathbf{d}} p(\mathbf{d})}{p(\mathbf{d})},$$

por (2.4.3) tenemos:

$$\begin{aligned}\frac{\partial}{\partial d_i} p(y_i | d_i) &= \frac{\partial}{\partial d_i} \sum_{j \in \mathcal{L}} \mathcal{N}(y_i - d_i | \mu_j, \sigma_j^2) \cdot p(x_i = j) \\ &= \sum_{j \in \mathcal{L}} \frac{\partial}{\partial d_i} [\mathcal{N}(y_i - d_i | \mu_j, \sigma_j^2) \cdot p(x_i = j)] \\ &= \sum_{j \in \mathcal{L}} p(x_i = j) \cdot \frac{\partial}{\partial d_i} \mathcal{N}(y_i - d_i | \mu_j, \sigma_j^2) \\ &= \sum_{j \in \mathcal{L}} p(x_i = j) \cdot \frac{\partial}{\partial d_i} \left[\frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{1}{2\sigma_j^2}(y_i - d_i - \mu_j)^2} \right] \\ &= \sum_{j \in \mathcal{L}} p(x_i = j) \mathcal{N}(y_i - d_i | \mu_j, \sigma_j^2) \cdot \left(\frac{y_i - d_i - \mu_j}{\sigma_j^2} \right).\end{aligned}$$

$$\frac{\partial}{\partial \mathbf{d}} \ln p(\mathbf{d} | \mathbf{y}) = \frac{\sum_{j \in \mathcal{L}} p(x_i = j) \mathcal{N}(y_i - d_i | \mu_j, \sigma_j^2) \cdot \left(\frac{y_i - d_i - \mu_j}{\sigma_j^2} \right)}{\sum_{j \in \mathcal{L}} \mathcal{N}(y_i - d_i | \mu_j, \sigma_j^2) p(x_i = j)} + \frac{\frac{\partial}{\partial d_i} p(\mathbf{d})}{p(\mathbf{d})}.$$

Escribimos

$$w_{i,j} = \frac{\mathcal{N}(y_i - d_i | \mu_j, \sigma_j^2) p(x_i = j)}{\sum_{j \in \mathcal{L}} \mathcal{N}(y_i - d_i | \mu_j, \sigma_j^2) p(x_i = j)}.$$

de modo que la condición de primer orden se puede expresar como:

$$\left[\sum_{j \in \mathcal{L}} w_{ij} \left(\frac{y_i - d_i - \mu_j}{\sigma_j^2} \right) + \frac{\frac{\partial}{\partial d_i} p(\mathbf{d})}{p(\mathbf{d})} \right]_{d=\hat{\mathbf{d}}} = 0, \quad \forall i = 1, 2, \dots, N.$$

la cual podemos reexpresar de la siguiente manera:

$$\left[\sum_{j \in \mathcal{L}} w_{ij} \sigma_j^{-2} (y_i - \mu_j) - \sum_{j \in \mathcal{L}} w_{ij} \sigma_j^{-2} d_i + \frac{\frac{\partial}{\partial d_i} p(\mathbf{d})}{p(\mathbf{d})} \right]_{d=\hat{\mathbf{d}}} = 0.$$

Denotamos por

$$R_i = \sum_{j \in \mathcal{L}} w_{ij} \sigma_j^{-2} (y_i - \mu_j)$$

y

$$\psi_i = \sum_{j \in \mathcal{L}} w_{ij} \sigma_j^{-2},$$

la condición de primer orden se puede escribir como:

$$\left[R_i - \psi_i d_i + \frac{\frac{\partial}{\partial d_i} p(\mathbf{d})}{p(\mathbf{d})} \right]_{d=\hat{\mathbf{d}}} = 0.$$

Empleando notación matricial

$$\mathbf{R} = (R_1, R_2, \dots, R_N)',$$

y denotando a ψ como la matriz cuyas entradas están definidas como:

$$\psi_{it} = \begin{cases} \sum_{j \in \mathcal{L}} w_{ij} \sigma_j^{-2}, & \text{si } i = t; \\ 0, & \text{en otros casos.} \end{cases}$$

Se obtiene:

$$p(\mathbf{d}) = \mathcal{N}(\mathbf{0}, \psi_{\mathbf{d}}),$$

$$p(\mathbf{d}) = \frac{1}{\sqrt{(2\pi)^N |\psi_{\mathbf{d}}|}} e^{-\frac{1}{2} \mathbf{d}^t \psi_{\mathbf{d}}^{-1} \mathbf{d}},$$

entonces

$$\frac{\frac{\partial}{\partial \mathbf{d}} p(\mathbf{d})}{p(\mathbf{d})} = -\psi_{\mathbf{d}}^{-1} \mathbf{d}.$$

Así, la condición de primer orden queda expresada:

$$[\mathbf{R} - \psi \mathbf{d} - \psi_{\mathbf{d}}^{-1} \mathbf{d}]_{d=\hat{\mathbf{d}}} = 0.$$

Finalmente, despejando \mathbf{d} se tiene que:

$$\hat{\mathbf{d}} = (\psi + \psi_{\mathbf{d}}^{-1})^{-1} \mathbf{R}, \quad (2.4.5)$$

con

$$R_i = \sum_{j \in \mathcal{L}} w_{ij} \sigma_j^2 (y_i - \mu_j), \quad (2.4.6)$$

y ψ tal que:

$$\psi_{it} = \begin{cases} \sum_{j \in \mathcal{L}} w_{ij} \sigma_j^2, & \text{si } i = t; \\ 0, & \text{en otros casos.} \end{cases}$$

2.5 Estimación de la distribución inicial

La estimación máximo verosímil se sabe es sensible a pequeñas perturbaciones en los datos, de aquí se puede decir que no sería, en principio, un método adecuado para el análisis de imágenes de ultrasonido. La teoría de campos aleatorios de Markov es una forma conveniente de representar la información contextual en una imagen, la cual es crucial en muchos casos para la interpretación del contenido de la imagen.

Supongamos que \mathbf{x}^* es la verdadera, pero desconocida etiquetación de la imagen y que $\hat{\mathbf{x}}$ es una estimación de \mathbf{x}^* . Consideremos que ambas representan una realización de un campo aleatorio \mathbf{X} , más aún, el logaritmo de la intensidad ideal \mathbf{y}^* de la imagen puede ser interpretado como la realización de la variable aleatoria \mathbf{Y}^* . El problema de clasificación se convierte en estimar \mathbf{x}^* , dado \mathbf{y}^* . Tal estimación puede ser obtenida a través del MAP, en efecto, de acuerdo con el teorema de Bayes

$$p(\mathbf{x} \mid \mathbf{y}^*) \propto p(\mathbf{y}^* \mid \mathbf{x}) p(\mathbf{x}), \quad (2.5.1)$$

donde

$$p(\mathbf{y}^* \mid \mathbf{x}) = \prod_{i=1}^S p(y_i^* \mid x_i), \quad (2.5.2)$$

y finalmente

$$\hat{\mathbf{x}} = \arg_{\mathbf{x}} \max p(\mathbf{x} \mid \mathbf{y}^*). \quad (2.5.3)$$

En un campo aleatorio de Markov, solamente observaciones cercanas (suponiendo que lo cercano tiende a estar correlacionado) tienen interacción directa, y tienden por lo tanto a tener la misma clase de etiqueta. De acuerdo al teorema de Hammersley-Clifford, la distribución de probabilidad de un campo aleatorio de Markov \mathbf{X} está dada por la distribución de Gibbs

$$p(\mathbf{x}) = \frac{1}{z} e^{-u(\mathbf{x})}, \quad (2.5.4)$$

donde la función de energía

$$u(\mathbf{x}) = \beta \sum_{c \in G} v_c(\mathbf{x}), \quad (2.5.5)$$

es la suma de los potenciales v_c sobre todos los cliques y β es un parámetro que representa el nivel de interacción (controla el tamaño del cluster).

Como ya hemos supuesto, dada la etiqueta x_i del pixel i , la distribución condicional de la intensidad y_i está dada por $\mathcal{N}(\mu_{x_i}, \sigma_{x_i}^2)$. Si suponemos que una vez conocido el

valor de x_i , los valores de y_i son independientes, entonces:

$$\begin{aligned}
 p(\mathbf{y} | \mathbf{x}) &= \prod_{i=1}^N p(y_i | x_i) \\
 &= \prod_{i=1}^N (2\pi\sigma_{x_i}^2)^{-\frac{1}{2}} e^{-\frac{1}{2\sigma_{x_i}^2}(y_i - \mu_{x_i})^2} \\
 &= \prod_{i=1}^N (2\pi)^{-\frac{1}{2}} e^{\left[-\frac{1}{2\sigma_{x_i}^2}(y_i - \mu_{x_i})^2 - \ln \sigma_{x_i}\right]} \\
 &= (2\pi)^{-\frac{N}{2}} e^{\sum_{i=1}^N \left(-\frac{1}{2\sigma_{x_i}^2}(y_i - \mu_{x_i})^2 - \ln \sigma_{x_i}\right)},
 \end{aligned}$$

es decir, $p(\mathbf{y} | \mathbf{x})$ la podemos escribir de la forma:

$$p(\mathbf{y} | \mathbf{x}) = \frac{1}{z} e^{-u(\mathbf{y}|\mathbf{x})},$$

donde

$$z = (2\pi)^{\frac{N}{2}}$$

y

$$u(\mathbf{y} | \mathbf{x}) = \sum_{i=1}^N \left[-\frac{1}{2\sigma_{x_i}^2} (y_i - \mu_{x_i})^2 - \ln \sigma_{x_i} \right].$$

De este modo

$$p(\mathbf{x} | \mathbf{y}) \propto e^{-u(\mathbf{y}|\mathbf{x})} \cdot e^{-u(\mathbf{x})},$$

es decir

$$\log p(\mathbf{x} | \mathbf{y}) = -u(\mathbf{y} | \mathbf{x}) - u(\mathbf{x}) + \text{constante}, \quad (2.5.6)$$

por lo que (2.5.3) puede replantearse como:

$$\hat{\mathbf{x}} = \arg_{\mathbf{x}} \text{mín} \{u(\mathbf{y} | \mathbf{x}) + u(\mathbf{x})\}. \quad (2.5.7)$$

Solución iterativa de (2.5.7), (solución glotona) en el k -ésimo paso: dadas las intensidades \mathbf{y} y considerando $\mathbf{x}_i^{(k)}$ (todas las etiquetas excepto las del i -ésimo pixel, el valor $\mathbf{x}_i^{(k)}$ se actualiza $\mathbf{x}_i^{(k+1)}$ de manera que se minimice $u(x_i | \mathbf{y}, \mathbf{x}_i^{(k)})$); en el caso de (2.5.5) β es una constante positiva, z una constante de normalización y

$$u(\mathbf{x}) = \sum_{i=1}^N \sum_{l \in V(i)} \beta [1 - \delta_{(x_i=x_l)}], \quad (2.5.8)$$

donde

$$\delta_{(x_i=x_l)} = \begin{cases} 1, & \text{si } x_i = x_l; \\ 0, & \text{en otro caso.} \end{cases}$$

y $\nu(i)$ denota los pixeles vecinos de i . Iterando unas cuantas veces este algoritmo glotón podemos considerar una etiquetación óptima $\hat{\mathbf{x}}$. El esquema estadístico de estimación está completo solo si conocemos, tanto la forma funcional como el valor de los parámetros. Debido a que tanto los parámetros como las etiquetas y más aún el número de clases involucradas es desconocido, podemos proceder como en el algoritmo EM:

1. Dar una aproximación a los parámetros desconocidos:

$$\theta_l = (\mu_l, \sigma_l^2), \quad l \in \mathcal{L}.$$

2. A través del algoritmo glotón estimar \mathbf{x} a través de $\hat{\mathbf{x}}$.
3. Con el vector de datos completo $(\hat{\mathbf{x}}, \mathbf{y})$, estimar $\theta_l^{(i)}$ maximizando

$$E[\log p(\hat{\mathbf{x}}, \mathbf{y} | \boldsymbol{\theta})]. \quad (2.5.9)$$

4. Repetir 2 y 3 hasta encontrar convergencia.

Queda ahora claro, que si conocemos $\hat{\mathbf{x}}$ a través de $\boldsymbol{\theta}$, podemos estimar entonces el campo de distorsión \mathbf{d} a través de (2.4.5) y (2.4.6). Sin embargo, es también necesario conocer de antemano el número de elementos que hay en \mathcal{L} , es decir, conocer de antemano el número de componentes en la imagen.

Al ser desconocido, nuestro problema se resuelve si hallamos un método de estimación también para k el número de componentes en la imagen.

Para $l = 1, 2, \dots, k$, la ecuación (2.4) implica conocer $\theta_l = (\mu_l, \sigma_l^2)$, denotemos por $\boldsymbol{\theta}$ el vector de todos los parámetros desconocidos $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$, si por el momento suponemos k desconocido, la distribución final de

$$p(\boldsymbol{\theta} | k, \mathbf{y}) = \frac{p(\mathbf{y} | \boldsymbol{\theta}, k) \cdot p(\boldsymbol{\theta} | k)}{p(\mathbf{y} | k)}, \quad (2.5.10)$$

con

$$p(\mathbf{y} | k) = \int p(\mathbf{y} | \boldsymbol{\theta}, k) \cdot p(\boldsymbol{\theta} | k) d\boldsymbol{\theta}, \quad (2.5.11)$$

al ser el valor de k desconocido, necesitamos calcular la distribución final

$$p(\boldsymbol{\theta}, k | \mathbf{y}) = p(\boldsymbol{\theta} | k, \mathbf{y}) \cdot p(k | \mathbf{y}) \quad (2.5.12)$$

y

$$p(k | \mathbf{y}) = \frac{p(\mathbf{y} | k) \cdot p(k)}{p(\mathbf{y})}, \quad (2.5.13)$$

combinando (2.5.10), (2.5.12) y (2.5.13) se sigue que:

$$p(\boldsymbol{\theta}, k | \mathbf{y}) = \frac{p(\mathbf{y} | \boldsymbol{\theta}, k) \cdot p(\boldsymbol{\theta} | k) p(k)}{p(\mathbf{y})}. \quad (2.5.14)$$

La idea es ahora obtener la distribución (2.5.14) o una muestra, para ello, suponemos que podemos plantear un número máximo de etiquetas $k_{\text{máx}}$ y que, para $l = 1, 2, \dots, k_{\text{máx}}$, podemos a través del método MCMC obtener una mezcla de tamaño n_l de (2.5.10), la cual denotamos por $\boldsymbol{\theta}^{(l)} = (\theta_1^{(l)}, \theta_2^{(l)}, \dots, \theta_{n_l}^{(l)})$; con el fin de obtener la muestra de

(2.5.14), nos fijamos en la ecuación (2.5.12), que conecta a la muestra que ya tenemos con la que queremos generar excepto por el factor $p(k | \mathbf{y})$.

Si suponemos por el momento que conocemos $p(k | \mathbf{y})$ para cada una de las etiquetas $k = 1, 2, \dots, k_{\text{máx}}$, entonces desde el punto de vista jerárquico elegimos k con probabilidad

$$p_k = \frac{p(k | \mathbf{y})}{\sum_{k'=1}^{k_{\text{máx}}} p(k' | \mathbf{y})} \quad (2.5.15)$$

y digamos para el valor k^* seleccionado a elegir un modelo $\theta^{(k)}$ de los generados a partir de (2.5.10) con distribución uniforme.

Regresamos ahora al problema de calcular $p(k | \mathbf{y})$ de la ecuación (2.5.13), combinando (2.5.13) con (2.5.15)

$$p_k = \frac{p(\mathbf{y} | k) p(k)}{\sum_{k'=1}^{k_{\text{máx}}} p(\mathbf{y} | k') p(k')} \quad (2.5.16)$$

Una vez conociendo el valor de (2.5.16) es posible entonces obtener una muestra análoga a la que obtendríamos vía RJMCMC, a través de un muestreo aleatorio, usando las ponderaciones p_k .

2.6 Algoritmo completo

El algoritmo completo que resuelve de manera simultánea el problema de la corrección de inhomogeneidades, la segmentación y el número de etiquetas k que se explican en las secciones anteriores se plantea como sigue:

En el paso uno, para obtener una etiquetación inicial de la imagen sin la necesidad de emplear un algoritmo tipo glotón para resolver la ecuación (2.5.7), se recurre al algoritmo EM con un número de etiquetas fijo y la etiquetación vía máximo a posteriori.

En el paso dos, se estima el campo de distorsión a través de las ecuaciones (2.4.5)-(2.4.6).

En el paso tres, se recuperan las intensidades reales eliminando el campo de distorsión; la etiquetación para cada pixel i se define como $x_i = \arg_{j \in \mathcal{L}} \max w_{ij}$ y, para encontrar el número de etiquetas se emplea la técnica mencionada de saltos transdimensionales y máximo a posteriori propuesta en la ecuación (2.5.14).

Con las estimaciones halladas repetimos desde el paso dos para volver a encontrar una nueva estimación de los parámetros del modelo y el número de etiquetas.

1. Ejecutar una segmentación de la imagen vía EM, por ejemplo con k fijo para obtener $(\boldsymbol{\theta}, \mathbf{x})$, i.e, los parámetros y una etiquetación inicial.
2. Estimar el campo de distorsión con las ecuaciones (2.4.5)-(2.4.6):

$$\hat{\mathbf{d}} = (\boldsymbol{\psi} + \boldsymbol{\psi}_{\mathbf{d}}^{-1})^{-1} \mathbf{R}$$

$$w_{ij} = \frac{\eta(y_i - d_i | \mu_{x_i} \sigma_{x_i}^2) p(x_i = j)}{\sum_{j \in \mathcal{L}} \eta(y_i - d_i | \mu_{x_i} \sigma_{x_i}^2) p(x_i = j)},$$

donde, por (2.5.4) y (2.5.5) obtenemos que:

$$\begin{aligned} p(x_i = j) &= \frac{1}{z} e^{-\sum_{c \in \mathcal{V}(i)} \beta (1 - \delta_{[x_i = x_l]})} \\ &= \frac{1}{z} e^{-\beta \sum (1 - \delta_{[\text{de vecinos de } i \text{ que tienen etiqueta } j]})}. \end{aligned}$$

3. Recuperar $y_i^* = y_i - \widehat{d}_i$, $x_i = \arg_{j \in \mathcal{Z}} \max w_{ij}$ y proponer k_{\max} para que con la técnica de remuestreo propuesta en (2.5.14), vía MAP, se estimen

$$(\boldsymbol{\theta}, k)^* = \arg_{\boldsymbol{\theta}, k} \max p(\boldsymbol{\theta}, k \mid \mathbf{y}^*).$$

4. Repetir desde el paso 2 con $y_i = y_i^*$.

2.7 Conclusiones

Los métodos bayesianos son quizá la única forma coherente de combinar la información obtenida de los datos con la información preliminar del investigador o de otras investigaciones, en este sentido, es valioso desarrollar sistemas de cómputo que sean capaces de combinar ambas fuentes de información.

La metodología propuesta se puede extender al caso de imágenes de ultrasonido 3D como sigue: Suponiendo que los pixeles están correlacionados espacialmente y que la imagen de ultrasonido 3D se genera a través de la toma de imágenes 2D a distancias pequeñas, entonces es posible aplicar la metodología propuesta en 2D y reconstruir un volumen a partir de la segmentación que se logre interpolando las secciones segmentadas. Se han aplicado técnicas de splines para lograr esto; sin embargo, a través de los métodos variacionales clásicos se pueden obtener superficies suaves con las que quizá se logren mejores resultados.

Bibliografía

- [1] Rueckert, D; Burger, P.; Forbat, S.M.; Mohiaddin, R.D.; Yang, G.Z. **Automatic tracking of the aorta in cardiovascular MR images using deformable models.** IEEE Transactions on Medical Imaging: vol. 16, núm. 5, 581- 590. 1997.
- [2] Tan A, Brian E. **Emergency departamento ultrasoundand echocardiography.** Emerg Med Clin N Am. 23, 1179-94, 2005.
- [3] S.L. Bridal, J.M. Correas, A. Saied, and P. Laugier. **Milestones on the road to higher resolution, quantitative, and functional ultrasonic imaging.** Proceedings of the IEEE, 91(10), 1543-1561, 2003.
- [4] J.A. Noble and D. Boukerroui. **Ultrasound image segmentation: a survey.** IEEE Transaction on Meddical Imaging, 25(8), 987-1010, 2006.
- [5] Hashimoto BE, Kramer DJ, Witala L. **Applications of musculoskeletal sonography.** J Clin Ultrasound, 27(6), 293-318, 1999.
- [6] Van Holsbeeck MT, Introcaso JH. **Physical principles of ultrasound imaging. Musculoskeletal ultrasound.** 2^a ed., St Louis Missouri, Mosby, 1-7, 2001.
- [7] Y. Li, Q. Huang, and L. Jin. **A parameter-automatically-optimizedgraph-based segmentation method for breast tumors in ultrasound images.** Control Conference (CCC), 31st Chinese, 2012.
- [8] H. Neemuchwala, A. Hero, and P. Carson. **Feature coincidence trees for registration of ultrasound breast images.** In International Conference on Image Processing, volume 3, 10-13 vol.3, 2001.