

Medida de Similitud Mixta para el Registro y Fusión de Imágenes

Juan Morales Sánchez, Rafael Verdú Monedero, José Luis Sancho Gómez
Departamento de las Tecnologías de la Información y las Comunicaciones
Universidad Politécnica de Cartagena
e-mail: juan.morales@upct.es; rafael.verdu@upct.es; josel.sancho@upct.es

Resumen. *En este trabajo se presenta un método de registro y fusión de imágenes que hace uso de una medida de similitud mixta, basada en la correlación cruzada normalizada y la diferencia cuadrática media. Históricamente se han utilizado muchas medidas de similitud, pero ninguna es válida para cualquier tipo de imágenes. El uso de medidas de similitud combinadas permite conjugar las propiedades de detección de patrones de diversas medidas de similitud en una sola medida. La medida combinada de similitud propuesta se puede utilizar para registro multimodal de imágenes médicas. Para este propósito, el algoritmo presentado calcula primero la relación entre las intensidades de píxel de cada imagen, partiendo del histograma conjunto de niveles de gris. Entonces se aplica el método de estimación de movimiento no rígido basado en modelos difusos. Los resultados de registro multimodal y fusión de imágenes de CT con PET mostrados prueban la validez del método propuesto.*

1. Introducción

El registro de imagen es el proceso de cálculo de la correspondencia entre dos o más imágenes de una misma escena tomadas generalmente en tiempos diferentes, desde distintas orientaciones y/o captadas por varios tipos de sensores. Esta es una operación realmente necesaria en muchas tareas de procesado y análisis de imágenes. Especial atención merece el registro de imágenes médicas donde las informaciones específicas de cada tipo de imagen (modalidad) se combina y fusionan en una sola aumentando así la capacidad, por parte del clínico, de mejorar la diagnosis. Por ejemplo, un buen registro y fusión de una imagen de tomografía computerizada (CT, "Computed Tomography") con una imagen de tomografía por emisión de positrones (PET, "positron emission tomography") se traduce en una aportación de información adicional de una modalidad respecto a la otra; en este caso, a la información de tipo funcional que aporta la imagen PET, se le añade la información anatómica detallada disponible en la imagen CT [1]. En el registro de imagen, la precisión de los resultados depende de la medida de similitud empleada. Gracias a ésta, se realizan las correspondencias entre las imágenes completas o entre bloques de ellas. A grandes rasgos, las distintas medidas de similitud empleadas hoy día se pueden dividir en dos conjuntos: medidas basadas en niveles de intensidad y medidas basadas en la teoría de la información. Medidas como la suma de las diferencias absolutas, el coeficiente de correlación cruzada y medidas basadas en distancias geométricas son ejemplos de medidas del primer conjunto, mientras que la información mutua lo es del segundo [2]. En [3], se introducen dos nuevas medidas de similitud,

una basada en la divergencia de Jensen y la otra en la entropía de Tsallis-Havrda-Charvát; se trata de dos medidas basadas en la teoría de la información. No existe una única medida de similitud que sea perfecta, por lo que la búsqueda de una medida apropiada para cada caso sigue constituyendo un punto central de investigación en el procesado digital de imagen. En este artículo, se presenta una medida de similitud combinada y basada en los niveles de intensidad.

El artículo se estructura de la siguiente forma. En la Sección 2, se describe la medida de similitud utilizada; en la Sección 3, se describe el algoritmo de registro de imagen multimodal que conjuga una estimación de movimiento no rígido con modelos difusos; la estimación de movimiento está, a su vez, basada en una estimación no lineal de la relación del nivel de intensidad de píxel en ambas imágenes. La Sección 4 muestra el registro de imágenes CT y PET obtenido con la medida de similitud presentada. Por último, se concluye el artículo con una breve discusión del método propuesto y con las referencias utilizadas en este trabajo.

2. Medida de Similitud Combinada

2.1. Medida de Similitud Empleada

La medida de similitud empleada se basa en la correlación cruzada normalizada y en la diferencia cuadrática media. Esta combinación permite conjugar las propiedades de detección asociadas a cada una. La correlación cruzada normalizada es especialmente robusta a algunas diferencias de intensidad entre las imágenes. Esto quiere decir que mientras la difer-

encia en intensidad entre imágenes aumente o disminuya conjuntamente, el coeficiente de correlación seguirá siendo elevado. Además, la correlación cruzada normalizada también es robusta al ruido. Sin embargo, presenta cierta debilidad cuando existe cierta rotación entre las imágenes. En este caso, la diferencia cuadrática media ofrece mejores prestaciones.

Sean I_1 e I_2 dos imágenes. Vamos a obtener la correspondencia entre todos los píxeles de ambas imágenes. Para cada píxel (x', y') en la imagen I_1 y (x, y) en la I_2 , se calcula la medida de similitud empleando bloques centrados en los píxeles (x', y') y (x, y) como sigue:

$$\gamma(x, y) = \frac{r(x, y)}{1 + \kappa d(x, y)} \quad (1)$$

donde κ es una constante positiva que denota la importancia de la diferencia cuadrática media, $d(x, y)$, y $r(x, y)$ es la correlación cruzada normalizada. Se calculan de la siguiente manera:

$$r(x, y) = \frac{\frac{1}{2L+1} \sum_{i,j=-L}^L (I_1^{x'y'}(i, j) - m_1^{x'y'}) (I_2^{xy}(i, j) - m_2^{xy})}{c_1^{x'y'} c_2^{xy}} \quad (2)$$

$$d(x, y) = \frac{\frac{1}{2L+1} \sum_{i,j=-L}^L (I_1^{x'y'}(i, j) - I_2^{xy}(i, j))^2}{(c_1^{x'y'})^2} \quad (3)$$

donde $I_1^{x'y'}$ e I_2^{xy} son bloques de píxeles de anchura $(2L+1)$ centrados en los píxeles (x', y') y (x, y) de las imágenes I_1 e I_2 , respectivamente, y $m_1^{x'y'}$, $c_1^{x'y'}$, m_2^{xy} y c_2^{xy} son el valor medio y la desviación típica de los bloques respectivos calculados según:

$$m_1^{x'y'} = \frac{1}{2L+1} \sum_{i,j=-L}^L I_1^{x'y'}(i, j) \quad (4)$$

$$c_1^{x'y'} = \sqrt{\frac{1}{2L+1} \sum_{i,j=-L}^L (I_1^{x'y'}(i, j) - m_1^{x'y'})^2} \quad (5)$$

Finalmente, la medida de similitud se umbraliza como sigue:

$$\rho(\mathbf{x}) = \begin{cases} \gamma(x, y), & \text{si } \gamma(x, y) > \lambda \\ 0, & \text{si } \gamma(x, y) \leq \lambda \end{cases} \quad (6)$$

donde el umbral $0 < \lambda < 1$ se usa para descartar valores bajos de similitud. La imagen resultante se llama mapa de similitud.

3. Método de registro multimodal y fusión

El método propuesto para registro de imagen multimodal y fusión de imagen se divide en dos etapas. Primero se calcula la relación entre las intensidades de píxel de las imágenes a registrar con un método de estimación no lineal basado en funciones de base

radial mediante el histograma conjunto de niveles de gris. Esta relación se utiliza para modificar los niveles de gris de ambas imágenes. La segunda etapa realiza el registro de imagen con el método de estimación de movimiento no rígido basado en modelos difusos [4].

3.1. Ajuste no lineal de la relación de intensidades de píxel

Dado que la fuente de las imágenes a registrar es diferente, se necesita conocer la relación entre sus intensidades de píxel para poder encontrar la correspondencia espacial entre ambas imágenes. Para obtener esta relación se efectúa un ajuste no lineal del histograma conjunto de niveles de gris. El ajuste se realiza mediante funciones de base radial. El primer paso consiste en obtener el histograma conjunto de niveles de gris, contabilizando para ello el número de ocurrencias de cada pareja de niveles de gris. Este histograma se umbraliza posteriormente, y tratándolo como una imagen, se procesa mediante operadores morfológicos de dilatación y erosión realizando una operación de cierre. El objetivo de este procesado es eliminar las combinaciones con un número escaso de ocurrencias, que no proporcionan información significativa, para así tener en cuenta sólo los niveles de gris de mayor probabilidad. Entonces se etiqueta cada una de las regiones presentes en el histograma procesado. Después, para cada región etiquetada por separado, se procede a promediar los valores en vertical —niveles de gris de $I_2(x, y)$ — para cada uno de los valores en horizontal —niveles de gris de $I_1(x, y)$ —, obteniendo una curva única y equivalente al histograma conjunto procesado.

En este artículo se utilizará el método presentado para registrar imágenes de CT y de PET. Así, la primera imagen es de CT y la segunda es una imagen de PET. Estas imágenes se procesarán como se ha comentado para obtener dos imágenes de niveles de gris modificados y de acuerdo a la estimación no lineal que se realice del histograma conjunto. Las imágenes de CT y de PET con niveles de gris modificados se utilizarán en la fase de estimación de movimiento no rígido que se describe a continuación.

3.2. Algoritmo de estimación de movimiento no rígido

El algoritmo de estimación de movimiento no rígido se presenta en [4], [5]. El objetivo de este método es estimar la correspondencia espacial entre pares de imágenes que contienen movimiento no rígido. El método se puede descomponer en dos etapas:

Estimación del emparejamiento difuso: Esta fase utiliza la medida de similitud presentada en la sección anterior. El objetivo de esta fase es calcular y parametrizar, para cada bloque centrado en cada píxel de la primera imagen, la zona de la segunda imagen que es más similar. Se necesita definir el tamaño

del bloque en la imagen original y el tamaño de la región de búsqueda en la segunda imagen (imagen con movimiento no rígido) para obtener el mapa de similitud. Posteriormente el campo de movimiento con la similitud más alta se parametriza como modelos punto, línea curva o indeterminado.

Regularización mediante modelos deformables: esta fase convierte la información difusa de cada píxel en vectores de correspondencia numéricos. Esto se consigue mediante interpolación no uniforme y aplicación de restricciones, un proceso iterativo que consiste en un filtrado no lineal discreto de mínima energía restringido a los modelos paramétricos de cada píxel.

3.3. Fusión de imágenes

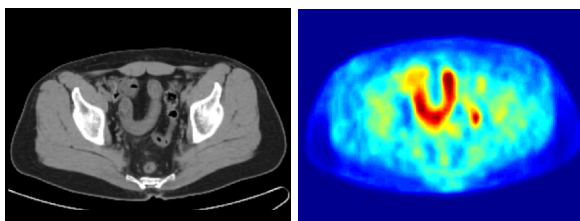
Para conseguir la visualización conjunta se aplica la técnica de la transparencia [6], donde se selecciona el canal de transparencia de acuerdo a la siguiente ecuación:

$$[r, g, b]_{fusion} = \alpha [r, g, b]_{CT} + (1-\alpha) [r, g, b]_{PET} \quad (7)$$

siendo r , g y b los canales de rojo, verde y azul respectivamente de la imagen fusionada. El valor del factor α puede estar entre 0 y 1.

4. Resultados

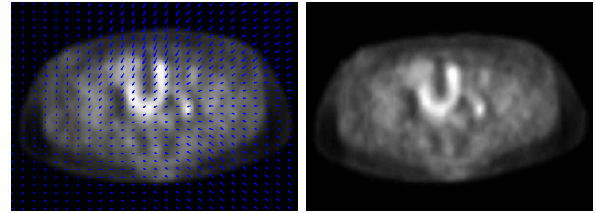
El tamaño de las imágenes de validación es de 540×720 píxeles. Ambas imágenes se muestran en la Fig. 1. El tamaño del bloque de la imagen $I_1(x, y)$ buscado en $I_2(x, y)$ se fija en 5×5 , mientras que la distancia de búsqueda en la imagen $I_2(x, y)$ es 7×7 , lo que produce un tamaño de 15×15 en el mapa de similitud. Los valores prácticos para las constantes κ y λ son 0.3 y 3.0 respectivamente.



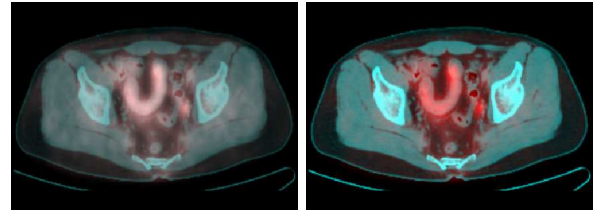
(a) Imagen de CT, $I_1(x, y)$. (b) Imagen de PET, $I_2(x, y)$.

Fig. 1. Imágenes de validación.

Las imágenes $I'_1(x, y)$ y $I'_2(x, y)$ —equivalentes a $I_1(x, y)$ y $I_2(x, y)$ respectivamente pero corregidas en intensidad— se utilizan con el algoritmo de estimación de movimiento no rígido para obtener el registro de las imágenes iniciales. Primero, el algoritmo calcula el mapa de similitud, que se parametriza mediante modelos difusos. Con la regularización paramétrica difusa se calculan los vectores de movimiento. Finalmente, la imagen $I_2(x, y)$ con movimiento no rígido se deforma con los vectores de movimiento obtenidos. Se utiliza la técnica de la transparencia con diferentes valores del factor α para fusionar ambas imágenes. En la Fig. 2 se muestran los resultados obtenidos.



(a) Vectores de movimiento sobre imagen de PET. (b) Imagen de PET registrada.



(c) Imágenes fusionadas con factor de transparencia $\alpha = 0,5$. (d) Imágenes fusionadas con factor de transparencia $\alpha = 0,1$.

Fig. 2. Resultados de fusión de CT y PET.

5. Conclusiones

En este trabajo se presenta un nuevo método de registro y fusión de imágenes multimodales. El método emplea una medida de similitud mixta. El método propuesto se divide en dos etapas. La primera etapa calcula una aproximación no lineal de la relación de intensidades de píxel. La segunda etapa registra las imágenes con un algoritmo de estimación de movimiento no rígido. El registro realizado con un par de imágenes de CT y PET muestra la bondad de la fusión lograda.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia y Tecnología a través del proyecto TIC2002-03033. Las imágenes utilizadas para la obtención de los resultados mostrados son cortesía de *MedPix Medical Database*.

Referencias

- [1] C. Studholme, Derek L. G. Hill, and D. J. Hawkes. Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures. *Medical Physics*, 24(1):25–35, January 1997.
- [2] L. Ding, A. Goshtasby, and M. Satter. Volumetric image registration by template matching. *Image and Vision Computing*, 19:821–832, 2001.
- [3] A. Bardera, M. Feixas, and I. Boada. Normalized similarity measures for medical image registration. *SPIE International Symposium in Medical Imaging*, February 2004.
- [4] J. Morales, R. Verdú, R. Berenguer, and L. Weruaga. Non-rigid motion estimation based on fuzzy models. In *IEEE International Conference on Digital Signal Processing DSP'02*, pages 559–562, Santorini, Greece, 2002.
- [5] J. Morales, R. Verdú, R. González, and L. Weruaga. Fuzzy non-rigid motion estimation robust to rotation. In *IEEE International Conference on Image Processing ICIP'04*, pages 1489–1492, Singapore, 2004.
- [6] T. Porter and T. Duff. Compositing digital images. *Computer Graphics*, 18(3):253–255, 1984.