

# Generación semi-automática de mosaicos. Aplicación de técnicas de fusión de imágenes

C. Miravet\*, J. Santamaría\*, E. Coiras\*, J. Ureña\* J. C. Escudero\*\*, A. Sarasúa\*\*

\* *División Aeroespacial. SENER Ingeniería y Sistemas, S.A.*

*Severo Ochoa 4, 28760, Tres Cantos (Madrid)*

\*\* *Centro de Estudios Ambientales, Ayuntamiento de Vitoria.*

*Armentia 23, 01007, Vitoria-Gasteiz*

## RESUMEN

Se presenta un sistema semi automático de generación de mosaicos que reduce la intervención humana a un llUnimo. El procedimiento esta constituido por los pasos de corrección geométrica, corrección radiométrica, determinación del contenido óptimo en la zona de solape mediante técnicas de fusión de imágenes y generación de la imagen compuesta. Se presentan resultados de su aplicación sobre imágenes utilizadas en aplicaciones reales de teledetección.

**PALABRAS CLAVE:** Generación de mosaicos, registro de imágenes, fusión de imágenes.

## ABSTRACT

A system for semiautomatic mosaic generation is presented. The system, which reduces the human interaction to a minimum, is composed of fue following processing steps: geometric correction, radiometric correction, determination of the optimal image content on overlapping areas by means of image fusion techniques and, finally, generation of the composed image. Results are presented on a real remote sensing application.

**KEY WORDS:** Mosaic generation, image registration, image fusion.

## INTRODUCCIÓN

La generación de una imagen mosaico a partir de un conjunto de imágenes parciales de una zona de interés, constituye una operación de uso frecuente en el ámbito de la teledetección. La obtención de mosaicos de alta calidad requiere frecuentemente la intervención humana especializada en la realización de tareas de tipo repetitivo, con el consiguiente impacto en el coste y laboriosidad del proceso.

En el presente trabajo se describe un sistema semiautomático de generación de mosaicos que reduce la intervención humana a un mínimo, y se presentan resultados de su aplicación sobre imágenes utilizadas en aplicaciones reales de teledetección.

El problema a tratar puede formularse como el de composición en una única imagen (mosaico) de la información proveniente de un conjunto de vistas parciales de una escena. Estas vistas pueden estar adquiridas por distintos sensores, en momentos diferentes o con geometrías de adquisición diferentes, lo que da lugar a la aparición de discontinuidades geométricas y radiométricas de la imagen resultante. El proceso de generación debe minimizar estas discontinuidades y debe seleccionar de forma óptima la información a presentar en las zonas de solapamiento entre distintas vistas, donde se dispone de fuentes de información diferentes relativas al valor radiométrico de cada punto de la escena.

Sin pérdida de generalidad, puede reducirse el problema al caso de la generación de un mosaico a partir de dos imágenes de entrada, pudiendo tratarse el problema general por aplicación iterativa del proceso sobre el mosaico generado en el paso anterior y cada nueva imagen de entrada.

En el siguiente punto se presenta una descripción básica del procedimiento desarrollado, y se exponen en detalle los módulos constituyentes del proceso. Finalmente, se presentan y analizan los resultados obtenidos.

## DESCRIPCION DEL PROCEDIMIENTO

El procedimiento desarrollado para la generación de un mosaico a partir de dos imágenes está constituido por los siguientes pasos:

### a) *Corrección Geométrica*

En este módulo se determinan las funciones de correspondencia geométrica entre las dos imágenes de entrada. El conocimiento de estas funciones permite representar ambas imágenes en un mismo sistema de referencia, asegurando la continuidad geométrica en las zonas de solape.

### b) *Corrección radiométrica*

Una vez conocidas la transformaciones geométricas, es necesario aplicar transformaciones radiométricas para evitar discontinuidades de intensidad y contraste en la imagen resultante. Estas transformaciones se basarán en el análisis de la zona de solape, que es la más sensible a estos

efectos, y en la que además se dispone de información en ambas imágenes, permitiendo imponer condiciones de igualdad en las distribuciones de nivel de gris.

c) *Determinación del contenido óptimo en la zona de solape*

En la zona de solape se dispone de dos valores diferentes de nivel de gris para cada punto de la escena, aún después de uniformizar intensidades y contrastes. En este módulo se calcula el resultado óptimo para cada pixel de forma automática, mediante un procedimiento de fusión de imágenes (Santamaría y Gómez, 1993) desarrollado en SENER y que ha sido aplicado con éxito en diversos entornos operativos (Duclos et al., 1997), (Desco et al., 1996).

d) *Generación de la imagen compuesta*

Los pasos anteriores proveen la zona de superposición. Es necesario entonces generar el resto de la imagen mediante la aplicación del registro según se ha determinado en el primer paso y de las correspondientes correcciones radiométricas.

Estos pasos se describen a continuación en detalle.

**Registro de las imágenes de entrada**

El proceso de registro permite determinar las leyes de correspondencia geométrica entre las dos imágenes de entrada a partir del conocimiento de las coordenadas de un conjunto de puntos homólogos (correspondientes al mismo punto de la escena) en ambas imágenes.

La selección de estos pares de puntos homólogos constituye la única intervención humana en el proceso, pudiendo obviarse si la correspondencia geométrica entre imágenes es conocida de antemano, tal como es el caso si se dispone de imágenes georreferenciadas.

En el sistema desarrollado, el tiempo requerido para la realización de esta operación se ha reducido al máximo mediante la implementación de un entorno gráfico adecuado y la integración de un sistema automático de refinamiento del registro que permite reducir el número de puntos a suministrar para obtener una estimación precisa de las funciones de registro.

Así, una vez determinados manualmente un conjunto de puntos homólogos, el sistema de refinamiento opera mediante extracción de una ventana centrada en las coordenadas del punto seleccionado en la primera imagen, y búsqueda, en un entorno del punto correspondiente de la segunda imagen, de la ventana que maximiza el grado de similitud con la información extraída. Las coordenadas refinadas del punto homólogo en la segunda imagen vienen dadas por la posición del punto central de la ventana seleccionada por este procedimiento.

En nuestro sistema se ha utilizado como medida de similitud el coeficiente de correlación lineal

entre ambas imágenes. Esta medida, que puede absorber variaciones lineales de contraste entre ambas vistas, viene dada, para ventanas locales de dimensión N x N, por:

$$C = \frac{\sum_{i=-N}^N \sum_{j=-N}^N (p_1(i,j) - \bar{p}_1)(p_2(i,j) - \bar{p}_2)}{\sqrt{\sum_{i=-N}^N \sum_{j=-N}^N (p_1(i,j) - \bar{p}_1)^2 \cdot \sum_{i=-N}^N \sum_{j=-N}^N (p_2(i,j) - \bar{p}_2)^2}} \quad (1)$$

Donde  $\bar{p}(i,j)$  representa el nivel de gris en las coordenadas (i, j),  $\bar{p}$  el valor medio de nivel de gris en la ventana seleccionada y el subíndice (1, 2), la imagen considerada.

Tras este proceso de refinamiento, el conjunto de puntos homólogos se utiliza en el ajuste de los parámetros de una función de registro preestablecida, que siguiendo un procedimiento de uso generalizado, se ha aproximado por una función polinómica. En caso de no existir distorsiones o deformaciones de perspectiva importantes, los polinomios de primer grado (transformación afín) permiten obtener correspondencias entre imágenes con un grado de precisión suficiente para esta aplicación. Esta transformación permite corregir efectos de traslación, rotación y escalado anamórfico entre imágenes, resultando en una transformación con seis coeficientes libres que se ajustan por mínimos cuadrados a las coordenadas de los puntos homólogos (Pratt, 1991).

**Nivelación radiométrica de las imágenes de entrada**

Una vez en disposición de las transformaciones geométricas que relacionan ambas imágenes, puede determinarse la zona de superposición entre ambas en la imagen mosaico. Como se ha mencionado anteriormente, la nivelación radiométrica de ambas imágenes se realizará utilizando esencialmente dicha zona de solape.

La nivelación radiométrica permite reducir la diferente percepción de intensidad y contraste globales en cada imagen. En nuestro sistema, este proceso se realiza mediante aplicación de transformaciones afines de nivel de gris que igualen parámetros estadísticos (media y desviación estándar) en los histogramas de las imágenes de entrada sobre la zona de solape. La aplicación directa de estas funciones puede causar saturaciones en los niveles de gris altos y bajos, por lo que se impone como condición adicional la limitación de los porcentajes de saturación de ambas imágenes a los niveles prefijados por el usuario (utilizándose usualmente un valor del 1 % de saturación en ambos extremos del histograma).

La aplicación de las condiciones anteriores resulta en las siguientes funciones de transformación,  $f_1$  y  $f_2$  para los niveles de gris,  $p(0 \leq p \leq 255)$ , de las dos imágenes de entrada, respectivamente:

$$\begin{aligned}
 f_1(p) &= m_1 \cdot p + b_1 \\
 f_2(p) &= m_2 \cdot p + b_2 \\
 m_1 &= \frac{255}{I_{MAX} - I_{MIN}} \\
 m_2 &= m_1 \cdot \frac{\sigma_1}{\sigma_2} \\
 b_1 &= -m_1 \cdot I_{MIN} \\
 b_2 &= m_1 \cdot \mu_1 + b_1 - m_2 \cdot \mu_2
 \end{aligned} \tag{2}$$

Donde  $\mu$  y  $\sigma$  representan, respectivamente, el valor medio y la desviación estándar en la región de intersección para la imagen especificada por el subíndice. Los parámetros ( $I_{MIN}$ ,  $I_{MAX}$ ) vienen dados por las expresiones:

$$\begin{aligned}
 I_{MIN} &= \text{MIN} \left( I_{1,MIN}, \left[ \frac{\sigma_1}{\sigma_2} \left[ I_{2,MIN} - \mu_2 \right] + \mu_1 \right] \right) \\
 I_{MAX} &= \text{MAX} \left( I_{1,MAX}, \left[ \frac{\sigma_1}{\sigma_2} \left[ I_{2,MAX} - \mu_2 \right] + \mu_1 \right] \right)
 \end{aligned} \tag{3}$$

Donde  $MIN$  y  $MAX$  representan, respectivamente, las operaciones de mínimo y máximo y  $[I_{1,2MIN}, I_{1,2MAX}]$  definen los intervalos de nivel de gris  $[0, I_{1,2MIN}]$  y  $[I_{1,2MAX}, 255]$  cuya frecuencia de aparición en la imagen corresponde a los porcentajes de saturación fijados.

### Fusión de imágenes

Una vez realizados los pasos anteriores, se puede proceder a la generación de un mosaico a partir de las dos imágenes de entrada mediante la nivelación radiométrica de ambas imágenes (utilizando las tablas de transformación calculadas) y su transformación geométrica a un sistema de referencia común (utilizando los polinomios de registro). La aplicación de los pasos anteriores en la zona de solape provee dos imágenes diferentes que hay que combinar de forma óptima.

Este proceso se ha realizado mediante un algoritmo de fusión de imágenes desarrollado en SENER. Este algoritmo, que está basado en el funcionamiento de los campos receptivos del sistema visual humano, permite generar, a partir de un conjunto de imágenes de la misma escena, una única imagen de salida que contiene el máximo de información, evitando redundancias.

El algoritmo de fusión de imágenes, opera sobre imágenes rectangulares, lo que imposibilita su aplicación directa a la zona de solape, que en el caso más general tendrá una forma poligonal.

Por ello, la generación de la imagen en la zona de solape se realiza en dos fases:

- Generación de imágenes rectangulares de entrada al algoritmo de fusión.

- Aplicación del algoritmo de fusión.

Estos pasos se describen en los puntos subsiguientes.

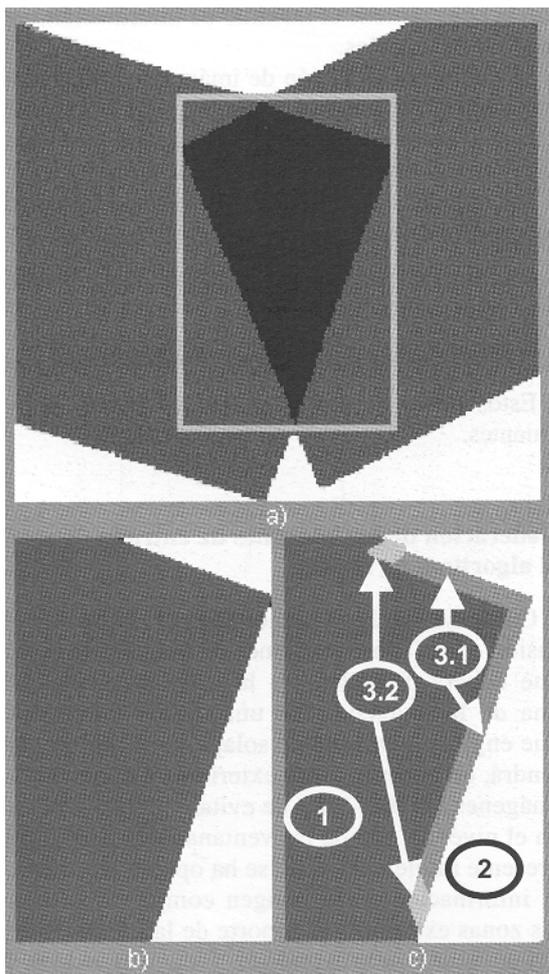
### Generación de las imágenes de entrada al algoritmo de fusión

Como ya se ha mencionado, el algoritmo de fusión opera sobre imágenes rectangulares, por lo que deberá procederse a la extracción, en cada una de las imágenes, de una región rectangular que englobe a la zona de solape. Esta región contendrá, en general, áreas exteriores a una o ambas imágenes a tratar. A fin de evitar cambios bruscos en el nivel de gris de las ventanas extraídas, en la presente implementación se ha optado por utilizar la información de la imagen complementaria en las zonas exteriores al soporte de la imagen sobre la que se realiza la extracción. Esta operación debe realizarse de forma que se eliminen eventuales discontinuidades radiométricas entre las imágenes que componen el mosaico. Estas discontinuidades residuales aparecen incluso después de haber aplicado la nivelación radiométrica de ambas imágenes.

En la Figura 1 se presenta de forma esquemática el proceso de generación de la imagen rectangular, correspondiente a la imagen 1, de entrada al algoritmo de fusión. En *a*) se muestra la zona de solape (en negro) entre ambas imágenes y el rectángulo (que la engloba (tono gris claro)). En *b*) se representan las zonas de existencia (gris) y no existencia (blanco) de datos de la primera imagen en el rectángulo definido. A fin de evitar saltos bruscos de nivel de gris, la información en las zonas de no existencia de datos, para cada imagen, se tomará de la otra imagen. Las discontinuidades radiométricas residuales a lo largo del límite de soporte de cada imagen se eliminan mediante la aplicación de un procedimiento de promediado con pesos variables, que asegura una transición suave entre la información de ambas imágenes.

A continuación se describe el procedimiento para la generación de la imagen de entrada rectangular correspondiente a la imagen 1. El proceso para la imagen 2 se realiza de forma análoga. En la aplicación de esta operación, se consideran cuatro zonas diferentes (Figura 1.c):

1. Puntos exclusivamente existentes en la imagen 1 o existentes en ésta y situados a una distancia de la línea de discontinuidad superior a un umbral predefinido, anchura de la zona de promediado, (región 1 en la Figura 1.c). Estos puntos se transfieren directamente de la imagen 1 a la imagen rectangular 1.
2. Puntos no existentes en la imagen 1 y existentes en la imagen 2 (región 2 en la Figura 1.c). Estos puntos se transfieren directamente de la imagen 2 a la imagen rectangular 1.



**Figura 1.** Proceso de generación de las imágenes de entrada al algoritmo de fusión. a) Zona de superposición entre imágenes (en negro) y rectángulo circunscrito. b) Zonas de existencia (en gris) y no existencia (en blanco) de datos de la primera imagen. c) Regiones de promediado (ver texto) para la primera imagen.

3. Puntos existentes en ambas imágenes y situados dentro de la zona de promediado (regiones 3.1 y 3.2 en la Figura 1 c). En estas regiones, se aplica la siguiente función de promediado:

$$p(i, j) = \frac{d1(i, j)}{D} \cdot p_1(i, j) + \left[ 1 - \frac{d1(i, j)}{D} \right] \cdot p_2(i, j) \quad (4)$$

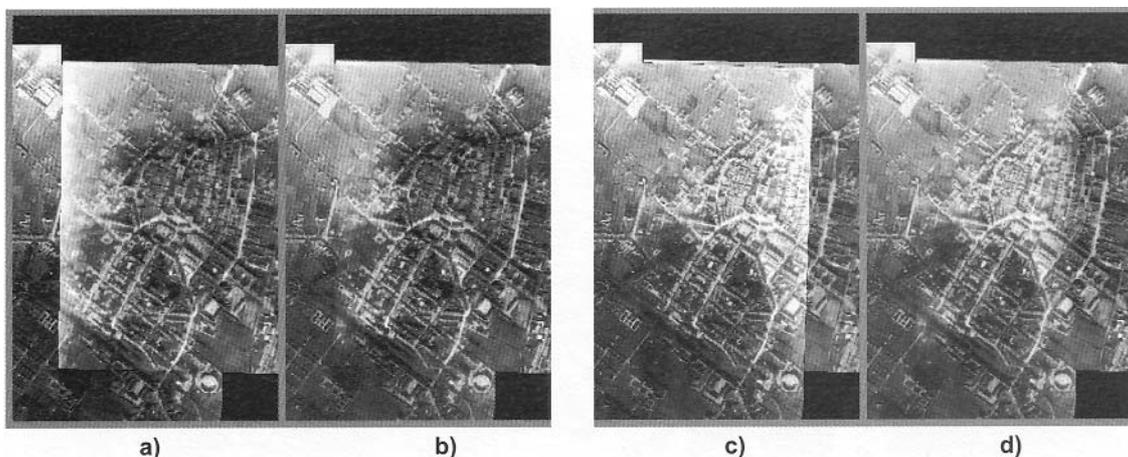
donde  $d1(i, j)$  representa la distancia del punto a la línea de discontinuidad y  $D$  la anchura de la zona de promediado. en nuestro sistema,  $d1(i, j)$  se determina mediante aplicación del algoritmo de distancia  $3/4$  de charnfer (Borgefors, 1988), que proporciona una estimación a la distancia euclídea usando aritmética entera.

La operación anterior garantiza una transición suave si se dispone de información de ambas imágenes en la anchura completa ( $D$ ) de la zona de promediado fijada. en la imagen 1.c) se representa, región 3.1, los puntos de la imagen sobre los que se aplica la función de promediado descrita. en los puntos cercanos al límite de soporte de la imagen 2, region 3.2, no es posible obtener información sobre la anchura completa de la zona de promediado, por lo que es necesario modificar los pesos para impedir la aparición de discontinuidades. en este caso, el nivel de gris de la imagen final se obtendrá por aplicación de los siguientes pesos de promediado:

$$p(i, j) = (1-k) \cdot p_1(i, j) + k \cdot p_2(i, j)$$

$$k = \left[ \frac{d1(i, j) \cdot d2(i, j) + d2(i, j)}{d1(i, j) + d2(i, j)} \right] \cdot \left[ 1 - \frac{d1(i, j)}{D} \right] \quad (5)$$

En la Figura 2 se presentan los resultados obtenidos en la generación de subimágenes correspondientes a la zona de superposición por el método descrito, y se comparan con los resultados obtenidos mediante el uso directo de la imagen complementaria en las regiones exteriores al soporte de la imagen a extraer. Como puede observarse, la aplicación del método descrito resulta en la desaparición casi completa de las discontinuidades radiométricas que resultan patentes en las imágenes generadas por el método directo.



**Figura 2.** Generación de imágenes de entrada al algoritmo de fusión. Generación de la primera imagen: a) método directo, b) método propuesto. Generación de la segunda imagen: c) método directo, d) método propuesto

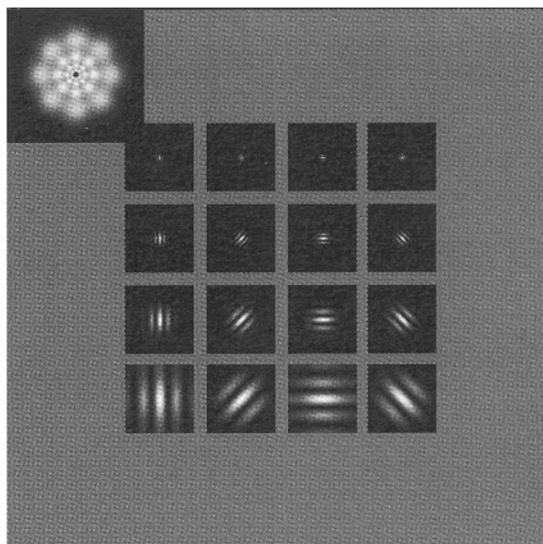
## Fusión de imágenes

Una vez en disposición de las dos imágenes correspondientes a la zona de superposición, estas se combinan de forma óptima en una única imagen mediante aplicación de un algoritmo de fusión basado en el funcionamiento de los campos receptivos del sistema visual humano (filtros de Gabor) (Navarro y Taberero, 1991), (Navarro et al., 1995). El algoritmo permite integrar, en una única imagen, y con máximo contraste, el conjunto completo de detalles contenidos en un conjunto de imágenes de la misma escena, correspondientes a la misma o distintas bandas espectrales.

Este algoritmo consta de los siguientes pasos de proceso:

1. Aplicación de los filtros de Gabor (con 4 niveles de resolución y 4 orientaciones) sobre las imágenes de entrada. Esta operación produce como resultado 16 imágenes que representan la información contenida en cada banda espectral y orientación de las imágenes de entrada. Para minimizar el costo computacional se utiliza un esquema piramidal de cálculo.
2. Comparación, pixel a pixel, de las descomposiciones obtenidas para las dos imágenes de entrada. Generación de la descomposición correspondiente a la imagen fundida.
3. Reconstrucción de la imagen fundida a partir de su descomposición de Gabor.

En la Figura 3 se representan los filtros de Gabor correspondientes a las 16 combinaciones nivel-orientación consideradas y el recubrimiento obtenido en el espacio de frecuencias. En la Figura 4 se presenta el resultado completo del procedimiento de generación de la zona del mosaico correspondiente a la región de solapamiento entre imágenes de entrada. En *a)* y *b)* se muestran, respectivamente, el contenido de información de las dos imágenes



**Figura 3.** Funciones de Gabor para los niveles de resolución (filas) y orientaciones (columnas) considerados. En la esquina superior izquierda se presenta el recubrimiento en el plano de Fourier.

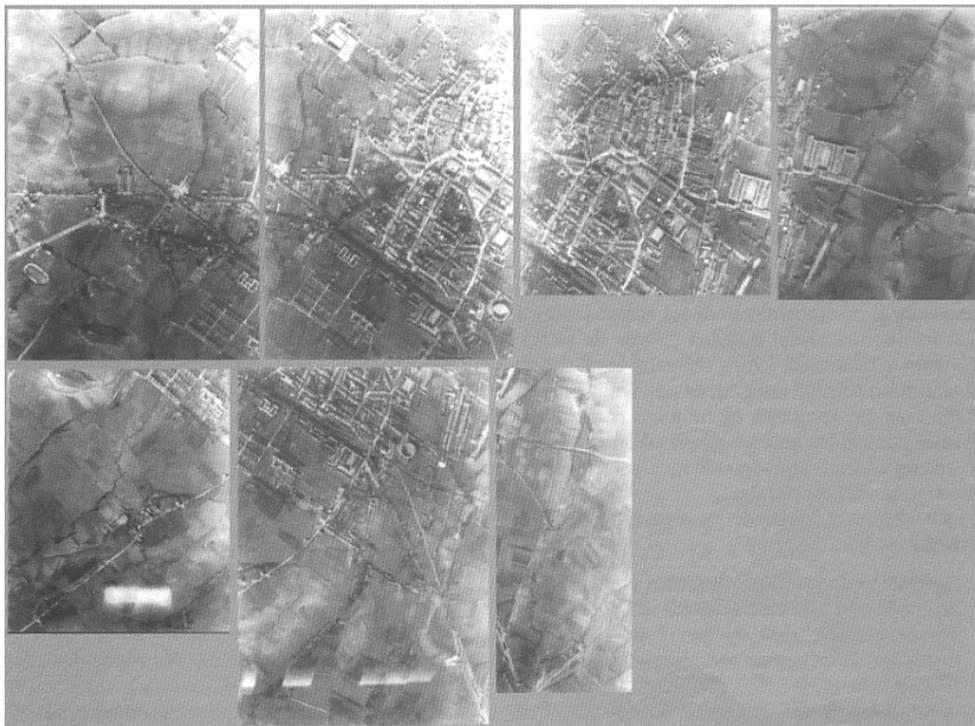
nes a componer en el entorno de la zona de superposición. En *c)* se presenta el resultado obtenido mediante el procedimiento presentado de fusión de imágenes. Como puede observarse, el procedimiento de fusión selecciona de forma automática los componentes de las imágenes que presentan un contenido mayor de información, reduciendo las bandas residuales observables en las imágenes suministradas como entrada al algoritmo (ver Figura 2).

## GENERACION DEL MOSAICO DE IMAGENES. RESULTADOS

Una vez generada la zona del mosaico correspondiente a la región de superposición de las imágenes de entrada, el mosaico en las zonas restantes se calcula mediante aplicación de las tablas



**Figura 4.** Generación de imagen mosaico en la zona de superposición. *a)* Información de la primera imagen en un entorno de la zona de superposición. *b)* Información correspondiente a la segunda imagen. *c)* Imagen fundida.



**Figura 5.** Imágenes de partida utilizadas en la generación del mosaico.



**Figura 6.** Mosaico obtenido mediante aplicación del procedimiento propuesto.

de nivelación de contraste sobre la imagen de entrada con soporte sobre la zona considerada. En las zonas con soporte de la segunda imagen se aplicará de forma previa un procedimiento de rectificación geométrica determinado por los polinomios de registro calculados.

El método presentado ha sido aplicado a un problema real de generación de un mosaico de imágenes a partir de un conjunto de 7 vistas parciales de la ciudad de Vitoria obtenidas durante los años 1932-33. En la Figura 5 se presenta una versión submuestreada de las imágenes utilizadas. Estas imágenes presentan variaciones considerables de contraste entre zonas adyacentes con condiciones cercanas a la saturación en algunos casos, lo que dificulta considerablemente el proceso de eliminación de discontinuidades radiométricas, constituyendo un test crítico de las prestaciones del método desarrollado. En la Figura 6 se presenta el mosaico obtenido con nuestro procedimiento.

## BIBLIOGRAFIA

BORGEFORS, G. 1988, Hierarchical Chamfer Matching: A parametric Edge Matching Algorithm. *IEEE*

*Transactions on Pattern Analysis and Machine Intelligence*, 10: 849-865.

DESCO, M., SANTAMARIA, J., BENITO, C., GÓMEZ, M. T., DOMÍNGUEZ, P., GARCIA-BARRENO, P. 1996. Image Fusion in Multimodality Imaging. *Visualization in Biomedical Computing*.

DUCLOS, SAMY, BONNET, HERAULT, LIVI, MAGRINI, LANDINI, MIRAVET, SANTAMARFA, GARRIGA, BORSKEN, GOUVERNEUR, DRUYTS, ROSENORM, LETH. 1997. Artificial Neural Networks Evaluation in Three Military Contexts: The EUCLID 6.2 HISPARS Project Demonstrators. *Applications and Science of Artificial Neural Networks III*, vol. 3077, SPIE AEROSENSE.

NAVARRO, R., TABERNERO, A. 1991. Gaussian Wavelet Transform: Two Alternative Fast Implementations for Images. *Multidimensional Systems and Signal Processing*, 2: 421-436.

NAVARRO, R., TABERNERO, A., CRISTOBAL, G. 1995. *Image Representation with Gabor Wavelets and its Applications*, Advances in Imaging and Electron Physics, Academic Press.

PRATT, W.K. 1991. *Digital Image Processing*, 2º Ed., John Wiley & Sons.