

# REGISTRO ELÁSTICO DE IMÁGENES DE SATÉLITE DE ALTA RESOLUCIÓN MEDIANTE OPTIMIZACIÓN DE REDES TRIANGULARES

Vicente Arévalo, Javier González, Gregorio Ambrosio  
Dpto. Ingeniería de Sistemas y Automática. Universidad de Málaga.  
Campus de Teatinos. 29071 Málaga.  
e-mail: {varevalo, jgonzalez, goyo}@ctima.uma.es

## Resumen

*El registro de imágenes es una etapa básica en multitud de procesos en el campo de la teledetección, entre otros para la fusión de imágenes de distintos sensores, para la detección de cambios temporales, para el “cosido” de imágenes, etc. Cuando las imágenes presentan distorsiones locales importantes (por ejemplo debidas a la captura desde distintos ángulos de observación) la transformación global de una sobre la otra produce errores considerables en el ajuste. Para resolver este problema se utilizan procedimientos más complejos basados en la deformación elástica de la imagen, como son la descomposición en elementos finitos o las funciones de base radial. Estas técnicas requieren un conjunto de puntos de control que representen adecuadamente las deformaciones locales, por lo que la localización y precisión de estos resultan cruciales en la exactitud del registro. En este trabajo se aborda el registro de imágenes de satélite de alta resolución ( $< 1$  m/píxel) mediante un procedimiento de optimización de redes triangulares adaptativas en dos fases: una que actúa sobre la topología de la red, y otra que modifica la posición de los vértices. El objetivo de ambas es maximizar una medida de consistencia del registro. Este proceso en dos etapas permite adaptar el número y posición de los vértices de la red, modelando las deformaciones geométricas que presentan las imágenes. La medida de consistencia utilizada está basada en el concepto de información mutua y es especialmente robusta cuando las imágenes presentan importantes diferencias radiométricas.*

**Palabras Clave:** Registro de imágenes, imágenes de satélite, teledetección, redes triangulares.

## 1 INTRODUCCIÓN

El registro de imágenes es el proceso de superponer dos o más imágenes de la misma escena tomadas en diferentes fechas, desde diferentes puntos de vista y/o

con sensores diferentes. En este proceso una de las imágenes permanece sin modificar (imagen de referencia) mientras que las restantes (imágenes a corregir) son deformadas hasta encajar con la primera. El registro de imágenes es una etapa crucial en todas las tareas de análisis de imágenes donde los resultados finales se obtienen de la combinación de varias fuentes de datos, como por ejemplo, fusión de imágenes, detección de cambios, detección de patrones, etc.

Normalmente, el proceso de registro se realiza en dos etapas. En la primera etapa, se localiza un conjunto de pares de correspondencias (denominados puntos de control) en ambas imágenes, y en la segunda, se utilizan estas correspondencias para estimar una función de transformación que nos permita superponer con precisión el resto de los puntos de la imagen a corregir. Si nos centramos en el tipo de transformación utilizada y el ámbito de aplicación de la misma (local o global), son numerosos los autores que utilizan funciones polinómicas globales (rígidas, afines, etc.) [3], funciones de base radial locales (por ejemplo “*thin plate spline*” [9]) o combinaciones de ambas [6] para abordar el proceso de registro. La utilización de uno u otro tipo de función depende fundamentalmente del tipo de deformación que presenta la imagen a corregir. En general, las funciones polinómicas no ofrecen buenos resultados en presencia de deformaciones locales, mientras que, las funciones de base radial son más flexibles en este sentido y ofrecen mejores resultados.

Sin embargo, estas técnicas no producen resultados totalmente satisfactorios en las imágenes satelitales de alta resolución disponibles comercialmente en los últimos años (QuickBird, Ikonos, Orbview, etc.) ya que presentan importantes deformaciones locales debidas, fundamentalmente, a la orografía del terreno y los distintos ángulos de adquisición. En la figura 1 se puede observar, de un modo ilustrativo, cómo la geometría de la imagen captada por el satélite en la posición 1 difiere sustancialmente de la captada por el satélite en la posición 2 (obsérvense las marcas de referencia). Estos efectos tratan de compensarse en el proceso de ortorectificación [12], sin embargo, la

ausencia, en muchos casos, de mapas digitales de elevación con suficiente resolución y/o precisión, así como la falta de un modelo exacto conocido del sensor hacen inviable el proceso de compensación, pudiendo incluso producir imágenes con mayor discrepancia respecto a la de referencia que las originales.

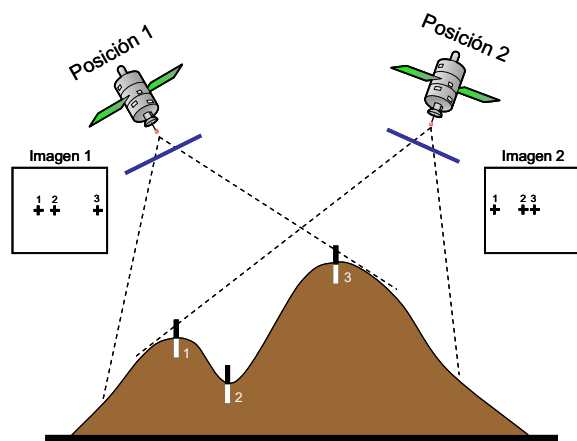


Figura 1: Deformaciones producidas por la orografía del terreno y los distintos ángulos de adquisición.

Para poder registrar imágenes que presentan este tipo de deformaciones surgen los procedimientos basados en elementos finitos triangulares o rectangulares. Este enfoque trata de abordar el proceso de registro mediante la subdivisión de las imágenes en regiones que pueden registrarse con funciones afines [4] o cúbicas [5]. Por otro lado, puesto que los parámetros de las funciones de transformación se determinan en base a un conjunto de pares de correspondencias, es indispensable que dichos pares se encuentren distribuidos convenientemente sobre las imágenes, de acuerdo con la deformación que se pretende corregir. Estos requisitos no quedan garantizados en los métodos que se emplean habitualmente para la búsqueda y emparejamiento automático de puntos.

En este trabajo se presenta un procedimiento para el registro de imágenes de satélite de alta resolución basado en la optimización de redes triangulares adaptativas mediante la maximización de la *información mutua*. Este procedimiento de registro permite adaptar el número y posición de los vértices de la red, modelando las deformaciones geométricas locales que presentan las imágenes.

A continuación se describe la notación utilizada a lo largo del trabajo y algunas definiciones previas. En la sección 3 se describe el método propuesto, sus fases y un esquema general de su funcionamiento. En la sección 4 se presenta los resultados obtenidos para un par de imágenes de alta resolución (0.6 m/píxel). Finalmente, se presentan algunas conclusiones.

## 2 NOTACIÓN Y DEFINICIONES PREVIAS

En esta sección se presenta la notación empleada a lo largo de este trabajo, así como algunas definiciones necesarias.

### 2.1 RED TRIÁNGULAR

Dado un conjunto  $V = \{v_i \mid i = 1, \dots, n\}$  de  $n$  vértices  $v_i = (x_i, y_i)$  se tienen las siguientes definiciones.

**Definición 2.1:** Una triangulación  $T$  de  $V$  es un conjunto de  $m$  triángulos:

$$T = \{t_i \mid i = 1, \dots, m\} \tag{1}$$

donde  $t_i = \{v_j \mid v_j \in V, j = 1, \dots, 3\}$ .

**Definición 2.2:** Una triangulación de Delaunay  $D$  de  $V$  es un conjunto de  $m$  triángulos:

$$D = \{t_i^D \mid i = 1, \dots, m\} \tag{2}$$

donde  $t_i^D = \{v_j \mid v_j \in V, j = 1, \dots, 3\}$  es un triángulo de Delaunay.

**Definición 2.3:** Un triángulo  $t^D$  es de Delaunay si y sólo si el círculo que lo circunscribe está *vacío*, es decir, no encierra ningún vértice de  $V$  (ver figura 2).

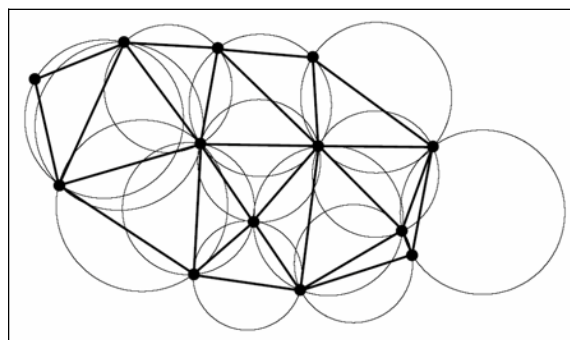


Figura 2: Triangulación de Delaunay.

El algoritmo propuesto por Delaunay [11] para la construcción de redes triangulares maximiza el ángulo mínimo de los triángulos. Esta propiedad es especialmente deseable en los procesos de registro basado en este tipo de redes, ya que triángulos con ángulos agudos no proporcionan resultados aceptables [4].

En general, toda triangulación verifica que:

- (a) Todo punto incluido en la componente convexa de  $V$  queda dentro de un triángulo.

- (b) Los triángulos se cortan sólo en los vértices y aristas compartidas.

**Definición 2.4:** Dada una triangulación  $T$  sobre una imagen  $M$ , se define el área de influencia  $M_v$  del vértice  $v$  como el conjunto de píxeles de  $M$  contenidos en los triángulos con vértice común  $v$  (ver figura 3):

$$M_v = \{p \in T_v\} \tag{3}$$

donde  $T_v = \{t_i \in T \mid v \text{ es vértice de } t_i\}$ .

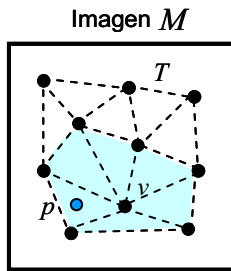


Figura 3: Área de influencia  $M_v$  de un vértice  $v$  para una triangulación  $T$ .

## 2.2 REGISTRO BASADO EN UNA RED TRIÁNGULAR

Sea

$$\{(p_i, p_i') \mid i=1, \dots, n\} \tag{4}$$

con  $p_i = (x_i, y_i)$  y  $p_i' = (x_i', y_i')$

el conjunto de  $n$  pares de correspondencias (puntos de control) localizados en la imagen de referencia y la imagen a corregir (ver figura 4).

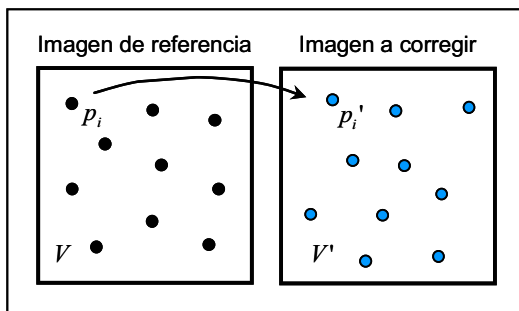


Figura 4: Pares de puntos de control localizados en las imágenes.

El proceso de registro basado en una red triangular se aborda del siguiente modo:

- 1) Se construye una red triangular  $D$  con el conjunto de puntos de control localizados en la imagen de referencia  $V$  (ver figura 5.a).

- 2) Se construye una red triangular  $D'$  con la misma topología que  $T$  pero sustituyendo sus vértices por los puntos de control localizados en la imagen a corregir  $V'$  (ver figura 5.b).

Con esta partición se trata de delimitar aquellas regiones de las imágenes con deformaciones geométricas similares.

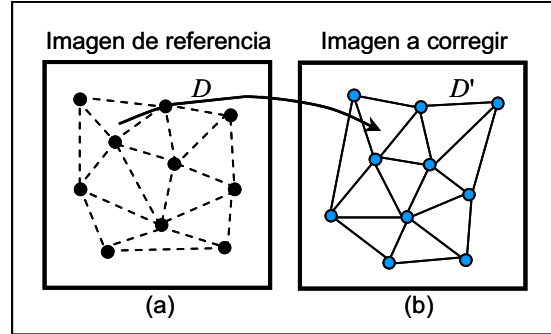


Figura 5: (a) Red de Delaunay  $D$  construida sobre la imagen de referencia. (b) Red de Delaunay  $D'$  construida sobre la imagen a corregir. Nótese que ambas redes comparten la misma topología.

Completada la partición, la siguiente etapa del proceso consiste en determinar, para cada una de las regiones triangulares, las  $m$  transformaciones  $f_i$  que relacionan los  $m$  triángulos de  $D$  y  $D'$ .

$$p = f(p'), f = \begin{cases} f_1(p') & p' \in t_1' \\ f_2(p') & p' \in t_2' \\ \vdots & \\ f_m(p') & p' \in t_m' \end{cases} \tag{5}$$

donde

$$f_i(p') = \begin{pmatrix} a_i & b_i & c_i \\ d_i & e_i & f_i \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix}$$

es un transformación afín con 6 grados de libertad, que puede estimarse a partir de los vértices correspondientes del par  $i$ -ésimo de triángulos [7].

El procedimiento expuesto proporciona resultados aceptables si la distribución y localización de los puntos modelan adecuadamente las deformaciones que presentan las imágenes. Sin embargo, los procedimientos automáticos de búsqueda de pares de puntos de control no garantizan estos prerrequisitos, y la búsqueda manual de pares puede asegurar la distribución pero no que estos sean lo suficientemente representativos de la deformación que se pretende modelar.

### 2.3 MEDIDA DE CONSISTENCIA DEL REGISTRO BASADA EN LA INFORMACIÓN MUTUA

La información mutua (MI) se utiliza por primera vez en el registro de imágenes médicas [10][13] y tiene su origen en la teoría de la información [2]. Desde un punto de vista estadístico, la información mutua mide la dependencia estadística o redundancia de información de dos variables aleatorias. A diferencia de otras medidas de consistencia tales como: la suma de diferencias al cuadrado (SSD), la suma de diferencias en valor absoluto (SAD) o la correlación cruzada normalizada (NCC), la MI no asume una relación funcional a priori entre las imágenes sino una relación estadística que puede medirse analizando su entropía conjunta (ver figura 6). Es decir, no asume que la naturaleza radiométrica de las imágenes sea la misma o muy similar, por tanto, se puede utilizar con imágenes pertenecientes a sensores diferentes o con propiedades radiométricas muy dispares [1].

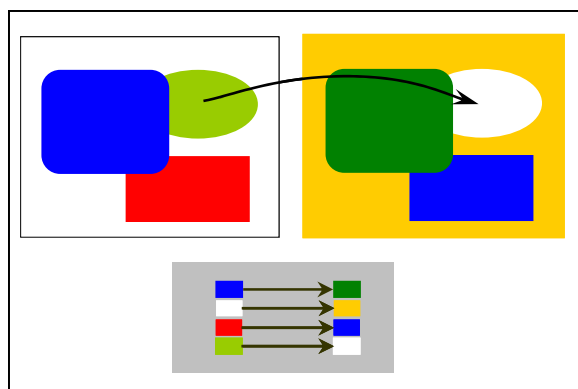


Figura 6: La información mutua asume una relación estadística no funcional.

Sean  $M$  y  $S$  las variables aleatorias correspondientes a los niveles de gris de las imágenes que intervienen en el análisis.

**Definición 2.5:** La información mutua de  $M$  y  $S$  viene dada por la siguiente expresión:

$$MI(M, S) = \sum_i \sum_j P_{M,S}(i, j) \log \left( \frac{P_{M,S}(i, j)}{P_M(i)P_S(j)} \right) \quad (6)$$

donde

$$P_M(i) = \frac{1}{N} \sum_j h_{M,S}(i, j),$$

$$P_S(j) = \frac{1}{N} \sum_i h_{M,S}(i, j),$$

$$P_{M,S}(i, j) = \frac{1}{N} \sum_i \sum_j h_{M,S}(i, j),$$

$N$  es el número de píxeles y  $h_{M,S}$  el histograma conjunto de ambas imágenes.

### 3 MÉTODO PROPUESTO

El método que se propone en este artículo asegura la distribución y localización óptima del conjunto de pares de puntos de control que modelan adecuadamente las deformaciones de las imágenes y, por ende, el registro de las mismas. Se trata de un procedimiento de optimización de redes triangulares basada en la maximización de la información mutua (6). El método requiere que ambas imágenes estén prealineadas, ya sea manual o mediante algún procedimiento automatizado.

El proceso de optimización combina dos fases que se ejecutan iterativamente: una primera fase en la que se modifica la estructura de la red mediante un conjunto de acciones topológicas; y una segunda fase, en la que se optimiza, mediante una técnica de relajación, la localización de los vértices de la red con objeto de determinar su localización óptima para la topología actual. Mediante estas dos etapas se actúa sobre el número, distribución y localización de los pares de correspondencia. El proceso finaliza cuando la función a maximizar alcanza un determinado valor o supera un número máximo de iteraciones. Un método similar para la minimización de una función de energía basada exclusivamente en propiedades topológicas (número de vértices, longitud de los ejes, etc.) se propone en [8].

#### 3.1 MODIFICACIÓN TOPOLÓGICA DE LA RED

En esta fase se altera el número y forma de los triángulos (la topología de la red) a la espera de que dicha modificación mejore la medida de consistencia.

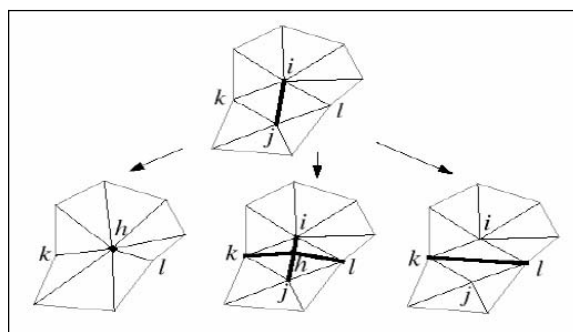


Figura 7: Acciones topológicas válidas.

En el proceso de modificación se emplean las siguientes acciones topológicas (ver figura 7):

- El colapso de aristas (no es aplicable a las aristas que forman parte de la componente convexa de  $V$ ).
- El cambio de orientación de las aristas.

c) La ruptura de aristas.

La elección de la arista o aristas sobre las que actuar y el número y tipo de acciones a ejecutar en cada iteración del algoritmo depende de la política que se establezca. En el trabajo que nos ocupa, el tipo y número de acciones seleccionadas en cada iteración es aleatoria, mientras que, el conjunto aristas sobre las que se actúa se determina en base a la medida de consistencia del registro de la zona de influencia de cada uno de los vértices de  $V$ . Es decir, el conjunto de aristas candidatas son aquellas en las que interviene el vértice cuya área de influencia presenta el menor valor de información mutua.

### 3.2 OPTIMIZACIÓN DE LA POSICIÓN LOS VÉRTICES

En esta fase se refina la localización de los vértices para mejorar la consistencia del registro en sus áreas de influencia. Esta fase no actúa sobre la topología de la red.

El proceso de optimización consiste en recorrer iterativamente aquellos vértices que se han visto afectados por el cambio de topología de la red, modificando su posición mediante incrementos en  $x$  e  $y$  (y siempre dentro de una región de búsqueda acotada), finalizando cuando sus áreas de influencia no mejoran la consistencia del registro (ver figura 8). En el trabajo que nos ocupa, el método de optimización utilizado en el refinamiento de los vértices es el de Levenberg-Marquardt utilizando como función a maximizar la información mutua de las áreas de influencia del vértice sobre el que se actúa.

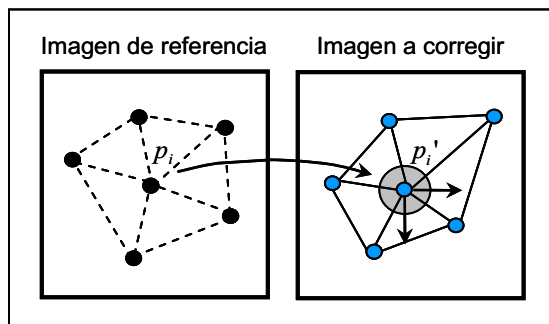


Figura 8: Proceso de optimización de los vértices de la red.

### 3.3 ALGORITMO

En la figura 9 se muestra, a manera de pseudocódigo, el funcionamiento general del método de registro elástico propuesto. Asimismo, el procedimiento se ilustra con el ejemplo de las figuras 10 y 11.

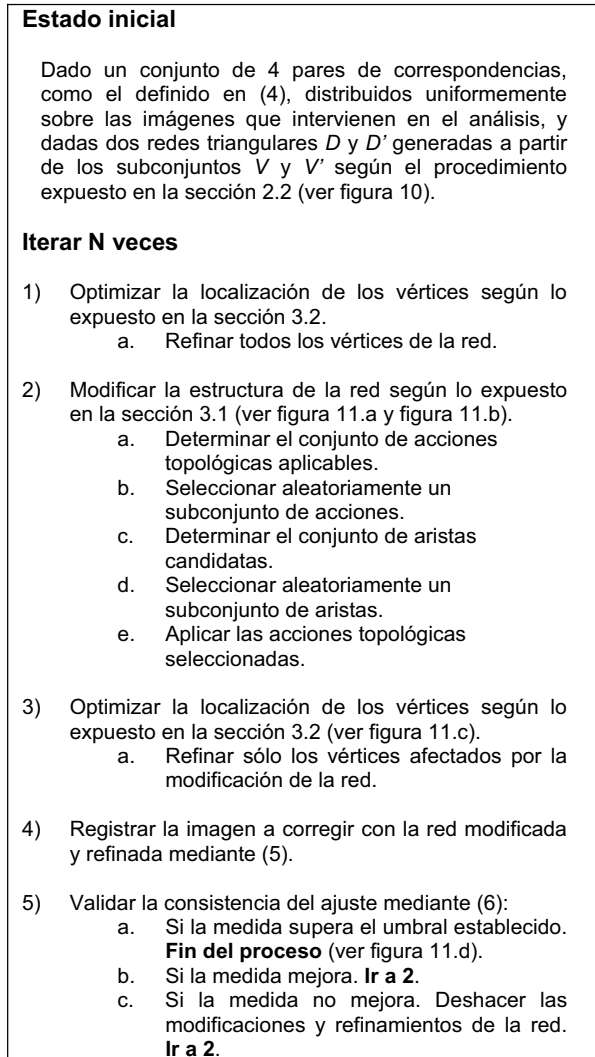


Figura 9: Esquema general del procedimiento de registro propuesto.

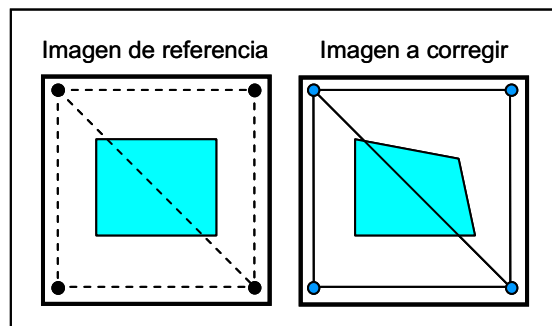


Figura 10: Redes triangulares  $D$  y  $D'$  generadas a partir del conjunto de correspondencias inicial.



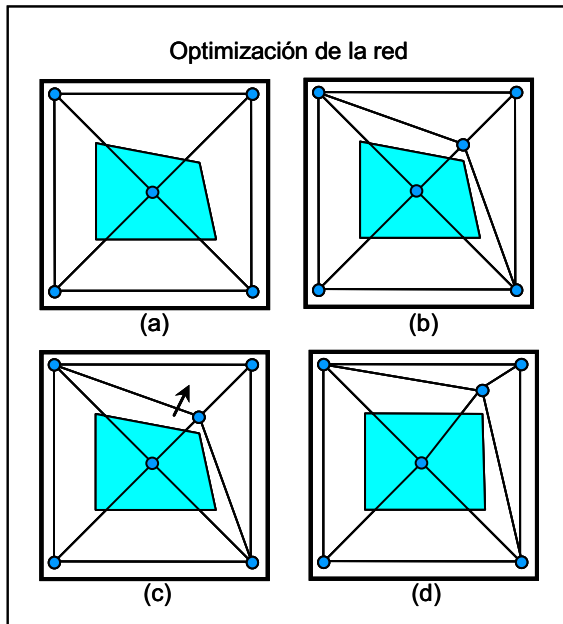


Figura 11: Modificaciones topológicas realizadas sobre la red  $D'$  inicial. (a) Ruptura de la arista compartida. (b) Ruptura de la arista en la que interviene el vértice con el menor valor de MI en su zona de influencia. (c) Refinamiento de los vértices afectados por el cambio de topología. (d) Imagen a corregir registrada.

#### 4 RESULTADOS

El algoritmo desarrollado en este trabajo ha sido probado tanto con imágenes sintéticas que ha sido distorsionadas a mano, como imágenes reales de satélite. A continuación se presentan, a modo de ilustración, resultados de un experimento realizado con imágenes QuickBird con una resolución espacial de 0.6 m/píxel (ver figura 12 y figura 13). Los fragmentos pertenecen a una zona montañosa del municipio malagueño del Rincón de la Victoria y fueron captadas en Julio del 2004 y Abril del 2005, respectivamente.



Figura 12: Imagen QuickBird de referencia (2004).



Figura 13: Imagen QuickBird a corregir (2005).

En la figura 14 se muestran ambas imágenes prealineadas mediante una transformación rígida y puntos de control localizados manualmente. Las flechas indican las distintas deformaciones locales que presentan las imágenes debidas a los diferentes ángulos de adquisición y a la orografía del terreno.



Figura 14: Imágenes QuickBird superpuestas antes de iniciar el proceso de registro.



Figura 15: Imágenes QuickBird superpuestas una vez finalizado el proceso de registro.

En la figura 15 se puede observar la imagen de Abril registrada después de 67 iteraciones. El proceso de registro se inició con un conjunto de 9 vértices distribuidos regularmente (8 triángulos) y finalizó con conjunto de 33 vértices organizados como muestra la figura 16. En ambas figuras se puede observar el proceso de optimización de la red triangular, y cómo este proceso permite modelar las deformaciones locales que presentaba la imagen a corregir.

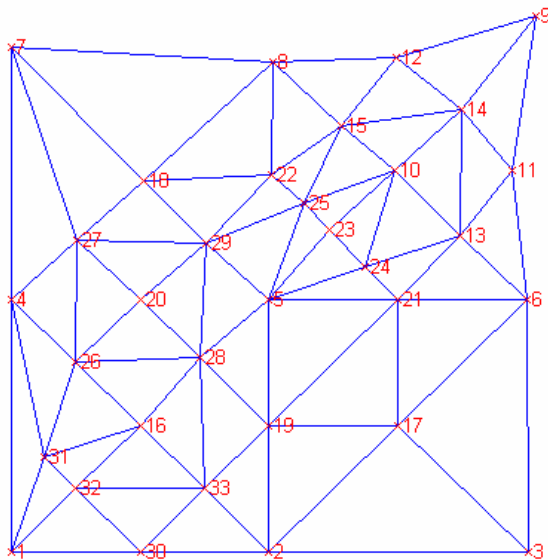


Figura 16: Topología de la red  $D'$  al finalizar el proceso.

## 5 DISCUSIÓN Y CONCLUSIONES

El registro de imágenes de satélite es un problema abierto y en continua evolución dada la proliferación de nuevos satélites capaces de generar imágenes de alta resolución o muy alta resolución (QuickBird, Ikonos, etc.). La posibilidad de adquirir imágenes de una misma zona desde distintas posiciones conlleva que las imágenes captadas presentan importantes deformaciones locales, requiriendo, por tanto, procedimientos de registro capaces de corregir con éxito los importantes desajustes geométricos entre imágenes. Este problema no es exclusivo de la teledetección sino que se presenta también en multitud de aplicaciones como en medicina, en la construcción de vistas panorámicas mediante cosido de imágenes, en reconocimiento de patrones, etc.

En este trabajo se presenta un procedimiento de registro de imágenes basado en la optimización de redes triangulares mediante la maximización de la información mutua. Como resultado del procedimiento conseguimos generar un conjunto de correspondencias representativas (en términos

geométricos) de las deformaciones locales presentes en la imagen, y como consecuencia de esto, la imagen a corregir registrada.

El método ha sido probado con imágenes de muy alta resolución (0.6 m/píxel) adquiridas por el QuickBird en distintas fechas y con unos perfiles radiométricos completamente distintos con unos resultados aceptables.

## Agradecimientos

Agradecer la ayuda y colaboración en el desarrollo de este trabajo a la empresa "DECASAT S.L." de Málaga.

## Referencias

- [1] Chen, H.M., Pramod, K.V., and Manoj, K.A. (2003). "Performance of Mutual Information Similarity Measure for Registration of Multitemporal Remote Sensing Images". *IEEE Transactions on Geosciences and Remote Sensing*, 41(11), pp. 2445.
- [2] Cover, T.M., and Thomas, J.A. (1991). "Elements of Information Theory", Wiley-Interscience, New York.
- [3] González, J., Ambrosio, G., and Arévalo, V. (2001). "Automatic Urban Changes Detection From the IRS-1D PAN", *IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion Over Urban Areas*, Roma.
- [4] Goshtasby, A. (1986). "Piecewise Linear Mapping Functions for Image Registration", *Pattern Recognition*, 19, pp. 459-466.
- [5] Goshtasby, A. (1987). "Piecewise Cubic Mapping Functions for Image Registration", *Pattern Recognition*, 20, pp. 525-533.
- [6] Goshtasby, A. (1988). "Registration of Images with Geometric Distortions". *IEEE Transactions on Geoscience and Remote Sensing*, 26, pp. 60-64.
- [7] Hartley, R.I., and Zisserman, A. (2000). "Multiple View Geometry in Computer Vision". Cambridge University Press.
- [8] Hoppe, H., DeRose, T., Duchamp, T., McDonald, J., and Stuetzle, W. (1993). "Mesh Optimization", *ACM SIGGRAPH*, pp. 19-26.
- [9] Likar, B., and Pernuš, F. (2001). "A Hierarchical Approach to Elastic Registration Based on Mutual Information". *Image and Vision Computing*, 19, pp. 33-44.
- [10] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P. (1997). "Multimodality Image Registration by Maximization of Mutual Information". *IEEE Transactions on Medical Imaging*, 16(2), pp. 187.
- [11] Shewchuk, J.R. (1999). "Lecture Notes on Delaunay Mesh Generation". *Department of Electrical Engineering and Computer Science*. University of California at Berkeley.
- [12] Toutin, T. (2004). "Review Article: A Geometric Processing of Remote Sensing Images: Models, Algorithms, and Methods". *International Journal of Remote Sensing*, 25(10), pp. 1893-1924.
- [13] Viola, P., and Wells W.M. (1997). "Alignment by Maximization of Mutual Information". *International Journal of Computer Vision*, 24(2), pp. 137-154.