

Análisis multirresolución y fusión de imágenes

MARÍA MONCAYO¹ Y JUAN FRANCISCO REINOSO²

1. Departamento de Matemática Aplicada y Estadística.
Universidad Politécnica de Cartagena.
2. Departamento de Ingeniería Cartográfica, Geodésica y Fotogrametría.
Universidad de Granada.

maria.moncayo@upct.es; jreinoso@ugr.es

Resumen

Una posible definición del término *fusion* podría ser la combinación simultánea de información procedente de fuentes distintas. El uso de datos que se complementan tiene como objetivo mejorar la calidad y la interpretabilidad del resultado una vez que los datos a fusionar se “mezclan”. En la última década se ha producido una cantidad ingente de diferentes técnicas para fusionar imágenes, uniéndose a la multitud de métodos existentes. Entre las técnicas recientes, destacan las descomposiciones multirresolución porque, entre otras cosas, minimizan la distorsión de color en la imagen final o fusionada. En nuestro análisis, del que haremos un esbozo en este artículo, hemos extendido la clase de filtros que tradicionalmente se emplean en los algoritmos de fusión a una clase más amplia que se define mediante filtros interpolatorios con origen en esquemas de subdivisión. Esto nos ha permitido emplear filtros que dependen de un parámetro. Este grado extra de libertad se ha traducido en la posibilidad de resolver problemas de optimización en términos del parámetro que dan lugar a imágenes fusionadas que cuentan con las mejores medidas de calidad. También hemos introducido transformadas no lineales que mejoran los clásicos algoritmos (lineales) de fusión.

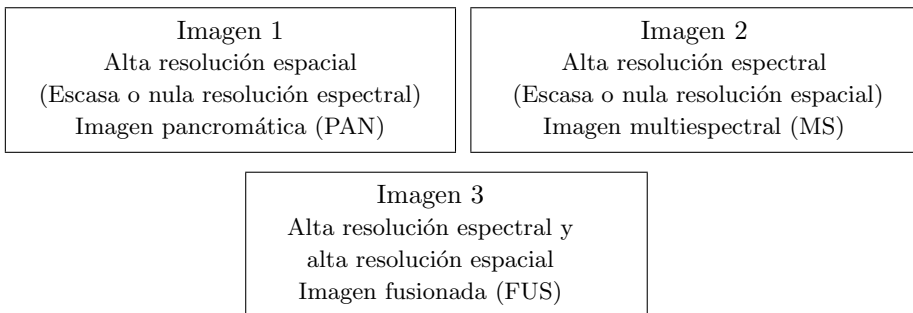
Proyecto/Grupo de investigación: Métodos numéricos y su relación con modelos matemáticos. Entidad financiadora: FUNDACIÓN SÉNECA. Plan de Ciencia y Tecnología de la Región de Murcia. Código: 08662/PI/08.

Líneas de investigación: *Fusión de imágenes; Análisis multirresolución; Teoría de aproximación; Esquemas de subdivisión.*

1. Introducción

En lo que sigue nos referiremos al problema de fusión de imágenes a aquel que consiste en componer una nueva imagen a partir de dos originales, ambas tomadas en una misma escena. Existen muchas situaciones, por ejemplo en control remoto, en las que se requiere que una misma imagen tenga una alta resolución espacial y *también* presente una alta resolución espectral (o calidad en el color). Sin embargo, los dispositivos que recogen imágenes y toman fotografías desde satélites nos ofrecen o bien imágenes con una alta calidad espectral pero nula información espectral o bien imágenes con gran calidad espectral pero sin resolución espacial. Las primeras se denominan *imágenes pancromáticas* (PAN), mientras que las segundas reciben el calificativo de *multiespectrales* (MS).

El objetivo, pues, es conseguir una imagen, la imagen fusionada (FUS), que presente tanto una alta calidad espacial como espectral. La imagen fusionada se obtiene a partir de la imagen PAN y la imagen MS:



Conviene tener presente que existe una relación inversa entre ambos tipos de resolución [17]. Esto implica el hecho de que no sea posible incrementar arbitrariamente ambos tipos de resoluciones. Ni tan siquiera es posible incrementar una de ellas sin que vaya en detrimento de la otra. En los trabajos [3], [10] podemos encontrar algunos métodos ya clásicos para llevar a cabo la fusión de imágenes. Estos métodos muestran degradación espectral y distorsión de color. De hecho, la distorsión de color aparece fundamentalmente porque nunca se consigue una separación perfecta entre la información espacial y la espectral.

Entre las técnicas más recientes que se han incorporado al problema de fusión se encuentran las transformadas multirresolución [1]. Estas transformadas se concretan básicamente mediante dos tipos de algoritmos: Los que tienen estructura *piramidal* y los que tienen estructura en *paralelepípedo*. Con la estructura piramidal, las imágenes o señales cambian de tamaño al aumentar o disminuir la resolución. Sin embargo, con la estructura en paralelepípedo, en cada etapa de descomposición o reconstrucción las imágenes presentan distinta resolución pero conservan el mismo tamaño. La forma piramidal corresponde a

transformadas ortogonales tales como el algoritmo de Mallat. El algoritmo de Mallat, aunque se ha empleado con éxito en compresión de señales, provoca artefactos no deseados en lo que se refiere a la fusión de imágenes. Estos artefactos se explican por el hecho de ser la transformada ortogonal una transformada no invariante por traslaciones.

Ambos algoritmos son discretizaciones diferentes de la misma transformada wavelet continua. En [4] se presenta una implementación de la transformada wavelet discreta que utiliza evaluaciones en lugar de integración. La correspondiente transformada es interpolatoria y los filtros que la definen tienen una forma característica (ver [4] y [11], thm. 7.19.). Los filtros interpolatorios resultantes son simétricos [14] y son iguales a cero en las componentes pares. Esto hace que se les denomine filtros con *agujeros* o filtros *à trous*. La correspondiente transformada es no ortogonal e invariante por traslaciones. Con la transformada paralelepípedo, el tamaño de los coeficientes de la imagen es constante en todos los niveles de descomposición y la fusión se realiza sin que aparezcan los artefactos comunes al emplear transformadas ortogonales. Por tanto, la transformada con estructura de paralelepípedo es una herramienta muy apropiada para abordar satisfactoriamente el problema de fusión.

En nuestro análisis hemos trabajado las siguientes ideas:

- Definición de nuevas transformadas paralelepípedo (lineales) haciendo uso de filtros interpolatorios con origen en esquemas de subdivisión.
- Utilización de filtros interpolatorios dependientes de un parámetro con idea de considerar el valor del parámetro que ofrezca mejores medidas de calidad espacial y espectral para la imagen fusionada.
- Definición de transformadas en paralelepípedo no lineales con las que se consigue mejorar las medidas de calidad que se obtienen en el caso lineal.

En este artículo se dará una visión general acerca de en qué consiste el análisis multirresolución y de la aplicación al problema de fusión.

2. Análisis multirresolución

Las técnicas de análisis multirresolución aparecen en los años 80 con idea de introducir una transformada que respondiera a ciertos problemas para los que la transformada de Fourier no resulta una herramienta del todo satisfactoria. Se trata de la transformada de ondículas o *wavelets*. El marco teórico en el que se inserta todo el engranaje de las pequeñas ondas, ondículas o *wavelets* es justamente el análisis multirresolución. En este marco teórico, que una señal tenga más o menos resolución es equivalente a que pertenezca a un espacio de aproximación u a otro. Más precisamente, consideremos el caso unidimensional y el espacio de señales de energía finita $L^2(\mathbb{R})$. En este espacio se define una cadena de espacios de aproximación de la siguiente forma:

$$\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots \subset V_{j+1} \subset V_j \subset \cdots \subset L^2(\mathbb{R}).$$

Se asume que

$$\overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R}), \quad \bigcap_{j \in \mathbb{Z}} V_j = \{0\}.$$

Los espacios V_j son espacios de aproximación en los que se proyectan señales que tienen resolución 2^{-j} , $j \in \mathbb{Z}$. A medida que j decrece, los espacios ganan en resolución. Empleamos notación decreciente para los espacios mayores o de mayor resolución. De este modo, si $f \in L^2(\mathbb{R})$, la proyección de f sobre V_j y que denotaremos f^j , tiene mayor resolución que f^{j+1} , es decir, que su proyección sobre V_{j+1} . Aunque no es el objetivo de estas páginas entrar minuciosamente en definiciones y teoremas, no podemos dejar de mencionar que la definición matemática de resolución tiene que ver con el número de elementos que conforman una base de cada espacio de aproximación.

En el análisis multirresolución tanto o más importantes que los espacios V_j son los espacios intermedios entre dos espacios de aproximación consecutivos. Son los denominados *espacios de detalle*. Es decir, los espacios W_j tales que

$$V_{j+1} + W_{j+1} = V_j. \quad (1)$$

En el caso de que la suma anterior sea suma directa, entonces el marco de análisis multirresolución que se define es ortogonal. En general W_j no tiene por qué ser el complemento ortogonal de V_{j+1} en V_j .

Hasta ahora, todo lo que se ha comentado no deja de ser un ejercicio puramente algebraico. Es muy relevante el hecho de que verdaderamente existan bases de funciones que generan a los espacios de aproximación y otras bases, distintas a las anteriores pero relacionadas entre sí, tales que generan a los espacios de detalle. En concreto, los espacios de aproximación V_j se generan mediante dilaciones y traslaciones de una función que se denomina función de escala, $\varphi(t)$ y que satisface la relación a dos escalas siguiente:

$$\varphi(t) = \sqrt{2} \sum_{n \in \mathbb{Z}} h_n \varphi(2t - n).$$

A partir de la función de escala se define una función $\psi(t)$, denominada función wavelet mediante $\psi(t) = \sqrt{2} \sum_{n \in \mathbb{Z}} g_n \varphi(2t - n)$, donde $g_n = (-1)^{1-n} h_{1-n}$. Los espacios de detalle W_j se generan mediante

$$W_j = \langle \psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k); k \in \mathbb{Z} \rangle.$$

A partir de las propiedades de la cadena de subespacios y de (1) obtenemos también que

$$L^2(\mathbb{R}) = V_j \cup \left(\bigcup_{j \in \mathbb{Z}} W_j \right). \quad (2)$$

Teniendo en cuenta (2) y las bases de los espacios V_j y W_j , se deduce que toda $f \in L^2(\mathbb{R})$ se puede descomponer como

$$f(t) = \sum_{k \in \mathbb{Z}} f_k^j \varphi_{j,k}(t) + \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} d_k^j \psi_{j,k}(t).$$

Los coeficientes f_k^j y d_k^j se denominan, respectivamente, coeficientes de aproximación y de detalle. De acuerdo con (1),

$$\begin{aligned} \{f^{j+1}, d^{j+1}\} & \text{ contiene la misma información que } f^j \\ \{f^j, d^j, \dots, d^1\} & \text{ contiene la misma información que } f^0. \end{aligned}$$

Dado $f^0 = (f_k^0)$ un vector original de datos, una representación multirresolución de f^0 a resolución $j \geq 1$ es una sucesión

$$\{f^j, d^j, d^{j-1}, \dots, d^2, d^1\}.$$

De modo genérico, un algoritmo multirresolución relaciona los coeficientes f_k^j y d_k^j con f_k^{j+1} . Teniendo en cuenta (1), el procedimiento para obtener f_k^j a partir de $\{f_k^{j+1}, d_k^{j+1}\}$ se denomina reconstrucción. Recíprocamente, la obtención de $\{f_k^{j+1}, d_k^{j+1}\}$ a partir de f_k^j se denomina descomposición. Esquemáticamente, tenemos los algoritmos de descomposición y reconstrucción como sigue:

$$\begin{aligned} f^j & \longrightarrow \{f^{j+1}, d^{j+1}\} & \text{Descomposición} \\ f^j & \longleftarrow \{f^{j+1}, d^{j+1}\} & \text{Reconstrucción.} \end{aligned}$$

2.1. Algoritmos multirresolución en paralelepípedo

En el caso de un análisis multirresolución ortogonal, los algoritmos de descomposición y reconstrucción se definen mediante algoritmos rápidos en términos de convoluciones con los filtros h and g . Estos algoritmos se conocen como los algoritmos de Mallat (ver [11], p. 255).

En el caso de análisis multirresolución no ortogonal (como, por ejemplo, de tipo interpolatorio), la construcción de wavelets interpolatorio aunque sigue la misma estructura formal que el caso ortogonal, los coeficientes se obtienen mediante evaluaciones y no mediante productos escalares de tipo integral. En este caso, la función de escala φ satisface una condición de tipo interpolatoria $\varphi(n) = 1$ si $n = 0$ y $\varphi(n) = 0$ si $n \neq 0$. La función de escala interpolatoria $\varphi(t)$ satisface la relación a dos escalas con $h_n = (h_{or} \star \bar{h}_{or})_n$, donde, h_{or} es un filtro asociado a un esquema multirresolución ortogonal y, para cualquier sucesión s , se denota $(\bar{s})_n = (s)_{-n}$. Como consecuencia h es un filtro simétrico y verifica

$$\begin{aligned} h_{2k} &= 0 & \text{si } k \neq 0, \\ h_0 &= 1/2. \end{aligned} \tag{3}$$

Estos filtros son los denominados filtros con *agujeros* filtros *à trous* [14]. Introducimos la siguiente notación para una sucesión arbitraria v . La notation

es $v_k^{(j)} = v_k$ si $\frac{k}{2^j} \in \mathbb{Z}$ y $v_k^{(j)} = 0$, en otro caso. En el caso del algoritmo en forma de paralelepípedo, las fórmulas de descomposición vienen dadas mediante convoluciones unidimensionales como sigue [8]:

$$\begin{aligned} f_k^{j+1} &= (\bar{h}^{(j)} \star f^j)_k \\ d_k^{j+1} &= (\bar{g}^{(j)} \star f^j)_k, \end{aligned} \quad (4)$$

donde

$$g_k = -h_k \quad \text{si } k \neq 0 \quad \text{and } g_0 = h_0. \quad (5)$$

La extensión a dimensión dos de (4) se consigue mediante convoluciones matriciales o dos dimensionales. Las convoluciones se realizan mediante matrices de datos (a posteriori imágenes) $f_{(k,l)}^j$ y filtros separables producto tensorial entre h y g . Más precisamente, si h y g son los filtros presentes en (4), entonces se define el filtro separable y matricial como $(hg)_{(k,n)} = h_k g_n$. La fórmula de descomposición en dimensión dos viene dada por:

$$\begin{aligned} f_{(k,l)}^{j+1} &= (\bar{h}^{(j)} \bar{h}^{(j)} \star f^j)_{(k,n)}, \\ dH_{(k,l)}^{j+1} &= (\bar{h}^{(j)} \bar{g}^{(j)} \star f^j)_{(k,n)}, \\ dV_{(k,l)}^{j+1} &= (\bar{g}^{(j)} \bar{h}^{(j)} \star f^j)_{(k,n)}, \\ dD_{(k,l)}^{j+1} &= (\bar{h}^{(j)} \bar{g}^{(j)} \star f^j)_{(k,n)}. \end{aligned}$$

Los detalles dH, dV y dD representan, respectivamente, a los detalles horizontales, verticales y diagonales. Esta fórmula es la que se emplea para descomponer imágenes y “separar” aproximaciones y detalles como en (1).

Existe una relación importante entre la función de escala interpolatoria, $\varphi(t)$, y los esquemas de subdivisión de tipo interpolatorio [2], [5]. Por ejemplo, si se define $h_{or} = db2$ asociado a la función wavelet de Daubechies de longitud 4 (ver [11]), obtenemos el filtro interpolatorio

$$h = \frac{1}{2}(-1/16, 0, 9/16, 1, 9/16, 0, -1/16). \quad (6)$$

La sucesión

$$(-1/16, 0, 9/16, 1, 9/16, 0, -1/16)$$

define el esquema de subdivisión interpolatorio basado en 4 puntos [6], [7]. Existe una gran variedad de esquemas de subdivisión de tipo interpolatorio definidos mediante sucesiones susceptibles de ser empleadas en el algoritmo (4). Esta idea nos permitió ampliar la familia de filtros empleada por excelencia en este tipo de algoritmos que son la familia de filtros definidas tomando como funciones de escala B-splines de orden n , (see [16], p. 208). También hemos podido utilizar filtros dependientes de un parámetro en el algoritmo en paralelepípedo y definir un proceso de fusión más dinámico por la posibilidad de variar el valor del parámetro.

3. Etapas en la fusión de imágenes

En esta sección nos volvemos a situar en el punto de partida. Dos imágenes iniciales (PAN y MS) que han de fusionarse para definir una tercera (FUS). Las imágenes PAN recorren una longitud de onda entre $465 - 680$ nm y son percibidas por los humanos como una imagen gris. Por el contrario, las imágenes MS recorren tres bandas que son percibidas mediante tres colores distintos: Roja ($610 - 660$ nm), verde ($535 - 585$ nm) y azul ($430 - 490$ nm). Estas bandas se representan mediante el modelo de coordenadas RGB. Para las imágenes MS , las bandas RGB se denotarán mediante MS_R, MS_G, MS_B , respectivamente. Cada pixel toma un valor que se relaciona con la intensidad en cada banda. En nuestro caso la intensidad varía desde 0 a 256. La imagen fusionada FUS se obtiene mediante el siguiente procedimiento:

1. A partir de la imagen PAN se obtienen tres imágenes PAN . Cada imagen PAN tiene un histograma acorde con uno de los tres colores del histograma RGB. Se denotan mediante PAN_R, PAN_G y PAN_B , respectivamente.
2. La fusión se realiza entre cada par del mismo color:

$$(PAN_R, MS_R), (PAN_G, MS_G) \text{ y } (PAN_B, MS_B).$$

Como consecuencia, la imagen FUS se define mediante su representación RGB (FUS_R, FUS_G, FUS_B), donde FUS_R se obtiene fusionando PAN_R y MS_R . Análogamente se obtienen FUS_R y FUS_B .

3. MS_R y PAN_R corresponden al nivel $j = 0$ para aplicar el algoritmo en paralelepípedo. El procedimiento de descomposición se aplica desde $j = 0$ a $j = 2$ tanto para la imagen MS_R como para la imagen PAN_R . De este modo, obtenemos la aproximación $f_{MS_R}^2$ y los detalles $d_{PAN_R}^2$.
4. FUS_R se obtiene mediante $FUS_R = f_{MS_R}^2 + d_{PAN_R}^2$.
5. Los pasos 3 y 4 se llevan también a cabo para las bandas verdes y azul con idea de obtener FUS_R y FUS_B .
6. Finalmente, FUS image se define mediante coordenadas RGB como

$$FUS = (FUS_R, FUS_G, FUS_B).$$

Nota 1

- (a) El nivel $j = 2$ viene motivado a raíz de los resultados experimentales.
- (b) El uso de $FUS_R = f_{MS_R}^2 + d_{PAN_R}^2$ en lugar de $FUS_R = d_{MS_R}^2 + f_{PAN_R}^2$ se debe al hecho de que las imágenes PAN images preservan mejor los detalles geométricos, mientras que la información espacial que proporciona $d_{MS_R}^2$ es pobre.

4. Aplicaciones

En esta sección se trabaja un ejemplo concreto de fusión de dos imágenes iniciales que fueron tomadas en el centro de Granada. El motivo principal de las imágenes es el anillo circular que define la plaza de toros de la ciudad. La Figura 1 muestra la imagen original multiespectral MS con su representación RGB.

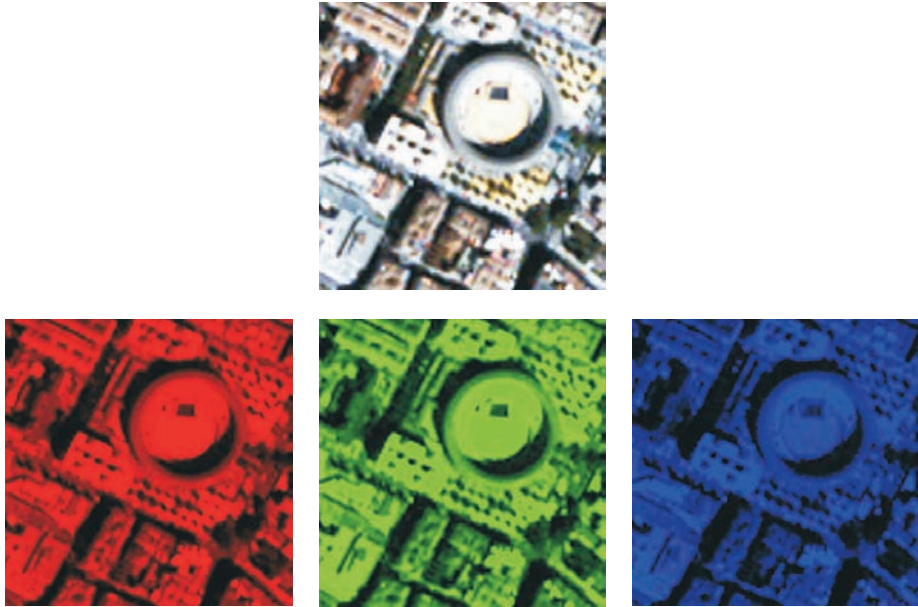


Figura 1: Arriba: Imagen multiespectral (MS). Abajo: Representación RGB (MS_R, MS_G, MS_B).

La Figura 2 corresponde a la imagen pancromática PAN . Las Figuras 3 y 4 representan la aproximación y los detalles obtenidos a partir de MS_R y PAN_R aplicando desde $j = 0$ a $j = 2$ el algoritmo de descomposición (4) en dimensión dos. Hemos utilizado el filtro interpolatorio (6) asociado con el esquema de subdivisión de 4 puntos. La Figura 5 muestra la representación o coordenada roja de la imagen fusionada.

Las representaciones verde y azul de la imagen fusionada, FUS_G and FUS_B , se obtienen de manera similar a FUS_R . Finalmente, la Figura 6 muestra la imagen fusionada completa FUS .

A partir de aquí hemos conseguido mejorar la calidad espectral y espacial de la imagen fusionada trabajando con filtros que dependen de un parámetro. El grado de libertad que supone disponer de este parámetro nos ha permitido resolver cierto problema de optimización y, de este modo, encontrar el valor del



Figura 2: Imagen pancromática (*PAN*).

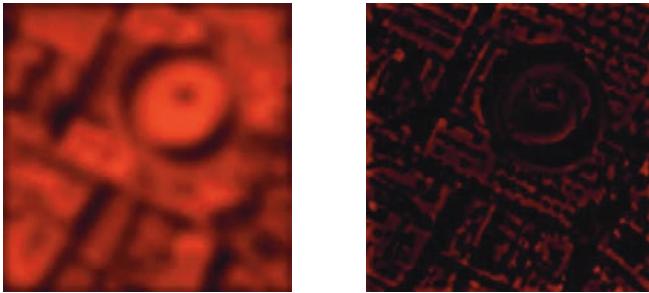


Figura 3: Izquierda: Aproximación f_{MSR}^2 . Derecha: Detalles d_{MSR}^2 .

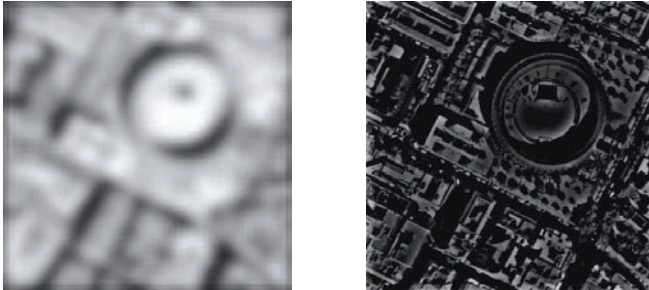


Figura 4: Izquierda: Aproximación f_{PANR}^2 . Derecha: Detalles d_{PANR}^2 .

parámetro que lo resuelve. La fusión con el parámetro óptimo nos ofrece medidas de calidad espacial y espectral realmente buenas. Por otro lado, la introducción de una transformada en paralelepípedo no lineal con la que realizar fusión nos ha permitido mejorar la fusión lineal con filtros interpolatorios. Más detalles de estas mejoras pueden consultarse en [12] y [13].

A la imagen fusionada se le aplicaron medidas de calidad (calidad universal,



Figura 5: Representation en rojo de la imagen fusionada $FUS_R = f_{MSR}^2 + d_{PANR}^2$.



Figura 6: Imagen fusionada FUS obtenida mediante el algoritmo de multirresolución.

ERGAS, etc. [15], [17]), que muestran una muy buena resolución tanto espacial como espectral y que confirmaron lo que nuestra propia visión nos indicaba en cuanto a la buena calidad de la imagen fusionada. Incluso podemos observar con precisión un escenario en el ruedo, indicador de que aquella noche hubo concierto en la ciudad del *embrujo*. Pero esto es otra historia.

Referencias

- [1] K. Amolins, Y. Zhang, P. Dare, Wavelet based image fusion techniques - An introduction, review and comparison. *Journal of Photogrammetry and Remote Sensing* 62, 4, 249-324 (2007).
- [2] K. Berkner, R. O. Wells, Jr., Smoothness Estimates for Soft-Threshold Denoising via Translation-Invariant Wavelet Transforms, *Appl. Comput. Harmon. Anal.* 12, 1-24 (2002).
- [3] P.S. Chavez, S.C. Sides and J. A. Anderson, Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic, *Photogram. Eng. Remote Sensing* 57, 3, 295-303 (1991).
- [4] D. Donoho, Interpolating wavelet transforms. Preprint, Department of Statistics, Stanford University, 1992.
- [5] G. Deslaurier and S. Dubuc, Symmetric iterative interpolation processes, *Constr. Approx.* vol.5, 49-68 (1989).
- [6] N. Dyn, *Interpolatory Subdivision Schemes*, Tutorials on Multiresolution in Geometric Modelling, A. Iske, E. Quak, M. S. Floater (eds.), Springer-Verlag Berlin Heidelberg, (2002), pp. 25-45.
- [7] N. Dyn, D. Levin and A. Gregory, A 4-point interpolatory subdivision scheme for curve design, *Comput. Aided Geom. Design* 4, 4, 257-268 (1987).
- [8] M. Holschneider, R. Kronland-Martinet, J. Morlet, and P. Tchamitchian, *A real-time algorithm for signal analysis with the help of the wavelet transform*, in *Wavelets: Time-Frequency Methods and Phase-Space*. New York: Springer-Verlag, 1989, pp. 286-297.
- [9] M. Lillo, C. Gonzalo: Spectral or spatial quality for fused satellite imagery? A tradeoff solution using the wavelet à trous algorithm, *Int. J. Remote Sens.* 27(7) 1453-1464 (2006).
- [10] S. Li, J. T. Kwok, Y. Wang, Using the discrete wavelet frame transform to merge Landsat 29 TM and SPOT panchromatic images. *Information Fusion* 3,1, 17-23 (2002).
- [11] S. Mallat, *A wavelet tour of signal processing*. Academic Press, 1999.
- [12] M. Moncayo and J. F. Reinoso, Optimal quality for image fusion with interpolatory parametric filters. *Math. Comput. Simul.* (2011), doi:10.1016/j.matcom.2010.12.004.
- [13] M. Moncayo, J. F. Reinoso and M. Pasadas, Image fusion by using a class of nonlinear discrete multiresolution transform. *Enviado a Applied Numerical Mathematics* (2010).
- [14] M. J. Shensa, The Discrete Wavelet Transform: Wedding the À trous and Mallat algorithms, *IEEE Trans. Signal Proc.* 40, 10, 2464-2482 (1992).
- [15] L. Wald, Quality of high resolution synthesized images: Is there a simple criterion?' in *Proc. Int. Conf. Fusion of Earth Data 2000*, Jan, 99-105.
- [16] D. F. Walnut, *An introduction to Wavelet Analysis*, Birkhäuser Boston, 2004.
- [17] Z. Wang, A. C. Bovik, A Universal Image Quality Index, *IEEE Signal Process. Lett.* 9, 3, 81-84 (2002).