

Medida de la calidad espectral de imágenes fusionadas usando el algoritmo CIEDE2000

Dionisio Rodríguez-Esparragón⁽¹⁾, Ángel García-Pedrero⁽²⁾, Consuelo Gonzalo-Martín⁽²⁾, Javier Marcello-Ruiz⁽¹⁾ y Francisco Eugenio-González⁽¹⁾

⁽¹⁾ Instituto de Oceanografía y Cambio Global (IOCAG), Universidad de Las Palmas de Gran Canaria (ULPGC), Campus Universitario de Tafira, 35017 Las Palmas de Gran Canaria, España. Emails: dionisio.rodriguez@ulpgc.es, javier.marcello@ulpgc.es, francisco.eugenio@ulpgc.es

⁽²⁾ Centro de Tecnología Biomédica (CTB), Universidad Politécnica de Madrid (UPM), Campus de Montegancedo, Pozuelo de Alarcón, 28223 Madrid, España. Emails: am.garcia@alumnos.upm.es, consuelo.gonzalo@upm.es

Resumen: La fusión de imágenes juega un papel importante en las aplicaciones de teledetección en las que se necesita un análisis visual. Debido a esto, la evaluación de la calidad espectral de las imágenes fusionadas es un tema fundamental para optimizar y comparar el resultado de distintos algoritmos. El objetivo de este trabajo es explorar el uso de la distancia CIEDE2000 para evaluar la calidad espectral de las imágenes fusionadas. Para ello, se creó una base de datos que contiene imágenes de satélite y sus productos de fusión. La calidad espectral de las imágenes en la base de datos se evaluó utilizando tanto índices cuantitativos comunes como la distancia CIEDE2000. Los resultados se compararon con la evaluación cualitativa obtenida mediante la aplicación de un test de opinión media.

Palabras clave: fusión de imágenes, calidad de imagen, calidad espectral, indicadores de calidad, CIEDE2000

Assessment of the spectral quality of fused images using the CIEDE2000 algorithm

Abstract: Image fusion plays an important role in remote sensing applications where a visual analysis is needed. Because of this, the evaluation of the spectral quality of pan-sharpened images is a fundamental subject to optimize and compare the results of different algorithms. The aim of this paper is to explore the use of distance CIEDE2000 to evaluate the spectral quality of the fused images. To do this, a database containing remote sensing imagery and its fusion products was created. The spectral quality of the imagery on the database was evaluated using both common quantitative indices and CIEDE2000. The results were compared to the qualitative assessment provided by a mean opinion score test.

Keywords: image fusion, image quality, spectral quality, quality measures

1. INTRODUCCIÓN

La fusión de imágenes a nivel de píxel es un proceso de combinación de las características relevantes de dos o más imágenes en una sola minimizando la introducción de distorsión o pérdida de información (Yong *et al.*, 2010). En el ámbito de la teledetección, la fusión de imágenes (pan-sharpening) proporciona recursos para realizar una mejor interpretación visual de las imágenes, por medio de la combinación de la alta resolución espacial de una imagen pancromática (PAN) y de las características de la resolución espectral de la correspondiente multispectral (MS) (Lillo-Saavedra y Gonzalo, 2006).

En este aspecto, el color resultante en el proceso de pan-sharpening juega un papel importante para aplicaciones de análisis visuales, tales como la detección de cambios, la identificación de objetos, y vigilancia de los riesgos (Alimuddin *et al.*, 2012). Es bien sabido que el ojo humano puede percibir sólo unas pocas docenas de niveles de gris; sin embargo, es capaz de distinguir entre miles de colores. Por lo tanto, una buena representación del color permite distinguir entre diferentes cubiertas y, por consiguiente, mejorar el análisis visual de la imagen.

Varios índices de calidad se han desarrollado con el fin de evaluar la calidad espectral de las imágenes fusionadas, la mayoría de ellos basados en un producto escalar o definidos a partir de una distancia de Minkowski entre las imágenes MS y fusionada (FUS) (Wang y Bovik, 2006). Sin embargo, no toman en cuenta la respuesta del sistema visual humano, arrojando resultados que la mayoría de las veces no están de acuerdo con la percepción visual del usuario.

En los últimos años, muchos de los esfuerzos en el área de la colorimetría se han dedicado a desarrollar una métrica de evaluación objetiva de la calidad que pueda reflejar con precisión la calidad subjetiva de la percepción humana (Yang *et al.*, 2012). Como resultado de estos esfuerzos, se ha formulado la medida CIEDE2000 como una evaluación de la distancia entre dos colores (Yang *et al.*, 2012), basado en el espacio de color Lab, pero adaptado a la percepción del sistema visual humano. En este sentido, parece que CIEDE2000 podría ser una medida apropiada para comparar el resultado de algoritmos de fusión de imágenes de teledetección, en aplicaciones donde la fidelidad espectral es primordial.

El objetivo de este trabajo es analizar el uso de la distancia CIEDE2000 como medida para evaluar la calidad espectral de imágenes fusionadas. Para lograr

esto, la hemos comparado con dos índices clásicos y ampliamente usados de calidad espectral: *Spectral Angle Mapper* (SAM) y error cuadrático medio (RMSE) (Shi *et al.*, 2014, Sun *et al.*, 2014, Chikr *et al.*, 2014). Finalmente, hemos contrastado estas medidas cuantitativas con una medida cualitativa (subjetiva) de la calidad usando un Test de Opinión Media (MOS) (ITUR, 2011).

2. DATOS Y MÉTODOS

2.1. Base de Datos de Imágenes

La Base de Datos creada para los experimentos descritos en este trabajo contiene 4 imágenes reales que se pueden observar en la Figura 1. Estas imágenes han sido seleccionadas de forma que representan diferentes cubiertas terrestres: vegetación urbana, ciudad y costa. Todas ellas pertenecen a zonas de la isla de Gran Canaria (España) y fueron capturadas por los sensores del satélite GeoEye (Figura 1(a)), Ikonos (Figura 1(b)) y QuickBird (Figuras 1(c), 1(d)). Para cada una de estas imágenes se dispone de la imagen MS de cuatro bandas (rojo, verde, azul y el infrarrojo cercano) y la PAN.

Además, cada una de las imágenes de la base de datos ha sido fusionada con diferentes algoritmos (Lillo-Saavedra *et al.* 2011, Marcello *et al.* 2013): Brovey, análisis de componentes principales (PCA), intensidad, tono y saturación (IHS), IHS extendida (E-IHS), Mallat (TWD), algoritmo aditivo Á trous (AW), combinación Á trous con IHS (AW-I), con PCA (AW-PCA) y con mapas de dimensión fractal (AW-FR). Por lo tanto, se generan un conjunto de 9 imágenes fusionadas por cada imagen original. En la Figura 2 se pueden observar las fusiones obtenidas a partir de la imagen Ikonos original.

2.2. Evaluación de la calidad espectral

Tal y como se comentó en la introducción, si existe una imagen de referencia, la evaluación espectral de la calidad de imagen se realiza mediante el cálculo de algún tipo de distancia. Las distancias más utilizadas se basan en el producto escalar o cualquier métrica de Minkowski (ecuación 1). Es por ello que para el propósito de este trabajo, hemos seleccionado los índices SAM y RMSE como las evaluaciones de referencia a efectos de comparación. Además, para contrastar los resultados de estos indicadores con los obtenidos por la distancia CIEDE2000, se realizó un Test de Opinión Media (MOS) con las imágenes de la base de datos. Dado que trabajamos en un espacio de color visible, sólo se consideran las bandas R, G y B. En la ecuación 1, x_i e y_i denotan el píxel i de cada imagen x e y respectivamente, mientras que n es el número de píxeles y p es cualquier número entero que indica el orden de la distancia de Minkowski.

$$M = \left(\sum |x_i - y_i|^p \right)^{1/p}, p \in [0, \infty] \quad (1)$$



(a) Geoeye

(b) Ikonos



(c) QuickBird (QB16)

(d) QuickBird (QB24)

Figura 1. Imágenes de la Base de Datos

El RMSE (ecuación 2) aparece ampliamente en la literatura relativa al pan-sharpening para comparar dos imágenes como la suma de las diferencias de píxeles. A mayor valor, mayor es la distancia entre las imágenes, y por lo tanto el error. La ecuación 2 deriva la ecuación 1 cuando p es 2, es decir, la distancia de Minkowski orden 2.

$$RMSE = \sqrt{\frac{\sum (x_i - y_i)^2}{n}} \quad (2)$$

El SAM cuantifica el valor absoluto del ángulo espectral entre dos vectores. SAM se expresa generalmente en grados y un valor igual a 0 indica ausencia de distorsión espectral, lo que significa que ambos vectores de píxeles son paralelos pero, ya que pueden tener diferentes longitudes, es posible que exista distorsión radiométrica. SAM se puede escribir como aparece en la ecuación 3, donde de nuevo x_i e y_i denotan el píxel i de cada imagen x e y respectivamente, mientras que n es el número de píxeles.

$$SAM = \cos^{-1} \frac{\sum x_i y_i}{\sqrt{\sum x_i^2} \sqrt{\sum y_i^2}} \quad (3)$$

La distancia CIEDE2000 (también identificada por el acrónimo alemán DE) evalúa la diferencia de color entre dos píxeles. Su expresión matemática fue publicado por la CIE en 2001. DE se basa en una transformación del espacio de color *Lab* a un espacio llamado *LCh* (luminosidad, croma y tono) con el fin de considerar las no uniformidades del sistema de percepción visual humana. La ecuación 4 muestra su expresión. Donde K_L ,

K_C y K_H ponderan el entorno de observación. A su vez, los factores de ponderación de Luminosidad (S_L), croma (S_C) y tono (S_H), describen respectivamente la acción de la percepción visual sobre esos tres atributos. Por último, el elemento de rotación R_T se utiliza para corregir una desviación en la región azul que se ha descrito en la percepción visual (Luo *et al.*, 2001).

$$DE = \sqrt{\left(\frac{\Delta L}{k_L S_L}\right)^2 + \left(\frac{\Delta C}{k_C S_C}\right)^2 + \left(\frac{\Delta H}{k_H S_H}\right)^2 + R_T \left(\frac{\Delta C}{k_C S_C}\right)^2 \left(\frac{\Delta H}{k_H S_H}\right)^2} \quad (4)$$

Finalmente, se llevó a cabo el test MOS mediante la recopilación de las opiniones de 10 personas (6 hombres y 4 mujeres) proyectando las imágenes en un monitor ASUS alta definición, modelo MK241H.

Todos los evaluadores estaban sentados a la misma distancia (1 metro) y ángulo del monitor y con las mismas condiciones de iluminación. Para este propósito, se desarrolló una interfaz, que mostraba la imagen de referencia a la izquierda y la imagen bajo test a la derecha. Se utilizó proyección a pantalla completa con el fondo negro. Las imágenes de prueba relativas a cada imagen de referencia se proyectaron en secuencia aleatoria. En cuanto a la evaluación de la calidad, se adoptó el modelo de clasificación recomendada por la Unión Internacional de Telecomunicaciones (UIT). Los observadores anotaron la calidad espectral de las imágenes como excelente (*excellent*), buena (*good*), aceptable (*fair*), regular (*poor*) o mala (*bad*). Se asignó una escala de 5-1 puntos de manera que 5 puntos corresponden a la mejor puntuación (*excellent*) y 1 a la peor (*bad*). Así, la media y la moda de las observaciones se utilizaron para establecer la calidad espectral.

3. RESULTADOS Y DISCUSIÓN

La Figura 3 muestra los resultados de la prueba de MOS (sólo para una de las imágenes de la base de datos debido a restricciones de espacio).

Como cabría esperar, se puede observar que las mejores anotaciones de calidad espectral (excelente calificación) se producen en imágenes fusionadas obtenidas aplicando las técnicas basadas en la transformada wavelet (TWD, AW, AW-PCA, AW-I, y AW-FR).

Por el contrario, los peores resultados son obtenidos por los métodos basados en la combinación lineal de imágenes (Brovey, PCA, IHS, E-IHS). Si se analiza la moda de los resultados, las imágenes fusionadas con el algoritmo PCA presentan los peores resultados en todos los casos. En cambio, es más difícil decidir qué imagen es la mejor subjetivamente, excepto en el caso de la imagen Ikonos, donde sólo el algoritmo AW-IHS se califica como excelente.

Los resultados de los índices RMSE, SAM y DE, así como la moda y los promedios de la prueba de MOS se resumen en la Tabla 1. Las mejores y peores puntuaciones para cada imagen y método de fusión se resaltan en letra negra.

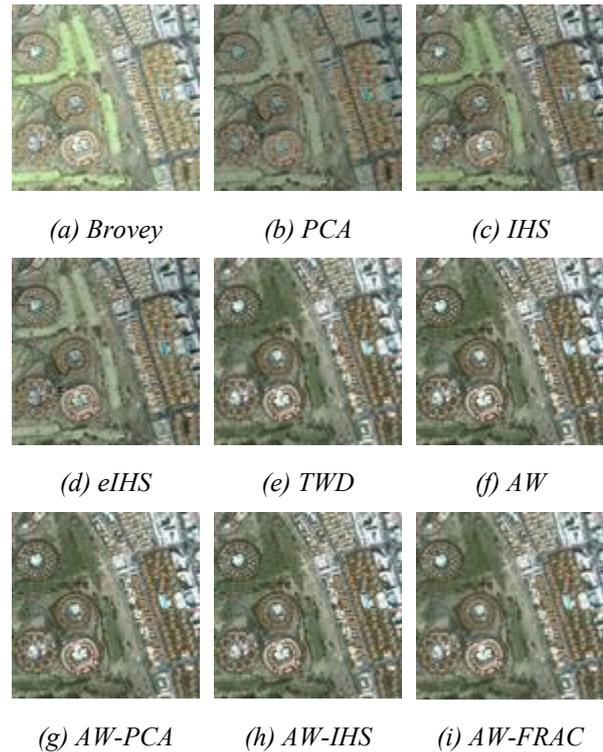


Figura 2. Fusiones correspondientes a las imagen de la Figura 1(b).

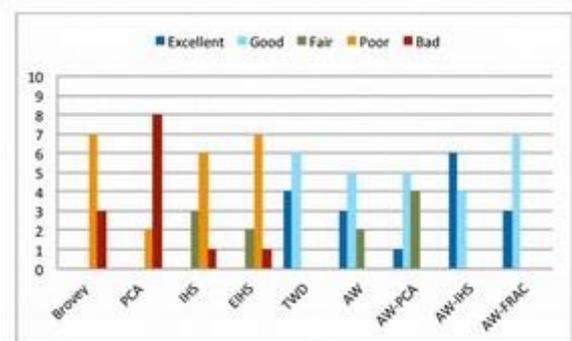


Figura 3. Resultados del test MOS para la imagen de la Figura 1(b).

Para las imágenes Ikonos y QB16, los resultados del RSME indican que las fusiones obtenidas por el algoritmo PCA producen las imágenes con menor diferencia de color con respecto a la MS, contrariamente al comportamiento esperado de este algoritmo. Sucede algo similar con el índice SAM, que, en este caso, califica a las imágenes obtenidas con el algoritmo IHS entre aquellas con mejor fidelidad espectral. Además, el índice de RMSE evalúa la imagen producida por el algoritmo de Mallat como la peor de las imágenes fusionadas en la secuencia QB16, mientras que la evaluación cualitativa la etiqueta como excelente. Estos resultados difieren de los percibidos subjetivamente. Por el contrario, las evaluaciones obtenidas con el índice DE corroboran las calificaciones buenas o excelentes de los participantes en la prueba MOS.

Tabla 1. Resultados de las medidas de moda y promedio del test MOS e índices RMSE, SAM y DE para la imagen de la Figura 1(b).

Algoritmo	Evaluación Subjetiva		Evaluación Objetiva		
	Moda	Media	RMSE	SAM	DE
Brovey	Pobre	1,7	0,38	0,02	18,83
PCA	Pobre	1,2	0,13	0,05	15,08
IHS	Acceptable	2,2	0,22	0,02	14,87
EIHS	Pobre	2,1	0,22	0,04	15,17
TWD	Buena	4,4	0,23	0,04	9,07
AW	Buena	4,1	0,16	0,02	8,55
AW-PCA	Buena	3,7	0,14	0,03	7,10
AW-IHS	Excelente	4,6	0,16	0,02	8,55
AW-FRAC	Buena	4,3	0,14	0,03	6,06

En la Tabla 2, se observan los resultados de la correlación entre los datos de los índices RMSE, SAM, y DE por separado con el promedio de los datos de la prueba MOS. El signo negativo indica que los índices RMSE, SAM, y DE son inversamente proporcionales al promedio MOS. Como se puede observar, la distancia DE produce la correlación más alta con la percepción visual.

Tabla 2. Correlaciones entre el promedio del test MOS y los índices RMSE, SAM y DE.

Método	Geor	Ikonos	QB16	QB24
Promedio MOS	1	1	1	1
RMSE	-0.795	-0.411	-0.077	-0.372
SAM	-0.502	-0.279	-0.653	-0.566
DE	-0.946	-0.905	-0.893	-0.809

4. CONCLUSIONES

Los resultados llevan a la conclusión de que las evaluaciones efectuadas con índices RMSE y SAM no coinciden con los obtenidos por la prueba MOS. Por otra parte, la correlación entre las medidas derivadas del índice DE y los proporcionados por la opinión subjetiva de calidad es más fuerte que la que se deriva de los otros dos índices.

Aunque este estudio no contempla todos los indicadores existentes, los resultados parecen concluir un rendimiento mejor de la distancia DE, esto es, una distancia que considere las características de la percepción humana frente a métodos derivados del producto escalar o distancia de Minkowski. Para la aplicación de este indicador a la evaluación del desempeño espectral de los algoritmos de fusión será necesario ampliar el campo de funcionamiento de esta técnica de medida a otros enfoques en los que no sólo es necesario el color sino también una comparación de múltiples (más de 3) bandas espectrales.

5. AGRADECIMIENTOS

Este trabajo ha sido apoyado por el proyecto ARTeMISat (CGL2013-46674-R), financiado por el Ministerio de Economía y Competitividad de España.

A. García-Pedrero agradece a CONACyT el apoyo proporcionado a través de la beca para estudios de doctorado número 216146.

6. BIBLIOGRAFÍA

- Alimuddin, I., *et al.* 2012. *Assessment of pan-sharpening methods applied to image fusion of remotely sensed multi-band data.* International Journal of Applied Earth Observation and Geoinformation, 18, 165-175.
- Chikr El-Mezouar, M., *et al.* 2014. *A Pan-Sharpener Based on the Non-Subsampled Contourlet Transform: Application to Worldview-2 Imagery.* IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7.5, 1806-1815.
- ITUR, "BT. 500-11, 2011. *Methodology for the subjective assessment of the quality of television pictures.* International Telecommunication Union.
- Lillo-Saavedra, M., *et al.* 2011. *Toward reduction of artifacts in fused images.* International Journal of Applied Earth Observation and Geoinformation, 13.3, 368-375.
- Lillo-Saavedra, M., y Gonzalo, C. 2013. *Spectral or spatial quality for fused satellite imagery? A trade-off solution using the wavelet à trous algorithm.* International Journal of Remote Sensing, 27.7, 1453-1464.
- Luo, M. R., *et al.* 2001. *The development of the CIE 2000 colour-difference formula: CIEDE2000.* Color Research & Application, 26.5, 340-350.
- Marcello, J., *et al.* 2013. *Evaluation of spatial and spectral effectiveness of pixel-level fusion techniques.* Geoscience and Remote Sensing Letters, IEEE, 10.3, 432-436.
- Shi, Y., *et al.* 2014. *Pan-sharpening of multispectral images using the nonseparable framelet lifting transform with high vanishing moments.* Information Fusion, 20, 213-224.
- Sun, W., *et al.* 2014. *Nearest-neighbor diffusion-based pan-sharpening algorithm for spectral images.* Optical Engineering, 53.1, 013107-013107.
- Wang, Z., y A.C. Bovik. 2006. *Modern image quality assessment.* Synthesis Lectures on Image, Video, and Multimedia Processing, 2.1, 1-156.
- Yang, Y., *et al.* 2012. *Color image quality assessment based on CIEDE2000.* Advances in Multimedia, 11.
- Yong, Y., *et al.* 2010. *Medical image fusion via an effective wavelet-based approach.* EURASIP Journal on Advances in Signal Processing 2010, 1, 579341.