

Super-resolución de imágenes multiespectrales e híperespectrales usando DWT con SWT y representación dispersa

Yeredith G. Mora-Martínez, Beatriz P. García-Salgado,
Clara Cruz-Ramos, Osvaldo López-García,
Volodymyr Ponomaryov, Rogelio Reyes-Reyes,

Instituto Politécnico Nacional,
México

{ymoram1500, bgarcias1404}@alumno.ipn.mx
{ccruzra, olopezg, vponomar, rreyesre}@ipn.mx

Resumen. Las imágenes multiespectrales (MS) e híperespectrales (HS) muestran representaciones en diferentes longitudes de onda de una misma escena que se utilizan en una amplia gama de aplicaciones, como la agricultura, medicina, astronomía y vigilancia, entre otros. Sin embargo, este tipo de imágenes suelen sufrir degradaciones provocadas por la baja resolución espacial del sensor que las registra. En este artículo se propone un método de superresolución de imágenes MS e HS para escenas urbanas a partir de una imagen de baja resolución (LR). Este método crea diccionarios de una base de datos de escenas naturales mediante k-Singular Value Decomposition (k-SVD) usando las sub-bandas de detalle de la Transformada Discreta Wavelet (DWT). Después, los diccionarios son empleados para reconstruir la imagen de entrada. Posteriormente, se realiza una interpolación de las sub-bandas de detalle de la DWT y la Transformada Estacionaria Wavelet (SWT) de la imagen de entrada junto con una reconstrucción de sus bordes. Finalmente, se utiliza la imagen reconstruida a partir de los diccionarios como la sub-banda de aproximación para el cálculo de la Transformada Discreta Wavelet Inversa (IDWT). El método propuesto se comparó con otros métodos similares del estado del arte en términos de Relación Media Señal-Ruido Pico Promedio (MPSNR), Índice de Similitud Estructural Promedio (MSSIM), Índice de Preservación de Bordes (EPI) y Mapeo de Ángulo Espectral (SAM) para el caso de las imágenes HS, demostrando una competitiva preservación de bordes al aumentar la resolución de las imágenes de prueba.

Palabras clave: Imágenes multiespectrales, imágenes híperespectrales, telede-tección, superresolución, representación dispersa, DWT, SWT.

Super-Resolution of Multispectral Images and Hyperspectral Using DWT with SWT and Scattered Representation

Abstract. Multispectral (MS) and hyperspectral (HS) images show different wavelength representations of a single scene employed in various applications,

such as agriculture, medicine, astronomy and surveillance, among others. Nevertheless, these kinds of images are prone to degenerate due to the sensor's low resolution that acquires them. This article proposes a superresolution method for urban MS and HS images using a low-resolution image (LR). First, the method creates dictionaries from a natural scenes database by k-Singular Value Decomposition (k-SVD) using the detail sub-bands of the Discrete Wavelet Transform (DWT). Then, the dictionaries are employed to reconstruct the input image. Subsequently, an interpolation of DWT and Stationary Wavelet Transform (SWT) sub-bands of the input image is performed in combination with edges' reconstruction. Finally, the image reconstructed by dictionaries is utilized as the approximation sub-band in the Inverse Wavelet Transform (IDWT) computation. The proposed method was compared with other similar state-of-the-art methods according to Mean Peak Signal-To-Noise Ratio (MPSNR), Mean Structural Similarity Index (MSSIM), Edge Preservation Index (EPI) and Spectral Angle Mapper (SAM) for the case of HS images, demonstrating competitive edge preservation while increasing the test images' resolution.

Keywords: Remote sensing, multispectral images, hyperspectral images, super-resolution, sparse representation, DWT, SWT.

1. Introducción

Los sensores de imágenes HS recopilan alrededor de 200 imágenes con un rango espectral que abarca las regiones de longitud de onda visible e infrarroja (400-2500 nm). Por otra parte, las imágenes MS son obtenidas con menor número de bandas abordando la longitud de onda visible y cercano infrarrojo. A pesar de las imágenes HS logran una alta resolución espectral, su resolución espacial es limitada.

Esto es debido a que los sistemas de teledetección necesitan varias exposiciones para la adquisición simultánea de varias bandas para que se garantice una relación señal/ruido (SNR) suficiente, lo que resulta en el sacrificio de la resolución espacial. Mientras que los sensores de imágenes MS capturan imágenes RGB con una resolución varias veces mayor que una imagen híper-espectral y con mayor SNR con el mismo tiempo de exposición.

Sin embargo, el registro de numerosas bandas espectrales de una misma escena con una alta resolución espacial es importante para las aplicaciones de teledetección [2] como la segmentación, seguimiento y reconocimiento de objetos [21,17,22]. En este artículo se presenta un método de superresolución mediante el uso de diccionarios entrenados con k-SVD [1], para reconstruir una imagen MS e HS a partir de imágenes RGB de alta resolución (HR). Las principales contribuciones de este trabajo son las siguientes:

Primero, se propone un algoritmo de multi-diccionarios donde su salida se utiliza como una nueva sub-banda de aproximación (LL) que puede ser utilizada tanto para imágenes HS como MS. Adicionalmente, se elabora una fusión de sub-bandas de detalle entre la DWT y SWT para mantener la información de alta frecuencia (HF) y mejorar la calidad de los bordes de la imagen. De este modo, se combinan dos enfoques utilizados en el estado del arte: aprendizaje de diccionarios y uso de transformaciones de dominio.

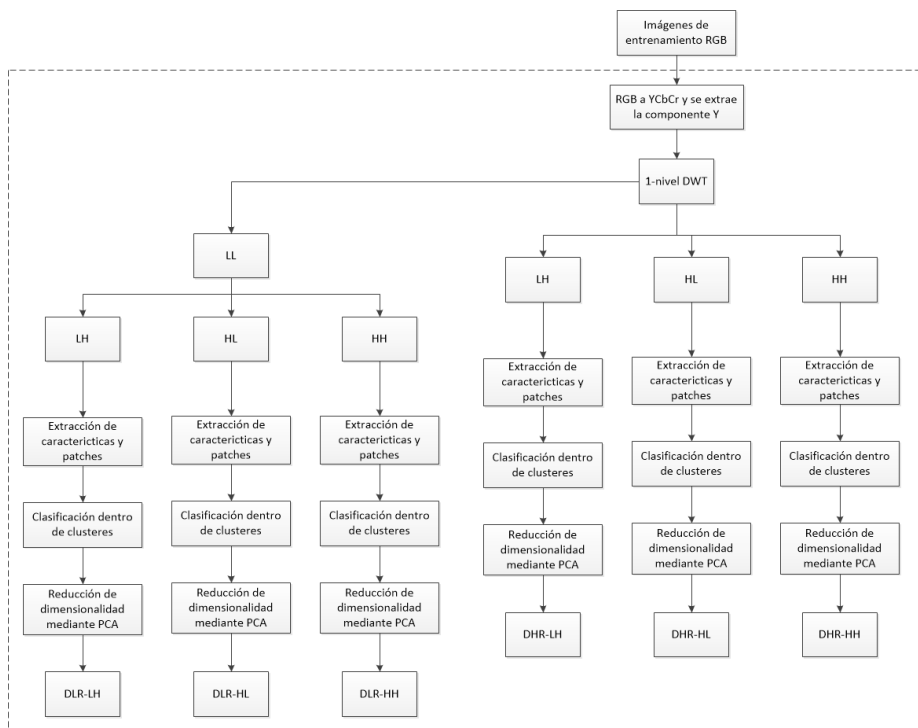


Fig. 1. Aprendizaje de sub-diccionarios en el dominio wavelet.

Para presentar el método propuesto, este documento está organizado de la siguiente forma: los trabajos relacionados con los algoritmos utilizados se presentan en la sección 2. Después, en la sección 3, se describen los pasos del método propuesto. En la sección 4 se establecen las condiciones de los experimentos y discusiones de los resultados. Finalmente, se presentan las conclusiones.

2. Trabajos relacionados

La superresolución puede abordarse de las siguientes formas: métodos basados en interpolación, métodos basados en reconstrucción y métodos basados en aprendizaje. Los métodos basados en interpolación [9,14,18,20], como la interpolación por el vecino más cercano (INN por sus siglas en ingles), interpolación bilineal e interpolación bicúbica, no ocupan muchos recursos computacionales, sin embargo, tienden a perder detalles en la imagen y dejar un efecto mosaico en ella.

Los métodos basados en reconstrucción [8,11,16], por lo general, asumen algunos modelos degradados y resuelven problemas inversos en la obtención de imágenes de HR. Este tipo de método funciona bien cuando el factor de ampliación es pequeño, pero su rendimiento requiere que la versión suavizada y muestreada de la imagen HR sea cercana a la imagen LR. Por último, los métodos basados en el aprendizaje [2,6,10,24] asumen que los detalles de alta frecuencia (HF) perdidos en la imagen LR pueden recuperarse aprendiendo las relaciones entre los parches de imágenes LR-HR.

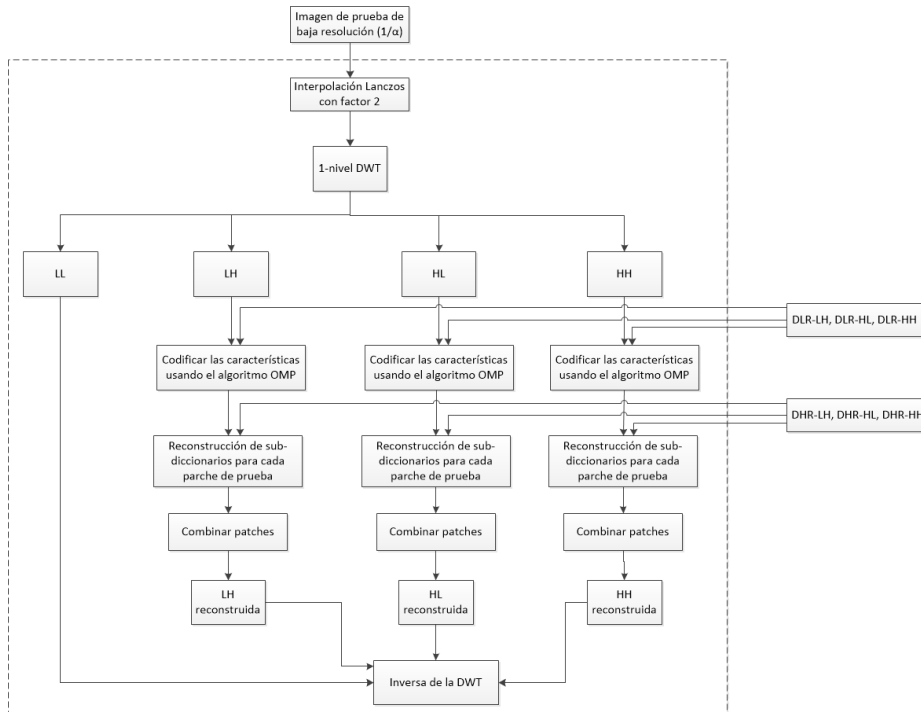


Fig. 2. Reconstrucción de la imagen.

Algunos ejemplos de este tipo de algoritmos se presentan a continuación. Alvarez-Ramos et al. [2] formulan el sistema Super-Resolution Via Wavelet Feature Extraction and Sparse Representation (SR-WAFE-SR) basado en la extracción de características wavelet y representación dispersa, donde utilizan un análisis de componentes principales (PCA) en cada sub-banda de detalle, reduciendo la dimensión de la información para obtener mejores resultados en el proceso de aprendizaje de diccionarios.

Wu et al. [24] proponen el algoritmo Wavelet Domain Multidictionary Learning (WDMML-SR) que realiza un aprendizaje multi-diccionario de las sub-bandas wavelet de una imagen de baja resolución mediante un PCA. De igual forma aplican una retroproyección iterativa (IBP) y la autosimilitud no local (NLM) para evitar perder los detalles en la superresolución. Por último, Dong et al. [10] plantean el método Non-Negative Structures Sparse Representation (NSSR) donde realiza la superresolución de imágenes LR a partir de un diccionario de imágenes y una imagen estimada RGB tomada de la misma escena.

Existen otros métodos que implementan transformadas en el dominio de la frecuencia para mejorar la reconstrucción de la imagen. Chavez-Roman et al. [7] proponen el método Super Resolution Wavelet Domain Interpolation Edge Extraction Sparse Representation (SR-WDIEE-SR) donde se interpolan las sub-bandas HF de la imagen en el dominio wavelet, junto con una representación dispersa para lograr una imagen con bordes más nítidos.

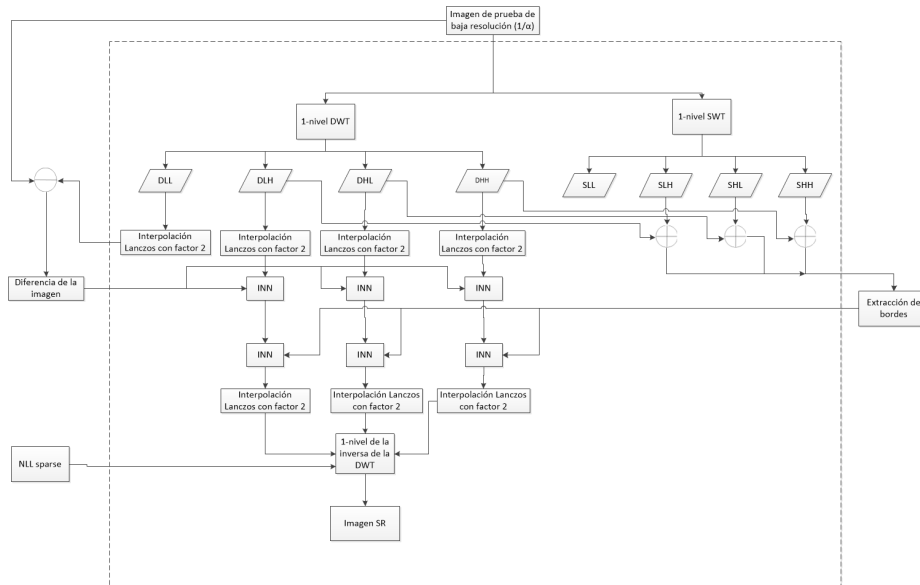


Fig. 3. Mejoramiento de bordes y obtención de imagen SR.

Por su parte, Feng et al. [11] plantean una detección de características de bordes basada en la descomposición y reconstrucción mediante la DWT tomando sólo las sub-bandas de detalle junto con una estimación de distribución de energía.

Estos enfoques mejoran la calidad visual de la imagen, sin embargo, no brindan resultados comparables con los métodos basados en aprendizaje [2,6,10,24]. Considerando esto, el método propuesto se basa en el aprendizaje de diccionarios junto con transformaciones en el dominio wavelet.

Su objetivo consta de recuperar los detalles de la imagen reconstruyendo la información sobre las sub-bandas con un diccionario y un algoritmo de preservación de bordes por medio de las transformaciones.

3. Método propuesto

En esta sección, se describe el proceso del algoritmo propuesto; y sus respectivos diagramas. El algoritmo consta de una fase de entrenamiento, una fase de reconstrucción y una etapa de mejora de bordes.

3.1. Aprendizaje de sub-diccionarios en el dominio wavelet

El aprendizaje de diccionarios, como se muestra en la Fig. 1 consta de encontrar diccionarios de las sub-bandas wavelet por medio de sub-bandas $y^{(i)}$, $i \in \{LH, HL, HH\}$ de imágenes HR, donde LL es la sub-banda de aproximación y LH, HL , junto con HH las de detalle. Primero, se convierten todas las imágenes de entrenamiento al espacio de color $YCbCr$ y se toma el canal de luminancia Y al ser la componente más sensible al sistema visual humano.

Tabla 1. Comparación de resultados cuantitativos de Pavia Center, Pavia University y Washington DC Mall.

Imagen		INN	Bilinear	Bicubic	SR-WDIEE-SR	WDML-SR	NSSR	Propuesto
		(1/4)	(1/4)	(1/4)	[8] (1/4)	[24] (1/4)	[10] (1/4)	(1/4)
Pavia Center	MPSNR	27.05	27.72	28.26	27.4	28.45	32.57	30.03
	MSSIM	0.92	0.91	0.93	0.91	0.94	0.98	0.97
	SAM	10.3	9.52	8.96	9.34	8.73	7.12	7.21
	EPI	0.06	0.19	0.22	0.19	0.22	0.5	0.41
	RMSE	11.61	10.73	10.07	10.87	9.63	5.99	8.13
Pavia University	MPSNR	26.25	27.04	27.41	27.91	28.25	30.33	30.26
	MSSIM	0.9	0.89	0.91	0.9	0.93	0.97	0.98
	SAM	8.18	7.45	7.13	6.11	6.47	5.36	5.09
	EPI	0.09	0.24	0.26	0.26	0.26	0.51	0.56
	RMSE	12.56	11.48	11.01	10.25	9.63	7.76	7.85
Washington DC Mall	MPSNR	43.08	43.64	44.04	42.98	44.27	42.6	45.62
	MSSIM	0.96	0.96	0.96	0.95	0.97	0.94	0.98
	SAM	3.53	3.53	3.37	3.37	3.39	3.97	2.72
	EPI	0.08	0.22	0.24	0.23	0.24	0.23	0.37
	RMSE	4.1	3.84	3.67	1.8	1.55	1.89	1.33

Luego, se aplica el primer nivel de la DWT en 2D a las imágenes para separarlas en sus sub-bandas obteniendo el conjunto $y^{(h)}$, $h \in \{LH, HL, HH\}$ como en [7,11,24] donde los autores probaron diferentes niveles de descomposición de la DWT siendo el primer nivel de descomposición el que mejor funcionó.

Seguidamente, a la representación LL se le aplica de nuevo la DWT para formar el conjunto $y^{(l)}$, $l \in \{LH, HL, HH\}$. El entrenamiento de los diccionarios consta de extraer parches superpuestos locales, con tamaño de $\sqrt{n} \times \sqrt{n}$ donde $n = 25$, de las sub-bandas de ambos conjuntos por separado. Estos parches se concatenan en un vector para cada banda definido como $p^{(j)}$, $j \in \{LH, HL, HH\}$.

Siguiendo la idea de Dong et al. [10], se clasifican los parches de entrenamiento en k -clústeres para cada sub-banda de los conjuntos $y^{(h)}$ y $y^{(l)}$. Posteriormente, se aplica el algoritmo k -SVD con el fin de construir un diccionario sobre-completo utilizando la siguiente función:

$$\min_{D^{(j)}, \{\gamma^{(j)}\}} \left\| p^{(j)} - D^{(j)} \gamma^{(j)} \right\|^2 \text{ donde } \left\| \gamma^{(j)} \right\|_0 \leq L. \quad (1)$$

Siendo L el nivel máximo de dispersión y $\gamma^{(j)}$ un vector de coeficientes de representación dispersa de modo que:

$$p^{(j)} = D^{(j)} \gamma^{(j)T} \left(\gamma^{(j)} \gamma^{(j)T} \right)^{-1}. \quad (2)$$

Finalmente, se aplica un análisis de componentes principales (PCA) [19], como método de reducción de dimensión a los parches extraídos para cada sub-banda.

3.2. Reconstrucción de la imagen

En la fase de reconstrucción se amplían las imágenes MS y HS, las cuales se consideran que originalmente están sub-muestreadas con un factor de $(1/\alpha)$, siendo α igual a 4.

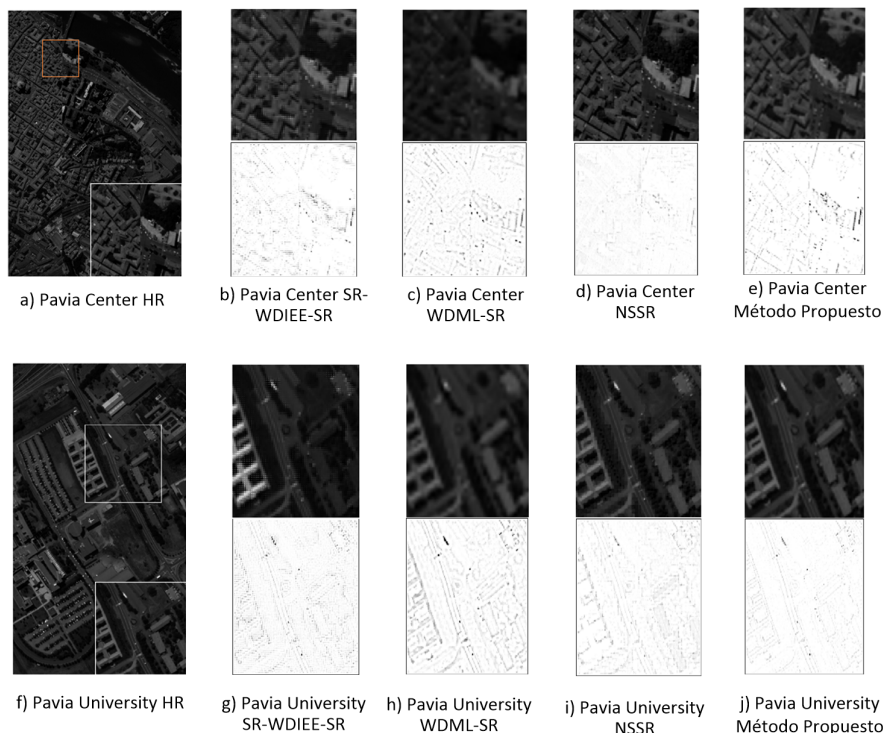


Fig. 4. Resultados visuales de Pavia Center y Pavia University.

Luego se aplica una interpolación Lanczos con factor 2 y se descompone la imagen en sub-bandas mediante la DWT. Después se toman las sub-bandas de detalle (LH , HL y HH) y se utilizan los diccionarios generados en la etapa de aprendizaje de sub-diccionarios, como se muestra en la Fig. 2, con el fin de reconstruir cada sub-banda. Consecutivamente, el algoritmo Orthogonal Matching Pursuit (OMP) [12] se aplica a los parches para encontrar los vectores de representación dispersa utilizando los diccionarios de las sub-bandas de detalle. Los vectores de representación dispersa se calculan como:

$$\gamma^{(j)} = \arg \min_{\gamma^{(j)}} \left\| p^{(j)} - D^{(j)} \gamma^{(j)} \right\|^2 \text{ donde } \left\| \gamma^{(j)} \right\|_0 \leq L. \quad (3)$$

Estos vectores son multiplicados por el diccionario $D^{(h)}$, que es resultante del conjunto $y^{(h)}$, reconstruyendo los parches de las sub-bandas mediante:

$$p^{(h)} = D^{(h)} \gamma^{(h)}. \quad (4)$$

Después, los parches de y_h se fusionan a través de un promedio del área superpuesta para crear las imágenes de sub-banda denominadas NLH , NHL , NHH . La imagen de salida de mediana resolución (NLL sparse) se reconstruye realizando IDWT con las sub-bandas reconstruidas NLH , NHL , NHH y la sub-banda de aproximación original LL .

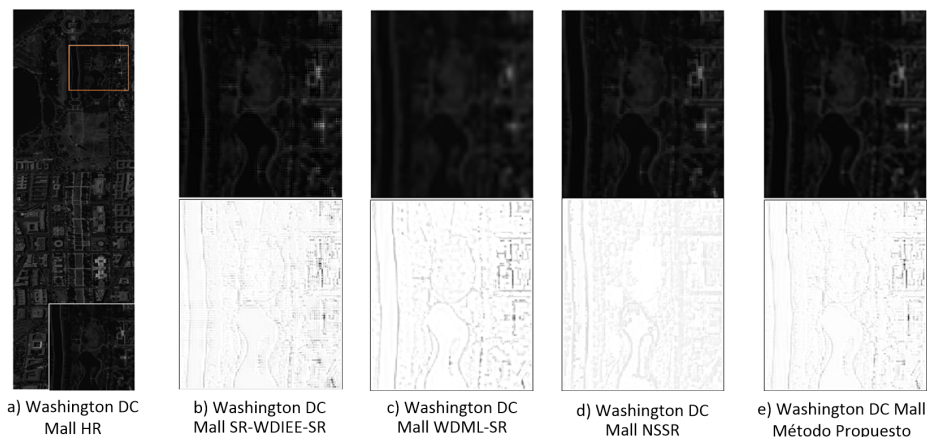


Fig. 5. Resultados visuales de Washington DC Mall.

3.3. Mejoramiento de bordes y obtención de imagen SR

La fase de mejoramiento de bordes toma una imagen LR y la descompone con el primer nivel de la DWT en 2D obteniendo las representaciones DHH , DHL , DLH y DLL . Se realiza el mismo procedimiento a la imagen LR pero con la SWT generando SHH , SHL , SLH y SLL .

Se aplica una interpolación Lanczos con factor 2 a las sub-bandas generadas por la DWT. Luego, se obtiene la diferencia entre DLL y la imagen LR para ajustar las bajas frecuencias. Posteriormente, las sub-bandas de detalles DHL , DLH y DHH son sumadas y normalizadas con las sub-bandas SHL , SLH y SHH .

Seguidamente, se extraen los bordes con estas sub-bandas de alta frecuencia de acuerdo con Feng et al. [11] mediante la siguiente ecuación:

$$S = \sqrt{(HH)^2 + (HL)^2 + (LH)^2}. \quad (5)$$

Por su parte, cada una de las sub-bandas DHL , DLH y DHH se interpola acorde con Chavez-Roman et al. [8] usando la imagen de diferencia e INN y las sub-bandas obtenidas se vuelven a interpolar con INN utilizando la imagen S . Después, cada una de estas sub-bandas es interpolada de nuevo por medio de lanczos con factor 2. Finalmente, las sub-bandas resultantes se emplean en junto con la imagen NLL sparse en el cómputo de la IDWT. La salida obtenida de este proceso resulta en una imagen de superresolución (SR) como se muestra en la Fig. 3.

4. Resultados

4.1. Métricas de evaluación

Para evaluar la calidad espacial de la reconstrucción de las imágenes SR obtenidas de las imágenes HS y MS se tomaron en consideración la métrica MPSNR que es la

Tabla 2. Comparación de resultados cuantitativos de la base de datos CAVE.

Imagen		INN (1/4)	Bilinear (1/4)	Bicubic (1/4)	SR-WDIEE-SR [8] (1/4)	WDML-SR [24] (1/4)	NSSR [10] (1/4)	Propuesto (1/4)
Balloons	MPSNR	35.74	38.39	39.37	38.99	39.33	41.88	42.93
	MSSIM	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	SAM	4.16	3.3	3.09	3.23	3.23	2.63	1.81
	EPI	0.12	0.38	0.4	0.33	0.4	0.58	0.64
	RMSE	4.21	3.11	2.78	2.86	2.75	2.05	1.81
Beads	MPSNR	24.9	26.05	26.9	27.94	26.84	38.4	31.15
	MSSIM	0.95	0.94	0.96	0.97	0.96	0.99	0.99
	SAM	17.83	15.52	14.31	13.96	14.41	3.52	8.84
	EPI	0.05	0.2	0.24	0.21	0.24	0.87	0.6
	RMSE	14.82	12.95	11.76	10.21	11.6	3.06	7.06
Cloth	MPSNR	25.93	26.24	26.69	29.63	26.64	38.67	32.2
	MSSIM	0.91	0.88	0.91	0.94	0.91	0.99	0.98
	SAM	16.49	16.77	15.04	14.66	15.12	2.94	7.95
	EPI	0.09	0.19	0.22	0.25	0.22	0.9	0.64
	RMSE	13.16	12.72	12.1	8.41	11.87	2.97	6.25
Face	MPSNR	35.02	36.77	37.69	35.57	37.63	44.79	42.47
	MSSIM	0.98	0.99	0.99	0.99	0.99	0.99	0.99
	SAM	10.05	8.89	8.2	8.83	8.29	3.32	4.1
	EPI	0.1	0.32	0.34	0.31	0.34	0.78	0.58
	RMSE	4.62	3.78	3.41	4.24	3.349	1.46	1.91
Fake and real peppers	MPSNR	33.62	35.71	37.08	35.34	36.97	43.81	38.93
	MSSIM	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	SAM	8.81	7.5	6.87	7.46	6.94	3.46	5.07
	EPI	0.06	0.3	0.36	0.35	0.36	0.71	0.65
	RMSE	5.4	4.29	3.7	4.36	3.61	1.64	2.88
Hairs	MPSNR	32.34	33.34	34.18	38.72	34.09	40.65	38.33
	MSSIM	0.98	0.97	0.98	0.99	0.98	0.99	0.99
	SAM	11.71	11.06	10.4	7.12	10.47	4.75	6.96
	EPI	0.17	0.24	0.29	0.54	0.29	0.79	0.53
	RMSE	6.26	5.6	5.1	2.95	5.03	2.36	3.09

Relación Media Señal-Ruido Pico (PSNR) promediada entre las bandas de diferente longitud de onda de las imágenes, MSSIM que indica el Índice de Similitud Estructural Promedio (SSIM) medio entre las bandas, ambos señalados en [15], el Índice de Preservación de Bordos (EPI) promediado mostrado en [13] y la Raíz del Error Cuadrático Medio (RMSE) promediado.

En el caso de las imágenes HS, también se utilizó el Mapeo del Ángulo Espectral (SAM) [15] para evaluar la calidad de la reconstrucción espectral. Cabe mencionar que MPSNR, MSSIM y EPI altos significan una mejor calidad visual mientras que valores de SAM bajos implican menor distorsión espectral y una mayor calidad de reconstrucción espectral.

4.2. Descripción de datos de prueba utilizados

Para el entrenamiento de los sub-diccionarios se utilizó un total de 91 imágenes RGB [5]. Para evaluar el método propuesto, se realizaron experimentos con el siguiente conjunto de datos de imágenes HS y MS.

1. Pavia Center : contiene 102 bandas espectrales con un tamaño de 1096×715 píxeles,
2. Pavia University [3]: contiene 103 bandas espectrales con un tamaño de 610×340 píxeles,

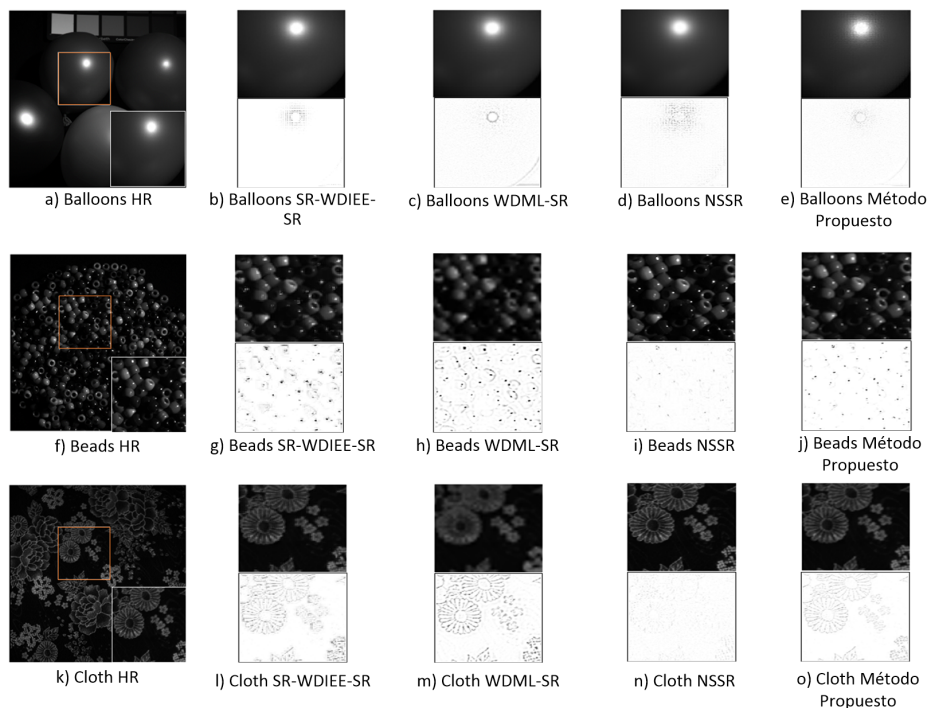


Fig. 6. Resultados visuales de las imágenes Balloons, Beads y Cloth (CAVE).

3. Washington DC Mall [4]: contiene 191 bandas espectrales con un tamaño de 1280×307 .
4. CAVE [25]: contiene 31 bandas espectrales con un tamaño de 512×512 píxeles donde se toman 6 datos de prueba,
5. Zurich [23]: contiene las bandas R,G,B y NIR para la experimentación se toman 4 imágenes las cuales son zh2 con tamaño de 833×881 , zh3 con tamaño de 926×943 , zh6 con tamaño de 812×984 y zh19 con tamaño de 729×1223 .

Los parámetros utilizados en el diccionario del método propuesto son los siguientes: el número de átomos para cada diccionario $D^{(j)}$ obtenido de los conjuntos $y^{(h)}$ y $y^{(l)}$ es 25×1024 , el número máximo de iteraciones son $t = 100$, con $\eta = 1e^{-3}$ siendo η un parámetro de regularización como en [2], para el entrenamiento del diccionario y reconstrucción de la imagen se ocupó la plataforma de programación MATLAB versión R2021a.

4.3. Resultados cuantitativos y cualitativos

En esta sección se presentan los resultados de los experimentos realizados. El método propuesto representa la combinación de los enfoques de aprendizaje de diccionarios por k-SVD, OMP y la transformación en el dominio a través de DWT y SWT. Por lo que, adicionalmente a la comparación con los métodos clásicos de

Tabla 3. Comparación de resultados cuantitativos de la base de datos Zurich.

Imagen		INN	Bilinear	Bicubic	SR-WDIEE-SR	WDML-SR	Propuesto
		(1/4)	(1/4)	(1/4)	[8] (1/4)	[24] (1/4)	(1/4)
Zh2	MPSNR	19.35	20.3	20.24	23.81	24.18	24.83
	MSSIM	0.78	0.79	0.8	0.95	0.92	0.96
	EPI	0.07	0.25	0.27	0.29	0.34	0.38
	RMSE	27.66	24.81	24.98	16.44	15.75	14.62
Zh3	MPSNR	24.75	25.99	26.39	27.58	27.97	29.56
	MSSIM	0.93	0.93	0.94	0.98	0.96	0.98
	EPI	0.09	0.27	0.31	0.31	0.34	0.55
	RMSE	14.94	12.98	12.41	10.65	10.18	8.48
Zh6	MPSNR	24	25.08	25.97	25.33	25.97	28.04
	MSSIM	0.93	0.92	0.94	0.98	0.94	0.99
	EPI	0.09	0.3	0.33	0.31	0.33	0.7
	RMSE	16.08	14.2	12.82	13.8	12.82	10.1
Zh19	MPSNR	20.07	21.13	21.28	23.23	23.8	24.55
	MSSIM	0.82	0.82	0.84	0.94	0.91	0.95
	EPI	0.07	0.25	0.28	0.3	0.32	0.37
	RMSE	25.29	22.38	22	17.58	16.46	15.1

interpolación (INN, Bilineal y Bicúbica), se contrastó su rendimiento con los algoritmos basados en aprendizaje NSSR [10] y WDML-SR [24], al igual que con el método SR-WDIEE-SR [8] que está fundamentado en las transformaciones de dominio. A cada imagen de prueba se le aplicó un sub-muestreo de 4 veces simulando una imagen LR para utilizarse como entrada de los algoritmos. La evaluación cualitativa de los métodos se realizó por medio de una imagen de error entre la imagen original Y_i y la imagen reconstruida $F(x)_i$ a través de la siguiente ecuación:

$$I_e = l_i - c |Y_i - F(x)_i|, \tag{6}$$

donde i corresponde a la i -ésima banda, l_i es el rango del valor máximo de las imágenes y $c = 3$ es una constante para amplificar el error. Se tomó en cuenta que un valor de EPI y SSIM con valor igual a 1 es una imagen idéntica a la original y que un valor SAM cercano a 0 tiene una relación espectral igual a la original.

En la Tabla 1 se observan los resultados cuantitativos obtenidos de la imagen Pavia Center, Pavia University y Washington DC Mall. Se puede observar en Pavia Center y Pavia University que los métodos clásicos de interpolación y el método SR-WDIEE-SR resultaron en valores de EPI cercanos a 0, indicando que no tienen una buena preservación de bordes, también, presentaron valores SAM altos que señalan la baja relación espectral entre las bandas de las imágenes en comparación con los otros métodos.

Adicionalmente, sus valores MPSNR y MSSIM fueron inferiores a los del método NSSR y el sistema propuesto. Para el caso de Pavia Center, NSSR proporciona un MPSNR de 32.57 y RMSE de 5.99, MSSIM de 0.98, SAM de 7.12 y EPI de 0.5 mostrando más texturas y detalles finos en la imagen de error (Fig. 5), mientras que Pavia University con el método propuesto muestra un SAM de 5.09, EPI de 0.56 y RMSE de 7.76 resultando en bordes más nítidos y percepción de mayores detalles en

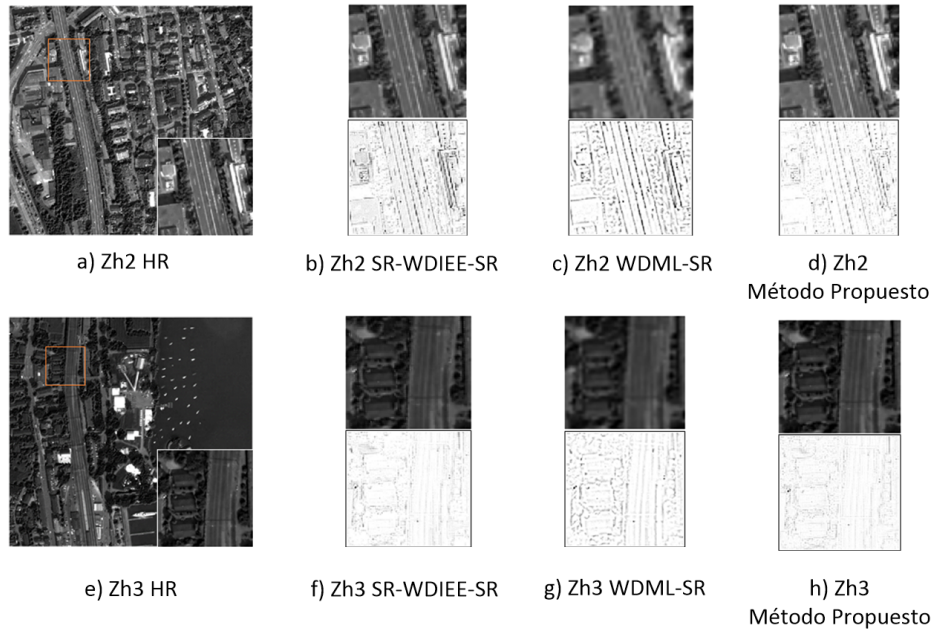


Fig. 7. Resultados visuales de las imágenes Zh2 y Zh3 (Zurich).

la imagen recortada. Por otro lado, en el caso de Washington DC Mall, los métodos clásicos, así como SR-WDIEE-SR y NSSR mostraron valores de MPSNR, RMSE, MSSIM y EPI menores al método propuesto. Esto se ve reflejado en la Fig. 6 donde la imagen de error (e) indica una reconstrucción con bordes más definidos incluso en comparación con el método WDML-SR.

En la Tabla 2 se muestran los valores promedio (MPSNR, MSSIM, EPI, SAM y RMSE) obtenidos de la base de datos CAVE. El método propuesto genera imágenes con menor error en los bordes para el caso de Balloons (Fig. 7 (e)) a comparación de los otros métodos. Sin embargo, para las demás imágenes, el método NSSR (Fig. 7 (i), (n)) disminuye de mejor forma este error, dándole nitidez a las imágenes reconstruidas.

No obstante, cabe mencionar que este método requiere ser entrenado de acuerdo con las características del sensor que capturó las imágenes híper-espectrales. Los resultados obtenidos en la base de datos de Zurich muestran que el método propuesto preserva la nitidez dado que, visualmente, la imagen reconstruida se parece en gran medida a la imagen original (Fig. 8 (d)) obteniendo RMSE de 14.62 el cual se ve reflejado en la imagen de error.

En cambio SR-WDIEE-SR (Fig. 8 (b)) muestra una imagen de error con zonas poco limpias mientras que el método WDML-SR presenta un emborronamiento en la imagen recortada. Cuantitativamente, los valores obtenidos en la Tabla 3 corroboran los resultados visuales, ya que los métodos clásicos así como SR-WDIEE-SR y WDML-SR obtienen resultados en MPSNR, MSSIM, EPI marginalmente bajos y RMSE altos con respecto al método propuesto. Los resultados experimentales obtenidos de las diferentes bases de datos de imágenes HS y MS demuestran que el método propuesto es efectivo para aumentar la calidad de las imágenes, preservando los detalles de los bordes y

resultando en imágenes claras. No obstante, cabe recalcar que el método propuesto se entrena con imágenes RGB para reconstruir imágenes de baja resolución a alta resolución, a diferencia del método NSSR, el cual es entrenado con imágenes HS de baja resolución e imágenes RGB tomadas por el mismo sensor.

Por lo que, NSSR toma en cuenta las características de las imágenes HS de un sensor específico por cada entrenamiento, es decir, se entrena un diccionario para cada base de datos de un determinado sensor y se ajustan los diferentes parámetros de regularización. En comparación, el método propuesto sólo necesita de un diccionario para el proceso de superresolución, permitiendo su uso en diferentes bases de datos y generando resultados competitivos para imágenes MS e HS como en el caso de las bases de datos de Zurich, Pavia University y Washington DC Mall.

5. Conclusiones

En este trabajo se propuso un método efectivo de superresolución de imágenes MS e HS basado en representación dispersa y reconstrucción mediante interpolación. El diccionario creado a partir de imágenes RGB mejora la precisión espacial de las imágenes MS e HS y la reconstrucción basada en interpolación de sub-bandas de alta frecuencia permite mantener la información de los bordes y características finas de la imagen.

Los resultados experimentales obtenidos de las diferentes bases de datos mostraron que este método puede lograr una menor tasa de error en reconstrucción para imágenes HS, como Pavia University y Washington DC Mall, así como en las imágenes MS utilizadas de la base de datos de Zurich, obteniendo una mejor calidad visual.

Sin embargo, los resultados obtenidos para las imágenes MS varían considerablemente de los registrados para imágenes HS y los valores SAM en imágenes HS distan de ser ideales. Por lo que es necesario mejorar el diccionario empleado con el fin de aumentar la calidad espacial-espectral en la reconstrucción de este tipo de imágenes.

Referencias

1. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on signal processing*, vol. 54, no. 11, pp. 4311–4322 (2006) doi: 10.1109/TSP.2006.881199
2. Alvarez-Ramos, V. Ponomaryov, V., Sadovnychiy, S.: Image super-resolution via wavelet feature extraction and sparse representation. *Radioengineering*, vol. 27, no. 2, pp. 602–609 (2018) doi: 10.13164/re.2018.0602
3. GIC: Hyperspectral remote sensing scenes. Grupo de Inteligencia Artificial (GIC), Universidad del País Vasco (2021)
4. Remote sensing laboratory: Remote sensing datasets. School of Surveying and Geospatial Engineering (2022)
5. Kaggle.: Starter: T91 image dataset 5362b8e1-6. Data Science platform (2019)
6. Ayas, S., Ekinci, M.: Single image super resolution based on sparse representation using discrete wavelet transform. *Multimedia Tools and Applications*, vol. 77, no. 13, pp. 16685–16698 (2019) doi: 10.1007/s11042-017-5233-5

7. Bioucas-Dias, J. M., Plaza, A., Camps-Valls, G., Scheunders, P., Nasrabadi, N. M., Chanussot, J.: Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 2, pp. 6–36 (2013) doi: 10.1109/MGRS.2013.2244672
8. Chavez-Roman, H., Ponomaryov, V.: Super resolution image generation using wavelet domain interpolation with edge extraction via a sparse representation. *IEEE Geoscience and remote sensing Letters*, vol. 11, no. 10, pp. 1777–1781 (2014) doi:10.1109/LGRS.2014.2308905
9. Dengwen, Z.: An edge-directed bicubic interpolation algorithm. In: 3rd International congress on image and signal processing, vol. 3, pp. 1186–1189 (2010) doi: 10.1109/CISP.2010.5647190
10. Dong, W., Fu, F., Shi, G., Cao, X., Wu, J., Li, G., Li, X.: Hyperspectral image super-resolution via non-negative structured sparse representation. *IEEE Transactions on Image Processing*, vol. 25, no. 5, pp. 2337–2352 (2016) doi: 10.1109/TIP.2016.2542360
11. Feng, L., Suen, C. Y., Tang, Y. Y., Yang, L. H.: Edge extraction of images by reconstruction using wavelet decomposition details at different resolution levels. *International journal of pattern recognition and artificial intelligence*, vol. 14, no. 6, pp. 779–793 (2000) doi: 10.1142/S0218001400000519
12. Goklani, H. S., Sarvaiya, J. N., Fahad, A. M.: Image reconstruction using orthogonal matching pursuit (OMP) algorithm. In: 2nd International Conference on Emerging Technology Trends in Electronics, Communication and Networking, pp. 1–5 (2014) doi: 10.1109/ET2ECN.2014.7044960
13. Justin, J., Sivaraman, J., Rajadurai, P., Simi, V. R.: An edge preservation index for evaluating nonlinear spatial restoration in MR images. *Current Medical Imaging*, vol. 13, no. 1, pp. 58–65 (2017)
14. Kirkland, E. J.: Bilinear interpolation. *Advanced Computing in Electron Microscopy*, 2nd edition. Springer, pp. 261–263. (2010) doi: 10.1007/978-3-030-33260-0
15. Mei, S., Yuan, X., Ji, J., Zhang, Y., Wan, S., Du, Q.: Hyperspectral image spatial super-resolution via 3D full convolutional neural network. *Remote Sensing*, vol. 9, no.11, pp. 1–22 (2017) doi: 10.3390/rs9111139
16. Nason, G. P., Silverman, B. W.: The stationary wavelet transform and some statistical applications. Antoniadis, A., Oppenheim, G., eds., *Wavelets and statistics*. Lecture Notes in Statistics, vol. 103, pp. 281–299 (1995) doi: 10.1007/978-1-4612-2544-7_17
17. Nguyen, H. V., Banerjee, A., Chellappa, R.: Tracking via object reflectance using a hyperspectral video camera. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, pp. 44–51 (2010) doi: 10.1109/CVPRW.2010.5543780
18. Parker, J. A., Robert V. K., Donald E. T.: Comparison of interpolating methods for image resampling. *IEEE Transactions on medical imaging*, vol. 2, no. 1, pp. 31–39 (1983) doi: 10.1109/TMI.1983.4307610
19. Partridge, M., Calvo, R. A.: Fast dimensionality reduction and simple PCA. *Intelligent data analysis*, vol. 2, no. 3, pp. 203–214 (1998) doi: 10.3233/IDA-1998-2304
20. Rukundo, O., Hanqiang, C.: Nearest neighbor value interpolation. *arXiv:1211.1768v2* (2012) doi: 10.48550/arXiv.1211.1768
21. Tarabalka, Y., Chanussot, J., Benediktsson, J. A.: Segmentation and classification of hyperspectral images using minimum spanning forest grown from automatically selected markers. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 5, pp. 1267–1279 (2010) doi: 10.1109/TSMCB.2009.2037132
22. Uzair, M., Mahmood, A., Mian, A. S.: Hyperspectral face recognition using 3D-DCT and partial least squares. *British Machine Vision Conference (BMVC)*, pp. 57.1–57.10 (2013) doi: 10.5244/C.27.57

23. Volpi, M., Ferrari, V.: Structured prediction for urban scene semantic segmentation with geographic context. Joint Urban Remote Sensing Event (JURSE), pp. 1–4 (2015) doi: 10.1109/JURSE.2015.7120490
24. Wu, X., Fan, J., Xu, J., Wang, Y.: Wavelet domain multidictionary learning for single image super-resolution. Journal of Electrical and Computer Engineering, Vol. 2015, pp. 1–12 (2015) doi: 10.1155/2015/526508
25. Yasuma, F., Mitsunaga, T., Iso, D., Nayar, S. K.: Generalized assorted pixel camera: Post-capture control of resolution, dynamic range and spectrum: Technical Report CUCS-061-08. Department of Computer Science, Columbia University (2008)