

**APRENDIZAJE DE DICCIONARIOS PARA TRANSFORMACIÓN DE IMÁGENES
ESPECTRALES EN SEÑALES DISPERSAS APLICADO A SU
RECONSTRUCCIÓN MEDIANTE LA TÉCNICA DE MUESTREO COMPRESIVO**

**OSCAR ENRIQUE HURTADO CAMACHO
JENNY LILIANA SÁNCHEZ RANGEL**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISICOMECÁNICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA
2015**

**APRENDIZAJE DE DICCIONARIOS PARA TRANSFORMACIÓN DE IMÁGENES
ESPECTRALES EN SEÑALES DISPERSAS APLICADO A SU
RECONSTRUCCIÓN MEDIANTE LA TÉCNICA DE MUESTREO COMPRESIVO**

**OSCAR ENRIQUE HURTADO CAMACHO
JENNY LILIANA SÁNCHEZ RANGEL**

**Trabajo de grado para optar por el título de:
Ingeniero de Sistemas**

**Director:
Ph.D. HENRY ARGUELLO FUENTES
Ingeniero electricista, Ph.D.**

**Codirectora:
Ph.D. (c) LAURA VIVIANA GALVIS CARREÑO
Ingeniera de sistemas, Ph.D. (c)**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISICOMECAÑICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA**

2015

Este proyecto se lo dedico a mi familia que siempre a pesar de la distancia, los errores y los malos momentos siempre han estado conmigo apoyándome.

A mis padres Marta y Luis que siempre han creído en mí, y que con su apoyo constante e incondicional y a pesar de mis errores hicieron de mí, a lo largo de todo este tiempo, una mejor persona.

A mi hermano Diego por su amistad incondicional, que está ahí en los mejores y los peores momentos de mi vida.

A mi abuelo Jorge por su apoyo, sus consejos, y la motivación para cada día ser una mejor persona.

A mis amigos los cuales son unas personas excepcionales que me brindaron su amistad, su compañía, sus consejos y sus locuras a cambio de nada, y con quienes compartí y seguiré compartiendo un pedacito de mi vida.

Oscar Enrique Hurtado Camacho

Dedico este proyecto a todas y cada una de las personas que día a día me han apoyado y acompañado en cada paso de mi vida, de manera especial a mis grandes amores, mis padres Magdalena y Aníbal, porque han consagrado su vida, dando lo mejor de sí para que sus hijos alcancen todo lo que se proponen sin importar las circunstancias o los obstáculos.

A mis hermanos Fredy y Eduard que a pesar de las diferencias siempre los llevo mi corazón.

A mi Juandes que siempre lo recuerdo y está en mis pensamientos y sé que él siempre está presente en cada paso de nuestras vidas a pesar de que no lo esté físicamente.

A Chavis por su cariño y sus buenos consejos.

A mi gran amigo y novio Juan por todo su apoyo incondicional, sus sabios consejos y por todos los momentos que hemos compartido.

Y por último a toda mi Familia y las grandes amistades que he construido a lo largo de mi vida.

Jenny Liliana Sánchez Rangel

AGRADECIMIENTOS

Los autores agradecen:

Al grupo de investigación en diseño de algoritmos y procesamiento de datos multidimensionales (HDSP) por acogernos y brindarnos el apoyo necesario para la realización de este proyecto.

Al profesor Henry Arguello Fuentes por su guía, consejos y apoyo para nuestro desarrollo profesional.

A nuestra codirectora Laura Viviana Galvis Carreño por todo el tiempo, la paciencia, sus consejos, su guía, y el esfuerzo invertido en el acompañamiento de este proyecto.

A todos nuestros amigos, compañeros de carrera y del grupo de investigación que directa e indirectamente nos ayudaron a la realización de este proyecto.

CONTENIDO

	pág.
INTRODUCCIÓN	20
1. OBJETIVOS	22
1.1 OBJETIVO GENERAL	22
1.2 OBJETIVOS ESPECÍFICOS	22
2. MARCO TEÓRICO	23
2.1 IMAGEN HIPERESPECTRAL	23
2.2 MUESTREO COMPRESIVO	25
2.2.1 Propiedad de dispersión	26
2.2.2 Proceso de muestreo	26
2.2.3 Reconstrucción	28
2.2.4 Aplicaciones	28
2.3 APRENDIZAJE DE DICCIONARIOS	30
2.3.1 Aprendizaje de diccionarios para representaciones dispersas	31
2.4 ESTADO DEL ARTE	32
3. DISEÑO DE LOS ALGORITMOS	34

3.1 APRENDIZAJE DE TRANSFORMACIONES PARA SEÑALES DISPERSAS	34
3.1.1 Codificación dispersa	35
3.1.2 Actualización del diccionario	36
3.1.3 Resumen	36
3.2 K-SVD Y OMP	36
3.2.1 K-SVD	37
3.2.2 OMP	38
3.2.3 Resumen	39
4. SIMULACIONES Y RESULTADOS	41
4.1 APRENDIZAJE DE TRANSFORMACIONES PARA SEÑALES DISPERSAS	42
4.1.1 Parches bidimensionales	42
4.1.2 Parches tridimensionales	46
4.2 K-SVD Y OMP	49
4.2.1 Parches bidimensionales	49
4.2.2 Parches tridimensionales	53
4.2.3 Parches espectrales	56
4.3 COMPARACIÓN DE LOS ALGORITMOS	60
5. CONCLUSIONES	63
6. RECOMENDACIONES	64
REFERENCIAS BIBLIOGRÁFICAS	65

LISTA DE FIGURAS

	pág.
Figura 1. Cubo de datos de una imagen espectral.	23
Figura 2. Espectro de un pixel de una imagen hiperespectral.	24
Figura 3. Ejemplo de la representación matricial de la teoría de muestreo compresivo.	27
Figura 4. Diagrama del proceso general del algoritmo de aprendizaje de transformaciones para señales dispersas.	35
Figura 5. Imágenes hiperespectrales original y obtenida mediante el algoritmo de aprendizaje de transformaciones para señales dispersas con parches bidimensionales mapeadas a un perfil RGB.	43
Figura 6. Representación dispersa de la imagen hiperespectral original y obtenida mediante el algoritmo de aprendizaje de transformaciones para señales dispersas con parches bidimensionales.	44
Figura 7. Firmas espectrales de la imagen hiperespectral obtenida mediante el algoritmo de aprendizaje de transformaciones para señales dispersas con parches bidimensionales.	45
Figura 8. Imágenes hiperespectrales original y obtenida mediante el algoritmo de aprendizaje de transformaciones para señales dispersas con parches tridimensionales mapeadas a un perfil RGB.	46
Figura 9. Representación dispersa de la imagen hiperespectral original y obtenida mediante el algoritmo de aprendizaje de transformaciones para señales dispersas con parches tridimensionales.	47
Figura 10. Firmas espectrales de la imagen hiperespectral obtenida mediante el algoritmo de aprendizaje de transformaciones para señales dispersas con parches tridimensionales.	48
Figura 11. Imágenes hiperespectrales original y obtenida mediante el algoritmo K-SVD y OMP con parches bidimensionales mapeadas a un perfil RGB.	50

Figura 12. Representación dispersa de la imagen hiperespectral original y obtenida mediante el algoritmo K-SVD y OMP con parches bidimensionales.	51
Figura 13. Firmas espectrales de la imagen hiperespectral obtenida mediante el algoritmo K-SVD y OMP con parches bidimensionales.	52
Figura 14. Imágenes hiperespectrales original y obtenida mediante el algoritmo K-SVD y OMP con parches tridimensionales mapeadas a un perfil RGB.	53
Figura 15. Representación dispersa de la imagen hiperespectral original y obtenida mediante el algoritmo K-SVD y OMP con parches tridimensionales.	54
Figura 16. Firmas espectrales de la imagen hiperespectral obtenida mediante el algoritmo K-SVD y OMP con parches tridimensionales.	55
Figura 17. Imágenes hiperespectrales original y obtenida mediante el algoritmo K-SVD y OMP con parches espectrales mapeadas a un perfil RGB.	57
Figura 18. Representación dispersa de la imagen hiperespectral original y obtenida mediante el algoritmo K-SVD y OMP con parches espectrales.	58
Figura 19. Firmas espectrales de la imagen hiperespectral obtenida mediante el algoritmo K-SVD y OMP con parches espectrales.	59

LISTA DE TABLAS

	pág.
Tabla 1. Descripción algoritmo K-SVD.	38
Tabla 2. Descripción algoritmo OMP.	39
Tabla 3. Parámetros de las simulaciones de los algoritmos de aprendizaje de transformaciones para señales dispersas y K-SVD y OMP.	60
Tabla 4. Desempeño de los algoritmos de aprendizaje de transformaciones para señales dispersas para diferentes niveles de dispersión de la matriz de coeficientes de representación.	60
Tabla 5. Desempeño de los algoritmos K-SVD y OMP para diferentes niveles de dispersión de la matriz de coeficientes de representación.	61

GLOSARIO

NORMA FROBENIUS: se conoce también como la norma euclidiana, es la norma de una matriz A de $m \times n$, se define como la raíz cuadrada de la suma de los cuadrados absolutos de sus elementos,

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}$$

También puede ser considerada como una norma vector

$$\|A\|_F = \sqrt{AA^*}$$

Donde A^* es la traspuesta de la matriz conjugada.

NORMA ℓ_p : una norma es una función, que se denota como $\|\cdot\|$, y satisface para un conjunto de objetos las condiciones de no negatividad, relación de la multiplicación escalar y desigualdad triangular. La norma ℓ_p es un tipo de norma definida para vectores que a menudo se denota como $\|\cdot\|_p$. La norma ℓ_p se define como:

$$\|x\|_p = \left(\sum_i |x_i|^p \right)^{\frac{1}{p}}$$

Esta norma satisface las condiciones de no negatividad, relación de multiplicación escalar.

QUASI NORMA ℓ_0 : se denota como $\|\cdot\|_0$ es una medida de la cardinalidad funcional es decir, el número de coeficientes distintos de cero en la representación dispersa, simplemente cuenta los elementos distintos de cero en un vector.

RELACIÓN SEÑAL A RUIDO DE PICO (PSNR): el término relación de pico de señal a ruido (PSNR) es una expresión para la relación entre el valor máximo posible de potencia de una señal y la potencia de distorsión de ruido que afecta a

la calidad de su representación. Debido a que muchas señales tienen un amplio rango dinámico, (relación entre los mayores y menores valores posibles de una cantidad variable) el PSNR es generalmente expresado en términos de la escala logarítmica en decibelios.

$$PSNR = 10 \log_{10} \frac{MAX_I^2}{MSE}$$

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2$$

$$MAX_I = 2^{B-1}$$

Donde B es el número de bits y MAX_I denota el máximo valor que puede tomar un pixel en la imagen.

Teniendo en cuenta que mejorar la calidad visual de una imagen digital puede ser subjetiva ya que varía de persona a persona, se hace necesario establecer medidas cuantitativas para comparar los efectos de los algoritmos en la mejora de la calidad de la imagen, es decir, el PSNR es una medida que permite la comparación de algoritmos. Los valores típicos para algoritmos de compresión con pérdida con una profundidad de color de 8 bits varía entre 30 y 50 dB, siendo mejor cuanto mayor sea este valor.

RESUMEN

TÍTULO: APRENDIZAJE DE DICCIONARIOS PARA TRANSFORMACIÓN DE IMÁGENES ESPECTRALES EN SEÑALES DISPERSAS APLICADO A SU RECONSTRUCCIÓN MEDIANTE LA TÉCNICA DE MUESTREO COMPRESIVO*.

AUTORES: OSCAR ENRIQUE HURTADO CAMACHO, JENNY LILIANA SÁNCHEZ RANGEL**.

PALABRAS CLAVE: Aprendizaje de diccionarios, Muestreo compresivo.

DESCRIPCIÓN:

El tratamiento de imágenes espectrales es un área que está siendo estudiada activamente, sin embargo la captura de estas imágenes así como su tratamiento son un gran desafío debido a la gran cantidad de información que contienen. Es por esto que trabajos recientes han aplicado la teoría de muestreo compresivo (Compressive Sensing - CS) a su tratamiento, por ejemplo en sistemas para su captura y reconstrucción, debido a que esta teoría afirma que se pueden recuperar escenas espectrales a partir de un número menor de muestras que las requeridas por esquemas convencionales de sensores basados en el criterio de Nyquist-Shannon. Diversas aplicaciones del procesamiento de señales e imágenes se han dedicado a explotar el concepto de dispersión de estas en un cierto dominio de transformación. Existen transformadas analíticas que permiten la representación dispersa de una señal como por ejemplo wavelets, DCT entre otras que han sido ampliamente utilizadas en los estándares de compresión y que han dado lugar a novedosas aplicaciones en diversos campos. Estos diccionarios, basados en modelos matemáticos, se caracterizan por una formulación analítica además de una rápida implementación implícita, a pesar de esto, existen algunas aplicaciones donde resulta ventajoso entrenar un diccionario que permita una mayor flexibilidad y una adaptación a los datos de una señal específica.

Este trabajo de investigación aborda el problema de encontrar un diccionario o matriz de transformación para la representación dispersa de una imagen hiperespectral por medio del aprendizaje de diccionarios. Para esto, se propone el desarrollo e implementación de algoritmos de aprendizaje de diccionarios que transformen una imagen hiperespectral en su equivalente dispersa con el fin de aplicar la técnica de muestreo compresivo y obtener como resultado la reconstrucción de la imagen original.

* Trabajo de grado

** Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería de Sistemas e informática.
Director: Ph.D Henry Arguello Fuentes. Codirectora: Ph.D. (c) Laura Viviana Galvis Carreño.

ABSTRACT

TITLE: SPARSE REPRESENTATION BY DICTIONARY LEARNING FOR COMPRESSIVE SENSING HYPERSPECTRAL IMAGES*.

AUTHORS: OSCAR ENRIQUE HURTADO CAMACHO, JENNY LILIANA SÁNCHEZ RANGEL**.

KEYWORDS: Dictionary learning, Compressive sensing.

DESCRIPTION:

The spectral images treatment is an area that is being actively studied, however these image's acquisition and treatment are a major challenge due to the large amount of information on it. For that reason, recent studies have applied the theory of compressive sensing (CS) for its treatment, for example in acquisition and reconstruction systems, because this theory asserts that spectral scenes can be recovered from a small number of samples than those required by conventional schemes based on the Nyquist-Shannon standard. Several signal and image processing applications have been dedicated to exploit the sparse concept in a certain transform domain. There are analytical transformations that allow the sparse representation of a signal such as wavelets, DCT among others that have been widely used in compression standards and have led to new applications in diverse fields. These dictionaries, based on mathematical models, are characterized by an analytical formulation further a quickly implicit implementation, despite this, there are some applications where it is advantageous to train a dictionary allowing greater flexibility and adaptability to a specific signal data.

This research addresses the problem of finding a dictionary or transformation matrix to the sparse representation for a hyperspectral image by dictionary learning. For this, it is proposed the development and implementation of dictionary learning algorithms to transform a hyperspectral image into its sparse equivalent in order to apply compressive sensing techniques and reconstruct the original image.

* Bachelor thesis

** Faculty of Physical-Mechanical Engineering. School of Systems and Computing Engineering. Director: Ph.D. Henry Arguello Fuentes. Co-director: Ph.D.(c) Laura Viviana Galvis Carreño.

INTRODUCCIÓN

El tratamiento de imágenes espectrales ha ganado popularidad recientemente, debido a la gran cantidad de información que se puede adquirir mediante la captura de este tipo de imágenes multidimensionales. Su captura así como su tratamiento continúa planteando un gran desafío; es por esto que sistemas para su captura han sido creados [1] y diversos trabajos se desarrollan actualmente en búsqueda de reconstruir la imagen o señal real [2].

La teoría de muestreo compresivo (Compressive sensing - CS) ha sido aplicada en estos sistemas de captura, para la reconstrucción de los cubos de datos espectrales. La teoría de CS, es un nuevo paradigma que va en oposición al tradicional teorema de muestreo en el área de procesamiento de señales [3]; básicamente CS es una técnica de procesamiento de señales para adquirir y comprimir de manera eficiente una señal que luego puede ser reconstruida mediante la búsqueda de soluciones a un sistema lineal indeterminado [4], y se caracteriza porque la reconstrucción de esta señal requiere una cantidad significativamente menor de muestras de las que se necesitan cuando se usa el teorema de muestreo de Nyquist-Shannon [3], [5]. Debido a esto, la teoría de CS se ha posicionado como un nuevo marco para la adquisición de señales y diseño de sensores, lo que ha permitido una gran reducción en el muestreo y los costos de computación tanto para el sensado como la compresión de señales que tienen una representación dispersa [6].

Uno de los conceptos claves de la teoría de CS, es la dispersión de una señal, que establece que la mayoría de la energía de dicha señal, se encuentra concentrada en un pequeño grupo de sus componentes [7], [8]. La mayoría de las señales naturales tales como sonidos o imágenes no son dispersas por sí mismas, sin embargo se puede encontrar una representación dispersa en una base o diccionario apropiados [3], [5]. La teoría de CS se basa en explotar estas

representaciones dispersas de las señales en algún dominio de transformación apropiado y la incoherencia de estas muestras en el dominio original de la señal [5]. Es por esto que la elección de una base o diccionario apropiado que cumpla estas condiciones es una tarea importante en el campo de CS, para la cual se puede contar con algunos diccionarios basados en modelos matemáticos o por medio del entrenamiento de un diccionario a partir de un gran conjunto de señales de ejemplo o entrenamiento, lo cual es conocido como aprendizaje de diccionario [9].

Actualmente en CS los diccionarios basados en modelos matemáticos tales como Wavelet, Fourier y DCT son ampliamente usados debido a que es bien conocido que proporcionan una representación dispersa de señales e imágenes naturales [10], [11] estas han sido aplicadas en áreas como compresión y eliminación de ruido [12], [13]. Los diccionarios basados en modelos matemáticos se caracterizan por una formulación analítica y una rápida implementación implícita, mientras que el aprendizaje de diccionarios permite una mayor flexibilidad y una adaptación a los datos de una señal específica [14].

Con este proyecto se busca tratar el problema de encontrar un diccionario o base apropiado para la representación dispersa de una imagen espectral a través del aprendizaje de diccionarios, un enfoque recientemente planteado en la literatura. Para esto, se propone el desarrollo e implementación de algoritmos de aprendizaje de diccionarios que transformen una imagen espectral en su equivalente dispersa con el fin de aplicar la técnica de muestreo compresivo y obtener como resultado la reconstrucción de la imagen original. Además de la comparación de estos algoritmos, para verificar su rendimiento.

1. OBJETIVOS

1.1 OBJETIVO GENERAL

Desarrollar e implementar un algoritmo de aprendizaje de diccionarios para transformación de imágenes espectrales en sus equivalentes señales dispersas y su posterior reconstrucción mediante la técnica de muestreo compresivo.

1.2 OBJETIVOS ESPECÍFICOS

1. Determinar el modelo matemático para la transformación de imágenes espectrales en señales dispersas utilizando la técnica de aprendizaje de diccionarios.
2. Desarrollar e implementar algoritmos para calcular una matriz de transformación mediante el aprendizaje de diccionarios que permitan representar imágenes espectrales en sus correspondientes señales dispersas y posteriormente obtener su respectiva reconstrucción.
3. Verificar el rendimiento del modelo propuesto a través de la comparación entre los resultados obtenidos mediante el uso de las matrices de transformación obtenidas para imágenes espectrales y bases de representación comúnmente utilizadas.
4. Aplicar y verificar las matrices de transformación obtenidas dentro de un sistema de muestreo compresivo de imágenes espectrales.

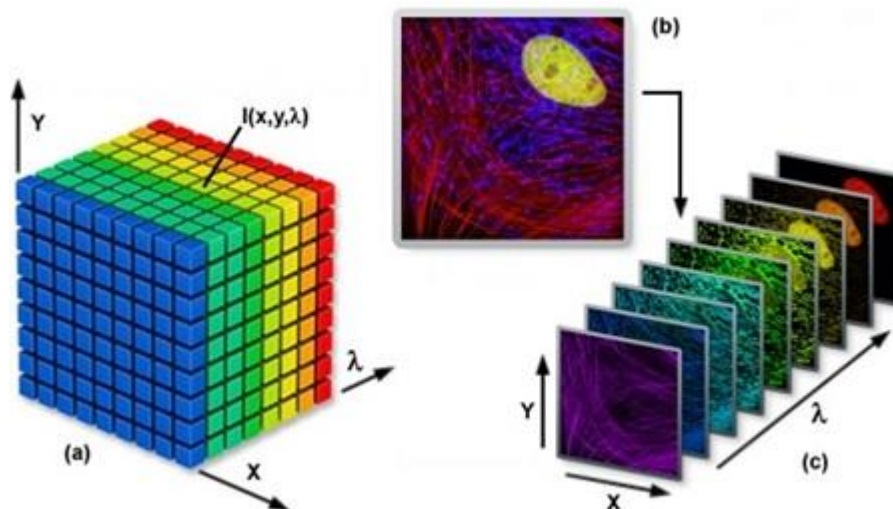
2. MARCO TEÓRICO

2.1 IMAGEN HIPERESPECTRAL

Una imagen hiperespectral es un modelo de la realidad que consiste en medir en un rango de frecuencias la variación de radiación electromagnética que es emitida, reflejada o absorbida, o lo que es también llamado espectro, de cada posición espacial de una escena.

Estas imágenes se capturan por medio de detectores sensibles a un rango de longitudes de onda del espectro electromagnético, los cuales forman una imagen tridimensional (x, y, λ) , es decir, un cubo de datos hiperespectral.

Figura 1. Cubo de datos de una imagen espectral



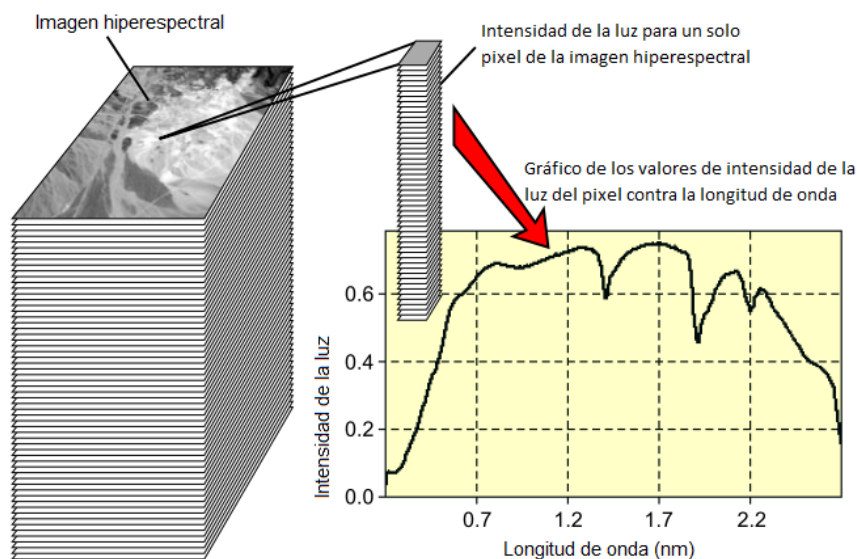
Fuente: [en línea] <http://zeiss-campus.magnet.fsu.edu/tutorials/spectralimaging/lambdastack/lambdastacktutorialfigure1.jpg>

En la figura 1 (a) se puede apreciar la estructura de un cubo de datos hiperespectral donde x e y representan dos dimensiones espaciales de la escena, y λ representa la dimensión espectral o dominio de longitud de onda; $I(x, y, \lambda)$ es la intensidad de la luz en un punto (x, y) , dada la longitud de onda λ . En (b) se muestra una imagen que representa un rango de longitudes de onda del espectro

electromagnético también conocido como banda espectral, y en (c) se muestran las diferentes bandas espectrales del cubo organizadas según λ .

Debido a que los elementos químicos tienen una distribución característica de la radiación electromagnética que emiten, reflejan o absorben a diferentes longitudes de onda, a lo cual se le conoce como la firma espectral, es posible identificar de forma única a cada uno de ellos, por lo cual se pueden obtener propiedades de los materiales que componen una escena de una imagen hiperespectral con una alta resolución espacial y espectral con mucha más precisión que con imágenes RGB convencionales [15].

Figura 2. Espectro de un pixel de una imagen hiperespectral.



Fuente: [en línea] www.microimages.com/documentation/Tutorials/hyprspec.pdf
modificada por los autores

En la figura 2 se puede apreciar una imagen hiperespectral de la cual se obtienen valores de la intensidad de la luz para un solo pixel, estos valores de intensidad pueden graficarse para las diferentes longitudes de onda, dejando ver claramente

el espectro del pixel, el cual puede ser usado para identificar los materiales presentes.

Al usarse sensores hiperespectrales en plataformas aerotransportadas o en el espacio se puede estudiar la superficie de la tierra en gran detalle, pero los sensores deben tener la capacidad de proveer cientos de bandas espectrales, dando así un espectro por pixel a lo largo de la imagen. Esta capacidad de los sensores permitiría determinar las firmas espectrales de los componentes de la superficie terrestre a partir de los datos obtenidos [16]. Sin embargo en la mayoría de los casos las escenas representadas por los pixeles observados en las imágenes no son homogéneos. Como resultado las firmas espectrales obtenidas por los sensores en cada pixel están formadas por una integración de firmas espectrales de componentes presentes en la escena. A las firmas espectrales a las cuales se les puede considerar puras a escala macroscópica son usualmente llamadas endmembers en la literatura [17].

Es por esto que el aprovechamiento de estas imágenes hiperespectrales de la superficie terrestre, a pesar de que representa un desafío para diversas aplicaciones en el área de sensado remoto [18], es objeto de estudio de numerosos trabajos.

Es por lo anterior que, las imágenes hiperespectrales son usadas en diversas aplicaciones importantes en diferentes áreas como por ejemplo en el monitoreo del medio ambiente [19] o vigilancia, detección y clasificación de sustancias [20].

2.2 MUESTREO COMPRESIVO

El muestreo compresivo (Compressive sensing - CS) es una nueva teoría que establece que se pueden muestrear señales e imágenes con un número menor de muestras que al usar el teorema de Nyquist-Shannon [3], [21]. La teoría de CS se basa en explotar la dispersión de las señales o imágenes en alguna base o diccionario apropiados para recuperar una señal de este tipo a partir de pocas

muestras lineales no adaptativas a través de un proceso de optimización [3], [5], [6], [21].

2.2.1 Propiedad de dispersión. Uno de los conceptos claves de la teoría de CS, es la dispersión de una señal, que establece que la mayoría de la energía de dicha señal en algún dominio se encuentra concentrada en un pequeño grupo de sus componentes [7], [8]. La mayoría de las señales naturales tales como sonidos o imágenes no son dispersas por sí mismas, sin embargo se puede encontrar una representación dispersa en una base o diccionario apropiados [3], [5].

La representación dispersa de una señal se puede ver como una combinación lineal de elementos que pertenecen a una base o diccionario. Estos elementos generalmente son vectores de norma unitaria llamados átomos y matemáticamente al tener una señal $\mathbf{y} \in \mathbb{R}^n$ se tiene que se puede representar en un diccionario $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_N]$ como:

$$\mathbf{y} = \mathbf{W}\mathbf{x} = \sum_{k=1}^N x_k \mathbf{w}_k \quad (1)$$

donde N es el tamaño del diccionario. Si $N = n$ y \mathbf{W} es de rango completo, se dice que el diccionario \mathbf{W} es una base. Si $N > n$ el diccionario es sobrecompleto y entonces \mathbf{x} no es único. El vector $\mathbf{x} = [x_1, x_2, \dots, x_N]$ está compuesto de k componentes distintas de cero, donde x_i representa la contribución del átomo \mathbf{w}_i a la definición de la señal \mathbf{y} . La representación de \mathbf{y} en \mathbf{W} se dice que es S -dispersa si es una combinación lineal de $S \ll n$ vectores \mathbf{w}_k de \mathbf{W} .

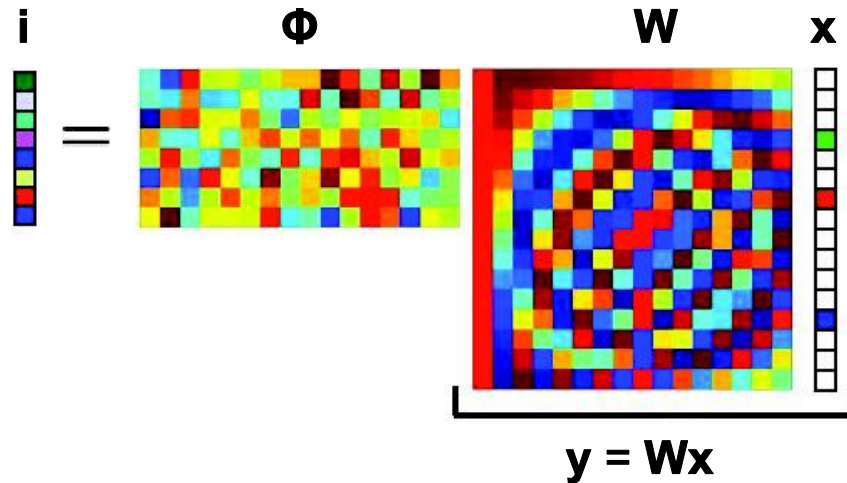
2.2.2 Proceso de muestreo. El proceso de muestreo en la teoría de CS se basa en adquirir m muestras de la señal $\mathbf{y} \in \mathbb{R}^n$ en donde generalmente $m \ll n$ a través de una matriz Φ y se puede expresar matemáticamente como $\mathbf{i} = \Phi\mathbf{y}$. El vector \mathbf{y}

puede expresarse en un diccionario o base como una señal dispersa de la siguiente manera:

$$\mathbf{i} = \Phi \mathbf{W} \mathbf{x} = \mathbf{A} \mathbf{x} \quad (2)$$

donde $\mathbf{i} \in \mathbb{R}^m$, $\Phi \in \mathbb{R}^{m \times N}$, $\mathbf{y} = \mathbf{W} \mathbf{x}$, y $\mathbf{A} = \Phi \mathbf{W}$.

Figura 3. Ejemplo de la representación matricial de la teoría de muestreo compresivo.



En la figura 3 se puede observar un ejemplo de la representación matricial de la ecuación (2).

La matriz de muestreo $\mathbf{A} = \Phi \mathbf{W}$ en este caso debería preservar la información importante en la señal de interés y esto se puede asegurar si la matriz \mathbf{A} cumple la propiedad de isometría restringida (Restricted Isometry Property - RIP) [5].

La propiedad de isometría restringida se define con una constante de isometría δ_s de una matriz la cual es el menor valor tal que la siguiente expresión

$$(1 - \delta_s) |\mathbf{x}|_2^2 \leq |\mathbf{A} \mathbf{x}|_2^2 \leq (1 + \delta_s) |\mathbf{x}|_2^2 \quad (3)$$

se mantiene para todos los vectores x S-dispersos [3], [5].

2.2.3 Reconstrucción. Para la reconstrucción de una señal a partir del vector de muestras $i \in \mathbb{R}^m$ se debe tener en cuenta que dado que $m \ll n$, hay infinitas soluciones x' que satisfacen la condición que $i = Ax$. Para resolver este problema se puede buscar el vector x más disperso que cumpla la anterior condición. Sin embargo esto llevaría a un problema combinatorio basado en la norma ℓ_0 , el cual es en general de la clase de complejidad NP-hard [22]. Es por esto que se usan algoritmos de optimización que intentan resolver el problema anterior y explotar el hecho de que la solución es dispersa, usualmente al minimizar la norma ℓ_1 sobre el espacio de la solución [5]. Estos algoritmos de optimización basados en la norma ℓ_1 intentan resolver el problema dado por

$$x = \arg \min_{x'} \|x'\|_1 \text{ sujeto a } Ax' = i \quad (4)$$

dando los mismos resultados bajo algunas circunstancias que al usar una minimización usando la norma ℓ_0 [5], [21].

En general, en la literatura existen una gran variedad de algoritmos de reconstrucción para la recuperación de señales dispersas en CS que exhiben diferentes características en términos de complejidad computacional entre los cuales se encuentran los algoritmos voraces iterativos, algoritmos basados en minimización no convexa, algoritmos iterativos de Bregman, entre otros [5].

2.2.4 Aplicaciones. Los trabajos en el área de CS han permitido diversas aplicaciones tales como las basadas en la arquitectura del sistema de adquisición de imágenes espectrales basado en apertura codificada de única captura (CASSI), que captura información espectral de una escena utilizando proyecciones

bidimensionales codificadas y acoge naturalmente el uso de los principios del muestreo compresivo [23], haciendo posible incursionar en otras áreas como super-resolución [24], detectores de un solo píxel [25], espectroscopía Raman [26], entre otras.

En la mayoría de las aplicaciones electrónicas para imágenes, las imágenes con alta resolución (HR) son preferidas o deseadas y requeridas a menudo. Alta resolución significa que la densidad de píxeles dentro de una imagen es alta, y por lo tanto, estas imágenes de alta resolución pueden ofrecer más detalles, lo cual es crítico en diversas aplicaciones. Algunos ejemplos de estas aplicaciones son las imágenes médicas en la realización de un diagnóstico correcto, el reconocimiento de patrones en visión artificial, entre otras.

En el caso de la super-resolución se pueden encontrar avances teniendo en cuenta que la resolución espacial en CASSI depende del detector utilizado; es decir imágenes de alta resolución requieren detectores de alta resolución que generalmente demandan altos costos. Para abordar este problema es necesario plantear modelos ópticos que puedan ser tratados por medio de la técnica de muestreo compresivo [24].

La espectroscopía Raman de imágenes ha sido ampliamente utilizada en el análisis de la composición química y la construcción de imágenes espectrales de compuestos orgánicos e inorgánicos, sin embargo se requieren largos periodos de tiempo tanto para la adquisición de datos como para el posterior tratamiento de imágenes espectrales. Por esta razón se han introducido modificaciones en la arquitectura óptica de espectroscopia Raman, por ejemplo con el uso del CASSI como aplicación a la técnica de muestreo compresivo ya que reduce el tiempo de adquisición al tiempo que realiza el proceso de detección y compresión de las señales espectrales subyacentes [26].

2.3 APRENDIZAJE DE DICCIONARIOS

El aprendizaje de diccionarios es un nuevo y prometedor campo de investigación que se enfoca en el desarrollo de innovadores algoritmos para construir diccionarios que provean representaciones eficientes de diferentes clases de señales [27].

Las técnicas de procesamiento de imágenes comúnmente requieren representaciones significativas que capturen las características útiles de la imagen, por ejemplo, para una aplicación de reconocimiento la representación debería resaltar características sobresalientes, para eliminación de ruido la representación debería separar eficientemente la señal del ruido, y para compresión esta representación debería buscar capturar la mayoría de la imagen con solo unos pocos coeficientes [14].

La representación de una señal involucra el uso de un diccionario el cual es un conjunto de señales elementales o átomos usado para descomponer la señal. Si el diccionario forma una base entonces la representación como una combinación lineal es única. En el caso en que el diccionario tiene más átomos que las dimensiones del espacio de la señal que representa se tiene un diccionario sobrecompleto el cual provee una ambigüedad en la representación de la señal.

Si se considera un diccionario sobrecompleto $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_N] \in \mathbb{R}^{N \times n}$ en donde las columnas representan los átomos del diccionario y $N \geq n$, existen dos formas de representar una señal $\mathbf{y} \in \mathbb{R}^n$ usando este diccionario, las cuales son mediante los denominados modelos de análisis y de síntesis [11]. La señal puede ser representada según el modelo de análisis por los productos internos con los átomos del diccionario

$$\mathbf{x}_a = \mathbf{W}^T \mathbf{y} \quad (5)$$

o según el modelo de síntesis como una combinación lineal de los átomos.

$$\mathbf{y} = \mathbf{W}\mathbf{x}_s \quad (6)$$

Si se tiene en cuenta el modelo de síntesis cuando el diccionario \mathbf{W} es sobrecompleto se tiene que la representación \mathbf{x}_s no es única, lo cual permite buscar la representación más significativa de la señal con respecto a una función de costo $C(\mathbf{x})$ [14], y se puede ver como el siguiente problema de minimización:

$$\mathbf{x}_s = \arg \min_{\mathbf{x}} C(\mathbf{x}) \text{ sujeto a } \mathbf{y} = \mathbf{W}\mathbf{x} \quad (7)$$

Esta función de costo, en el caso de buscar diccionarios que obtengan una representación dispersa, debe promover la búsqueda de un vector disperso y que contenga un número pequeño de coeficientes significativos, mientras que el resto de los coeficientes sea igual o cercano a cero. Estas opciones para la representación de señales (5) y (6), y específicamente el problema (7) han sido estudiadas extensamente lo que ha llevado al desarrollo de nuevos algoritmos para el procesamiento de señales que utilizan diccionarios sobrecompletos [14].

2.3.1 Aprendizaje de diccionarios para representaciones dispersas. Para la utilización de los anteriores conceptos en diferentes aplicaciones se debe tener en cuenta el desafío que significa la elección de un diccionario apropiado, es por esto que una gran variedad de diccionarios han sido desarrollados de acuerdo a las necesidades, los cuales provienen de dos fuentes, los diccionarios basados en modelos matemáticos, y los diccionarios basados en señales de entrenamiento, lo que es conocido como “aprendizaje de diccionarios” [9], [14]. Entre los diccionarios basados en modelos matemáticos que se utilizan para la representación dispersa de señales se encuentran la transformada de Fourier, la transformada discreta de coseno (DCT), wavelet, entre otros, debido a que es bien sabido que las señales naturales e imágenes son dispersas en ellos [10], [11].

El problema de aprendizaje de diccionarios para representaciones dispersas (mediante el uso del modelo de síntesis) es entrenar un diccionario basado en N señales de entrenamiento \mathbf{y} , lo cual puede ser visto como una matriz $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]$ en donde las columnas representan las señales de entrenamiento, y puede ser formulado como:

$$\min_{\mathbf{W}, \mathbf{X}} \|\mathbf{Y} - \mathbf{W}\mathbf{X}\|_F \text{ sujeto a } \|\mathbf{x}_i\|_0 \leq s \quad \forall i \quad (8)$$

en donde la matriz \mathbf{X} contiene en sus columnas el vector de coeficientes para la i -ésima señal de entrenamiento [9]. Este problema de optimización es combinatorio y no convexo lo cual dificulta su solución exacta, pero existen algoritmos que buscan aproximaciones por medio de mínimos locales que lo resuelvan, tales como el método de direcciones óptimas MOD [14], K-SVD [28], entre otros.

El desempeño de estos algoritmos en términos de la calidad de la aproximación y la dispersión de los coeficientes de la señal no depende solamente de la señal en sí misma, sino también del diccionario obtenido, además de que este desempeño no es el mismo al ser usados para una clase específica de señales \mathbf{Y} , ya que existen diccionarios que es más probable que lleven a soluciones dispersas que otros, y esto es lo que busca el aprendizaje de diccionarios, encontrar dichos diccionarios [27].

2.4 ESTADO DEL ARTE

El muestreo compresivo es una técnica que ha surgido recientemente para el procesamiento de señales que sirve para adquirir y comprimir de manera eficiente una señal que luego puede ser reconstruida con una cantidad significativamente menor de muestras de las que se necesitan cuando se usa el teorema de muestreo de Nyquist-Shannon [3], [5]. Debido a esto la teoría de muestreo compresivo se ha posicionado como un nuevo marco para la adquisición de

señales y diseño de sensores [6]. Dentro de las aplicaciones de la teoría de muestreo compresivo se encuentran los sistemas de captura de imágenes hiperespectrales [2], [23], [24], esto debido a la necesidad de superar el desafío que plantea el procesamiento estas imágenes por la gran cantidad de información que contienen.

Dentro de la teoría de muestreo compresivo una tarea importante es la elección de un diccionario. Tradicionalmente dentro de la comunidad de procesamiento de señales es muy común el uso de los diccionarios de Fourier y Wavelet, los cuales tienen un muy buen desempeño para señales unidimensionales, sin embargo estos diccionarios no son los más adecuados para representar señales naturales complejas, especialmente de alta dimensionalidad [9], [29]. Es por esto que una gran variedad de diccionarios han sido desarrollados en respuesta a las necesidades crecientes.

Los diccionarios pueden emerger de dos formas, ya sea mediante un modelo matemático de los datos [30]–[36], o por medio del aprendizaje de un diccionario a partir de datos de entrenamiento [28], [37]–[39]. Los diccionarios del primer tipo se conocen como diccionarios analíticos, porque se caracterizan por una formulación analítica y proporcionan una rápida implementación implícita. En contraste, los diccionarios del segundo tipo ofrecen mayor flexibilidad y poseen la capacidad de adaptarse a los datos de señales específicas. La principal ventaja de estos diccionarios entrenados radica en el rendimiento que logran alcanzar al ser aplicados en diversas áreas del procesamiento de señales [14].

Dentro de la literatura a pesar de que existen diversos métodos que abordan el problema de representación dispersa de imágenes de dos dimensiones, el problema de representación dispersa de imágenes hiperespectrales ha recibido poca atención.

3. DISEÑO DE LOS ALGORITMOS

3.1 APRENDIZAJE DE TRANSFORMACIONES PARA SEÑALES DISPERSAS

El problema de la representación de un conjunto de señales $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N] \in \mathbb{R}^{n \times N}$ en un diccionario de análisis $\mathbf{W} \in \mathbb{R}^{m \times n}$ como un conjunto de señales dispersas $\mathbf{X} \in \mathbb{R}^{m \times N}$ puede verse como el siguiente problema de minimización:

$$\min_{\mathbf{W}, \mathbf{X}} \|\mathbf{WY} - \mathbf{X}\|_F^2 \text{ sujeto a } \|\mathbf{x}_i\|_0 \leq s \quad \forall i \quad (9)$$

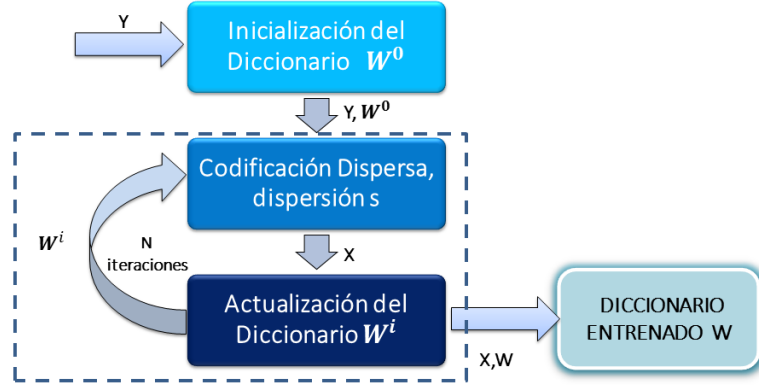
Este problema de minimización es no convexo y muy costoso de resolver en términos de complejidad computacional. Es por esto que una formulación alternativa del problema es usada, la cual es basada en el trabajo de Ravishankar y Bresler, en la cual se busca que el problema tenga un buen condicionamiento, evite las soluciones triviales y que busque que el diccionario a obtener sea de rango completo [10]. Esta formulación alternativa del problema se muestra a continuación:

$$\min_{\mathbf{X}, \mathbf{W}} \|\mathbf{WY} - \mathbf{X}\|_F^2 - \lambda \log \det \mathbf{W} + \mu \|\mathbf{W}\|_F^2 \text{ sujeto a } \|\mathbf{x}_i\|_0 \leq s \quad \forall i \quad (10)$$

En donde $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N] \in \mathbb{R}^{n \times N}$, $\mathbf{W} \in \mathbb{R}^{n \times n}$, y $\mathbf{X} \in \mathbb{R}^{n \times N}$. Este problema de minimización puede ser resuelto por medio de dos pasos o etapas. El primer paso consiste en la codificación dispersa o la representación dispersa y el segundo se encarga de la actualización del diccionario.

El algoritmo desarrollado alterna entre estos dos procesos o pasos, lo cual se puede observar en la figura 4, con el fin de encontrar el diccionario que mejor se ajuste a las señales o imágenes que se quieran representar.

Figura 4. Diagrama que muestra el proceso iterativo del algoritmo de aprendizaje de transformaciones para señales dispersas que lo componen los pasos de codificación dispersa y actualización del diccionario.



3.1.1 Codificación dispersa. En este paso se resuelve el problema de minimización de la ecuación (10) con W fijo. La norma ℓ_0 en la restricción de dispersión se relaja en la norma ℓ_1 . Esta a su vez, se agrega en el problema de minimización como una penalización. Por lo que el problema se replantea de la siguiente forma:

$$\min_X \|WY - X\|_F^2 + \eta \sum_{i=1}^N \|x_i\|_1 \quad (11)$$

donde η busca garantizar la dispersión dada por s . La solución para X , se puede calcular exactamente por medio de la umbralización (thresholding) de la siguiente manera:

$$X_{ij} = \begin{cases} (WY)_{ij} - \frac{\eta}{2}, & (WY)_{ij} \geq \frac{\eta}{2} \\ (WY)_{ij} + \frac{\eta}{2}, & (WY)_{ij} < -\frac{\eta}{2} \\ 0, & \text{otro} \end{cases} \quad (12)$$

3.1.2 Actualización del diccionario. En este paso se resuelve el siguiente problema de minimización con \mathbf{X} fijo:

$$\min_{\mathbf{W}} \|\mathbf{W}\mathbf{Y} - \mathbf{X}\|_F^2 - \lambda \log \det \mathbf{W} + \mu \|\mathbf{W}\|_F^2 \quad (13)$$

Este problema se puede resolver usando el método de gradiente conjugado con backtracking line search [40], pero en la práctica, usando valores de paso fijo también se obtienen buenos resultados. El criterio de terminación utilizado dentro del algoritmo de gradiente conjugado es un número de iteraciones fijo, esto debido a que este método converge generalmente de forma rápida y probó ser útil para esta aplicación.

3.1.3 Resumen. Al evaluar la complejidad computacional del algoritmo se tiene en cuenta el costo total por iteración de los pasos de codificación dispersa y actualización del diccionario, lo cual requiere $O(Nn^2)$ operaciones. En donde N es el número de señales de entrada, y n es el tamaño de las señales de entrada.

Con base en la formulación presentada para este algoritmo se planteó el entrenamiento de un diccionario \mathbf{W} a partir de $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]$ señales de entrada con el fin de obtener $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$ representaciones dispersas de las mismas en dicho diccionario.

3.2 K-SVD Y OMP

El problema de encontrar una representación dispersa para un conjunto de señales de entrada mediante un diccionario de síntesis puede verse como el siguiente problema de minimización:

$$\min_{\mathbf{W}, \mathbf{X}} \|\mathbf{Y} - \mathbf{W}\mathbf{X}\|_F^2 \text{ sujeto a } \|\mathbf{x}_i\|_0 \leq s \quad \forall i \quad (14)$$

En donde \mathbf{W} es el diccionario, \mathbf{Y} es la matriz de señales de entrada, \mathbf{X} la matriz de coeficientes de representación y s es el máximo número de coeficientes diferentes de cero de cada vector de representación x_i .

Este problema de minimización es no convexo y muy costoso en términos de complejidad computacional de resolver. Es por esto que se desarrolló un algoritmo basado en los algoritmos K-SVD [28] y Orthogonal Matching Pursuit (OMP) que trata este problema de minimización.

3.2.1 K-SVD. K-SVD es un algoritmo de aprendizaje de diccionarios que permite la creación de un diccionario para la representación dispersa de señales. El algoritmo K-SVD es un método iterativo que alterna entre dos pasos: la codificación dispersa que consiste en encontrar los coeficientes de representación de las señales de entrada basándose en el diccionario actual, y una actualización de los átomos del diccionario para ajustarse a las señales de entrada [28].

Para resolver el problema (14) el algoritmo realiza primero el paso de codificación dispersa en donde mantiene fijo \mathbf{W} para así buscar la mejor matriz de coeficientes \mathbf{X} que pueda ser encontrada. Dado que encontrar la matriz \mathbf{X} óptima es imposible, se puede usar cualquier algoritmo pursuit para aproximar la solución de este problema, en este caso se usó el algoritmo Orthogonal Matching Pursuit (OMP). Luego se realiza el paso de actualización del diccionario en búsqueda de uno mejor, en este proceso se mantiene fija la matriz de coeficientes \mathbf{X} , y se actualiza de a una columna del diccionario cada vez, manteniendo fijas todas las columnas en el diccionario \mathbf{W} a excepción de una, \mathbf{w}_k , y los coeficientes que le corresponden, es decir la k -ésima fila en \mathbf{X} , la cual se denota como x_T^k , con el objetivo de encontrar una nueva columna $\tilde{\mathbf{w}}_k$ que reduzca el error cuadrático medio.

Tabla 1. Descripción algoritmo K-SVD.

Algoritmo K-SVD	
Entradas: Señales de entrada $Y = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N] \in \mathbb{R}^{n \times N}$	
1:	Inicializar el diccionario $W^0 \in \mathbb{R}^{n \times K}$ con columnas normalizadas.
2:	Asignar $j = 1$
3:	repetir
3:	Codificación dispersa: Usar el algoritmo OMP para obtener los vectores x_i para cada señal y_i aproximando la solución de:
	$\min_{x_i} \ \mathbf{y}_i - Wx_i\ _2^2 \text{ sujeto a } \ x_i\ _0 \leq s \quad \forall i$
3:	Actualización del diccionario: Actualizar cada columna $k = 1, 2, \dots, K$ en $W^{(j-1)}$ al:
	<ul style="list-style-type: none"> • Definir el conjunto de señales que usa este átomo, $d_k = \{i 1 \leq i \leq N, x_T^k(i) \neq 0\}$ • Calcular la matriz de error de representación general, E_K
	$E_K = Y - \sum_{j \neq k} d_k x_T^j$
	<ul style="list-style-type: none"> • Limitar a E_K al escoger solo las columnas correspondientes a d_k para así obtener E_k^R. • Aplicar la descomposición en valores singulares a E_k^R para obtener $E_k^R = U\Delta V^T$. Luego se actualiza la columna del diccionario con $w_k = u_1$, y el vector de coeficientes con $x_T^k = v_1 \Delta(1,1)$.
4:	Fin repetir
Salida: Diccionario W^k	

La Tabla 1 describe el algoritmo K-SVD en detalle.

3.2.2 OMP. Orthogonal Matching Pursuit (OMP) es un algoritmo que busca encontrar la representación dispersa, x , de una señal y , dado un diccionario W [41]. Este algoritmo construye una aproximación a través de un proceso iterativo

en el cual en cada iteración se calcula la solución óptima local. Esto se hace con el fin de que esta secuencia de soluciones localmente óptimas conduzcan a la solución óptima global. Esto en general no es el caso pero bajo ciertas condiciones resulta en la solución óptima. La Tabla 2 describe el algoritmo K-SVD en detalle.

Tabla 2. Descripción algoritmo OMP.

Algoritmo OMP
<p>Entradas: Señal de entrada \mathbf{y}, diccionario \mathbf{W}, valor límite de error ϵ_0.</p> <ol style="list-style-type: none"> 1: Inicializar $k = 0$ 2: Asignar la solución inicial $\mathbf{x}^0 = 0$ 3: Asignar el residual inicial $\mathbf{r}^0 = \mathbf{y} - \mathbf{W}\mathbf{x}^0 = \mathbf{y}$ 4: Asignar el soporte de la solución inicial $S^0 = \text{soporte}\{\mathbf{x}^0\} = \emptyset$ 5: repetir 5: Asignar $k = k + 1$ 5: Calcular los errores $\epsilon(j) = \min_{z_j} \ \mathbf{a}_j z_j - \mathbf{r}^{k-1}\ _2^2$ para todo j usando $z_j^* = \mathbf{a}_j^T \mathbf{r}^{k-1} / \ \mathbf{a}_j\ _2^2$. 5: Encontrar el minimizador de j_0, de $\epsilon(j)$: $\forall j \notin S^{k-1}, \epsilon(j_0) \leq \epsilon(j)$ y actualizar $S^k = S^{k-1} \cup \{j_0\}$. 5: Calcular $\mathbf{x}^k = \text{argmin}_x \ \mathbf{W}\mathbf{x} - \mathbf{y}\$ sujeto a $\text{Soporte}\{\mathbf{x}\} = S^k$. 5: Actualizar el residual como $\mathbf{r}^k = \mathbf{y} - \mathbf{W}\mathbf{x}^k$. 6: Hasta que $\ \mathbf{r}^k\ ^2 \leq \epsilon_0$ <p>Salida: Representación dispersa \mathbf{x}^k</p>

3.2.3 Resumen. Al evaluar la complejidad computacional del algoritmo se tiene en cuenta el costo total por iteración de los algoritmos OMP y K-SVD, lo cual requiere $O(sNn^2)$ operaciones. En donde s es el nivel de dispersión de la matriz de coeficientes de representación, N es el número de señales de entrada, y n es el tamaño de las señales de entrada.

Con base en la formulación presentada para este algoritmo se planteó el entrenamiento de un diccionario W a partir de $Y = [y_1, y_2, \dots, y_N]$ señales de entrada con el fin de obtener $X = [x_1, x_2, \dots, x_N]$ representaciones dispersas de las mismas en dicho diccionario.

4. SIMULACIONES Y RESULTADOS

Con base en la formulación de los algoritmos desarrollada en el capítulo anterior se planteó el entrenamiento de diccionarios para la representación de imágenes hiperespectrales de diversas formas tomando en cuenta las limitaciones de cada uno de los algoritmos. Es por esto que con el fin obtener señales de entrada para los algoritmos que permitieran lidiar con problemas tales como los relacionados con velocidad de los algoritmos, tamaño de memoria requerido y otros, se tomaron parches de las imágenes hiperespectrales en las simulaciones realizadas.

Debido a la naturaleza de las imágenes hiperespectrales en las cuales existen tres dimensiones para representarlas, se pueden escoger diferentes tipos de parches a partir de una imagen de este tipo. En este trabajo se usan tres tipos de parches de imágenes hiperespectrales denominados parches bidimensionales, tridimensionales, y espectral. Los parches bidimensionales consisten en tomar un tamaño de píxeles predefinido a lo largo de las dimensiones espaciales de cada una de las bandas espectrales de la imagen hiperespectral. Los parches tridimensionales son aquellos en los cuales se toman un tamaño de píxeles predefinido a lo largo de las tres dimensiones de la imagen hiperespectral. Por último el parche espectral consiste en tomar todos los píxeles a lo largo de la dimensión espectral de un punto fijo en las dimensiones espaciales.

La implementación de todos los algoritmos se realizó en Matlab v.8.0.0.783 (R2012b). Las simulaciones realizadas fueron ejecutadas en un procesador Intel Core i7-4770 a 3.40 GHz con una memoria RAM de 32 GB, haciendo uso de un sistema operativo Windows 7 de 64 bits.

Para las pruebas de los algoritmos se usó una sección de una imagen hiperespectral de la escena Cuprite tomada por el sensor hiperespectral AVIRIS (Airborne Visible Infrared Imaging Spectrometer)[42], en la cual a su vez cada

banda espectral se compone del promedio de 4 bandas espectrales adjuntas de la imagen hiperespectral original de la escena. Esta imagen hiperespectral de prueba se compone de 64x64x52 pixeles.

Con el fin de comparar los resultados con otras bases de representación comunes, se utilizaron las transformadas de coseno discreta (DCT) bidimensional (2D), Wavelet Symmlet 2D, y el producto Kronecker de una base DCT con Wavelet Symmlet 2D. A continuación se presentan resultados numéricos de las simulaciones realizadas para los algoritmos propuestos.

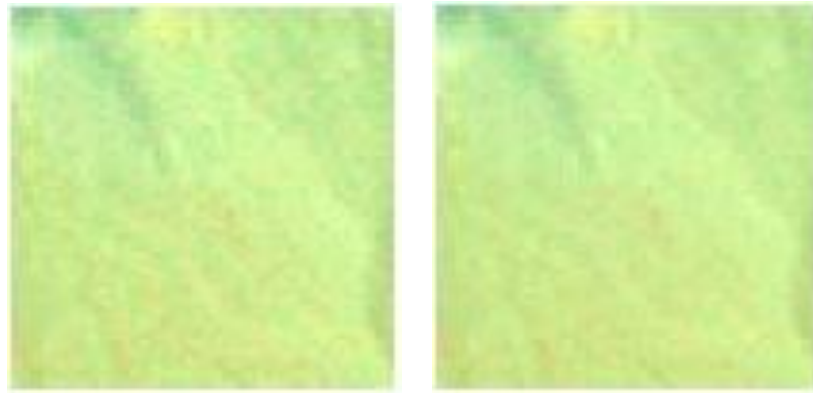
4.1 APRENDIZAJE DE TRANSFORMACIONES PARA SEÑALES DISPERSAS

Las simulaciones realizadas con este algoritmo fueron realizadas con parches de imágenes hiperespectrales como señales de entrenamiento. El algoritmo realiza 128 iteraciones, dentro de las cuales también realiza 128 iteraciones del algoritmo de gradiente conjugado, en cada paso de actualización del diccionario se utiliza un paso fijo de 10^{-8} . El algoritmo se inicializa utilizando como matriz de transformación inicial W^0 , una matriz de transformación ampliamente conocida, la transformada DCT 2D.

4.1.1 Parches bidimensionales. Dentro de esta simulación se extraen parches a todas las bandas espectrales de la imagen hiperespectral de prueba, los cuales no se superponen, de 8x8 pixeles. Las medias de los parches se restan para el entrenamiento y luego se suman sólo para efectos de visualización. En este caso la matriz $Y = [y_1, y_2, \dots, y_N]$ tiene $N = 3584$ señales de entrenamiento en donde cada parche se representa como un vector y_i para $1 \leq i \leq N$.

El problema de la ecuación (10) se resuelve con los parámetros $\lambda = \mu = 1$, y un nivel de dispersión $s = 13$ de las columnas de la matriz de coeficientes de representación. El diccionario W se inicializó con una matriz DCT 2D de 64x64.

Figura 5. (a) Imagen hiperespectral original mapeada a un perfil RGB. (b) Imagen hiperespectral obtenida por medio del diccionario entrenado con el algoritmo de aprendizaje de transformaciones para señales dispersas con parches bidimensionales mapeada a un perfil RGB.



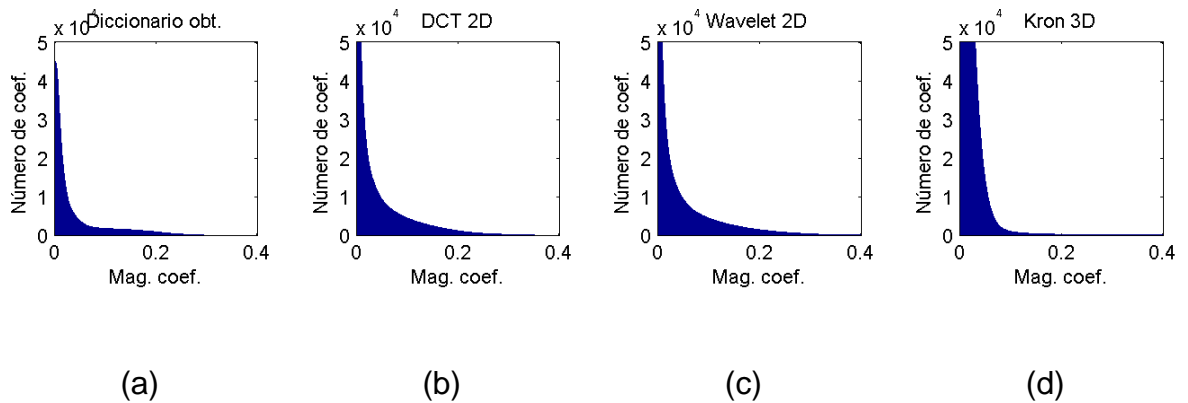
(a)

(b)

En la figura 5 (a) se muestra la imagen hiperespectral original usada, en (b) se muestra la imagen hiperespectral obtenida por medio del diccionario W entrenado, y se puede observar la similitud entre las mismas. Además el desempeño calculado del algoritmo mediante el PSNR espacial promedio de la imagen hiperespectral obtenida es de 36.33 dB y el PSNR espectral promedio es de 31.55 dB.

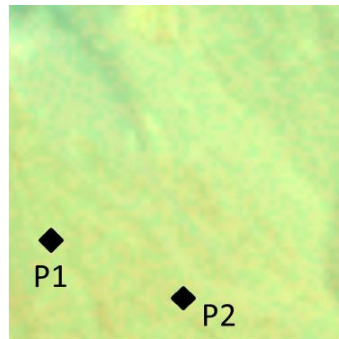
El tiempo de ejecución promedio que se obtuvo para la simulación propuesta con el algoritmo fue de 71.28 segundos.

Figura 6. Representación dispersa de los coeficientes de, (a) la imagen hiperespectral obtenida mediante el algoritmo de aprendizaje de transformaciones para señales dispersas con parches bidimensionales, (b)-(d) la imagen hiperespectral original representada en DCT 2D, Wavelet Symmlet 2D, y el producto Kronecker de una base DCT con Wavelet Symmlet 2D respectivamente.

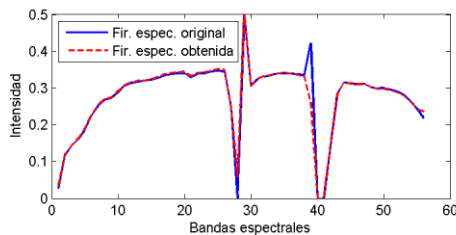


En la figura 6 (a) se puede apreciar el conteo del número de coeficientes con determinada magnitud para la representación de la imagen obtenida a través del diccionario entrenado. En (b), (c) y (d) se puede observar el conteo de coeficientes de representación de la imagen original en el dominio de las transformadas DCT 2D, Wavelet Symmlet 2D, y el producto Kronecker de una base DCT con Wavelet Symmlet 2D respectivamente. A partir de esta figura se puede observar que existen pocos coeficientes de representación de la imagen de prueba en el diccionario entrenado que tienen una alta magnitud, y por lo tanto contienen la mayoría de la energía de la imagen, esto en un nivel superior a los coeficientes de representación del resto de las transformadas consideradas.

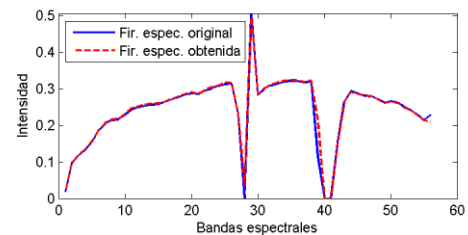
Figura 7. (a) Imagen hiperespectral original mapeada a un perfil RGB que muestra los puntos P1 y P2 seleccionados, (b) y (c) firma espectral de los puntos P1 y P2 de la imagen original y de la obtenida mediante el algoritmo de aprendizaje de transformaciones para señales dispersas con parches bidimensionales.



(a)



(b)



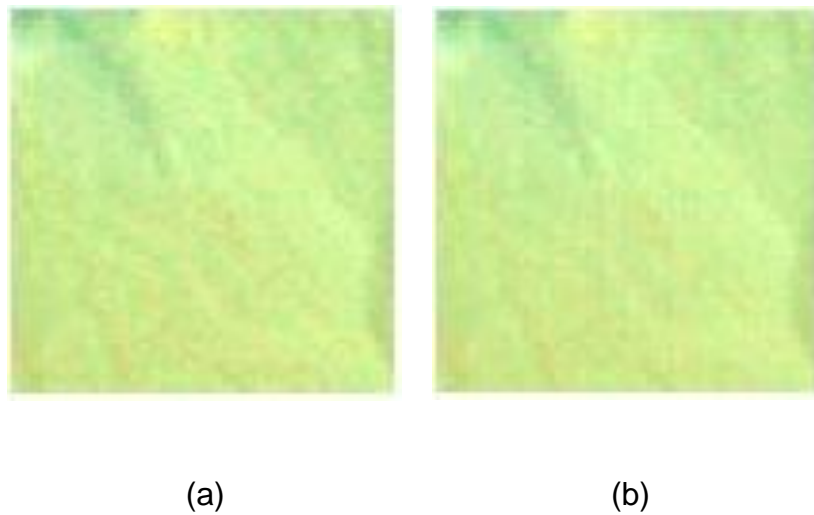
(c)

En la figura 7 (a) se puede apreciar la imagen hiperespectral original con 2 puntos aleatorios (P1 y P2) a los cuales se les obtiene la firma espectral. En (b) y (c) se pueden apreciar las firmas espectrales de la imagen hiperespectral original y la de la imagen hiperespectral obtenida mediante el diccionario algoritmo entrenado de los puntos P1 y P2 respectivamente, y se puede observar a partir de la figura que las firmas espectrales de los dos puntos tomados de la imagen hiperespectral reconstruida se asemejan a las pertenecientes a la imagen hiperespectral original.

4.1.2 Parches tridimensionales. Dentro de esta simulación se extraen parches de la imagen hiperespectral de prueba que no se superponen de $4 \times 4 \times 4$ píxeles. Las medias de los parches se restan para el entrenamiento y luego se suman sólo para efectos de visualización. En este caso la matriz $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]$ tiene $N = 3584$ señales de entrenamiento en donde cada parche se representa como un vector \mathbf{y}_i para $1 \leq i \leq N$.

El problema de la ecuación (10) se resuelve con los parámetros $\lambda = \mu = 1.6 \times 10^5$, y un nivel de dispersión $s = 13$ de las columnas de la matriz de coeficientes de representación. El diccionario \mathbf{W} se inicializó con una matriz DCT de 64×64 .

Figura 8. (a) Imagen hiperespectral original mapeada a un perfil RGB. (b) Imagen hiperespectral obtenida por medio del diccionario entrenado con el algoritmo de aprendizaje de transformaciones para señales dispersas con parches tridimensionales mapeada a un perfil RGB.

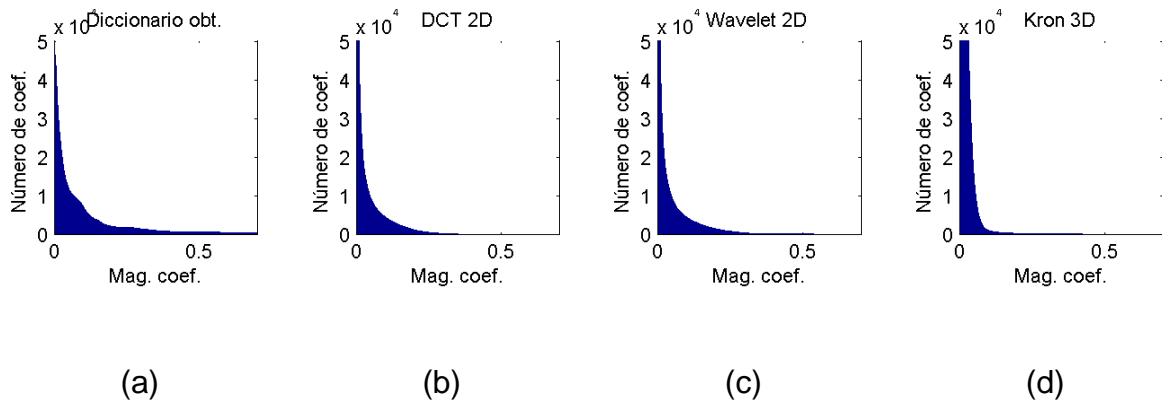


La figura 8 (a) muestra la imagen hiperespectral original usada, en (b) se muestra la imagen hiperespectral obtenida por medio del diccionario \mathbf{W} entrenado, y se puede observar la similitud entre las mismas. Además el desempeño calculado del

algoritmo mediante el PSNR espacial promedio de la imagen hiperespectral obtenida es de 31.79 dB y el PSNR espectral promedio es de 28.32 dB.

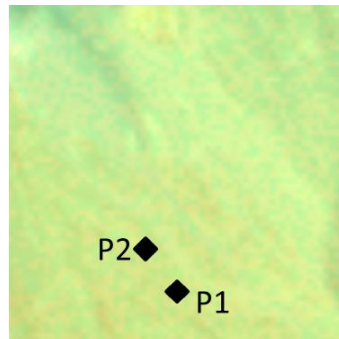
El tiempo de ejecución promedio que se obtuvo para la simulación propuesta con el algoritmo fue de 75.74 segundos.

Figura 9. Representación dispersa de los coeficientes de, (a) la imagen hiperespectral obtenida mediante el algoritmo de aprendizaje de transformaciones para señales dispersas con parches tridimensionales, (b)-(d) la imagen hiperespectral original representada en DCT 2D, Wavelet Symmlet 2D, y el producto Kronecker de una base DCT con Wavelet Symmlet 2D respectivamente.

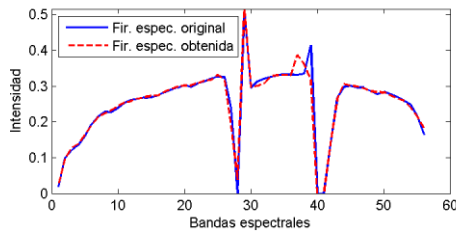


En la figura 9 (a) se puede apreciar el conteo del número de coeficientes con determinada magnitud para la representación de la imagen obtenida a través del diccionario entrenado. En (b), (c) y (d) se puede observar el conteo de coeficientes de representación de la imagen original en el dominio de las transformadas DCT 2D, Wavelet Symmlet 2D, y el producto Kronecker de una base DCT con Wavelet Symmlet 2D respectivamente. A partir de esta figura se puede observar que existen pocos coeficientes de representación de la imagen de prueba en el diccionario entrenado que tienen una alta magnitud, y por lo tanto contienen la mayoría de la energía de la imagen, esto en un nivel similar a los coeficientes de representación del resto de las transformadas consideradas.

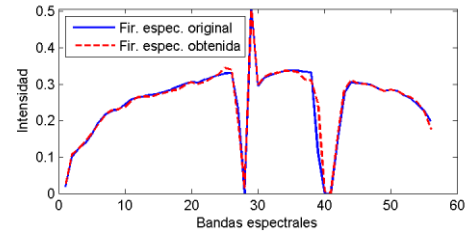
Figura 10. (a) Imagen hiperespectral original mapeada a un perfil RGB que muestra los puntos P1 y P2 seleccionados, (b) y (c) firma espectral de los puntos P1 y P2 de la imagen original y de la obtenida mediante el algoritmo de aprendizaje de transformaciones para señales dispersas con parches tridimensionales.



(a)



(b)



(c)

En la figura 10 (a) se puede apreciar la imagen hiperespectral original con 2 puntos aleatorios (P1 y P2) a los cuales se les obtiene la firma espectral. En (b) y (c) se puede apreciar las firmas espectrales de la imagen hiperespectral original y la de la imagen hiperespectral obtenida mediante el diccionario algoritmo entrenado de los puntos P1 y P2 respectivamente, y se puede observar a partir de la figura que las firmas espectrales de los dos puntos tomados de la imagen hiperespectral reconstruida se asemejan, a pesar de algunos errores de aproximación, a las pertenecientes a la imagen hiperespectral original.

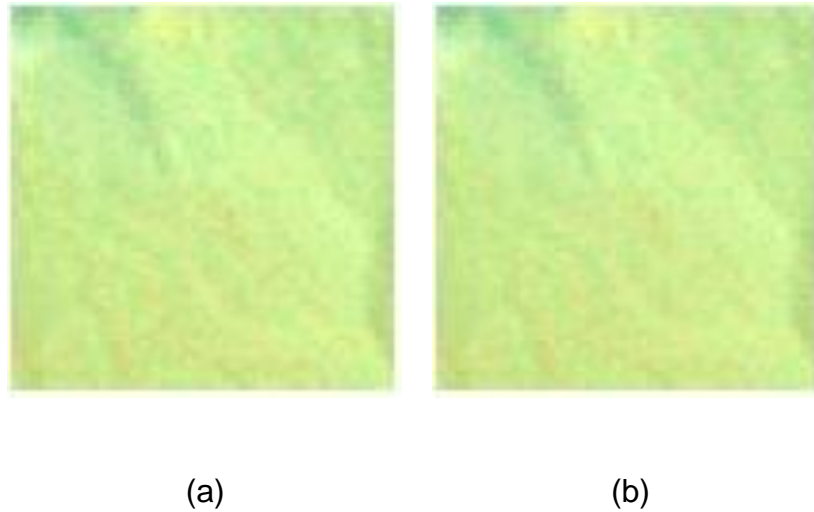
4.2 K-SVD Y OMP

Las simulaciones realizadas con este algoritmo fueron realizadas de dos formas: con parches bidimensionales y tridimensionales de imágenes hiperespectrales como señales de entrenamiento, y con parches espectrales de firmas espectrales de materiales. El número máximo de iteraciones del algoritmo fue establecido en 128. Las dimensiones del diccionario son $n \times K$ en donde $K = 64$, y n es el tamaño de las señales de entrenamiento, además el diccionario fue inicializado con las señales de entrada.

4.2.1 Parches bidimensionales. Dentro de esta simulación se extraen parches de la imagen hiperespectral de prueba que no se sobreponen de 8x8 píxeles. Es por esto que en este caso la matriz $Y = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]$ tiene $N = 3584$ señales de entrenamiento en donde cada parche se representa como un vector \mathbf{y}_i para $1 \leq i \leq N$.

El problema de la ecuación (14) se resuelve con un nivel de dispersión $s = 13$ de las columnas de la matriz de coeficientes de representación.

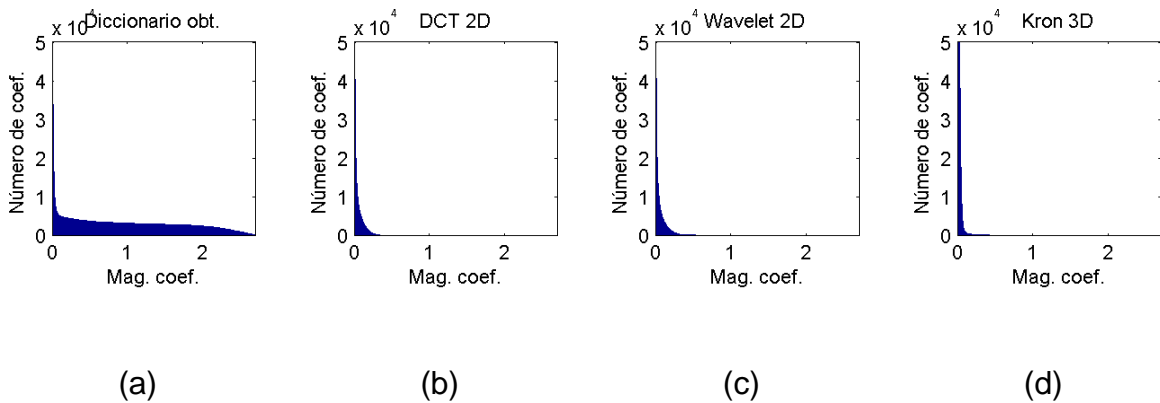
Figura 11. (a) Imagen hiperespectral original mapeada a un perfil RGB. (b) Imagen hiperespectral obtenida por medio del diccionario entrenado con el algoritmo K-SVD y OMP con parches bidimensionales mapeada a un perfil RGB.



La figura 11 (a) muestra la imagen hiperespectral original usada, en (b) se muestra la imagen hiperespectral obtenida por medio del diccionario W entrenado, y se puede observar la similitud entre las mismas. Además el desempeño calculado del algoritmo mediante el PSNR espacial promedio de la imagen hiperespectral obtenida es de 35.53 dB y el PSNR espectral promedio es de 35.72 dB.

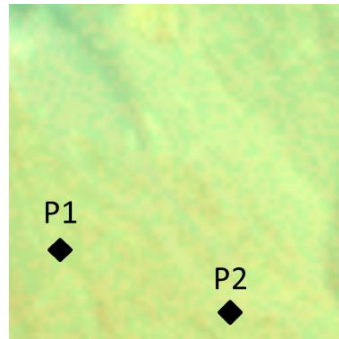
El tiempo de ejecución promedio que se obtuvo para la simulación propuesta con el algoritmo fue de 266.42 segundos.

Figura 12. Representación dispersa de los coeficientes de, (a) la imagen hiperespectral obtenida mediante el algoritmo K-SVD y OMP con parches bidimensionales, (b)-(d) la imagen hiperespectral original representada en DCT 2D, Wavelet Symmlet 2D, y el producto Kronecker de una base DCT con Wavelet Symmlet 2D respectivamente.

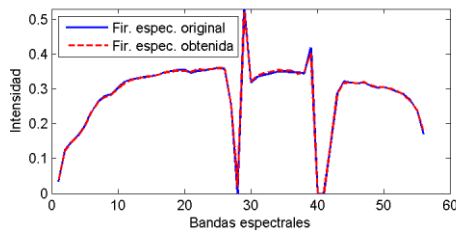


En la figura 12 (a) se puede apreciar el conteo del número de coeficientes con determinada magnitud para la representación de la imagen obtenida a través del diccionario entrenado. En (b), (c) y (d) se puede observar el conteo de coeficientes de representación de la imagen original en el dominio de las transformadas DCT 2D, Wavelet Symmlet 2D, y el producto Kronecker de una base DCT con Wavelet Symmlet 2D respectivamente. A partir de esta figura se puede observar que existen muchos coeficientes de representación de la imagen de prueba en el diccionario entrenado que tienen una alta magnitud, y por lo tanto contienen la mayoría de la energía de la imagen y son indispensables para su representación, lo cual no ocurre en el resto de las transformadas consideradas en donde la mayoría de la energía de la imagen está concentrada en muchos menos coeficientes.

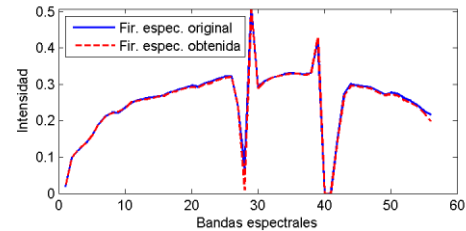
Figura 13. (a) Imagen hiperespectral original mapeada a un perfil RGB que muestra los puntos P1 y P2 seleccionados, (b) y (c) firma espectral de los puntos P1 y P2 de la imagen original y de la obtenida mediante el algoritmo K-SVD y OMP con parches bidimensionales.



(a)



(b)



(c)

En la figura 13 (a) se puede apreciar la imagen hiperespectral original con 2 puntos aleatorios (P1 y P2) a los cuales se les obtiene la firma espectral. En (b) y (c) se puede apreciar las firmas espectrales de la imagen hiperespectral original y la de la imagen hiperespectral obtenida mediante el diccionario algoritmo entrenado de los puntos P1 y P2 respectivamente. A partir de la figura se puede observar que las firmas espectrales de los dos puntos tomados de la imagen hiperespectral reconstruida se asemejan, a pesar de algunos errores de aproximación, a las pertenecientes a la imagen hiperespectral original.

4.2.2 Parches tridimensionales. Dentro de esta simulación se extraen parches de la imagen hiperespectral de prueba que no se superponen de $4 \times 4 \times 4$ píxeles. Es por esto que en este caso la matriz $Y = [y_1, y_2, \dots, y_N]$ tiene $N = 3584$ señales de entrenamiento en donde cada parche se representa como un vector y_i para $1 \leq i \leq N$.

El problema de la ecuación (14) se resuelve con un nivel de dispersión $s = 19$ de las columnas de la matriz de coeficientes de representación.

Figura 14. (a) Imagen hiperespectral original mapeada a un perfil RGB. (b) Imagen hiperespectral obtenida por medio del diccionario entrenado con el algoritmo K-SVD y OMP con parches tridimensionales mapeada a un perfil RGB.



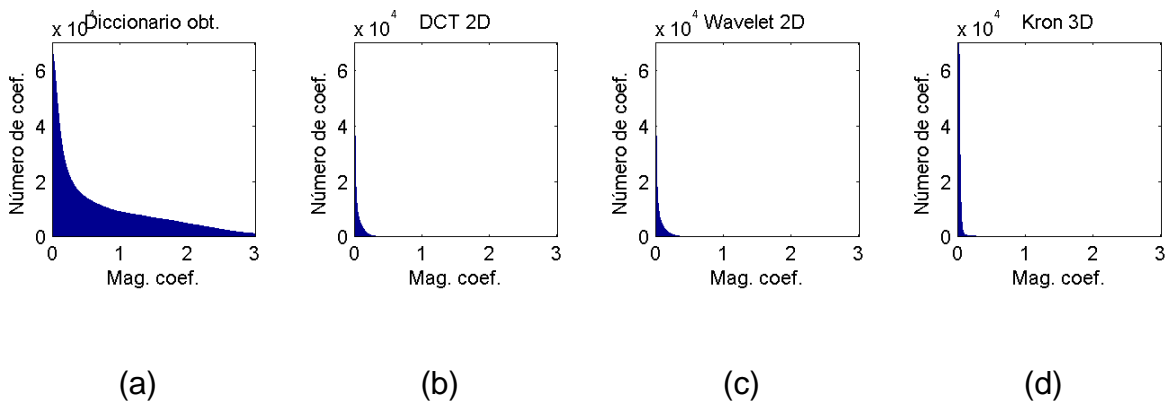
(a)

(b)

La figura 14 (a) muestra la imagen hiperespectral original usada, en (b) se muestra la imagen hiperespectral obtenida por medio del diccionario W entrenado, y se puede observar la similitud entre las mismas. Además el desempeño calculado del algoritmo mediante el PSNR espacial promedio de la imagen hiperespectral obtenida es de 37.82 dB y el PSNR espectral promedio es de 37.68 dB.

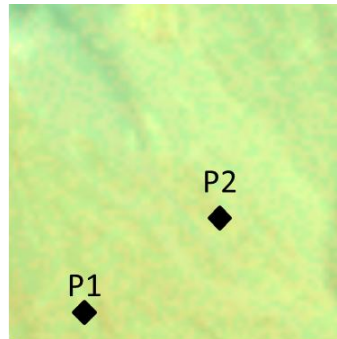
El tiempo de ejecución promedio que se obtuvo para la simulación propuesta con el algoritmo fue de 735.84 segundos.

Figura 15. Representación dispersa de los coeficientes de, (a) la imagen hiperespectral obtenida mediante el algoritmo K-SVD y OMP con parches tridimensionales, (b)-(d) la imagen hiperespectral original representada en DCT 2D, Wavelet Symmlet 2D, y el producto Kronecker de una base DCT con Wavelet Symmlet 2D respectivamente.

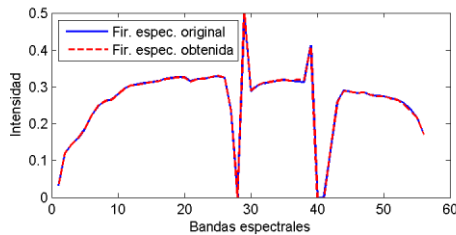


En la figura 15 (a) se puede apreciar el conteo del número de coeficientes con determinada magnitud para la representación de la imagen obtenida a través del diccionario entrenado. En (b), (c) y (d) se puede observar el conteo de coeficientes de representación de la imagen original en el dominio de las transformadas DCT 2D, Wavelet Symmlet 2D, y el producto Kronecker de una base DCT con Wavelet Symmlet 2D respectivamente. A partir de esta figura se puede observar que existen muchos coeficientes de representación de la imagen de prueba en el diccionario entrenado que tienen una alta magnitud, y por lo tanto contienen la mayoría de la energía de la imagen y son indispensables para su representación, lo cual no ocurre en el resto de las transformadas consideradas en donde la mayoría de la energía de la imagen está concentrada en muchos menos coeficientes.

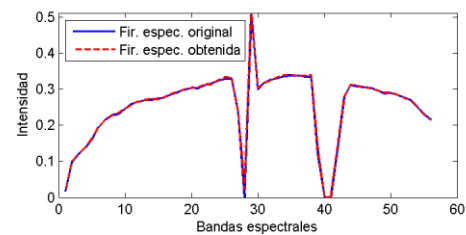
Figura 16. (a) Imagen hiperespectral original mapeada a un perfil RGB que muestra los puntos P1 y P2 seleccionados, (b) y (c) firma espectral de los puntos P1 y P2 de la imagen original y de la obtenida mediante el algoritmo K-SVD y OMP con parches tridimensionales.



(a)



(b)



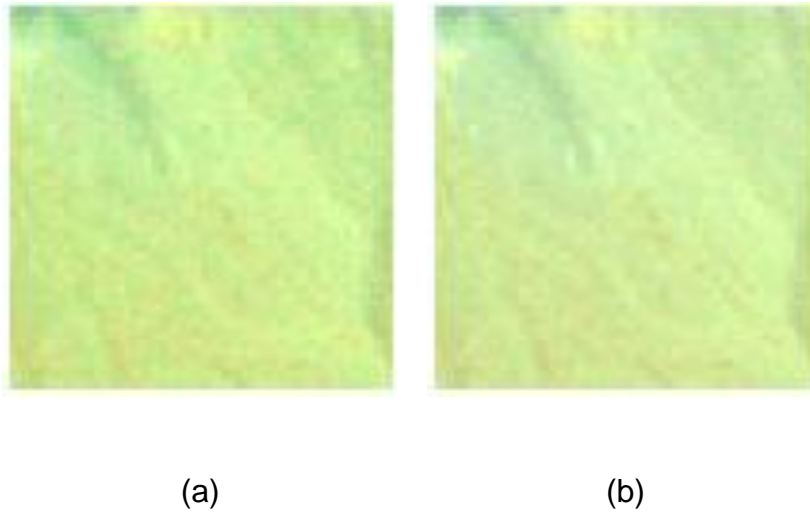
(c)

En la figura 16 (a) se puede apreciar la imagen hiperespectral original con 2 puntos aleatorios (P1 y P2) a los cuales se les obtiene la firma espectral. En (b) y (c) se puede apreciar las firmas espectrales de la imagen hiperespectral original y la de la imagen hiperespectral obtenida mediante el diccionario algoritmo entrenado de los puntos P1 y P2 respectivamente. A partir de la figura se puede observar que las firmas espectrales de los dos puntos tomados de la imagen hiperespectral reconstruida son muy parecidas a las pertenecientes a la imagen hiperespectral original.

4.2.3 Parches espectrales. Esta simulación se realizó basada en el trabajo de Chuarong Li et al. [43], en la cual se entrenó el diccionario a partir de firmas espectrales puras promediadas, endmembers, de minerales conocidos.

En la imagen hiperespectral de la escena Cuprite se encuentran identificados varios endmembers de minerales conocidos [44], para así entrenar un diccionario a partir de endmembers como señales de entrenamiento. Se usaron endmembers de la librería espectral del servicio geológico de los estados unidos (USGS)[45], que contiene endmembers de minerales tales como los Sulfatos, los Carbonatos, arcillas, el grupo de las arcillas de Caolinita, y otros minerales, entre los cuales se incluyen aquellos que fueron identificados como presentes en la escena Cuprite. Los endmembers se componen de $1 \times 1 \times 56$ pixeles, en donde en cada banda espectral se compone del promedio de 4 bandas espectrales adjuntas del endmember original obtenido de la librería espectral. Para la obtención de la representación dispersa de la imagen hiperespectral de la escena Cuprite en el diccionario entrenado se utilizó el algoritmo OMP con un nivel de dispersión $s = 13$.

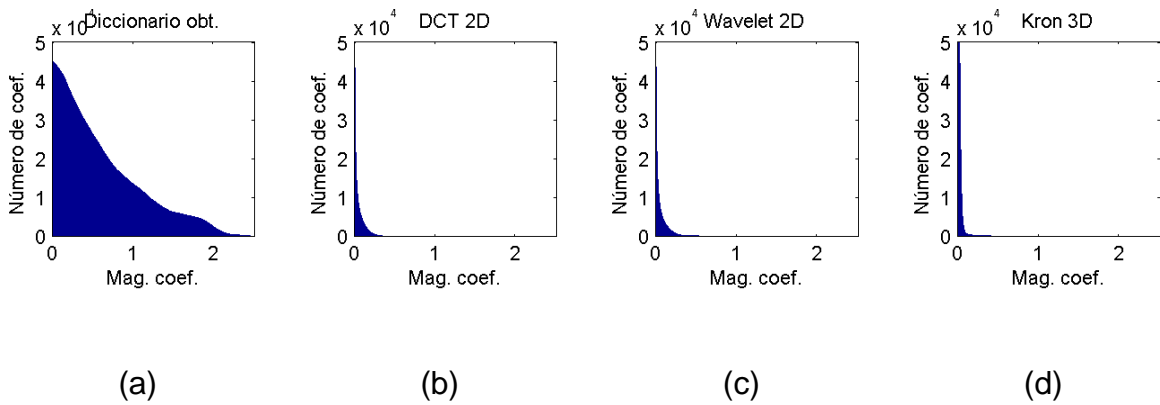
Figura 17. (a) Imagen hiperespectral original mapeada a un perfil RGB. (b) Imagen hiperespectral obtenida por medio del diccionario entrenado con el algoritmo K-SVD y OMP con parches espectrales mapeada a un perfil RGB.



La figura 17 (a) muestra la imagen hiperespectral original usada, en (b) se muestra la imagen hiperespectral obtenida por medio del diccionario W entrenado, y se puede observar la similitud entre las mismas. El desempeño del algoritmo calculado mediante el PSNR espacial promedio de la imagen hiperespectral obtenida es de 25.96 dB y el PSNR espectral promedio es de 13.41 dB.

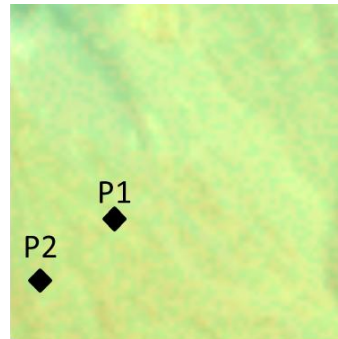
El tiempo de ejecución promedio que se obtuvo para la simulación propuesta con el algoritmo fue de 52.75 segundos.

Figura 18. Representación dispersa de los coeficientes de, (a) la imagen hiperespectral obtenida mediante el algoritmo K-SVD y OMP con parches espectrales, (b)-(d) la imagen hiperespectral original representada en DCT 2D, Wavelet Symmlet 2D, y el producto Kronecker de una base DCT con Wavelet Symmlet 2D respectivamente.

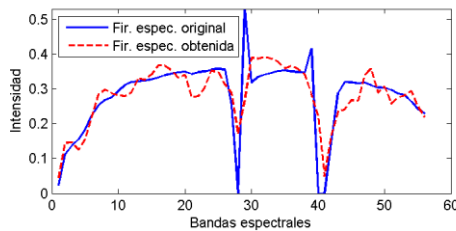


En la figura 18 (a) se puede apreciar el conteo del número de coeficientes con determinada magnitud para la representación de la imagen obtenida a través del diccionario entrenado. En (b), (c) y (d) se puede observar el conteo de coeficientes de representación de la imagen original en el dominio de las transformadas DCT 2D, Wavelet Symmlet 2D, y el producto Kronecker de una base DCT con Wavelet Symmlet 2D respectivamente. A partir de esta figura se puede observar que existen muchos coeficientes de representación de la imagen de prueba en el diccionario entrenado que tienen una alta magnitud, y por lo tanto contienen la mayoría de la energía de la imagen y son indispensables para su representación, lo cual no ocurre en el resto de las transformadas consideradas en donde la mayoría de la energía de la imagen está concentrada en muchos menos coeficientes.

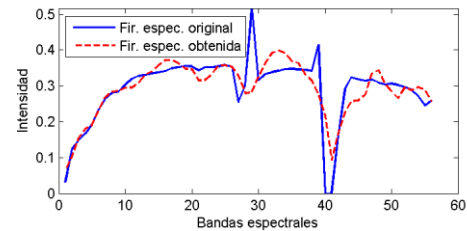
Figura 19. (a) Imagen hiperespectral original mapeada a un perfil RGB que muestra los puntos P1 y P2 seleccionados, (b) y (c) firma espectral de los puntos P1 y P2 de la imagen original y de la obtenida mediante el algoritmo K-SVD y OMP con parches espectrales.



(a)



(b)



(c)

En la figura 19 (a) se puede apreciar la imagen hiperespectral original con 2 puntos aleatorios (P1 y P2) a los cuales se les obtiene la firma espectral. En (b) y (c) se puede apreciar las firmas espectrales de la imagen hiperespectral original y la de la imagen hiperespectral obtenida mediante el diccionario algoritmo entrenado de los puntos P1 y P2 respectivamente. A partir de la figura se puede observar que las firmas espectrales de los dos puntos tomados de la imagen hiperespectral reconstruida son diferentes de las pertenecientes a la imagen hiperespectral original.

4.3 COMPARACIÓN DE LOS ALGORITMOS

Los algoritmos anteriormente expuestos fueron ejecutados para diferentes niveles de dispersión de la matriz de coeficientes de representación, con el objetivo de compararlos en términos de desempeño, y de nivel de energía contenida en los coeficientes de representación de los diferentes diccionarios obtenidos mediante los algoritmos y de otras transformadas.

Los parámetros de las simulaciones realizadas para tal fin se muestran en la tabla 3.

Tabla 3. Parámetros de las simulaciones de los algoritmos de aprendizaje de transformaciones para señales dispersas y K-SVD y OMP.

	Aprendizaje de transformaciones para señales dispersas		K-SVD & OMP		
	Parches bidimensionales	Parches tridimensionales	Parches bidimensionales	Parches tridimensionales	Parches espectrales
Iteraciones	128(Actualización de diccionario)	128(Actualización de diccionario)	128	128	128
	128(Proceso general)	128(Proceso general)			
Tamaño de señal de entrenamiento (n)	8x8=64	4x4x4=64	8x8=64	4x4x4=64	1x1x56=56
λ	1	160000	-	-	-
μ	1	160000	-	-	-
K	-	-	64	64	56

Tabla 4. Desempeño de los algoritmos de aprendizaje de transformaciones para señales dispersas para diferentes niveles de dispersión de la matriz de coeficientes de representación.

Nivel de dispersión (s)	Aprendizaje de transformaciones para señales dispersas									
	Parches bidimensionales					Parches tridimensionales				
	10%	20%	30%	40%	50%	10%	20%	30%	40%	50%
Porcentaje energía del 10% de los coeficientes del diccionario	100	82,85	75,16	70,34	67,68	100	89,4	82,03	76,05	72,01
PSNR espacial promedio (dB)	33,19	36,33	39,92	42,48	45,42	29,39	31,79	35,75	39,15	42,14
PSNR espectral promedio (dB)	29,03	31,55	33,50	36,07	38,52	25,49	28,32	30,36	32,95	35,41

Tabla 5. Desempeño de los algoritmos K-SVD y OMP para diferentes niveles de dispersión de la matriz de coeficientes de representación.

Nivel de dispersión (s)	K-SVD Y OMP														
	Parches bidimensionales					Parches tridimensionales					Parches espectrales				
	10%	20%	30%	40%	50%	10%	20%	30%	40%	50%	10%	20%	30%	40%	50%
Porcentaje energía del 10% de los coeficientes del diccionario	100	97,20	94,73	92,15	90,01	100	97,26	85,93	83,16	75,99	98,77	80,68	65,78	56,06	65
PSNR espacial promedio (dB)	30,76	35,53	38,42	39,72	44,12	19,13	13,48	37,82	42,51	44,35	24,12	25,96	24,84	24,53	23,72
PSNR espectral promedio (dB)	32,59	35,72	38,26	40,9	43,52	8,46	6,75	37,67	39,99	40,83	12,10	13,41	14,20	17,25	17,60

El desempeño de los algoritmos se detalla en las tablas 4 y 5 para diferentes niveles de dispersión de la matriz de coeficientes, en donde se busca que para esta matriz en cada una de sus columnas se garantice que contenga al menos un porcentaje de coeficientes entre el 10% y el 50% diferentes de cero. El desempeño de los algoritmos se mide en este caso mediante el cálculo de la energía contenida en el 10% de los coeficientes de la matriz de coeficientes de representación, además del PSNR espacial promedio, y el PSNR espectral promedio de la imagen hiperespectral obtenida. Se realizó a su vez el cálculo de la energía contenida en el 10% de los coeficientes de la imagen hiperespectral de

prueba en el dominio de las transformadas DCT 2D, Wavelet Symmlet 2D, y el producto Kronecker de una base DCT con Wavelet Symmlet 2D, y se obtuvieron valores de 73.73%, 78.95% y 39.44% respectivamente.

En estas tablas de las simulaciones realizadas se puede observar que el desempeño medido de acuerdo a las métricas establecidas del algoritmo K-SVD con parches bidimensionales tiene un comportamiento superior al del resto de algoritmos, esto debido a que la matriz de coeficientes de dicho diccionario entrenado tiene un alto nivel de energía a lo largo de los diferentes niveles de dispersión de la matriz, y a su vez un alto PSNR espacial promedio, y PSNR espectral promedio de la imagen hiperespectral obtenida.

5. CONCLUSIONES

Por medio de la investigación realizada de los diferentes métodos de aprendizaje de diccionarios en la literatura, se determinó para cada método propuesto en este trabajo de grado el modelo matemático para la transformación de imágenes espectrales en señales dispersas.

Partiendo del modelo matemático determinado para cada método se desarrollaron e implementaron los algoritmos correspondientes, que calculan una matriz de transformación para la representación dispersa de imágenes espectrales, y que permiten obtener su respectiva reconstrucción mediante el uso de métodos utilizados en la teoría de muestreo compresivo como el algoritmo OMP.

Se realizó la comparación de los resultados obtenidos al aplicar los diccionarios entrenados a imágenes espectrales con bases de representación como DCT, Wavelet, entre otras, y se encontró que el algoritmo K-SVD con parches bidimensionales ofrece un buen balance entre la concentración de la energía de los coeficientes en los diccionarios entrenados y el valor del PSNR promedio espacial y espectral con respecto a la imagen original.

El rendimiento obtenido a través del uso de los diccionarios entrenados por medio de los algoritmos propuestos demuestran su utilidad y un comportamiento comparable al obtenido tradicionalmente mediante transformaciones analíticas tradicionales como wavelet y DCT.

6. RECOMENDACIONES

Para trabajos futuros se recomienda en el desarrollo de algoritmos de aprendizaje de diccionarios para la representación dispersa de imágenes espectrales la optimización para el trabajo con señales de entrenamiento de alta dimensionalidad, el uso de eficiente de memoria y velocidad.

REFERENCIAS BIBLIOGRÁFICAS

- [1] A. Wagadarikar, R. John, R. Willett, y D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Appl. Opt.*, vol. 47, no. 10, pp. B44–B51, 2008.
- [2] H. Arguello y G. R. Arce, "Code aperture optimization for spectrally agile compressive imaging," *J. Opt. Soc. Am. A*, vol. 28, no. 11, pp. 2400–2413, Nov. 2011.
- [3] E. Candès y M. Wakin, "An Introduction To Compressive Sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, 2008.
- [4] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal l_1 -norm solution is also the sparsest solution," *Commun. Pure Appl. Math.*, vol. 59, no. 6, pp. 797–829, 2006.
- [5] S. Qaisar, R. M. Bilal, W. Iqbal, M. Naureen, y S. Lee, "Compressive sensing: From theory to applications, a survey," *J. Commun. Networks*, vol. 15, no. 5, pp. 443–456, 2013.
- [6] M. Davenport, M. Duarte, Y. Eldar, y G. Kutyniok, "Introduction to Compressed Sensing," in *Compressed Sensing: Theory and Applications*, Y. Eldar y G. Kutyniok, Eds. Cambridge University Press, 2012, pp. 1–68.
- [7] M. F. Duarte y R. G. Baraniuk, "Kronecker Compressive Sensing," *Image Process. IEEE Trans.*, vol. 21, no. 2, pp. 494–504, Feb. 2012.
- [8] E. J. Candes y T. Tao, "Decoding by linear programming," *Inf. Theory, IEEE Trans.*, vol. 51, no. 12, pp. 4203–4215, 2005.

- [9] C. Caiafa y A. Cichocki, "Multidimensional compressed sensing and their applications," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 3, no. 6, pp. 355–380, 2013.
- [10] S. Ravishankar y Y. Bresler, "Learning Sparsifying Transforms," *Signal Process. IEEE Trans.*, vol. 61, no. 5, pp. 1072–1086, 2013.
- [11] S. Ravishankar y Y. Bresler, "Learning sparsifying transforms for image processing," in *Image Processing (ICIP), 2012 19th IEEE International Conference on*, 2012, pp. 681–684.
- [12] M. Lustig, D. Donoho, y J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging.," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, Dec. 2007.
- [13] M. Lustig, J. M. Santos, D. L. Donoho, y J. M. Pauly, "k-t SPARSE: High frame rate dynamic MRI exploiting spatio-temporal sparsity," in *Proceedings of the 13th Annual Meeting of ISMRM*, 2006, p. 2420.
- [14] R. Rubinstein, A. M. Bruckstein, y M. Elad, "Dictionaries for Sparse Representation Modeling," *Proc. IEEE*, vol. 98, no. 6, pp. 1045–1057, 2010.
- [15] R. M. Willett, M. F. Duarte, M. A. Davenport, y R. G. Baraniuk, "Sparsity and Structure in Hyperspectral Imaging : Sensing, Reconstruction, and Target Detection," *Signal Process. Mag. IEEE*, vol. 31, no. 1, pp. 116–126, 2014.
- [16] P. J. Martínez, R. M. Pérez, A. Plaza, P. L. Aguilar, M. C. Cantero, y J. Plaza, "Endmember extraction algorithms from hyperspectral images," *Ann. Geophys.*, vol. 49, no. 1, 2006.
- [17] C. A. Bateson, G. P. Asner, y C. A. Wessman, "Endmember bundles: a new approach to incorporating endmember variability into spectral mixture

- analysis,” *Geosci. Remote Sensing, IEEE Trans.*, vol. 38, no. 2, pp. 1083–1094, Mar. 2000.
- [18] A. S. Charles, B. A. Olshausen, y C. J. Rozell, “Learning Sparse Codes for Hyperspectral Imagery,” *Sel. Top. Signal Process. IEEE J.*, vol. 5, no. 5, pp. 963–978, Sep. 2011.
- [19] W. L. Smith, D. K. Zhou, F. W. Harrison, H. E. Revercomb, A. M. Larar, H.-L. Huang, y B. Huang, “Hyperspectral remote sensing of atmospheric profiles from satellites and aircraft,” in *Proc. SPIE*, 2001, vol. 4151, pp. 94–102.
- [20] G. A. Shaw y H. K. Burke, “Spectral Imaging for Remote Sensing,” *Lincoln Lab. J.*, vol. 14, no. 1, pp. 3–28, 2003.
- [21] R. G. Baraniuk, “Compressive Sensing,” *Signal Process. Mag. IEEE*, vol. 24, no. 4, pp. 118–121, 2007.
- [22] M. Fornasier y H. Rauhut, “Compressive Sensing,” in *Handbook of Mathematical Methods in Imaging*, O. Scherzer, Ed. Springer, 2011, pp. 187–228.
- [23] G. R. Arce, D. J. Brady, L. Carin, H. Arguello, y D. S. Kittle, “Compressive Coded Aperture Spectral Imaging: An Introduction,” *Signal Process. Mag. IEEE*, vol. 31, no. 1, pp. 105–115, 2014.
- [24] H. Arguello, H. F. Rueda, y G. R. Arce, “Spatial super-resolution in code aperture spectral imaging,” in *Proc. SPIE*, 2012, vol. 8365, p. 83650A–83650A–6.
- [25] M. B. Wakin, J. N. Laska, M. F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. F. Kelly, y R. G. Baraniuk, “An Architecture for Compressive Imaging,” in

Proceedings of the International Conference on Image Processing (ICIP), 2006, pp. 1273–1276.

- [26] D. Galvis Carreño y H. Arguello Fuentes, “Transmittance analysis in coded aperture compressive Raman spectroscopy imaging,” in *Image, Signal Processing, and Artificial Vision (STSIVA), 2013 XVIII Symposium of*, 2013, pp. 1–5.
- [27] I. Tasic y P. Frossard, “Dictionary Learning,” *Signal Process. Mag. IEEE*, vol. 28, no. 2, pp. 27–38, 2011.
- [28] M. Aharon, M. Elad, y A. Bruckstein, “K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation,” *Signal Process. IEEE Trans.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
- [29] G. Chen y D. Needell, “Compressed sensing and dictionary learning,” in *Finite Frame Theory: A Complete Introduction to Overcompleteness*, 2015.
- [30] J. B. Allen y L. Rabiner, “A unified approach to short-time Fourier analysis and synthesis,” *Proc. IEEE*, vol. 65, no. 11, pp. 1558–1564, Nov. 1977.
- [31] S. G. Mallat, “A theory for multiresolution signal decomposition: the wavelet representation,” *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.
- [32] I. Daubechies, *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, 1992.
- [33] M. J. Bastiaans, “Gabor’s expansion of a signal into Gaussian elementary signals,” *Proc. IEEE*, vol. 68, no. 4, pp. 538–539, 1980.
- [34] A. J. E. M. Janssen, “Gabor representation of generalized functions,” *J. Math. Anal. Appl.*, vol. 83, no. 2, pp. 377–394, 1981.

- [35] E. J. Candès y D. L. Donoho, “New tight frames of curvelets and optimal representations of objects with piecewise C2 singularities,” *Commun. Pure Appl. Math.*, vol. 57, no. 2, pp. 219–266, 2004.
- [36] E. Candès, L. Demanet, D. Donoho, y L. Ying, “Fast Discrete Curvelet Transforms,” *Multiscale Model. Simul.*, vol. 5, no. 3, pp. 861–899, 2006.
- [37] M. Zhou, H. Chen, L. Ren, G. Sapiro, L. Carin, y J. W. Paisley, “Non-Parametric Bayesian Dictionary Learning for Sparse Image Representations,” in *Advances in Neural Information Processing Systems 22*, Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, y A. Culotta, Eds. Curran Associates, Inc., 2009, pp. 2295–2303.
- [38] J. Mairal, F. Bach, J. Ponce, y G. Sapiro, “Online Dictionary Learning for Sparse Coding,” in *Proceedings of the 26th Annual International Conference on Machine Learning*, 2009, pp. 689–696.
- [39] J. Mairal, F. Bach, J. Ponce, y G. Sapiro, “Online Learning for Matrix Factorization and Sparse Coding,” *J. Mach. Learn. Res.*, vol. 11, pp. 19–60, 2010.
- [40] R. Pytlak, *Conjugate Gradient Algorithms in Nonconvex Optimization*. Springer, 2009, p. 478.
- [41] Y. C. Pati, R. Rezaifar, y P. S. Krishnaprasad, “Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition,” in *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*, 1993, pp. 40–44 vol.1.
- [42] “AVIRIS Airborne Visible Infrared Imaging Spectrometer.” [En línea]. Disponible: <http://aviris.jpl.nasa.gov/>. [Último acceso:: 02-Feb-2015].

- [43] C. Li, L. Ma, Q. Wang, Y. Zhou, y N. Wang, "Construction of sparse basis by dictionary training for compressive sensing hyperspectral imaging," in *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International*, 2013, pp. 1442–1445.
- [44] R. N. Clark, G. A. Swayze, K. E. Livo, R. F. Kokaly, S. J. Sutley, J. B. Dalton, R. R. McDougal, y C. A. Gent, "Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems," *J. Geophys. Res. Planets*, vol. 108, no. E12, 2003.
- [45] R. N. Clark, G. A. Swayze, R. Wise, E. Livo, T. Hoefen, R. Kokaly, y S. J. Sutley, "USGS Digital Spectral Library 06," *U.S. Geological Survey, Digital Data Series 231*, 2007. [En línea]. Disponible: <http://speclab.cr.usgs.gov/spectral.lib06>. [Último acceso: 02-Feb-2015].

BIBLIOGRAFÍA

AHARON, M.; ELAD, M. y BRUCKSTEIN, A. K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation. En: IEEE Transactions on Signal Processing. 2006. vol. 54, no. 11, p. 4311–4322.

CANDÈS, E. y WAKIN, M. An Introduction To Compressive Sampling. En: IEEE Signal Processing Magazine. 2008. vol. 25, no. 2, p. 21–30.

CLARK, R.N., et al. Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems. En: Journal of Geophysical Research: Planets. 2003. vol. 108, no. E12, p. 1–44.

CLARK, R.N., et al. USGS Digital Spectral Library 06. U.S. Geological Survey [en línea]. Digital Data Series 231. [citado en 2015-02-02]. Disponible en internet: <URL:<http://speclab.cr.usgs.gov/spectral.lib06>>.

DAVENPORT, M., et al. Introduction to Compressed Sensing. En: Y. ELDAR y G. KUTYNIOK (eds.). Compressed Sensing: Theory and Applications. Cambridge University Press, 2012. pp. 1–68.

DUARTE, M.F. y BARANIUK, R.G. Kronecker Compressive Sensing. En: IEEE Transactions on Image Processing. 2012. vol. 21, no. 2, p. 494–504.

LI, C., et al. Construction of sparse basis by dictionary training for compressive sensing hyperspectral imaging. En: 2013 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). p. 1442–1445.

PATI, Y.C.; REZAIIFAR, R. y KRISHNAPRASAD, P.S. Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. En: Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on. p. 40–44.

PYTLAK, R. Conjugate Gradient Algorithms in Nonconvex Optimization. Springer, 2009. ISBN 978-3-540-85634-4.

RAVISHANKAR, S. y BRESLER, Y. Learning Sparsifying Transforms. En: IEEE Transactions on Signal Processing. 2013. vol. 61, no. 5, p. 1072–1086.

RUBINSTEIN, R., BRUCKSTEIN, A.M. y ELAD, M. Dictionaries for Sparse Representation Modeling. En: Proceedings of the IEEE. 2010. vol. 98, no. 6, p. 1045–1057.

TOSIC, I. y FROSSARD, P. Dictionary Learning. En: IEEE Signal Processing Magazine. 2011 vol. 28, no. 2, p. 27–38.

ZHOU, M., et al. Non-Parametric Bayesian Dictionary Learning for Sparse Image Representations. En: BENGIO, Y., et al. (eds.). Advances in Neural Information Processing Systems 22. Curran Associates, Inc., 2009. p. 2295–2303.