

**ALGORITMO PARA LA DETECCIÓN DE OBJETIVOS  
EN IMÁGENES ESPECTRALES UTILIZANDO  
MEDIDAS COMPRESIVAS**

DAVID ALBERTO BOADA SUPELANO

UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FISICO-MECÁNICAS  
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA  
BUCARAMANGA  
2015

# ALGORITMO PARA LA DETECCIÓN DE OBJETIVOS EN IMÁGENES ESPECTRALES UTILIZANDO MEDIDAS COMPRESIVAS

DAVID ALBERTO BOADA SUPELANO

Trabajo de grado para optar al título de :  
Ingeniero de sistemas

Director:

*Ms.(c)*, HÉCTOR MIGUEL VARGAS GARCÍA

Codirector:

*Ph.D.(c)*, *MPE*. HENRY ARGUELLO FUENTES

UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FÍSICO-MECÁNICAS  
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA  
BUCARAMANGA

2015



*No importa que tan talentosos, disciplinados, inteligentes o suertudos seamos, las cosas grandes nunca se logran sin el apoyo de alguien mas, por eso dedico mi trabajo a aquellas personas que a lo largo del trayecto estuvieron ahí, ayudándome, motivándome o simplemente acompañándome.*

*A mi madre, Lina Maria, precursora de mi existencia, mujer de la que aprendí que sin amor, trabajo, esfuerzo y perseverancia no se logran nuestros objetivos. Mis logros siempre serán suyos.*

*A mis hermanos, Pablo por ser mi amigo, mi apoyo, mi cómplice, y en algunos casos mi rival, no podría tener mejor compañero de armas. Y Julián, por ser la sonrisa que me espera al llegar a casa, mi gran motivación y lo mas cercano a un hijo que puedo tener. Todo mi amor para él.*

*A mi padrino Luis Eduardo, por ser el apoyo e inspiración que en momentos difíciles mas necesite, por estar siempre ahí, acompañándome, esto también es para él.*

*A mis abuelos y familia, sobretodo doña Flor, sin su cariño y apoyo nada habría sido posible.*

*A mis amigos y compañeros de estudio, en especial Miguel, Camilo y Kareth, por ser mi apoyo, motivación y una gran compañía, y a los demás muchachos del grupo, a Yuri y Ariolfo por ser un equipo de trabajo formidable y unos grandes amigos.*

*A todos ellos y a las demás personas que de una u otra forma me aconsejaron, guiaron, apoyaron y acompañaron. A ellos, toda mi gratitud.*

---

## **Agradecimientos**

El autor expresa su agradecimiento a Héctor Miguel Vargas, director de este proyecto por su asesoría académica, sus recomendaciones y su gran apoyo como tutor.

Al Profesor Henry Arguello Fuentes, por su acompañamiento y enseñanzas constantes, y por la oportunidad de trabajar en este proyecto.

# Índice general

<b>Introducción</b>	<b>14</b>
<b>1. CONCEPTOS BÁSICOS</b>	<b>16</b>
1.1. Imágenes hiperespectrales . . . . .	16
1.2. Espectrómetros . . . . .	17
1.3. Variabilidad espectral y píxeles mezclados . . . . .	18
1.4. Detección de objetivos en imágenes hiperespectrales . . . . .	20
1.5. Medidas de desempeño . . . . .	22
1.5.1. Curvas ROC ( <i>Receiver Operator Characteristics</i> ) . . . . .	22
1.6. Dispersión y teoría de muestreo compresivo . . . . .	23
<b>2. MÉTODOS TRADICIONALES DE DETECCIÓN DE OBJETIVOS</b>	<b>27</b>
2.1. Detectores de anomalías . . . . .	27
2.1.1. Detector de anomalías Reed-Xiaoli (RX) . . . . .	27
2.1.2. Detección de anomalías basada en subespacios . . . . .	28
2.1.3. Detección de anomalías Kernel Reed-Xiaoli (KRX) . . . . .	28
2.2. Detección con filtros ajustados ( <i>signature-based target detection</i> ) . . . . .	30
2.3. Detectores adaptativos basados en subespacios . . . . .	32
2.4. Detectores de coeficiente de probabilidad ( <i>Likelihood ratio test-based detectors</i> ) . . . . .	32
2.5. Detectores no lineales ( <i>Nonlinear detectors</i> ) . . . . .	33
2.6. Representación Dispersa para detección de objetivos . . . . .	34
2.6.1. Modelo de dispersión . . . . .	34
2.6.2. Asunciones de dispersión . . . . .	36
2.6.3. Construcción de diccionarios . . . . .	37
<b>3. MÉTODO PROPUESTO</b>	<b>38</b>
3.1. Detección de objetivos en imágenes hiperespectrales utilizando un modelo de dispersión . . . . .	38
3.2. Problema inverso . . . . .	42

---

<b>4. SIMULACIONES Y RESULTADOS</b>	<b>44</b>
4.1. Diseño de las aperturas codificadas . . . . .	46
4.2. Dataset RIT-CIS-DIRT . . . . .	47
4.3. Imagen del sensor Hyperion . . . . .	50
<b>5. Conclusiones, recomendaciones y trabajo futuro</b>	<b>53</b>
5.1. Conclusiones . . . . .	53
5.2. Recomendaciones y trabajo futuro . . . . .	53
<b>Referencias</b>	<b>54</b>

## Índice de figuras

1.1. Ejemplo de una imagen hiperespectral y su respectivo esquema tridimensional. . . . .	16
1.2. Representación del espectro electromagnético. . . . .	17
1.3. Proceso de adquisición de una imagen hiperespectral por un espectrómetro aéreo. . . . .	18
1.4. Pixeles mezclados y variabilidad espectral. . . . .	19
1.5. Funcionamiento de un detector espectral estándar. . . . .	20
1.6. Tipos de curvas ROC. . . . .	23
1.7. Representación matricial de la teoría muestreo compresivo. . . . .	24
1.8. Arquitectura CASSI . . . . .	26
2.1. Resultados de detección algoritmos RX y KRX . . . . .	30
2.2. El proceso SRC para detección de objetivos . . . . .	36
4.1. Imagen RGB del dataset self test. . . . .	44
4.2. Mapeo RGB de la imagen del sensor Hyperion. . . . .	45
4.3. Códigos de apertura . . . . .	47
4.4. Resultados de la detección D para la imagen RIT-CIS-DIRT . . . . .	49
4.5. Curvas ROC para la imagen RIT-CIS-DIRT . . . . .	49
4.6. Resultados de la detección D para la imagen del sensor Hyperion . . . . .	51
4.7. Curvas ROC para la imagen del sensor Hyperion . . . . .	52

# Índice de tablas

- 1.1. Matriz de confusión . . . . . 22
- 4.1. Desempeño detector tradicional . . . . . 46
- 4.2. Resultados experimentales para la imagen RIT-CIS-DIRT. . . . . 50
- 4.3. Resultados experimentales para la imagen Hyperion. . . . . 52

---

# Resumen

TITULO: ALGORITMO PARA LA DETECCIÓN DE OBJETIVOS EN IMÁGENES ESPECTRALES UTILIZANDO MEDIDAS COMPRESIVAS<sup>1</sup>

AUTOR: DAVID ALBERTO BOADA SUPELANO <sup>2</sup>

PALABRAS CLAVE: Imágenes espectrales, procesamiento de imágenes, muestreo compresivo, detección de objetivos.

## DESCRIPCIÓN:

La adquisición y procesamiento de imágenes espectrales involucra el manejo de grandes cantidades de información espectral multidimensional. En cada píxel de estas imágenes se representa la información del comportamiento espectral en una ubicación espacial, referente a varias longitudes de onda. Esta información puede ser usada para tareas como detección, clasificación, y extracción de características presentes en una escena.

En los últimos años se han desarrollado arquitecturas ópticas para la adquisición de información espectral de forma comprimida usando un conjunto reducido de mediciones codificadas por un modulador espacial. Estos avances han facilitado el proceso de muestreo y almacenamiento de información espectral. Sin embargo, el proceso de reconstrucción de las imágenes es una operación costosa tanto en tiempo como en recursos computacionales.

Aplicar el muestreo compresivo a la hora de procesar imágenes espectrales, disminuyendo el costo de adquisición y almacenamiento, es la principal motivación de este trabajo. Específicamente, se busca el diseño de un modelo de detección escasa que use mediciones comprimidas para detectar objetivos en escenas espectrales usando diccionarios preentrenados y las transformada wavelet. Un algoritmo de optimización heurístico es utilizado para la evaluación del modelo, con lo cual se determina si el píxel evaluado es o no un píxel objetivo

---

<sup>1</sup>Trabajo de Investigación

<sup>2</sup>Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería de Sistemas e Informática. Director, Héctor Miguel Vargas Garcia. Codirector, Profesor Henry Arguello Fuentes

---

## Abstract

TITLE: ALGORITHM FOR SPECTRAL IMAGE TARGET DETECTION FROM COMPRESSIVE MEASUREMENTS USING A JOINT SPARSITY MODEL AND A CASSI ARCHITECTURE <sup>1</sup>

AUTHOR: DAVID ALBERTO BOADA SUPELANO <sup>2</sup>

KEYWORDS: Spectral imaging, image processing, compressive sampling, target detection.

### DESCRIPTION:

Hyperspectral imaging entails data typically spanning hundreds of contiguous wavebands in a certain spectral range. Each spatial point in hyperspectral images is therefore represented by a vector whose entries correspond to the intensity on each spectral band.

These images enable object and feature detection, classification, or identification based on their spectral characteristics. Novel architectures have been developed for the acquisition of compressive spectral images with just a few coded aperture focal plane array measurements. However, spectral image reconstruction is a highly expensive computational process.

This work focuses on the development of a target detection approach in hyperspectral images, using directly compressive spectral measurements without first reconstructing the full data cube that represents the real image, using trained background and target dictionaries and the wavelet transform. Specifically, a sparsity-based target detection model that uses compressive measurement for the detection task is designed and tested using an optimization algorithm. Simulations show that it is possible to perform certain transformations to the dictionaries used in traditional target detection, in order to achieve an accurate image representation in the compressed subspace. Results suggest that the proposed model achieves a probability of detection between 98.92 % and 99.97 % using just the 40 % of the spectral information.

---

<sup>1</sup>Research Work

<sup>2</sup>School of Physical-Mechanical Engineering. Department of Systems Engineering and Informatics. Advisor, Héctor Miguel Vargas García, Co-advisor, Professor Henry Arguello Fuentes

## Introducción

El desarrollo de la óptica en las últimas décadas ha permitido que áreas como el sensado remoto suministren gran riqueza de información espacial y espectral con alta resolución. El aumento de la resolución espectral de las imágenes hiperespectrales y de los sensores infrarrojos ha impulsado al surgimiento de nuevos campos de aplicación e impuesto nuevos desafíos metodológicos a la hora de analizar la gran cantidad de información que se puede llegar a obtener. Las imágenes hiperespectrales, permiten la caracterización de objetos de interés presentes en una escena.

La clasificación de imágenes es uno de los problemas más importantes para diversas disciplinas científicas como aprendizaje de máquinas, procesamiento de imágenes, y visión artificial, sin embargo, en el procesamiento de imágenes hiperespectrales se deben tener en cuenta varios aspectos críticos. Por ejemplo, la gran cantidad de bandas espectrales y la falta de muestras de entrenamiento etiquetadas conducen al problema conocido como el efecto Hughes [1], el cual muestra que cuando aumenta la dimensionalidad de la información, la cantidad de espacio necesario para el almacenamiento y procesamiento de los datos aumenta exponencialmente. Se han propuesto numerosos aportes con el fin de tratar los problemas que conllevan la gran dimensionalidad de los datos, la variabilidad espacial de la información espectral, el alto costo de etiquetamiento y para mejorar la estabilidad numérica y algorítmica [2]. Estos trabajos aprovechan el hecho de que las señales naturales presentan cierta dispersión en una base específica, por tanto, pueden ser representadas por muy pocos coeficientes que contienen la información más relevante [3].

Las representaciones dispersas en visión artificial y reconocimiento de patrones pueden ser halladas en varios campos, incluyendo segmentación de movimiento [2], superresolución de imágenes [4], restauración de imágenes [5], [6], y tareas discriminativas como reconocimiento facial [7], reconocimiento de iris [8], clasificación de tumores [9], y clasificación de imágenes [10]. En estas aplicaciones, el uso de la dispersión de las señales como parámetro inicial ha producido resultados satisfactorios. De igual manera, la naturaleza dispersa de las imágenes espectra-

---

les puede explotarse en la clasificación de imágenes adquiridas mediante sistemas de muestreo compresivo. Estos sistemas requieren una cantidad mucho menor de mediciones para reconstruir la señal, en comparación con los sensores espectrales utilizados tradicionalmente [11], [12], [13].

El uso de mediciones comprimidas para procesar y analizar la información espectral y espacial de una imagen reduce el costo computacional y de almacenamiento. Es posible efectuar tareas de procesamiento de imágenes utilizando muestreo compresivo, evitando la costosa etapa de reconstrucción de la imagen, de esta forma, aplicaciones como la detección de objetivos podrían aprovechar las ventajas que sugiere la información espectral comprimida.

Con la detección de objetivos en imágenes hiperespectrales se busca determinar si un objeto particular, con una firma espectral específica, está o no presente en la escena, por lo general estos objetos, denominados objetivos, solo están presentes en un número significativamente menor de píxeles que los demás. La detección de objetivos por lo general implica la comparación de los datos obtenidos del detector contra un umbral específico, si el umbral es superado el objetivo estará presente en la imagen.

Este trabajo está enfocado en el diseño de un modelo de detección de objetivos en imágenes espectrales, que estima la representación dispersa de los píxeles de la imagen usando diccionarios preentrenados de imágenes. Adicionalmente, se implementa un algoritmo que determina si el píxel evaluado es o no un píxel de objetivo. El algoritmo propuesto está basado en un modelo de dispersión conjunta (*Joint sparsity model*), en donde cada píxel es aproximadamente representado por un pequeño conjunto de firmas espectrales de entrenamiento, pertenecientes a las firmas del diccionario, el cual está compuesto por dos subdiccionarios de firmas espectrales de objetivo y de fondo. El conjunto de coeficientes dispersos que representan a cada píxel en el diccionario de entrenamiento, pueden estimarse a partir de mediciones comprimidas adquiridas por un sistema de muestreo compresivo, resolviendo un problema de optimización con restricciones dispersas.

# Capítulo 1

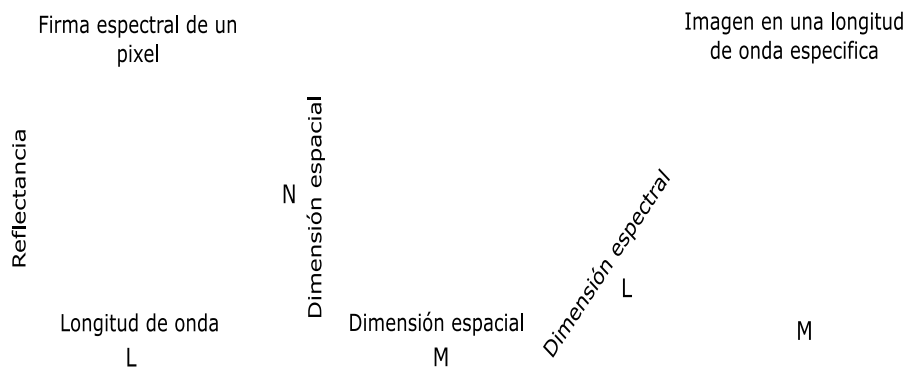
## CONCEPTOS BÁSICOS

### 1.1. Imágenes hiperespectrales

Una imagen hiperespectral se puede definir como un conjunto de imágenes bidimensionales concatenadas, que contienen información de una escena a lo largo del espectro electromagnético. Las imágenes hiperespectrales permiten la identificación de las propiedades de los elementos presentes en ésta. Debido a esto, se utilizan imágenes hiperespectrales en numerosas aplicaciones científicas y comerciales, tales como sensado remoto, manejo de recursos naturales, agricultura, explotación minera, y monitoreo medio ambiental. Sin embargo, el uso efectivo de las imágenes hiperespectrales requiere cierto conocimiento de sus limitaciones así como de ciertas técnicas de procesamiento.

Formalmente, se define a las imágenes hiperespectrales como una concatenación de imágenes bidimensionales que contienen información de la irradiancia a lo largo del espectro electromagnético. Matemáticamente, una imagen hiperespectral se puede expresar como  $f \in R^{(N \times M \times L)}$  donde  $N \times M$  representa la dimensión espacial y  $L$  es la profundidad espectral del cubo. La figura 1.1 ilustra esta definición.

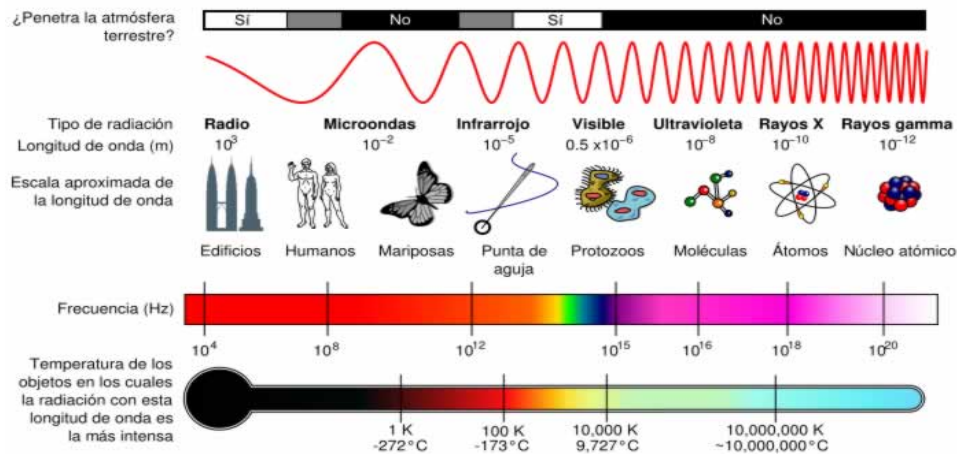
**Figura 1.1:** Ejemplo de una imagen hiperespectral y su respectivo esquema tridimensional.



Las imágenes hiperespectrales son captadas por espectrómetros, que son sen-

sores capaces de obtener información de una escena en un intervalo del espectro electromagnético, cuya representación se observa en la Figura 1.2. Estas mediciones permiten construir un espectro continuo para cada pixel de la imagen captada, proporcionando una cantidad más detallada de información que las imágenes tradicionales RGB [14], [15] y [16].

**Figura 1.2:** Representación del espectro electromagnético.



Después de ciertos ajustes y calibraciones atmosféricas y geométricas, las firmas espectrales de los pixeles de la imagen pueden ser comparadas con información espectral obtenida en el terreno o en laboratorio. De esta manera, las imágenes hiperespectrales permiten clasificar y caracterizar objetos de una escena basándose en sus propiedades espectrales [3], [17] y [18].

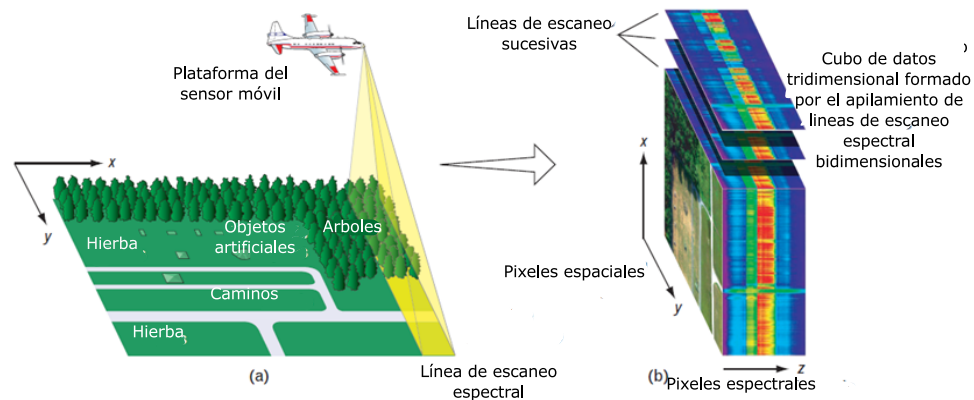
## 1.2. Espectrómetros

El desarrollo de los espectrómetros surgió de la mezcla de dos tecnologías distintas, la espectrometría y la adquisición de imágenes satelitales. La espectroscopia estudia el comportamiento de la luz que es emitida o reflejada por los materiales, y su variación de energía con respecto a la longitud de onda. En el campo de la óptica y el sensado remoto, la espectroscopía trabaja con el espectro de la luz solar que es reflejado por los materiales de la superficie terrestre.

Los espectrómetros (o espectro-radiómetros) son utilizados para hacer mediciones de campo o en laboratorio, de la luz reflejada por un material de prueba. Un espectrómetro contiene un elemento dispersor, un prisma o rejilla, que descompone la luz en muchas bandas espectrales angostas. La energía de cada banda puede ser medida por un detector.

Los sensores remotos están diseñados para medir la radiación reflejada desde distintas regiones adyacentes de la superficie terrestre. En muchos sensores digitales, se efectúan medidas de pequeñas áreas de forma secuencial, siguiendo un patrón geométrico constante, a medida que el sensor se desplaza, por lo cual se requiere una etapa de procesamiento posterior para integrar la información adquirida. La figura 1.3 muestra el proceso de adquisición de una imagen espectral por un sensor remoto.

**Figura 1.3:** Proceso de adquisición de una imagen hiperespectral por un espectrómetro aéreo.



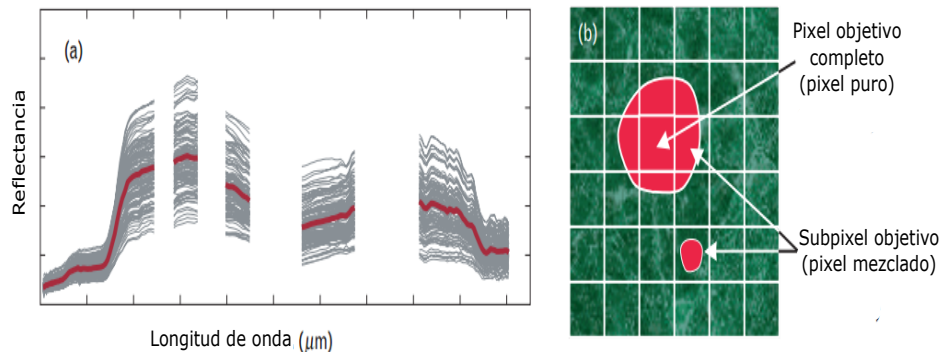
### 1.3. Variabilidad espectral y píxeles mezclados

La tarea principal en las aplicaciones basadas en imágenes hiperespectrales es la identificación de diferentes materiales a partir de su reflectancia, por ello, la idea de extraer una firma espectral que caracterice de manera única a cada material ampliamente aplicada. Sin embargo, los materiales en su forma natural no exhiben una firma espectral bien definida. La información espectral obtenida de muestras del mismo material nunca será idéntica, incluso en experimentos controlados, debido a variaciones en la superficie del material. La variabilidad es más profunda en apli-

caciones de sensado remoto, debido a los constantes cambios en las condiciones atmosféricas, el ruido inherente a los sensores, la composición de los materiales, la localización, los materiales circundantes, entre otros factores. Adicionalmente, materiales totalmente distintos pueden tener firmas espectrales muy parecidas y la variación en la iluminación y la calibración de los instrumentos son aspectos que no se solucionan con los códigos de corrección atmosférica. Algunos ejemplos de firmas espectrales de distintos materiales se muestran en la Fig. 1.4, ilustrando la variabilidad espectral que puede tener una imagen.

A pesar de la variabilidad espectral intrínseca y la ocasional falta de diferenciabilidad entre firmas espectrales de materiales distintos, las firmas espectrales son usadas ampliamente en sensado remoto y espectroscopía [19].

**Figura 1.4:** Pixeles mezclados y variabilidad espectral.



Otra complicación importante surge de la relación entre la resolución espacial de los sensores y la variabilidad espacial presente en la escena escaneada. El sensor integra la energía reflejada de todos los materiales en la superficie terrestre en una imagen, por tanto, dependiendo de la resolución espacial del sensor y de la variabilidad espacial de los materiales en cada pixel de la imagen, el resultado será una imagen hiperespectral compuesta de pixeles “puros” y pixeles “mezclados”. Los pixeles mezclados representan un desafío adicional a la hora de trabajar con información espectral, pues sus firmas espectrales no corresponden a ningún material bien definido.

El manejo de la variabilidad espectral y la composición de los pixeles mezclados, son de los problemas más desafiantes, tanto teórica como prácticamente en el cam-

po del procesamiento de imágenes hiperespectrales.

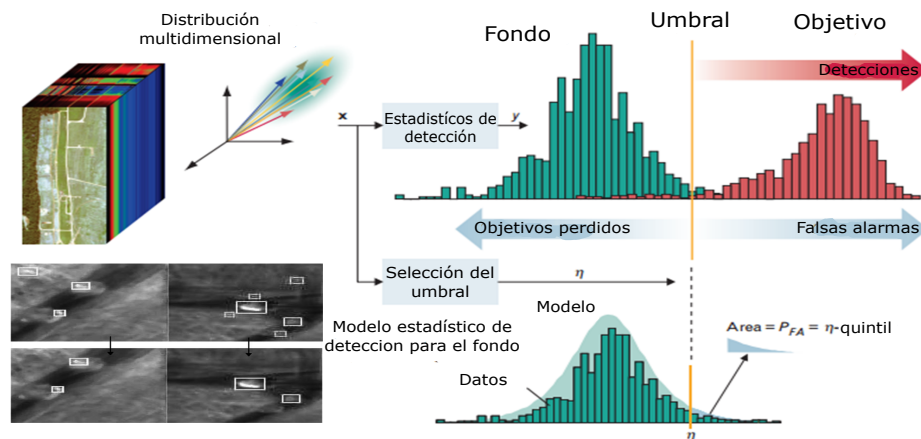
## 1.4. Detección de objetivos en imágenes hiperespectrales

En la detección de objetivos en imágenes hiperespectrales se busca determinar si un objeto, con una firma espectral específica, está o no presente en la escena. Por lo general, estos objetos solo están presentes en un número significativamente menor de píxeles que los demás [20].

El proceso de identificación y detección de un objetivo en una imagen hiperespectral puede ser visto como una tarea de dos fases, la primera fase es la detección de anomalías, que consiste en identificar cuáles píxeles tienen una diferencia espectral significativa con sus vecinos, el segundo paso es determinar si la anomalía es o no un objetivo, esto se puede lograr si la firma espectral del objetivo es conocida, es decir, si se puede encontrar en una librería espectral o de un conjunto de datos de entrenamiento, que también puede ser generado sintéticamente.

Por otro lado, un detector es fundamentalmente una función matemática que asigna un peso o valor numérico a cada píxel de la imagen, si este peso es mayor que un umbral establecido, el píxel evaluado será etiquetado como un píxel objetivo, de lo contrario, será catalogado como píxel de fondo. La figura 1.5 ilustra este concepto.

**Figura 1.5:** Funcionamiento de un detector espectral estándar.



---

En teoría, el diseño y evaluación de algoritmos de detección se facilita asumiendo cierta distribución de probabilidad para la información espectral de objetivo y de fondo. Las características principales de los modelos adoptados y los criterios de optimalidad utilizados determinarán la naturaleza y desempeño del detector producido.

La mayoría de las técnicas clásicas de detección de objetivos [21], [22], [23], están basadas en procesos lineales. También han sido desarrolladas técnicas de detección no lineales avanzadas basadas en métodos de kernel y teoría de aprendizaje [24], [16], estas técnicas explotan indirectamente estadísticas de mayor orden entre las bandas espectrales a través de una función de kernel.

---

## 1.5. Medidas de desempeño

### 1.5.1. Curvas ROC (*Receiver Operator Characteristics*)

En problemas de decisión binarios, un clasificador etiqueta muestras u objetos como positivos o negativos, la decisión tomada por el clasificador puede ser representada en una estructura conocida como matriz de confusión o tabla de contingencia, descrita en la Tabla 1.1. La matriz de confusión está compuesta de cuatro categorías: los verdaderos positivos (TP, *True positives*) que son las muestras correctamente etiquetadas como positivas. Los falsos positivos (FP *False positives*), que son las muestras negativas etiquetadas como positivas. Los verdaderos negativos (TN *True negatives*), muestras negativas etiquetadas correctamente como negativas. Y falsos negativos (FN *False negatives*) que son las muestras positivas etiquetadas como negativas [22].

**Tabla 1.1:** Matriz de confusión

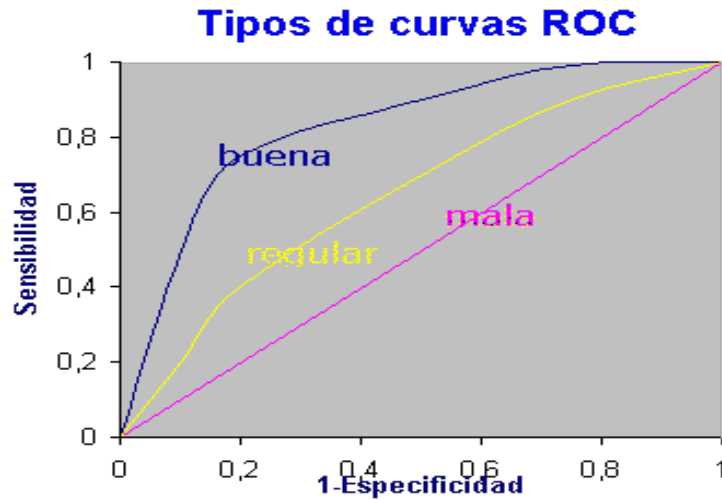
	Positivo	Negativo
Etiquetado Positivo	TP	FP
Etiquetado Negativo	FN	TN

La matriz de confusión puede ser usada para definir las métricas en las que se basan las curvas de características de funcionamiento del receptor o ROC, las cuales son herramientas esenciales para la medición del desempeño de los algoritmos de detección de objetivos. La curva ROC describe la probabilidad de detección (PD), como función de la probabilidad de falsa alarma (PDA). Una estimación segura de la PD requiere de aproximadamente 50 a 100 píxeles objetivo, este requisito tan costoso ha limitado la disponibilidad de conjuntos de datos para el desarrollo de algoritmos de detección de objetivos. Por tanto, se suelen usar modelos de predicción de desempeño, o se selecciona un umbral que permite la detección de todos los objetivos y se cuenta el número de falsas alarmas resultantes. La curva ROC se construye graficando en el eje x la tasa de falsos positivos (FPR *False positive rate*), y en el eje y la tasa de verdaderos positivos (TPR *True positive rate*). El FPR mide la cantidad de elementos negativos clasificados erróneamente como positivos, y el TPR la fracción de muestras positivas etiquetadas correctamente como positivas. La Figura 1.6 muestra los tipos de curvas ROC, donde la sensibilidad es la fracción de

---

verdaderos positivos y la especificidad la fracción de verdaderos negativos.

**Figura 1.6:** Tipos de curvas ROC.



Para una comparación completa y justa de los algoritmos de detección se debe tener en cuenta el punto de vista del usuario final y cómo el uso de cada detector en aplicaciones prácticas afecta su desempeño y resultados. Dos requisitos prácticos importantes son el buen desempeño en ambientes con gran variabilidad de objetivos y fondos y la facilidad operacional para el usuario. Además, la complejidad computacional, la robustez numérica, la robustez en el manejo de parámetros libres y la operación sin intervención humana son factores críticos para aplicaciones de vigilancia.

## 1.6. Dispersión y teoría de muestreo compresivo

La gran cantidad de información espectral poco significativa y la poca información conocida sobre la escena espectral (muestras etiquetadas en las distintas clases en problemas de clasificación, o información espectral del objetivo en problemas de detección) son factores que con frecuencia originan el conocido efecto Hughes [1].

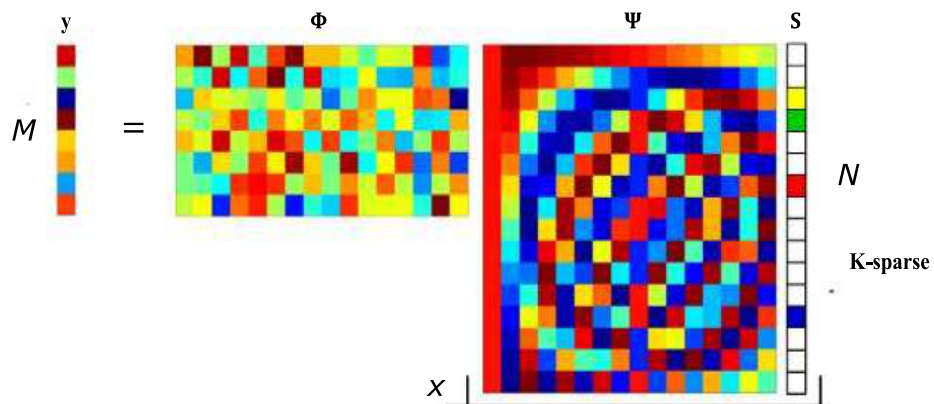
Para tratar el problema que genera la gran dimensionalidad de los datos, la variabilidad espacial de la información espectral, el alto costo de etiquetado de las muestras, y para mejorar la estabilidad numérica y algorítmica, una gran cantidad de enfoques han sido propuestos en diversos campos como segmentación de movimiento [2],

superresolución de imágenes [25], restauración de imágenes [5], [6] y tareas discriminatorias como reconocimiento facial [7], reconocimiento de iris [8], clasificación de tumores [9], y clasificación en imágenes hiperespectrales [10]. Estos avances aprovechan el hecho de que la mayoría de señales naturales son compresibles, es decir, tienen una representación dispersa cuando se expresan en una base apropiada, pueden ser representadas o reconstruidas usando solo una pequeña cantidad de coeficientes que contienen la información más relevante [3], a esta característica se le conoce como dispersión. Recientemente han surgido modelos matemáticos que permiten aprovechar esta característica al momento de la adquisición de señales multidimensionales, especialmente de imágenes hiperespectrales.

La teoría de muestreo compresivo garantiza que ciertas señales pueden ser reconstruidas usando una cantidad mucho menor de muestras, que las usadas por los métodos de adquisición tradicionales. Para hacer esto posible, el muestreo compresivo (CS de su sigla en inglés) se apoya en dos principios fundamentales, la dispersión inherente a las señales, y la incoherencia inherente a la modalidad de muestreo.

La teoría de CS aprovecha el hecho de que la dispersión de muchas señales naturales permite que estas sean compresibles, es decir, que tienen una representación en una base  $\Psi$  apropiada, en la que sus coeficientes representativos son escasos y la incoherencia, que expresa la idea de que los objetos que tienen una representación  $\Psi$ , deben estar distribuidos uniformemente en el dominio en el que son adquiridos.

**Figura 1.7:** Representación matricial de la teoría muestreo compresivo.



---

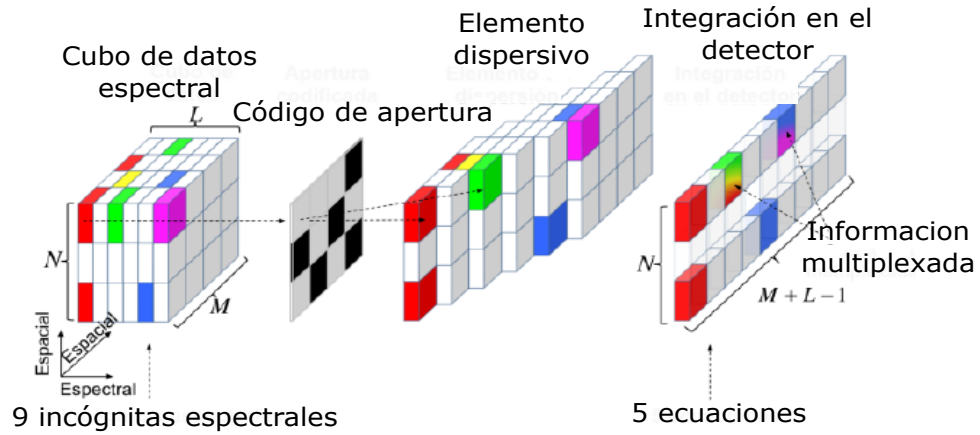
La importancia del CS es que se pueden diseñar protocolos eficientes de muestreo, que capturen la información útil contenida en las señales aprovechando su dispersión, y condensarla en una pequeña cantidad de datos. Estos protocolos requieren correlacionar la señal con un conjunto de ondas de alta incoherencia en la base de representación. Lo que es más relevante de estos protocolos es que permiten al sensor captar de forma eficiente la información de una señal con un alto nivel de dispersión. Incluso es posible utilizar optimización numérica para reconstruir el 100 % de la señal utilizando la pequeña cantidad de datos adquiridos, en otras palabras, CS es un protocolo de adquisición de señales que muestrea a una tasa menor que los modelos tradicionales, y usa algoritmos computacionales para reconstruir las señales usando un conjunto aparentemente incompleto de mediciones.

La naturaleza dispersa de las señales también puede ser aprovechada usando imágenes adquiridas por sistemas de muestreo espectral compresivos, que requieren una cantidad mucho menor de mediciones que los sensores espectrales tradicionales [11], [12].

Los principales elementos de un sistema de muestreo compresivo son los códigos de apertura, un elemento dispersor, y el detector encargado de captar la energía codificada de la escena. Las aperturas codificadas, son arreglos matriciales compuestos de elementos ópticos traslúcidos, que permiten o no el paso de la luz a través del sistema. Los elementos dispersores (por lo general prismas, o redes de difracción) se encargan de descomponer la luz en sus longitudes de onda. La calidad de las imágenes adquiridas por el sistema dependerá de tres factores principales: el porcentaje de elementos traslúcidos que dejan pasar la luz en las aperturas (comúnmente conocido como transmitancia), el tamaño del cubo de datos de la imagen espectral y el nivel de compresión.

El sistema CASSI (*Coded-aperture snapshot spectral imager*) es uno de los sistemas de muestreo espectral compresivos. Este sistema, mostrado en Fig. 1.8, muestrea simultáneamente información espacial y espectral de la escena en un pequeño conjunto de mediciones codificadas (*Focal Plane Array, FPA*) [11]

**Figura 1.8:** Arquitectura CASSI



Matemáticamente, las proyecciones medidas por CASSI en la  $i$ -ésima captación son expresadas como  $y_i = \mathbf{H}_i \mathbf{f} + \mathbf{w}$ , donde  $\mathbf{H}$  es una matriz de  $N(N+L-1)(N^2L)$  cuya estructura es determinada por las aperturas codificadas y el efecto del elemento dispersor. La arquitectura CASSI admite múltiples captaciones, cada una utilizando una apertura codificada distinta, lo cual genera un problema inverso bien condicionado y mejora la calidad de las señales reconstruidas [2]. El conjunto de  $K$  captaciones está definido como  $\mathbf{y} = \mathbf{H}\mathbf{F} + \mathbf{w}$  donde  $\mathbf{H} = [\mathbf{H}_0^T, \dots, \mathbf{H}_{K-1}^T]^T$  es la matriz de muestreo construida con las matrices de muestreo correspondientes a cada adquisición,  $[\mathbf{y}_0^T, \dots, \mathbf{y}_{K-1}^T]^T$  y  $\mathbf{w}$  es el ruido inherente al sistema. El cubo de datos espectral se puede estimar con  $\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} \|\mathbf{y} - \mathbf{H}\boldsymbol{\Psi}\boldsymbol{\theta}\| + \lambda \|\boldsymbol{\theta}\|_1$ ,  $\boldsymbol{\theta}$  es una representación dispersa de  $\mathbf{f}$  en la base  $\boldsymbol{\Psi}$ ,  $\lambda$  es una constante de regularización [26].

El uso de mediciones comprimidas para procesar y analizar información de forma directa es una forma ágil y útil en términos de costo temporal y computacional, de esta forma, es posible realizar las tareas de procesamiento utilizando muestreo compresivo, evadiendo el costoso proceso de reconstrucción [3], es por eso que en la teledetección o detección de objetivos se puede aprovechar el uso de información espectral adquirida usando muestreo compresivo.

## Capítulo 2

# MÉTODOS TRADICIONALES DE DETECCIÓN DE OBJETIVOS

## 2.1. Detectores de anomalías

Los detectores de anomalías son esquemas estadísticos de reconocimiento de patrones, usados para detectar objetos en fondos heterogéneos, no uniformes. En los algoritmos de detección espectral, los píxeles cuya firma espectral difiere significativamente de la de sus píxeles vecinos, son identificados como anomalías espectrales. En estos algoritmos no se conoce o asume información previa sobre la firma espectral del objetivo. Algunos de estos algoritmos son el detector de anomalías Reed-Xiaoli (RX), el detector kernel RX, y el detector de anomalías basado en subespacios.

### 2.1.1. Detector de anomalías Reed-Xiaoli (RX)

El algoritmo de detección de anomalías desarrollado en [27] detecta objetos en fondos heterogéneos, con características espectrales desconocidas. Este algoritmo conocido como el detector de anomalías RX ha sido ampliamente utilizado en numerosas aplicaciones hiperespectrales y es tenido como marco de referencia para algoritmos de detección de anomalías en datos hiperespectrales. El RX es un detector de anomalías adaptativo basado en una tasa de falsas alarmas constante (*CFAR constant false alarm rate*), derivado del test de tasa de posibilidad generalizado (GLRT generalizad likelihood ratio test). La CFAR permite al detector usar un solo umbral para mantener un nivel de falsas alarmas deseado, sin importar la variabilidad del fondo. Esta idea se basa en la asunción de que las características del fondo pueden ser modeladas como una distribución gaussiana multivariable, cuya media y covarianza son estimadas de los píxeles de la imagen. Asumiendo que cada píxel de la imagen está representado como un vector  $x = [x_1 x_2 \dots x_p]^T \in \mathbb{R}^p$

---

compuesto por  $p$  bandas espectrales, la salida del algoritmo RX esta dada por

$$RX(x) = (x - \hat{\boldsymbol{\mu}}_b)^T \hat{\mathbf{C}}_b^{-1} (x - \hat{\boldsymbol{\mu}}_b) \quad (2.1)$$

donde  $\hat{\boldsymbol{\mu}}_b$  es la media estimada del fondo y  $\hat{\mathbf{C}}_b$  la covarianza de fondo estimada, en la imagen evaluada

### 2.1.2. Detección de anomalías basada en subespacios

En detectores de anomalías basados en subespacios, la información espectral de entrada se proyecta en un subespacion en el cual existen bases de representación definidas para algunos vectores de proyección. La proyección de cada pixel  $x$  se calcula como  $s' = (\mathbf{x} - \hat{\boldsymbol{\mu}}_b)^T (\mathbf{I} - \mathbf{W}\mathbf{W}^*) (\mathbf{x} - \hat{\boldsymbol{\mu}}_b)$  donde  $\mathbf{W} = [w_1 w_2 \dots w_m]$  es una matriz cuyas columnas son las  $m$  proyecciones obtenidas de las muestras de fondo,  $\hat{\boldsymbol{\mu}}_b$ , es la media estimada de las muestras de fondo, y  $\mathbf{W}^* = (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T$  es la pseudoinversa de  $\mathbf{W}$ . El producto  $\mathbf{W}^T \mathbf{W}$ , conocido como el operador de proyección, representa el subespacio de fondo. Una anomalía se detecta si la proyección  $s'$  es mayor que algún umbral.

Existen varios metodos para generar las proyecciones y obtener los subespacios, generalmente se usan descomposiciones en valores propios de la matriz de covarianza del fondo, y seleccionar los vectores propios más significativos, que mejor representen las características del subespacio de fondo.

### 2.1.3. Detección de anomalías Kernel Reed-Xiaoli (KRX)

En general, el fondo en una imagen no está distribuido de forma gaussiana, y no es práctico tratar de modelarlo como una mezcla de distribuciones gaussianas, dada la cantidad insuficiente de información de entrenamiento y el desconocimiento del número de distribuciones. Para tratar este problema, una versión no lineal del algoritmo RX (KRX) es desarrollada en [28] el concepto clave del KRX es expresar el modelo RX original en un espacio de mayores dimensiones  $F$ , cuyas regiones de decisión son más complejas que las del espacio de entrada original. Dada una función de mapeo no lineal  $\psi(x)$ , que mapea el vector de entrada  $x$  a un espacio de características de potencialmente mucha mayor dimensionalidad, la salida del

algoritmo RX en este nuevo espacio está dada por

$$RX(\Phi(\mathbf{x})) = (\Phi(\mathbf{x}) - \hat{\boldsymbol{\mu}}_{b\phi})^T \hat{\mathbf{C}}_{b\phi}^{-1} (\Phi(\mathbf{x}) - \hat{\boldsymbol{\mu}}_{b\phi}) \quad (2.2)$$

donde  $\hat{\mathbf{C}}_{b\phi}^{-1}$  y  $\hat{\boldsymbol{\mu}}_{b\phi}$  son la covarianza y media de las muestras de fondo en el nuevo subespacio, respectivamente.

La implementación del algoritmo KRX en el espacio descrito en la Eq. 2.2 no es posible dada la no linealidad del mapeo en  $\Phi$ , lo cual genera datos de gran dimensionalidad. Para evitar una implementación explícita de Eq. 2.2, es necesario efectuar una kernelización, que se refiere a expresar la Eq. 2.2 en forma de productos punto y evaluándola usando el concepto de kernel trick, que permite el cálculo de productos punto en un espacio determinado usando kernels definidos positivos

$$k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle \quad (2.3)$$

La Eq. 2.3 muestra que los productos punto en  $F$  pueden ser reemplazados como una función de kernel  $k$ , una función no lineal que puede ser fácilmente calculada sin identificar el mapeo no lineal  $\Phi$ . Una función comúnmente utilizada es el kernel de base radial (RBF)  $k(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / (\sigma^2))$  donde  $\sigma > 0$  es el parámetro de ancho de banda del kernel. Información detallada de las distintas clases de kernels y teoría de aprendizaje basada en kernels puede ser encontrada en [29], [30], [24], [31], [32].

La versión kernelizada del algoritmo RX en el espacio de representación de la Eq. 2.2 es mostrada en [28] como

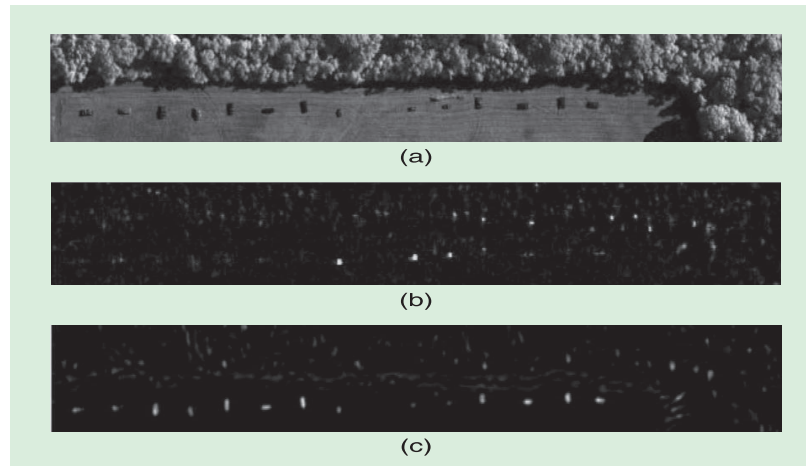
$$RX_k(x) = (\mathbf{k}_x - \mathbf{k}_{\hat{\boldsymbol{\mu}}_b})^T \hat{\mathbf{K}}_b^{-2} (\mathbf{k}_x - \mathbf{k}_{\hat{\boldsymbol{\mu}}_b}), \quad (2.4)$$

donde  $\hat{\mathbf{K}}_b = \Phi(\mathbf{X}_b)^T \Phi(\mathbf{X})$  denota la matriz de kernel centrada para los píxeles de fondo  $\Phi(\mathbf{X}_b)$  en el nuevo espacio de representación,  $\hat{\mathbf{K}}_x = \Phi(\mathbf{X}_b)^T \Phi(x)$ , representa el llamado mapa de kernel empírico del píxel de prueba  $\Phi(x)$  y  $\mathbf{k}_{\hat{\boldsymbol{\mu}}_b} = \Phi(\mathbf{X}_b)^T \Phi(\hat{\boldsymbol{\mu}}_b)$  es el mapa de kernel empírico del fondo  $\Phi(x)$ . La ecuación 2.4 puede ser imple-

mentada sin conocer la función de mapeo  $\Phi$ . El único requisito es la buena elección de la función de kernel  $k$ , que pueda producir una matriz Gram definida positiva [33].

La Figura 2.1 muestra resultados de detección de los algoritmos RX y KRX.

**Figura 2.1:** Imagen hiperespectral HYDICE con mapas de detección del algoritmo RX y KRX. (a) Imagen de prueba (Banda 48); resultados de detección basada en estadísticos del fondo para el algoritmo RX y el algoritmo RX kernelizado (Kernel RBF  $\sigma = 40$ )



## 2.2. Detección con filtros ajustados (*signature-based target detection*)

En situaciones en las que se tiene información previa de las características espectrales de los objetivos deseados estas pueden ser definidas o simplemente por el espectro del objetivo o por un subespacio de la señal. Aplicaciones como estas se pueden ver en el detector GLRT que usando información espectral de objetivos simples, es también conocido como filtro espectral ajustado (SMF) y la máxima probabilidad de abundancia de objetivo de un pixel de prueba y está expresada en [13] como

$$D_{SMF}(\mathbf{y}) = \frac{\mathbf{s}^T \hat{\mathbf{C}}^{-1} \mathbf{y}}{\mathbf{s}^T \hat{\mathbf{C}}^{-1} \mathbf{s}}, \quad (2.5)$$

donde  $s$  es la firma espectral del objetivo y  $\hat{\mathbf{C}}$  representa la matriz de covarianza estimada para los datos de observación. El modelo SMF está basado en la asun-

---

ción de que el ruido del fondo de la imagen tiene una distribución gaussiana  $N(0, \hat{C})$  y que la distribución de los objetivos también es gaussiana  $N(as, \hat{C})$  con la misma covarianza pero con media  $as$ , siendo  $a$  un escalar representando el nivel de presencia del objetivo.

En el caso en el que las características de los pixeles de objetivo y de fondo son modeladas por dos subespacios lineales, el detector anterior sufre algunos cambios y es llamado detector de subespacios ajustado MSD, por sus siglas en ingles. El GLRT para un modelo de dos subespacios esta expresado en [34].

$$D_{MSD}(\mathbf{y}) = \frac{\mathbf{y}^T(1 - \mathbf{P}_B)\mathbf{y}}{\mathbf{y}^T(1 - \mathbf{P}_{TB})\mathbf{y}}, \quad (2.6)$$

siendo  $P_B$  la matriz de proyecciones asociada al subespacio de fondo  $\langle B \rangle$  y  $P_{TB}$  es la matriz de proyecciones asociada al subespacio de fondo y de objetivos  $\langle TB \rangle$ . Usualmente, los eigenvectores de las matrices de covarianza de fondo y objetivo son utilizados para generar las columnas de B y T respectivamente.

Usando algunas variaciones de los modelos en SMF y MSD, investigadores han desarrollado diferentes detectores, tales como el detector de subespacios adaptativos (ASD, de su sigla en inglés) [35] y la proyección de subespacios ortogonal (OSP) [22]. En el ASD la firma del objetivo está representada por un subespacio lineal y los estadísticos de fondo por una distribución gaussiana de media cero, cuya covarianza es estimada desde la información espectral de la imagen. En el caso del OSP la firma del objetivo se representa por un único vector espectral  $s$  y el fondo se representa por las firmas no deseadas de objetos de fondo B. El algoritmo OSP está basado en la maximización de la relación señal a ruido (SNR) en el subespacio ortogonal al subespacio de fondo que está expresado como  $D_{OSP} = s^T \mathbf{P}_B^\perp \mathbf{y}$  donde  $\mathbf{P}_B^\perp = \mathbf{I} - \mathbf{B}^*$  es llamado el operador de rechazo de fondo y  $(.)^*$  denota la matriz seudoinvertida.

---

## 2.3. Detectores adaptativos basados en subespacios

Las hipótesis  $H_0$  y  $H_1$  para el detector ASD son:

$$\begin{aligned} H_0 : x &= n, \text{Objetivo ausente} \\ H_1 : x &= \mu S\Theta + n, \text{Objetivo presente,} \end{aligned}$$

donde  $S$  representa el subespacio del objetivo;  $\Theta$  es un vector desconocido cuyas entradas son los coeficientes correspondientes a cada columna de  $S$ , la señal objetivo  $S\Theta$  es escalada por el factor  $\mu$ ;  $n$  representa el ruido gaussiano distribuido como  $N(0, \sigma^2 C_b)$  donde  $C_b$  es el ruido del fondo obtenido de los datos de entrenamiento y  $C_b$  representa un valor escalar asociado a los datos evaluados [33].

El GLRT para el problema descrito, está dado en [23] como

$$D_{\mu ASD}(\mathbf{x}) = \frac{\mathbf{x}^T \hat{C}_b^{-1} S (S^T \hat{C}_b^{-1} S)^{-1} S^T \hat{C}_b^{-1} x}{\mathbf{x}^T \hat{C}_b^{-1} x} \underset{H_0}{\overset{H_1}{\geq}} \eta_{ASD}, \quad (2.7)$$

donde  $\hat{C}_b$  es la máxima estimación de probabilidad de covarianza,  $C_b$  de los datos y  $\eta_{ASD}$  representa un umbral, la ecuación (2.7), tiene una propiedad CFAR, también llamada CFAR ASD por subespacio CFAR o CFAR multirango en [36]. Cuando la señal es coherente ( $S$  es de rango 1) el detector descrito en la ecuación (2.7) es definido como el estimador coseno\coherencia adaptativo (*ACE adaptive coherence cosine estimator*).

## 2.4. Detectores de coeficiente de probabilidad (*Likelihood ratio test-based detectors*)

Por teoría de decisión estadística, se conoce que las decisiones basadas en la prueba de coeficiente de probabilidad (LRT), son óptimas para un amplio rango de criterios de desempeño [37], entre estos criterios se encuentra el de Neyman-pearson (NP) que maximiza la probabilidad de detección para cualquier probabilidad de falsa alarma. El LRT está definido como

$$\Lambda(x) = \frac{f(x | \text{Objetivo presente})}{f(x | \text{Objetivo ausente})} \stackrel{\Delta}{=} \frac{f(x | \mathbf{H}_1)}{f(x | \mathbf{H}_0)} \underset{\mathbf{H}_0}{\underset{\mathbf{H}_1}{\geq}} \eta, \quad (2.8)$$

siendo  $f(x | \mathbf{H}_i)$  la función de densidad de probabilidad (PDF) condicional para cada hipótesis. Intentos por simplificar el problema incluyen el uso de una única clase para el fondo [38], o agrupando los datos y para hacer la detección a cada clase [39]. A pesar de su atractivo, las técnicas de detección basadas en agrupamiento no muestran un desempeño más sorprendente que el de los algoritmos de filtros ajustados usados en la actualidad [40].

Una métrica esencial para el desempeño de los algoritmos de detección de objetivos es la curva de características de funcionamiento del receptor o ROC. La curva ROC describe la probabilidad de detección (PD), como función de la probabilidad de falsa alarma (PDA). Una estimación segura de la PD requiere de aproximadamente 50 a 100 pixeles objetivo, este requisito tan costoso ha limitado la disponibilidad de conjuntos de datos para el desarrollo de algoritmos de detección de objetivos. Por tanto, se suelen usar modelos de predicción de desempeño, o se selecciona un umbral que permite la detección de todos los objetivos y se cuenta el número de falsas alarmas resultantes.

Para una comparación completa y justa de los algoritmos de detección se debe tener en cuenta el punto de vista del usuario final y cómo el uso de cada detector en aplicaciones prácticas afecta su desempeño y resultados. Dos requisitos prácticos importantes son el buen desempeño en ambientes con gran variabilidad de objetivos y fondos y la facilidad operacional para el usuario. Además, la complejidad computacional, la robustez numérica, la robustez en el manejo de parámetros libres y la operación sin intervención humana son factores críticos para aplicaciones de vigilancia.

## 2.5. Detectores no lineales (*Nonlinear detectors*)

La mayoría de detectores de anomalías y de objetivos solo utilizan estadísticos de primer y segundo orden. La teoría de kernel machine learning [16] ha surgido co-

---

mo una alternativa de aprendizaje no lineal que desarrolla y extiende los algoritmos de reconocimiento de patrones clásicos. La explotación implícita de características no lineales a través de kernels proporciona información crucial del conjunto de datos, que en general los métodos basados en modelos lineales no pueden conseguir. El algoritmo de detección de anomalías RX, los algoritmos de detección estadística y los basados en modelos dispersos ya han sido extendidos a sus versiones no lineales [41] usando ideas de la teoría de kernel machine learning. Los resultados experimentales de estos estudios muestran que por lo general los algoritmos basados en métodos de kernel superan el desempeño de sus versiones lineales.

## 2.6. Representación Dispersa para detección de objetivos

Los clasificadores basados en dispersión (SRC) expresan una señal como una combinación lineal de muy pocos átomos de un diccionario sobrecompleto, conformado por un conjunto de datos de entrenamiento de todas las clases, este modelo revela que la información de señales de diferentes clases, pertenece a subespacios diferentes.

### 2.6.1. Modelo de dispersión

En el modelo de dispersión para la clasificación de imágenes hiperespectrales, se asume que la firma espectral de píxeles pertenecientes a la misma clase, reposa en un subespacio de baja dimensionalidad. Si se tienen  $K$  clases distintas, y la  $K$ -ésima clase tiene  $N_k$  muestras de entrenamiento  $\{\mathbf{a}_j^k\}_{j=1,2,\dots,N_k}$ . Sea  $x$  un pixel hiperespectral de prueba  $p$ -dimensional. Si  $x$  pertenece a la  $k$ -ésima clase, entonces su espectro reside aproximadamente en un subespacio de baja dimensionalidad compuesto por las firmas espectrales de la  $k$ -ésima clase.

Basándose en dicha suposición de dispersión, en SRC [7] un pixel de prueba desconocido es modelado para residir en la unión de  $K$  subespacios asociados a  $K$  clases. Combinando los subdiccionarios de las clases  $\{\mathbf{A}^k\}_{k=1,2,\dots,K}$  un pixel  $x$  puede ser reescrito como una combinación lineal dispersa de todas las muestras de

entrenamiento, tal que

$$\mathbf{x} = \mathbf{A}^1 \boldsymbol{\alpha}^1 + \mathbf{A}^2 \boldsymbol{\alpha}^2 + \dots + \mathbf{A}^K \boldsymbol{\alpha}^K = \underbrace{\begin{bmatrix} \mathbf{A}^1 & \dots & \mathbf{A}^K \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \boldsymbol{\alpha}^1 \\ \vdots \\ \boldsymbol{\alpha}^K \end{bmatrix}}_{\boldsymbol{\alpha}} = \mathbf{A} \boldsymbol{\alpha}, \quad (2.9)$$

Donde  $\mathbf{A}$  es un diccionario estructurado de tamaño  $p \times N$ , compuesto por muestras de entrenamiento de todas las clases con  $N = \sum_{k=1}^K N_k$  y  $\boldsymbol{\alpha}$  es un vector disperso  $N$ -dimensional formado por la concatenación de los vectores dispersos  $\{\boldsymbol{\alpha}^k\}_{k=1,2,\dots,K}$  asociados a cada subdiccionario de clase. En condiciones ideales, si  $\mathbf{x}$  pertenece a la  $k$ -ésima clase, entonces  $\alpha^j, \forall j = 1, \dots, K, j \neq k$ .

Dado un diccionario  $\mathbf{A}$ , el vector  $\boldsymbol{\alpha}$  que satisface  $\mathbf{x} = \mathbf{A} \boldsymbol{\alpha}$  puede obtenerse del siguiente problema de optimización:

$$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha}} \lambda \|\boldsymbol{\alpha}\|_0 \quad \text{sujeto a} \quad \|\mathbf{x} - \mathbf{A} \boldsymbol{\alpha}\|_2 \leq \epsilon_0, \quad (2.10)$$

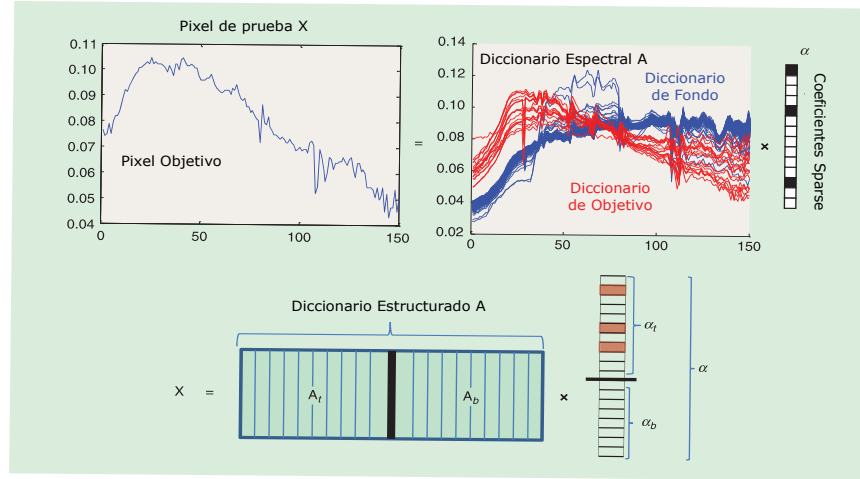
donde  $\|\boldsymbol{\alpha}\|_0$  denota la norma  $l_0$  que está definida como el número de entradas diferentes de cero en el vector  $\boldsymbol{\alpha}$  y  $\epsilon_0$  es la tolerancia al error de aproximación. El problema de optimización planteado en la Eq 2.10 puede ser resuelto aproximadamente por cualquier algoritmo de búsqueda voraz [42]. A pesar de que el problema de optimización en la ecuación (2.10) es NP-complejo, puede ser relajado, reemplazando la norma  $l_0$  con la norma  $l_1$ , donde  $\|\boldsymbol{\alpha}\|_1 = \sum_{i=1}^N |\alpha_i|$  y resolviendo con una técnica de optimización estándar [43]

La etiqueta de clase para el pixel  $X$  se obtiene directamente de las características del vector  $\hat{\boldsymbol{\alpha}}$  reconstruido, que se determina por la clase que tenga el menor error residual

$$r^k(x) = \arg \min_{k=1,\dots,K} \|\mathbf{x} - \mathbf{A}^k \hat{\boldsymbol{\alpha}}^k\|_2, \quad (2.11)$$

donde  $\hat{\boldsymbol{\alpha}}^k$  representa la porción de los coeficientes reconstruidos correspondientes a los átomos de la  $k$ -ésima clase.

**Figura 2.2:** El proceso SRC para detección de objetivos, muestra el diccionario estructurado, compuesto de subdiccionarios de fondo y de objetivo.



En el caso de la detección de objetivos, el diccionario para SRC consiste típicamente en muestras de entrenamiento provenientes de subdiccionarios de objetivo y de fondo, representados por  $\mathbf{A} = [\mathbf{A}_t \ \mathbf{A}_b]$ , como se muestra en la figura (2.2). El vector disperso  $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_t^T \ \boldsymbol{\alpha}_b^T]^T$  que satisface  $\mathbf{x} = \mathbf{A}\boldsymbol{\alpha}$  se obtiene resolviendo el problema de optimización anterior, donde  $\boldsymbol{\alpha}_t$  y  $\boldsymbol{\alpha}_b$  representan los coeficientes correspondientes a los subdiccionarios de objetivo y de fondo respectivamente. Una vez se obtiene el vector de coeficientes disperso, la clase del pixel  $\mathbf{x}$  es determinada comparando los errores residuales  $r_t(x) = \|\mathbf{x} - \mathbf{A}_t \hat{\boldsymbol{\alpha}}_t\|_2^2$  y  $r_b(x) = \|\mathbf{x} - \mathbf{A}_b \hat{\boldsymbol{\alpha}}_b\|_2^2$

## 2.6.2. Asunciones de dispersión

El desempeño en la clasificación de los SRC puede ser mejorado incorporando información contextual o haciendo asunciones sobre el comportamiento de los pixeles vecinos del pixel evaluado incorporándola al criterio de decisión del clasificador. Si el objetivo está comprendido por varios pixeles, la información contextual puede ser incluida haciendo asunciones más específicas, o imponiendo restricciones estructurales al problema de optimización. Uno de los enfoques más simples es adoptar un modelo de dispersión conjunta o colaborativa, como se estudia en [17], asumiendo que los vectores dispersos asociados a los pixeles en un vecindario, comparten un patrón de dispersión. En el modelo de dispersión conjunta, un

---

vecindario de T pixeles  $X = [\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_T]$  es un diccionario específico como

$$\hat{\Omega} = \arg \min \|\Omega\|_{row,0} \quad \text{sujeto a} \quad \|\mathbf{X} - \mathbf{A}\Omega\|_F^2 \leq \epsilon_0, \quad (2.12)$$

donde  $\|\cdot\|_F$  representa la norma Frobenius,  $\|\Omega\|_{row,0}$  el número de filas diferentes de cero de  $\Omega$  y  $\epsilon_0$  es un error de tolerancia. Los algoritmos voraces simultáneos [25] son útiles en la aproximación de resultados para la ecuación (2.12). La norma  $\|\Omega\|_{row,0}$  puede también ser reemplazada por que es una norma  $\|\Omega\|_{1,2}$  definida como la suma de normas  $l_{1,2}$  de las filas de  $\Omega$ , para convertir el problema NP-complejo de la ecuación (2.12) en un problema de optimización convexa.

### 2.6.3. Construcción de diccionarios

Otro aspecto importante que requiere especial atención, es cómo construir los diccionarios  $\mathbf{A}_b$  y  $\mathbf{A}_t$ . Los diccionarios globales para el objetivo y el fondo pueden diseñarse usando información de entrenamiento, sin embargo, en problemas de detección de objetivos no siempre hay información de entrenamiento disponible, en especial información del objetivo. Usando modelos físicos y el programa de modelado atmosférico MODTRAN [44] pueden generarse firmas espectrales de objetivo que describen el comportamiento del objetivo en un gran rango de condiciones atmosféricas [21], [44].

Estas firmas espectrales sintéticas pueden usarse en la construcción de diccionarios de objetivo redundantes que permanezcan invariantes a los cambios atmosféricos. Los diccionarios de fondo son modelados usualmente seleccionando aleatoriamente pixeles de la propia imagen, además, si se tienen suficientes muestras de entrenamiento de fondo y objetivo, se pueden usar técnicas especiales para el diseño de diccionarios, como el algoritmo K-SVD [42], que minimiza la dispersión de la representación, y actualiza los átomos del diccionario para un mejor entrenamiento de la información.

## Capítulo 3

### MÉTODO PROPUESTO

El modelo propuesto busca aprovechar las ventajas de la teoría de muestreo compresivo para lograr efectuar la tarea de detección de objetivos, usando una cantidad significativamente menor de información espectral de la imagen. Por esto, se buscó partir de un modelo de detección de objetivos que al igual que un sistema de muestreo compresivo, utilizara la noción de dispersión con el fin de aprovechar esta similitud entre ambos modelos y lograr el objetivo planteado. El sistema de muestreo compresivo utilizado para este trabajo es el sistema *Coded-aperture snapshot spectral imaging System (CASSI)*, presentado en la sección 1.7

#### 3.1. Detección de objetivos en imágenes hiperespectrales utilizando un modelo de dispersión

Sea  $\mathbf{f}_i \in \mathbb{R}^L$  un pixel hiperespectral de dimensión  $L$  cuyas entradas corresponden a las bandas espectrales de un cubo de datos. El espectro de  $\mathbf{f}_i$  se modela como la unión de dos subespacios de baja dimensionalidad, compuestos de muestras de entrenamiento de fondo  $\{\mathbf{m}_i^b\}_{i=1,2,\dots,N_b}$  y objetivo  $\{\mathbf{m}_i^t\}_{i=1,2,\dots,N_t}$  respectivamente, así pues,  $\mathbf{f}_i$  puede escribirse como una combinación lineal de todos los pixeles de entrenamiento [18], como se muestra en la ecuación (3.1).

$$\begin{aligned}
 \mathbf{f}_i &= (\alpha_1^b \mathbf{m}_1^b + \dots + \alpha_{N_b}^b \mathbf{m}_{N_b}^b) + (\alpha_1^t \mathbf{m}_1^t + \dots + \alpha_{N_t}^t \mathbf{m}_{N_t}^t) \\
 &= \underbrace{\begin{bmatrix} \mathbf{m}_1^b & \dots & \mathbf{m}_{N_b}^b \end{bmatrix}}_{\mathbf{M}^b} \underbrace{\begin{bmatrix} \alpha_1^b \\ \vdots \\ \alpha_{N_b}^b \end{bmatrix}}_{\boldsymbol{\alpha}^b} + \underbrace{\begin{bmatrix} \mathbf{m}_1^t & \dots & \mathbf{m}_{N_t}^t \end{bmatrix}}_{\mathbf{M}^t} \underbrace{\begin{bmatrix} \alpha_1^t \\ \vdots \\ \alpha_{N_t}^t \end{bmatrix}}_{\boldsymbol{\alpha}^t} \\
 &= \mathbf{M}^b \boldsymbol{\alpha}^b + \mathbf{M}^t \boldsymbol{\alpha}^t = \begin{bmatrix} \mathbf{M}^b & \mathbf{M}^t \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}^b \\ \boldsymbol{\alpha}^t \end{bmatrix} = \mathbf{M} \boldsymbol{\alpha}, \tag{3.1}
 \end{aligned}$$

donde  $\mathbf{M}^b$  y  $\mathbf{M}^t$  son subdiccionarios compuestos de  $N_b$  muestras de fondo y  $N_t$

---

muestras de objetivo (también llamados átomos) donde  $\alpha^b$  y  $\alpha^t$  son vectores cuyas entradas corresponden a los pesos de los átomos en  $M^b$  y  $M^t$ , respectivamente. La matriz  $M \in \mathbb{R}^{L \times N}$  con  $N = N_b + N_t$  es el diccionario de entrenamiento, compuesto por ambos subdiccionarios de fondo y de objetivo y  $\alpha \in \mathbb{R}^N$  es la concatenación de los vectores  $\alpha^b$  y  $\alpha^t$ .

Teniendo en cuenta la correlación espacial existente entre cada pixel y su vecindario, la ecuación (3.1) puede ser escrita como

$$\mathbf{F} = \mathbf{MA}, \quad (3.2)$$

donde  $\mathbf{F}$  es el vecindario de  $T$  pixeles y  $\mathbf{A}$  es la matriz dispersa representando el  $\{\mathbf{f}_i\}_{i=1,2,\dots,T}$  pixel en el vecindario.

Dado el diccionario  $M$  la matriz  $A$  se obtiene al resolver el problema de optimización de dispersión conjunta dado por

$$\begin{array}{ll} \text{minimizar} & \|\mathbf{F} - \mathbf{MA}\|_F \\ \text{sujeto a} & \|\mathbf{A}\|_{\text{fila},0} \leq K_0 \end{array} \quad (3.3)$$

siendo  $\|\mathbf{A}\|_{\text{fila},0}$  el número de filas diferentes de cero de la matriz  $\mathbf{A}$ . La solución al problema en la ecuación (3.3) es una matriz dispersa de tamaño  $N \times T$  con muy pocas filas distintas de cero. Sin embargo, el problema de reconstrucción en la ecuación (3.3) es NP-complejo y puede ser resuelto por algoritmos voraces [45], [46], o relajado con programación convexa [15].

La matriz  $\mathbf{A}$  es reconstruida usando el algoritmo SOMP (*Simultaneous Orthogonal Matching Pursuit*), un algoritmo voraz cuyos detalles de implementación se explican detalladamente en [45]. El SOMP reconstruye el conjunto de soporte  $\Gamma_K$  de manera secuencial, (los átomos del diccionario  $M$  son seleccionados secuencialmente). En cada iteración, el átomo seleccionado, conlleva a la mejor aproximación para todos los vectores residuales seleccionados. El algoritmo SOMP termina cuando el error residual  $\|\mathbf{F} - \mathbf{MA}\|_F$  es lo suficientemente pequeño (controlado por el número de

---

iteraciones) o cuando se alcanza el nivel de dispersión deseado, resolviendo así el problema de la ecuación (3.3).

Luego de eso, las clases o etiquetas de las muestras evaluadas, son determinadas con base en las características de los coeficientes dispersos, y calcula y compara el error residual total entre las muestras de prueba evaluadas y las aproximaciones obtenidas de los subdiccionarios de fondo y objetivo. La salida del detector de objetivos basado en dispersión del que parte el modelo propuesto, se define como la diferencia entre los errores residuales de todos los pixeles en el vecindario [17].

$$D(\mathbf{f}_i) = \|\mathbf{F} - \mathbf{M}^b \mathbf{A}^b\|_F - \|\mathbf{F} - \mathbf{M}^t \mathbf{A}^t\|_F, \quad (3.4)$$

donde  $\mathbf{A}^b$  son las primeras  $N_b$  filas reconstruidas de la matriz  $\mathbf{A}$ , correspondientes al diccionario de fondo  $\mathbf{M}^b$  y,  $\mathbf{A}^t$  son las  $N_t$  filas restantes de  $\mathbf{A}$ , correspondientes al subdiccionario de objetivo  $\mathbf{M}^t$ . Si el resultado es mayor que un umbral establecido, el pixel evaluado es marcado como objetivo, de lo contrario, será marcado como pixel de fondo.

Para poder utilizar la teoría de muestro compresivo en un detector con las características anteriores, es necesario efectuar una transformación en el diccionario  $\mathbf{M}$ . Ésta puede ser una transformación de componentes principales, de tal modo que la base de representación  $\Psi$  es obtenida del conjunto de  $N_t$  vectores de entrenamiento. La base ortonormal  $\Psi$  es formada por el conjunto de vectores propios de la matriz de correlación  $\mathbf{C}$  dado por

$$\mathbf{C} = (\mathbf{M} - \{E\})(\mathbf{M} - \{E\})^T = \mathbf{V} \sum \mathbf{V}^T, \quad (3.5)$$

Tal que  $\Psi = \mathbf{V}^T$ . El análisis de componentes principales, o de valores propios (PCA) es una de las técnicas de reducción de dimensionalidad y compresión en imágenes hiperespectrales más usadas [22], [47], [48], ya que preserva la información más significativa de la imagen usando solo un pequeño conjunto de sus componentes. Por tanto, el vector disperso original  $\alpha$  que representa el pixel  $\mathbf{f}_i$  se puede transformar en un vector disperso  $\theta_i \in \mathbb{R}^L$  que representa el pixel en la base ortogonal  $\Psi$ . El pixel

evaluado  $\mathbf{f}_i$  puede ser representado como

$$\mathbf{f}_i \approx \begin{bmatrix} \boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_L \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta}_{i1}, \dots, \boldsymbol{\theta}_{iL} \end{bmatrix}^T = \boldsymbol{\Psi} \boldsymbol{\theta}_i. \quad (3.6)$$

Usando la representación dispersa de los pixeles en la base de entrenamiento  $\boldsymbol{\Psi}$  las mediciones comprimidas del sistema CASSI se reescriben de la forma:

$$\mathbf{y} = \mathbf{H} \bar{\boldsymbol{\Psi}} \boldsymbol{\theta} + \mathbf{w}, \quad (3.7)$$

donde  $\bar{\boldsymbol{\Psi}} = \boldsymbol{\Psi} \otimes \mathbf{I}$  con  $\mathbf{I}$  siendo una matriz identidad de tamaño  $N^2 \times N^2$ ,  $\otimes$  el operador de producto kronecker, de igual forma,  $\boldsymbol{\theta} = \left[ \boldsymbol{\theta}_{\{1,1\}}, \dots, \boldsymbol{\theta}_{\{N^2,1\}}, \dots, \boldsymbol{\theta}_{\{1,L\}}, \dots, \boldsymbol{\theta}_{\{N^2,L\}} \right]^T$  donde  $\boldsymbol{\theta}_{\{i,L\}}$  corresponde al  $L$ -ésimo coeficiente PCA del  $i$ -ésimo pixel, y el vector disperso  $\boldsymbol{\theta}_i$  que representa al  $i$ -ésimo pixel esta dado por  $\boldsymbol{\theta}_i = \left[ \boldsymbol{\theta}_{\{i,1\}}, \boldsymbol{\theta}_{\{i,2\}}, \dots, \boldsymbol{\theta}_{\{i,L\}} \right]^T$ . En la ecuación (3.7),  $\mathbf{w}$  es el ruido inherente al sistema, y  $\mathbf{H}$  es la matriz de muestreo CASSI. El algoritmo propuesto primero estima el vector disperso  $\boldsymbol{\theta}$  directamente de las mediciones comprimidas representadas por el vector  $\mathbf{y}$ , resolviendo el problema de optimización con restricciones de dispersión dado por

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} \|\mathbf{y} - \mathbf{H} \bar{\boldsymbol{\Psi}} \boldsymbol{\theta}\|_2 + \lambda \|\boldsymbol{\theta}\|_1, \quad (3.8)$$

en el cual, la norma  $\ell_1$  implica una restricción de dispersión y la norma  $\ell_2$  encuentra el vector disperso que más aproxime las medidas del sistema CASSI. La correlación espacial entre pixeles se tiene en cuenta en el problema la ecuación (3.8), reemplazando la base de representación  $\bar{\boldsymbol{\Psi}} = \boldsymbol{\Psi} \otimes \mathbf{I}$  por  $\bar{\boldsymbol{\Psi}} = \boldsymbol{\Psi} \otimes \Omega_{2D}^T$  donde  $\Omega_{2D}^T$  es el diccionario de la base wavelet  $2D$ . Sin embargo, si la correlacion espacial se tiene en cuenta en la base de representación, entonces el vector disperso estimado  $\hat{\boldsymbol{\theta}}$  se obtiene haciendo la transformacion wavelet inversa en  $2D$  a la solución del problema planteado en la ecuación (3.8). Numerosos algoritmos han sido propuestos para resolver problemas como el planteado en la ecuación(3.8), incluyendo el método de solución de producto interior por mínimos cuadrados [49], o el método de proyección de gradiente para reconstrucciones disperso (GPRS) [50].

---

## 3.2. Problema inverso

El modelo de detección de objetivos propuesto en este trabajo, requiere que el algoritmo de detección descrito en la ecuación (3.4) trabaje en el dominio compresivo ilustrado en las ecuaciones (3.7) y en (3.8), de tal forma que se logre efectuar la tarea de detección utilizando las mediciones comprimidas del sistema CASSI, este modelo se explica en el Algoritmo 1

---

**Algoritmo 1** Detección de objetivos espectrales usando mediciones compresivas

---

**Entrada:**  $M^t \in \mathbb{R}^{L \times N_t}$  (Firmas espectrales del objetivo)

$M^b \in \mathbb{R}^{L \times N_b}$  (Firmas espectrales del fondo)

$\Psi \in \mathbb{R}^{L^2}$  (Base ortogonal)

$\hat{\theta} \in \mathbb{R}^{L \times N^2}$  (Vector disperso estimado para todos los pixeles)

**Salida:**  $D$  (Mapa de Detección)

Crear diccionarios de Objetivo y de fondo con base  $\Psi$

$\bar{M}^t = \Psi^T M^t$

$\bar{M}^b = \Psi^T M^b$

Crear una representación en base  $\Psi$  de la imagen

$$\bar{F} = (\mathbf{I} \otimes \Psi_{2D}^T) \hat{\theta}$$

**para**  $i = 1$  **hasta**  $N^2$  **hacer**

$$\hat{F} = [\bar{f}_1, \bar{f}_2, \dots, \bar{f}_T]$$

con,  $\bar{f}_i \in \mathbb{R}^L$

Resolver el problema de optimización

minimizar

$$\|\bar{M}\bar{A} - \hat{F}\|_F$$

sujeto a

$$\|\bar{A}\|_{\text{row},0} \leq K_0$$

Calcular el mínimo residuo total

$$D(f_j) = \|\hat{F} - \bar{M}^b \bar{A}^b\|_F - \|\hat{F} - \bar{M}^t \bar{A}^t\|_F$$

**fin para**

---

donde en la ecuación  $\bar{\mathbf{F}} = (\mathbf{I} \otimes \Psi_{2D}^T) \hat{\boldsymbol{\theta}}$ ,  $\mathbf{I} \in \mathbb{R}^{N^2 \times N^2}$  es una matriz identidad,  $\otimes$  es el operador del producto kronecker,  $\hat{\boldsymbol{\theta}}$  es el vector disperso obtenido en 3.8 y  $\bar{\mathbf{F}}$  es la matrix dispersa correspondiente a todas las bandas espectrales en la base  $\Psi$ .

El algoritmo de detección propuesto se basa en un modelo de dispersión conjunta para la extracción de información contextual en las imágenes hiperespectrales. Se asume que los pixeles del mismo material en una región, comparten un patrón común de dispersión, entonces, los pixeles que compartan cierta similitud espectral y que se encuentren cercanos espacialmente, pueden ser representados como una combinación lineal de unos cuantos átomos comunes, de la forma:

$$\hat{\mathbf{F}} = [\bar{\mathbf{f}}_1, \bar{\mathbf{f}}_2, \dots, \bar{\mathbf{f}}_T] \quad (3.9)$$

Con  $\hat{\mathbf{F}} \in \mathbb{R}^L$  como la representación dispersa de los pixeles que forman el vecindario espacial  $T$  del  $j$ -ésimo pixel de prueba. Partiendo de la matriz  $\hat{\mathbf{F}}$ , el problema de dispersión conjunta se formula como:

$$\begin{array}{ll} \text{minimizar} & \|\bar{\mathbf{M}}\bar{\mathbf{A}} - \hat{\mathbf{F}}\|_F \\ \text{subjeto a} & \|\bar{\mathbf{A}}\|_{\text{fila},0} \leq K_0 \end{array} \quad (3.10)$$

donde  $\bar{\mathbf{A}}$  es una matriz dispersa con solo  $K$  filas distintas de cero,  $K_0$  es el nivel de dispersión y  $\|\cdot\|_F$  la norma Frobenius. Una vez se obtiene la matriz  $\bar{\mathbf{A}}$ , la clase o etiqueta del pixel evaluado  $\bar{\mathbf{f}}_i$  se determina por el error residual mínimo total

$$\mathbf{D}(\mathbf{f}_j) = \|\hat{\mathbf{F}} - \bar{\mathbf{M}}^b \bar{\mathbf{A}}^b\|_F - \|\hat{\mathbf{F}} - \bar{\mathbf{M}}^t \bar{\mathbf{A}}^t\|_F \quad (3.11)$$

En el que  $\bar{\mathbf{A}}^b$  y  $\bar{\mathbf{A}}^t$  son formadas por las  $N_b$  y  $N_t$  filas de  $\bar{\mathbf{A}}$  asociadas a los subdiccionarios de fondo  $\bar{\mathbf{M}}^b$  y objetivo  $\bar{\mathbf{M}}^t$ .

## Capítulo 4

### SIMULACIONES Y RESULTADOS

El desempeño del modelo propuesto es evaluado efectuando detección de objetivos sobre las medidas compresivas de dos imágenes adquiridas mediante la simulación del sistema de muestreo compresivo CASSI. La primera imagen, es el conjunto de datos self test del proyecto RIT-CIS-DIRT [51], los datos de esta imagen, ilustrados en la Figura. 4.1, fueron adquiridos como parte de un experimento de campo llevado a cabo en Julio de 2006, en las inmediaciones del pueblo de Cooke City, en Montana, Estados Unidos. Las imágenes hiperespectrales fueron adquiridas por el sensor HyMap de HyVista, con aproximadamente 3 metros de resolución espacial. El sensor genera 128 bandas espectrales en el rango de  $0.45\text{-}2.5\ \mu\text{m}$ , con cobertura espectral continua (exceptuando las bandas de vapor atmosférico) y ancho de banda de  $15 - 20\ \mu\text{m}$ . Como objetivo se usó un pequeño panel de tela, su reflectancia espectral fue medida por un espectrómetro Cary 500 en condiciones de laboratorio.

**Figura 4.1:** Imagen RGB del dataset self test.



El segundo experimento se realizó usando la imagen hiperespectral EO1H-0070552014301110-PF-SG1-01, adquirida por el sensor EO-1 Hyperion el 28 de Octubre de 2014 en la región de Mogotes, Santander, Colombia. La imagen tiene una resolución espacial de aproximadamente 30 metros, una pequeña región de la imagen fue utilizada para la realización de los experimentos. La imagen está compuesta en su mayoría por terreno cultivado y listo para cultivar. El sensor es capaz de muestrear 220 bandas espectrales (de  $0.4$  a  $2.5\ \mu\text{m}$ ).

---

En ambas imágenes, se utiliza el modelo del sistema multicaptura CASSI para la simulación del sistema de adquisición de las mediciones compresivas, utilizando distintos niveles de compresión y de transmitancia, con el fin de establecer los valores que arrojen el mejor resultado de detección.

**Figura 4.2:** Mapeo RGB de la imagen del sensor Hyperion.



El detector propuesto, descrito en el Algoritmo 1, trabaja sobre la representación  $\bar{\mathbf{F}}$  del cubo de datos, el vector disperso  $\hat{\theta}$  que resuelve el problema de optimización de la ecuación (3.8), es obtenido utilizando el algoritmo GPSR propuesto en [50]. El problema de dispersión restringida de la ecuación (3.10) es resuelto por el algoritmo SOMP, con un nivel de dispersión fijo de  $K_0 = 4$  y un vecindario de dispersión conjunta de  $T = 9$ . Se usa la estimación de la matriz  $\bar{\mathbf{A}}$  para el cálculo de la matriz de puntajes  $\mathbf{D}$ , cuyas entradas representan la probabilidad de que el pixel

---

correspondiente sea un pixel objetivo.

**Tabla 4.1:** Desempeño detector tradicional

Desempeño del método de detección basado en dispersión tradicional	
Dataset RIT-CIS-DIRT	Imagen del sensor Hyperion
99.9595 %	99.7049 %

El desempeño del detector basado en dispersión para ambas imágenes se presenta en la tabla 4.1, que muestra la probabilidad de detección de éste para ambos experimentos.

Los resultados obtenidos se analizan cuantitativa y cualitativamente utilizando las curvas ROC, que describen la probabilidad de detección (PD) en función de la probabilidad de falsa alarma (PDA). Para la construcción de las curvas ROC, se escogieron miles de valores de umbral de entre el valor mínimo y máximo arrojado por el detector. Las etiquetas de los pixeles evaluados fueron asignadas para cada valor de umbral utilizado.

## 4.1. Diseño de las aperturas codificadas

La calidad de las imágenes reconstruidas en el proceso de muestreo compresivo depende en gran medida del diseño de las aperturas codificadas. En el sistema CASSI se han empleado distintos tipos de códigos de apertura, incluyendo códigos aleatorios, códigos booleanos, códigos binarios, en escala de grises y de Hadamard [52].

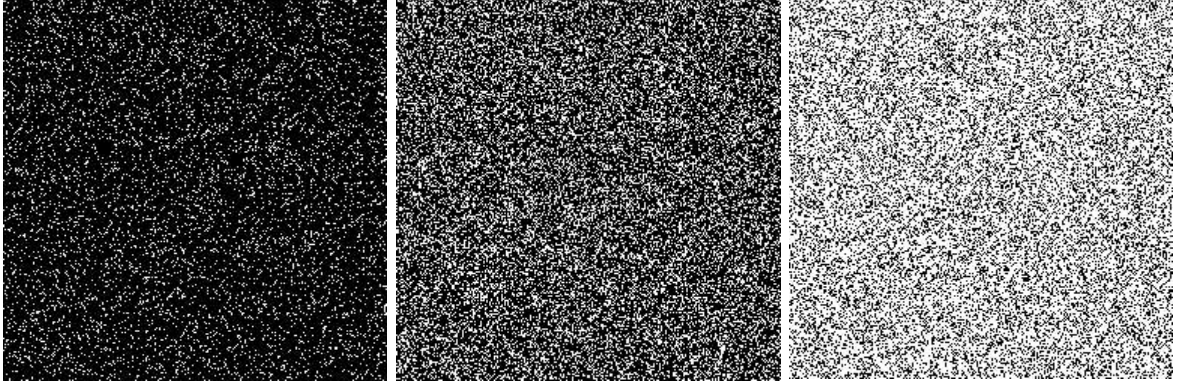
Para el desarrollo de este trabajo se utilizaron códigos de apertura aleatorios, cuyos elementos traslucidos satisfacen la condición  $t_{i,j}^\ell \in \{0, 1\}$  donde  $\ell$  indica el número de la proyección,  $t_{i,j}^\ell = 1$  representando un elemento que permite el paso de la luz y  $t_{i,j}^\ell = 0$  un elemento que la bloquea.

La transmitancia de los códigos de apertura está dada por

$$Tr = \sum_{i=0}^{N-1} \sum_{j=1}^{N-1} \frac{t_{i,j}}{N^2}, \quad (4.1)$$

donde  $N^2$  representa el tamaño del código de apertura. En la figura 4.3 se muestran

**Figura 4.3:** Códigos de apertura



(a) 10% Compresión, 10% Re-reflectancia (b) 30% Compresión, 30% Re-reflectancia (c) 30% Compresión, 70% Re-reflectancia

distintos códigos de apertura con niveles de transmitancia, que reflejan el porcentaje de elementos ópticos que permiten el paso de luz.

Asimismo, el número de proyecciones captadas, o captaciones producidas por el detector, afectan la calidad de las imágenes reconstruidas, pues representan el porcentaje de información utilizada, el nivel de compresión, expresado de la forma:

$$Tr = \frac{K(N(N + L - 1))}{N^2L}, \quad (4.2)$$

donde  $K$  representa el número de captaciones, y  $N$  y  $L$  la dimensión espacial y espectral del cubo de datos, la ecuación (4.2) puede ser vista como la razón entre el número de mediciones y el número de píxeles del cubo de datos reconstruido.

## 4.2. Dataset RIT-CIS-DIRT

La primera imagen hiperespectral usada para evaluar el desempeño del clasificador es el dataset self test de la ciudad de Cooke [51]. Para las pruebas el número de bandas espectrales fue reducido a 90(3-46,49,51-62,66,69-72,68-122),

---

se eliminaron 38 bandas correspondientes a ruido atmosférico, absorción y bajo SNR. La imagen cuenta con una resolución espacial de 3m por pixel y dimensiones de  $800 \times 200$  pixeles. Sin embargo, solo  $256 \times 256$  pixeles son usados en los experimentos, debido a que se requieren matrices de dimensiones diádicas para la implementación rápida de la transformada wavelet en 2D.

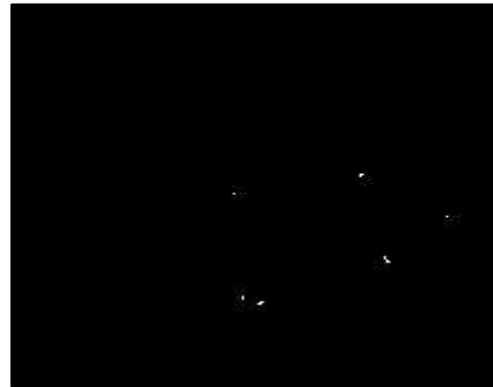
El paquete self test contiene un conjunto de librerías espectrales correspondientes a los diferentes objetivos presentes en la imagen; para este experimento, se usó la información espectral contenida en el archivo `SPL F1_f.txt` para la construcción del diccionario de entrenamiento de objetivo. Un total de  $N_t = 16$  pixeles que mostraron la mayor correlación con la firma espectral del archivo `F1_f.txt`, fueron extraídos de la imagen para formar el diccionario de representación  $\bar{M}^t$ , como se muestra en el Algoritmo.1. De los pixeles de la imagen que no fueron usados en el experimento, se seleccionaron de manera aleatoria  $N_b = 350$  y con ellos se construyó el diccionario de entrenamiento de fondo  $\bar{M}^b$ .

Cualitativamente, los resultados experimentales se observan en la figura 4.4, en donde los puntos blancos representan los pixeles etiquetados como objetivo, y los negros los pixeles etiquetados como fondo.

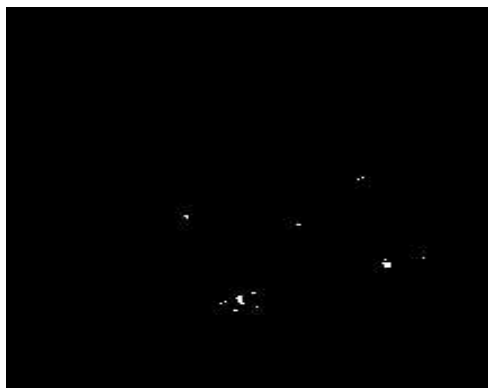
**Figura 4.4:** Resultados de la detección D para la imagen RIT-CIS-DIRT



(a) Porcion de la imagen evaluada

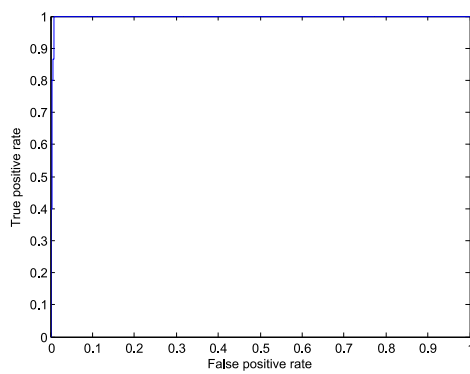


(b) Verdad de terreno o Ground Truth del objetivo en la imagen

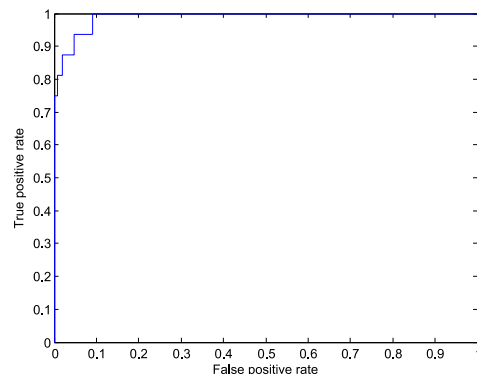


(c) Resultados de la deteccion con un 40 % de compresion,y un 20 % Reflectancia

**Figura 4.5:** Curvas ROC para la imagen RIT-CIS-DIRT



(a) Imagen sin comprimir



(b) 40 % Compresion, 20 % Reflectancia

Cuantitativamente, los resultados de las pruebas realizadas sobre esta imagen,

se presentan en la Figura 4.5 y en la tabla 4.2. La figura 4.5 muestra 2 curvas ROC generadas por el proceso de detección tradicional a la imagen sin comprimir, es decir, adquirida por un sensor hiperespectral y cuyo proceso de detección utilizó toda la información espectral muestreada, y por el proceso de detección propuesto, utilizando el sistema de muestreo compresivo CASSI, con el 40 % de la información espectral y códigos de apertura con una transmitancia del 20 %.

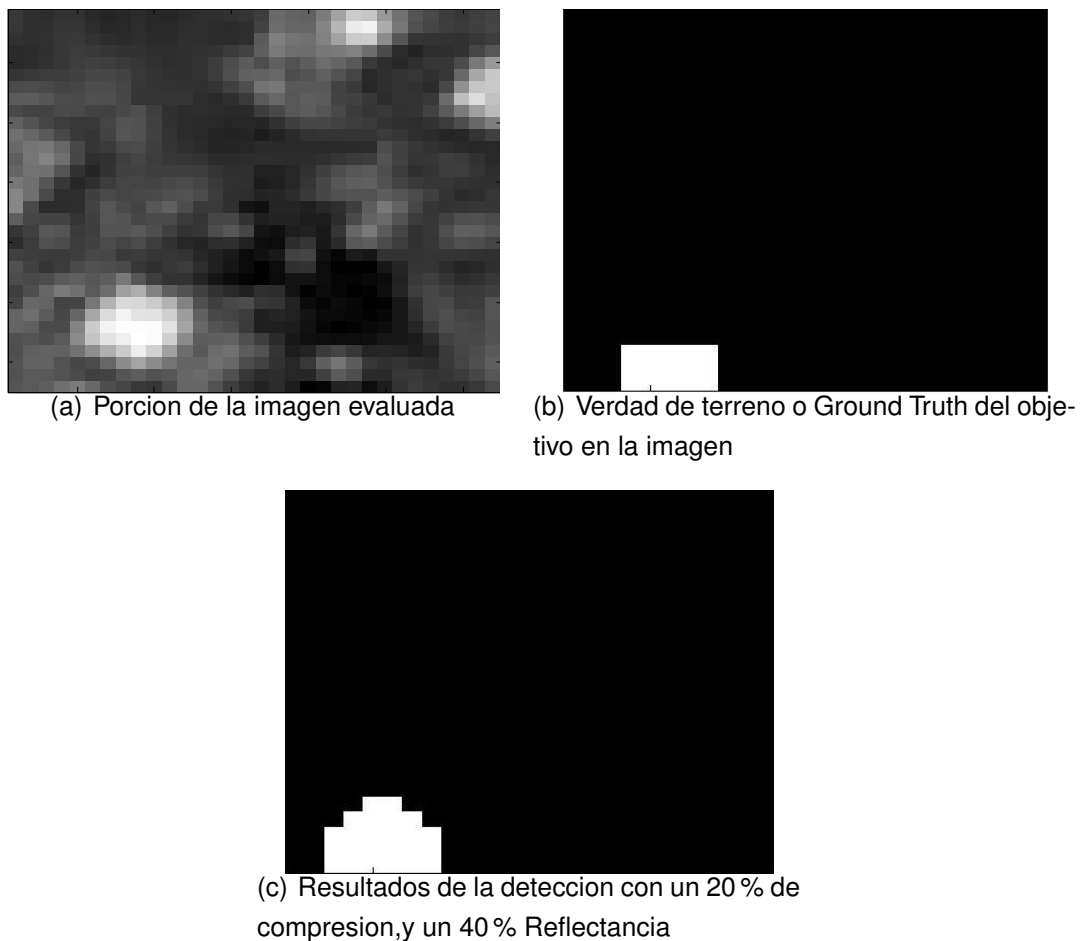
**Tabla 4.2:** Resultados experimentales para la imagen RIT-CIS-DIRT. En la tabla se muestran las probabilidades de detección obtenidas por el detector para diferentes niveles de compresión y transmitancia.

Probabilidad de detección (AUC) en función de nivel de compresión y transmitancia					
		Porcentaje de información espectral usada			
		40 %	30 %	20 %	15 %
Nivel de Transmitancia	50 %	85.3417 %	80.7505 %	81.3288 %	81.2644 %
	40 %	97.6642 %	85.7213 %	84.1170 %	84.2186 %
	30 %	95.4088 %	89.9378 %	84.8412 %	83.3338 %
	20 %	98.9259 %	95.4751 %	88.3528 %	88.3329 %
	10 %	98.0153 %	93.0168 %	94.3981 %	88.3216 %

### 4.3. Imagen del sensor Hyperion

Para los experimentos realizados con esta imagen, se utilizó un recorte de  $32 \times 32$  píxeles. Un conjunto de 10 píxeles de objetivo fueron tomados de la imagen y utilizados como muestras de entrenamiento. 206 píxeles del resto de la imagen fueron utilizados como muestras del fondo para el entrenamiento de los diccionarios. Así como en los experimentos anteriores, el algoritmo utilizado para resolver el problema de la ecuación (3.8), es el GPRS [50].

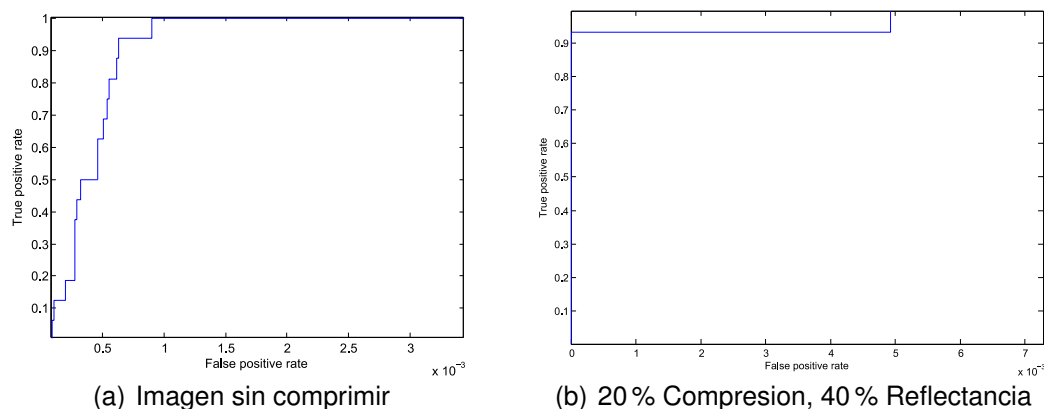
**Figura 4.6:** Resultados de la detección D para la imagen del sensor Hyperion



Al igual que con la primera imagen, cualitativamente, los resultados experimentales se observan en la figura 4.6, en donde los puntos blancos representan los píxeles etiquetados como objetivo, y los negros los píxeles etiquetados como fondo.

Cuantitativamente, los resultados de las pruebas realizadas se presentan en la Figura 4.7 y en la tabla 4.3. La figura 4.7 muestra 2 curvas ROC generadas por el proceso de detección tradicional a la imagen sin comprimir, es decir, adquirida por un sensor hiperespectral y cuyo proceso de detección utilizó toda la información espectral muestreada, y por el proceso de detección propuesto, utilizando el sistema de muestreo compresivo CASSI, con el 20 % de la información espectral y códigos de apertura con una transmitancia del 40 %.

**Figura 4.7:** Curvas ROC para la imagen del sensor Hyperion



La tabla 4.3 muestra las distintas probabilidades de detección obtenidas variando los parámetros de tasa de compresión y nivel de transmitancia de las aperturas codificadas.

**Tabla 4.3:** Resultados experimentales para la imagen Hyperion. En la tabla se muestran las probabilidades de detección obtenidas por el detector para diferentes niveles de compresión y transmitancia.

Probabilidad de detección (AUC) en función de nivel de compresión y transmitancia					
		Porcentaje de información espectral usada			
		40 %	30 %	20 %	15 %
Nivel de Transmitancia	50 %	99.9781 %	99.8797 %	99.8907 %	99.6612 %
	40 %	99.6502 %	99.9234 %	99.9672 %	99.8688 %
	30 %	99.7158 %	99.9016 %	99.8251 %	96.5136 %
	20 %	99.6612 %	99.4863 %	99.3661 %	99.1366 %
	10 %	99.8907 %	99.5519 %	94.7759 %	85.8469 %

## Capítulo 5

### Conclusiones, recomendaciones y trabajo futuro

#### 5.1. Conclusiones

- Se determina un modelo matemático de detección de objetivos en imágenes hiperespectral que utiliza mediciones adquiridas mediante un sistema de muestreo compresivo en lugar de información espectral pura.
- El modelo de detección propuesto logra una probabilidad de detección de hasta el 98.92 % utilizando el 40 % de la información de la escena espectral para la imagen RIT-CIS-DIRT, y una probabilidad de detección de hasta el 99.97 % utilizando el 40 % de la información de la escena espectral para la imagen del sensor Hyperion, en comparación con la probabilidad de detección del 99.95 % utilizando el 100 % de la información espectral.
- Se descubrió que las aperturas codificadas que más benefician el proceso de detección son las que tienen entre un 10 % y un 20 % de transmitancia.
- La diferencia entre la probabilidad de detección del método tradicional usando el 100 % de la información espectral y del método propuesto usando el 40 % de la información es del 1 %.

#### 5.2. Recomendaciones y trabajo futuro

- Como trabajo futuro se plantea el estudio de nuevas formas para construir los diccionarios de representación  $\bar{M}^b$  y  $\bar{M}^t$  de forma que se disminuya su dependencia de la información espectral conocida de la imagen.

## Referencias

- [1] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Transactions on Information Theory*, vol. 14, no. 1, 1968.
- [2] S. Rao, R. Tron, R. Vidal, and Y. Ma, "Motion segmentation in the presence of outlying, incomplete, or corrupted trajectories," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 10, pp. 1832–1845, 2010.
- [3] E. Candes and M. Wakin, "An Introduction To Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.
- [4] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches," *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2008.
- [5] H. Li and F. Liu, "Image denoising via sparse and redundant representations over learned dictionaries in wavelet domain," *Proceedings of the 5th International Conference on Image and Graphics, ICIG 2009*, vol. 15, no. 12, pp. 754–758, 2010.
- [6] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration." *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, vol. 17, no. 1, pp. 53–69, 2008.
- [7] J. Wright, a. Y. Yang, a. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation." *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 2, pp. 210–227, 2009. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/21646680>
- [8] J. K. Pillai, V. M. Patel, and R. Chellappa, "Sparsity inspired selection and recognition of iris images," *IEEE 3rd International Conference on Biometrics: Theory, Applications and Systems, BTAS 2009*, 2009.
- [9] X. Hang and F.-X. Wu, "Sparse representation for classification of tumors using gene expression data." *Journal of biomedicine & biotechnology*, vol. 2009, p. 403689, 2009.

- 
- [10] Y. Chen, N. M. Nasrabadi, T. D. Tran, and S. Member, "Hyperspectral Image Classification Using Dictionary-Based Sparse Representation," *October*, vol. 49, no. 10, pp. 3973–3985, 2011.
- [11] A. Wagadarikar, R. Jhon, R. Willet, and Brady.D, "Single disperser design for coded aperture snapshot spectral imaging," *Appl.Opt.*, vol. 47, no. 10, pp. B44–B5, 2008.
- [12] T. Sung and K. F. Kelly, "Compressive sensing hyperspectral imager," in *Comput.Opt.Sens.Imag.,..* San Jose, CA,USA: OSA, 2009.
- [13] M. a. Davenport, P. T. Boufounos, M. B. Wakin, and R. G. Baraniuk, "Signal processing with compressive measurements," *IEEE Journal on Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 445–460, 2010.
- [14] G. Camps-Valls, D. Tuia, L. Bruzzone, and J. A. Benediktsson, "Advances in hyperspectral image classification: Earth monitoring with statistical learning methods," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 45–54, 2014.
- [15] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images," *SIAM Review*, vol. 51, no. 1, pp. 34–81, 2009.
- [16] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 6, pp. 1351–1362, 2005.
- [17] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Simultaneous joint sparsity model for target detection in hyperspectral imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 4, pp. 676–680, 2011.
- [18] —, "Sparse Representation for Target Detection in Hyperspectral Imagery," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 5, no. 3, pp. 629–640, 2011.
- [19] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *Signal Processing Magazine, IEEE*, vol. 19, no. 1, pp. 29–43, 2002.
- [20] —, "Detection algorithms for hyperspectral imaging applications," *Signal Processing Magazine, IEEE*, vol. 19, no. 1, pp. 29–43, 2002.

- 
- [21] G. Healey, "Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 6, pp. 2706–2717, 1999.
- [22] J. C. Harsanyi and C. I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 4, pp. 779–785, 1994.
- [23] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 1, pp. 208–216, 1992.
- [24] H. Kwon and N. M. Nasrabadi, "A comparative analysis of kernel subspace target detectors for hyperspectral imagery," *Eurasip Journal on Advances in Signal Processing*, vol. 2007, 2007.
- [25] A. Ramirez, G. R. Arce, and B. M. Sadler, "Spectral Image Unmixing From Optimal Measurements," vol. 53, no. 1, pp. 405–415, 2015.
- [26] A. Ramirez, S. Member, H. Arguello, S. Member, G. R. Arce, and B. M. Sadler, "Spectral Image Classification From Optimal Coded-Aperture Compressive Measurements," vol. 52, no. 6, pp. 3299–3309, 2014.
- [27] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1760–1770, 1990.
- [28] H. Kwon and N. M. Nasrabadi, "Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 2, pp. 388–397, 2005.
- [29] L. Gómez-Chova, G. Camps-Valls, L. Bruzzone, and J. Calpe-Maravilla, "Mean map kernel methods for semisupervised cloud classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 1, pp. 207–220, 2010.
- [30] Y. Gu, C. Wang, D. You, Y. Zhang, S. Wang, and Y. Zhang, "Representative multiple kernel learning for classification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 7 PART 2, pp. 2852–2865, 2012.

- 
- [31] J. Li, P. R. Marpu, A. Plaza, J. M. Bioucas-Dias, and J. A. Benediktsson, "Generalized composite kernel framework for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 9, pp. 4816–4829, 2013.
- [32] a. Rakotomamonjy, F. Bach, S. Canu, and Y. Grandvalet, "SimpleMKL," *Journal of Machine Learning Research*, vol. 9, pp. 2491–2521, 2008.
- [33] N. M. Nasrabadi, "Hyperspectral Target Detection : An Overview of Current and Future Challenges," *Signal Processing Magazine, IEEE*, vol. 31, no. 1, pp. 34–44, 2014.
- [34] L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 2146–2156, 1994.
- [35] S. Kraut and L. L. Scharf, "Adaptive subspace detectors," *IEEE Transactions on Signal Processing*, vol. 49, no. 1, pp. 1–16, 2001.
- [36] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Kernel sparse representation for hyperspectral target detection," *International Geoscience and Remote Sensing Symposium (IGARSS)*, no. 2, pp. 7484–7487, 2012.
- [37] W. Verhoef and H. Bach, "Simulation of hyperspectral and directional radiance images using coupled biophysical and atmospheric radiative transfer models," *Remote Sensing of Environment*, vol. 87, no. 1, pp. 23–41, 2003.
- [38] A. Mueller, G. Aerospace, C. Dlr, S. Isis, and T. Committee, "International Spaceborne Imaging," no. march, pp. 18–21, 2014.
- [39] a. Plaza, J. Plaza, A. Paz, and S. Sánchez, "Parallel Hyperspectral Image and Signal Processing," *IEEE Signal Processing Magazine*, vol. 28, no. May, pp. 119–126, 2011. [Online]. Available: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=5753072](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5753072)
- [40] S. Lopez, T. Vladimirova, C. Gonzalez, J. Resano, D. Mozos, and A. Plaza, "The promise of reconfigurable computing for hyperspectral imaging onboard systems: A review and trends," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 698–722, 2013.

- 
- [41] H. Goldberg and N. M. Nasrabadi, "A comparative study of linear and nonlinear anomaly detectors for hyperspectral imagery," in *SPIE*, Orlando, FL, 2007.
- [42] M. Elad, "Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing," in *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York: Springer Science & Business Media, 2010.
- [43] S. Boyd and L. Vandenberghe, *Convex Optimization*, 2010, vol. 25, no. 3. [Online]. Available: [https://web.stanford.edu/~boyd/cvxbook/bv\\_cvxbook.pdf](https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf)
- [44] E. J. Lentilucci and P. Bajorski, "Stochastic modeling of physically derived signature spaces," *Journal of Applied Remote Sensing*, vol. 2, no. 1, p. 023532, 2008.
- [45] J. a. Tropp, "Algorithms for simultaneous sparse approximation. Part II: Convex relaxation," *Signal Processing*, vol. 86, no. 3, pp. 589–602, 2006.
- [46] S. F. Cotter, B. D. Rao, K. Engan, K. Kreutz-delgado, and S. Member, "Sparse solutions to linear inverse problems with Multiple Measurement Vectors," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2477–2488, 2005.
- [47] Q. Du and J. E. Fowler, "Hyperspectral image compression using JPEG2000 and principal component analysis," *Geoscience and Remote Sensing Letters, IEEE*, vol. 4, no. 2, pp. 201–205, 2007.
- [48] M. Pahor, "Principal Component Analysis," vol. 18, no. Figure 1, pp. 2230–2242, 2009.
- [49] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An Interior-Point Method for Large-Scale L1-Regularized Least Squares," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 606–617, 2007. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4407767>
- [50] M. Figueiredo, R. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems, IEEE J," *Sel. Top. Signa*, vol. vol, no. 1, pp. 1pp586–597, 2007.

- 
- [51] D. Snyder, J. Kerekes, I. Fairweather, R. Crabtree, J. Shive, and S. Hager, "Development of a web-based application to evaluate target finding algorithms," *International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2, no. 1, pp. 915–918, 2008.
- [52] H. B. Arguello and G. Arce, "Restricted isometry property in coded aperture compressive spectral imaging," *2012 IEEE Statistical Signal Processing Workshop, SSP 2012*, vol. 0, pp. 716–719, 2012. [Online]. Available: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84868243435&partnerID=40&md5=1b9c4f025b899481ee8b3da251efffe8>