
**IDENTIFICACIÓN E IMPLEMENTACIÓN DE ALGORITMOS DE
DETECCIÓN DE PATRONES EN IMÁGENES HIPERESPECTRALES
ADQUIRIDAS POR SENSORES REMOTOS, APLICADAS AL SECTOR
DE LA AGRICULTURA COLOMBIANA**

ARIOLFO CAMACHO VELASCO

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISICOMECÁNICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
MAESTRÍA EN INGENIERÍA DE SISTEMAS
BUCARAMANGA
2016**

**IDENTIFICACIÓN E IMPLEMENTACIÓN DE ALGORITMOS DE
DETECCIÓN DE PATRONES EN IMÁGENES HIPERESPECTRALES
ADQUIRIDAS POR SENSORES REMOTOS, APLICADAS AL SECTOR
DE LA AGRICULTURA COLOMBIANA**

ARIOLFO CAMACHO VELASCO

**Trabajo de Investigación para optar por el título de:
Magíster en Ingeniería de Sistemas e Informática**

Director:

HENRY ARGUELLO FUENTES
Ingeniero Electricista. PhD.

Co-director:

CESAR AUGUSTO VARGAS GARCÍA
Ingeniero de Sistemas. PhD. (c)

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISICOMECAÑICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
MAESTRÍA EN INGENIERÍA DE SISTEMAS
BUCARAMANGA
2016**

DEDICATORIA

A mi esposa, Licie Katerine.

Por tu compañía y apoyo en este proyecto, desde el inicio donde era un sueño. Solo tú y yo sabemos lo que ha significado este cambio de estilo de vida. Gracias por ayudarme a enfrentar y superar los obstáculos y limitantes, gracias por no dejarme desfallecer con tus palabras y amor, justo en los momentos difíciles. Todo este trabajo y esfuerzo a ti amada esposa por permitirme y ayudarme a ser padre y maestro.

A mi hija Isabel Sofia.

Amada hija, cuando brillan tus ojos y sonrisa, brillan las vidas de máma y pápa, iluminando con un amor infinito nuestro dulce hogar. Isabel Sofia, tú eres esperanza, alegría, sinceridad, paz y sencillez; haces que todo esfuerzo y sacrificio tengan un sentido.

A mis Familiares.

Padres, han realizado mucho con tan poco, los admiro, gracias por su amor y dedicación. A mi hermana Diana Isabel y mi sobrina Valentina por su paciencia, hermanita todo es posible con disciplina y trabajo. A mi hermano Pedro Nel y su esposa, Yudi, en sus manos está la posibilidad de desarrollar el más alto potencial de mis sobrinos, basados en una educación temprana y un excelente ejemplo de vida.

A Dios.

Por haberme bendecido con los dones que me ha otorgado, por permitir llegar al camino de la ciencia y la educación nuevamente. Por haberme dado salud para lograr mis objetivos, además de su infinita bondad y amor.

RECONOCIMIENTO

Agradezco a todos mis compañeros y profesores que creyeron en mi capacidad, capacidad que tenemos todos, es grato contar con un grupo de trabajo fuerte y disciplinado como lo es HDSP. Agradezco a su director el profesor Henry por su formación, no fue nada fácil, sin embargo, ha dejado una gran huella de cambio.

Agradezco la colaboración sobresaliente y la amabilidad del Dr. Stuard Frye, funcionario del centro de vuelos espaciales Goddard de la NASA, quien dio un apoyo permanente para la adquisición de las imágenes hiperespectrales sobre el territorio Colombiano.

A la Universidad Industrial de Santander y a cada profesor que entendió mis orígenes y competencias, en especial a la visión de inclusión de la profesora Sonia Gamboa, en su periodo como directora de la Maestría de Ingeniería de Sistemas e Informática.

CONTENIDO

INTRODUCCIÓN	5
1. TEORÍA DE SENSADO REMOTE	15
1.1. SENSADO REMOTO	15
1.2. IMAGENES HIPERESPECTRALES	15
1.3. SENSORES REMOTOS HIPERESPECTRALES	17
2. SENSADO REMOTO HIPERESPECTRAL APLICADO A LA AGRICULTURA	18
2.1. PROPIEDADES ESPECTRALES DE LA VEGETACIÓN	18
2.2. SENSADO REMOTO HIPERESPECTRAL APLICADO A LA AGRICULTURA COLOMBIANA	19
3. ALGORITMOS DE DETECCIÓN DE OBJETIVOS EN IMÁGENES HIPERESPECTRALES	21
3.1. ALGORITMO DE FILTRO DE EMPAREJADO (MF)	22
3.2. ALGORITMO DE MINIMIZACIÓN DE ENERGÍA RESTRINGIDA (CEM)	22
3.3. ALGORITMO ESTIMACIÓN DE COHERENCIA ADAPTATIVO (ACE)	23
3.4. ALGORITMO MAPEO DEL ÁNGULO ESPECTRAL (SAM)	23
3.5. ALGORITMO DE PROYECCIÓN SUBESPACIO ORTOGONAL (OSP)	24
4. MÉTRICAS DE DESEMPEÑO DE ALGORITMOS DE DETECCIÓN DE OBJETIVOS EN HSI	25
4.1. MÉTRICA POR PÍXEL Y MATRICES DE CONFUSIÓN	25
4.2. MÉTRICA VISUAL	26
4.3. CURVAS ROC (RECEIVER OPERATOR CHARACTERISTICS)	27
5. MÉTODO Y RESULTADOS	28
5.1. CONJUNTO DE DATOS DE IMÁGENES HIPERESPECTRALES REALES Y ÁREAS DE ESTUDIO	28
5.2. PREPROCESAMIENTO DE IMÁGENES HIPERESPECTRALES	29
5.3. CONJUNTO DE DATOS SINTÉTICOS E IMPLANTACIÓN DE OBJETIVOS FRACCIONADOS	30
5.4. EVALUACIÓN DE DESEMPEÑO DE ALGORITMOS DE DETECCIÓN DE OBJETIVOS CON DATOS SINTÉTICOS	32
5.4.1 Experimentos y resultados con datos sintéticos sin ruido	32
5.4.2 Experimentos y resultados con datos sintéticos con SNR=10 dB, SNR=15 dB, SNR=20 dB, SNR=25 dB y SNR=30 dB	35

5.5	EVALUACIÓN DE DESEMPEÑO DE ALGORITMOS DE DETECCIÓN DE OBJETIVOS CON DATOS REALES	37
5.5.1	Resultados de algoritmos de detección de objetivos en HSI del área de estudio de Puerto Wilches (PW)	37
5.5.2	Resultados de algoritmos de detección de objetivos en HSI del área de estudio de San Alberto (SA)	43
5.5.3	Resultados de algoritmos de detección de objetivos en HSI del área de estudio de Rojo Negro (RL)	47
5.5.4	Resultados de algoritmos de detección de objetivos en HSI del área de estudio de Mogotes (MG)	49
5.5.5	Resultados de algoritmos de detección de objetivos en HSI del área de estudio de Guadalupe (GL)	51
5.6.	RESUMEN DE RESULTADOS	53
6.	LONGITUDES DE ONDA CON MAYOR IMPORTANCIA EN EL ESTUDIO ESPECTRAL DE LA VEGETACIÓN	55
6.1.	ANÁLISIS DE LONGITUDES DE ONDA A PARTIR DE LAS MATRICES DE CORRELACIÓN Y LOS VALORES PROPIOS	56
6.2.	ANÁLISIS DE LA DISTRIBUCIÓN ESPACIAL Y ESPECTRAL DE LA VEGETACIÓN EN UNA HSI	57
6.3.	ESTADO DEL ARTE DE LONGITUDES DE ONDA Y BANDAS ESPECTRALES OPTIMAS PARA EL ESTUDIO DE LA VEGETACIÓN	59
7.	VIABILIDAD TÉCNICA Y CIENTÍFICA PARA LA CREACIÓN DE UN PROTOTIPO DE LABORATORIO DE SENSADO REMOTO HIPERESPECTRAL PARA EL NORORIENTE COLOMBIANO	61
7.1.	JUSTIFICACIÓN DE UN PROTOTIPO DE LABORATORIO DE SENSADO REMOTO HIPERESPECTRAL PARA EL NORORIENTE COLOMBIANO	61
7.2.	ESTRUCTURA DE UN LABORATORIO DE SENSADO REMOTO	62
7.3.	PROTOTIPO DE LABORATORIO DE SR PLANTEADO	63
7.4.	SELECCIÓN DE EQUIPOS BÁSICOS PARA EL PROTOTIPO DEL LABORATORIO DE SR PLANTEADO	64
8.	CONCLUSIONES, CONTRIBUCIONES Y TRABAJOS FUTUROS	67
8.1.	CONCLUSIONES	67
8.2.	CONTRIBUCIONES	68
8.3.	TRABAJOS FUTURO	69
	CITAS	70
	BIBLIOGRAFÍA	75

ÍNDICE DE FIGURAS

Figura 1.	Esquema general de un sistema de SR	15
Figura 2.	Concepto de HSI y firma espectral	16
Figura 3.	Clasificación de imágenes según su resolución espectral	16
Figura 4.	Características espectrales de la vegetación.	18
Figura 5.	Métrica por píxel de objetivos y fondo	26
Figura 6.	Imágenes bidimensionales (x, y) de resultados de algoritmos de detección de objetivos	27
Figura 7.	Curvas ROC, métrica de desempeño que muestra la P_D frente a la P_{FA} de uno o varios detectores	27
Figura 8.	Áreas de estudio ubicadas en los departamentos de Santander y Cesar, Colombia	28
Figura 9.	Flujo de trabajo de pre-procesamiento de HSI	30
Figura 10.	HSI del área de estudio MG con objetivos implantados	30
Figura 11.	Firmas espectrales de fondo y objetivos implantados a diferentes niveles de mezcla.	31
Figura 12.	Imágenes de banda espectral 29 que corresponde a la HSI sintética con los objetivos implantados, y diferentes niveles de SNR.	32
Figura 13.	Resultado de algoritmos con un objetivo implantado del 10% de la Firma pura de la alunita en HSI de prueba sin ruido	33
Figura 14.	Curva ROC, desempeño de los algoritmos ACE, CEM, OSP, MF y SAM con un objetivo del 10% de la firma pura de alunita	33
Figura 15.	Curvas ROC desempeño de detección de objetivo implantado con diferentes fracciones de la firma pura de alunita en HSI sin ruido	34
Figura 16.	Curvas ROC y respuesta de detección para objetivos puros en HSI Sin ruido	34
Figura 17.	Resultado de respuesta de los detectores sobre HSI sintética con SNR=10 dB y con una firma mezclada con fracción $f=10\%$ como objetivo de tipo sub-píxel a detectar	35
Figura 18.	Número de FA a diferentes niveles de ruido	36
Figura 19.	Curvas ROC de resultado de diferentes algoritmos al detectar un Objetivo tipo sub-píxel creado con una forma mezclada al 10% a diferentes niveles de ruido	36
Figura 20.	Curvas ROC de resultado de diferentes algoritmos al detectar un Objetivo de tipo píxel completo a diferentes niveles de ruido	37
Figura 21.	HSI captadas por el sensor Hyperion para el área de estudio PW en color verdadero (bandas 29=rojo, 20=verde, y 12=azul)	38

Figura 22.	Firmas espectral de cultivos característicos en área de estudio PW	39
Figura 23.	Coeficientes de determinación R2 de píxeles frente a la media de cultivo de palma africana enferma en PW	40
Figura 24.	Resultados de los algoritmos en la detección de palma africana sana	40
Figura 25.	Resultado de los algoritmos en la detección de palma africana Enferma	41
Figura 26.	Resultado de los algoritmos en la detección de cultivo de caucho	42
Figura 27.	Curvas ROC de desempeño de detección para los objetivos en PW	42
Figura 28.	HSI del área de estudio de San Alberto (SA), Cesar	43
Figura 29.	Firmas espectral de cultivos característicos en área de estudio SA	44
Figura 30.	Resultado de los algoritmos en la detección de cultivo o áreas dedicadas a pastos para ganado en SA	45
Figura 31.	Resultado de los algoritmos en la detección de cultivos de palma Africana adulta.	45
Figura 32.	Resultado de los algoritmos en la detección de cultivos de palma africana en vivero	45
Figura 33.	Curvas ROC de desempeño en la detección de cultivos en SA	
Figura 34.	PD de los algoritmos con una $P_{FA} = 0,009$ para los objetivos de pasto, palma africana adulta y palma africana en vivero	45
Figura 35.	Desempeño de los algoritmos basados en umbrales para objetivo de palma africana en vivero y sus matrices de confusión.	47
Figura 36.	HSI del área de estudio de Río Negro (RL), Santander	48
Figura 37.	Firma espectral de cultivos de cítricos en RL	48
Figura 38.	Curvas ROC de desempeño en cultivos de cítricos en RL.	49
Figura 39.	HSI del área de estudio de Mogotes (MG), Santander	49
Figura 40.	Firmas espectral de cultivos característicos en área de MG	50
Figura 41.	Curvas ROC de desempeño de detección de objetivos en MG	51
Figura 42.	HSI del área de estudio de Guadalupe (GL), Santander	51
Figura 43.	Firmas espectral de cultivos característicos en área de GL	52
Figura 44.	Curvas ROC de desempeño para cultivos de pasto tipo sabana como objetivo en GL y número de FA para una $P_{FA} = 0,004$	53
Figura 45.	Matrices de correlación en áreas de estudio PW, SA, RL, y MG	56
Figura 46.	Comportamiento de los valores propios de cada banda espectral	57
Figura 47.	Dispersión de espacial y espectral del NIR y el rojo en HSI de SA	58
Figura 48.	Dispersión de espacial y espectral del NIR y el rojo en HSI de MG	58
Figura 49.	Equipos básicos para un laboratorio de SR	64
Figura 50.	Componentes del sistema de SR de bajo costo propuesto para el Prototipo de laboratorio de SR	66

ÍNDICE DE TABLAS

Tabla 1.	Áreas de estudio, cultivos y principales características	29
Tabla 2.	Número de falsas alarmas frente a probabilidades de detección del 90% y 80 %.	42
Tabla 3.	Longitudes de onda seleccionadas por análisis de las matrices de correlación	56
Tabla 4.	Resumen de longitudes de onda y bandas espectrales según estado del arte	60
Tabla 5.	Áreas de aplicación de prototipo de laboratorio de SR	63
Tabla 6.	Estudio comparativos de costos y características técnicas de Equipos de prototipo de laboratorio de SR	65

RESUMEN

TITULO: IDENTIFICACIÓN E IMPLEMENTACIÓN DE ALGORITMOS DE DETECCIÓN DE PATRONES EN IMÁGENES HIPERESPECTRALES ADQUIRIDAS POR SENSORES REMOTOS, APLICADAS AL SECTOR DE LA AGRICULTURA COLOMBIANA ¹

AUTOR: ARIOLFO CAMACHO VELASCO ²

PALABRAS CLAVE: Sensado remoto, imagen hiperespectral, algoritmo de detección de objetivos en imágenes hiperespectrales, propiedades espectrales de la vegetación.

La técnica de sensado remoto hiperespectral permite adquirir información espacial, temporal y espectral de un objeto, o escena. Las imágenes hiperespectrales (HSI) captan información en alta resolución espectral, acerca de las características espectrales de materiales o vegetación presente sobre la superficie de la Tierra. La aplicación e investigación de las HSI es de importancia en diferentes áreas, dentro de las cuales se destaca el área agrícola. Existen diversos algoritmos que permiten aprovechar la información contenida en una HSI, tales como algoritmos de clasificación, detección de anomalías y algoritmos de detección de objetivos. Específicamente la tarea de un algoritmo de detección de objetivos en HSI es detectar un píxel deseado (objetivo) en una escena HSI con el menor número de falsas alarmas (FA). Los algoritmos de detección de objetivos usados en HSI en las dos últimas décadas han presentado mejoras, modificaciones y nuevas versiones. Sin embargo, la tarea de encontrar un algoritmo de detección que cuente con un desempeño óptimo y que este libre de ajustes previos para diferentes escenarios y objetivos, aun es materia de investigación. Colombia posee condiciones geográficas y climáticas particulares, generando escenarios diversos, permitiendo evaluar el desempeño de los algoritmos seleccionados para diferentes tipos de escenarios, relieves, climas y cultivos agrícolas. La presente investigación realiza la identificación e implementación de algoritmos de detección de objetivos en HSI adquiridas por sensores remotos, sobre el territorio colombiano.

En la evaluación de los algoritmos y el desarrollo de la presente investigación se trabajó con 20 HSI reales; 6 HSI sintéticas, 1 HSI sintética sin ruido y 5 con ruido, se implantaron 5 objetivos sintéticos, creados a partir de la firma espectral pura de la alunita a diferentes grados de mezcla, y se extrajeron más de 115 firmas espectrales reales, de las cuales 11 fueron usadas como objetivo, permitiendo la caracterización de cinco cultivos agrícolas del nororiente colombiano en 5 diferentes áreas de estudio. Los resultados muestran que el algoritmo ACE presenta un mejor desempeño para diferentes HSI y objetivos de tipo agrícola, tanto en imágenes sintéticas como reales.

¹Trabajo de Investigación

²Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería de Sistemas e informática. Director, Henry Arguello Fuentes

ABSTRACT

TITLE: Identification and implementation of algorithms target detection in hyperspectral imagery acquired by remote sensing, Applied to the colombian agriculture ³

AUTHOR: ARIOLFO CAMACHO VELASCO ⁴

KEYWORDS: Remote sensing, hyperspectral imagery, target detection in hyperspectral imagery, spectral properties of vegetation.

Hyperspectral remote sensing techniques allow spatial, spectral and temporal information from an object or a spectral scene. Hyperspectral Images (HSI) yield high spectral resolution information from materials and vegetation features of the surface. HSI applications and research are of great importance in diverse fields, in which the agricultural area is highlighted. There are several HIS processing tasks such as classification and anomaly and target detection. The aim in a target detection algorithm is to label each desired pixel as a target with the lesser false alarm rate (FA), thus estimating the target presence in the image. In the last few decades, Target detection algorithms have been modified, updated and improved. However, finding a target detection algorithm with an optimal performance and independent from external presettings for different targets and scenes is a working research area. Colombia has particular climatic and geographical conditions, providing diverse landscapes that allow to test the selected algorithms in numerous reliefs, climates and crops. This research looks for identification and implementation of target detection algorithms over hyperspectral images acquired by remote sensors in Colombian territory, allowing the technological and scientific flow at a national level, through appropriation and research in HIS use and application.

For the experimental and testing labor, several datasets were used: 20 real HSI, 6 synthetic HIS, 1 noise free and 5 with noisy synthetic HIS, 5 synthetic targets were used created from an alunite pure spectral signature with different mixture levels and more than 115 real spectral signatures, from which 11 were used as target signatures, yielding the characterization of 5 kinds of crops from north-east Colombia in 5 different locations. Results exhibit that the ACE algorithm presents a better performance for different HIS and agricultural targets as much as synthetic and real images.

³Research Work

⁴School of Physical-Mechanical Engineering. Departament of Systems Engineering and Informatics. Advisor, Henry Arguello

INTRODUCCIÓN

El Sensado Remoto (SR) se refiere generalmente al uso de tecnologías de sensores aéreos en la detección o clasificación de objetos en la Tierra (en la superficie, en la atmósfera y los océanos) a través del análisis de señales propagadas como la radiación electromagnética. La técnica de SR permite adquirir información espacial, temporal y espectral de un objeto, o escena, sin entrar en contacto físico [1]. Los datos adquiridos van desde imágenes multiespectrales hasta imágenes hiperespectrales. Gracias a la evolución electrónica y óptica de los sensores, actualmente existen plataformas aeroespaciales, aerotransportadas o terrestres que poseen sensores con resolución espectral que va desde 11 bandas, como el sensor multiespectral (de banda ancha y discontinua) Landsat 8 [2], hasta 220 bandas del sensor hiperespectral (de banda angosta y contigua) Hyperion [3]; y resoluciones espaciales desde los 30 centímetros o pocos metros (Quickbird, Ikonos o WorldView-3) [4].

En los últimos años, las HSI han ganado importancia en el mundo científico, debido a su potencial, el número de publicaciones y citas científicas ha aumentado [5]. Su uso y aplicación se ha dado principalmente en países desarrollados, así mismo, el desarrollo y creación de los algoritmos de detección aplicados a HSI se inició principalmente en aplicaciones militares [5], sin embargo, en los últimos años las HSI han iniciado su aplicación en diferentes áreas, tales como medio ambiente, minería, geología, caracterización del suelo, y la caracterización de coberturas vegetales, dentro de las cuales se destaca el área agrícola con el monitoreo, seguimiento y control de la salud de los cultivos e identificación de plagas, enfermedades o deficiencias de nutrientes [6, 7, 8, 9].

Las imágenes hiperespectrales (HSI) captan información en alta resolución espectral, acerca de las características espectrales de materiales o vegetación presente sobre la superficie de la tierra [10]. La información contenida en una HSI corresponde a datos de radiancia espectral o datos de reflectancia de la superficie, que pueden ser observados como una dispersión de puntos en un espacio Euclidiano K -dimensional, donde K es el número de bandas espectrales. Cada banda corresponde a un eje del espacio y todos los ejes son ortogonales entre sí. Dado que cada píxel en una HSI puede ser considerado como un vector $x = [x_1, x_2, \dots, x_k]^T$, que contiene uno o varios espectros mezclados (firmas espectrales) de diferentes materiales presentes en una escena. Es posible medir la similitud entre dos espectros, o detectar un objetivo a través de la explotación de los datos contenidos en el espacio K -dimensional o HSI a través del uso de algoritmos de detección de objetivos. La principal tarea de un algoritmo de detección de objetivos en HSI es detectar un píxel deseado (objetivo) en una escena HSI o fondo con el menor número de falsas alarmas (FA).

Los algoritmos de detección consisten en determinar si una hipótesis acerca de la presencia de un objetivo es verdadera. Dado un espectro observado, el algoritmo decide entre las dos siguientes hipótesis [11, 10]:

$$H_0 : \text{Objetivo Ausente}$$
$$H_1 : \text{Objetivo Presente}$$

Los algoritmos de detección de objetivos han evolucionado continuamente para proporcionar mejores resultados de detección. Sin embargo, de acuerdo a la manera como se

aborda el problema de modelar la variabilidad espectral, principalmente del fondo, existen dos enfoques, los modelos de aproximación geométrica y los modelos de aproximación estadística. Según estos dos enfoques y una revisión del estado del arte de algoritmos de detección aplicados a HSI, la presente investigación realiza una selección de cinco algoritmos con el mayor número de citas científicas. Estos algoritmos son: Estimador de Coherencia Adaptativo - ACE [10], Algoritmo de Minimización de Energía Restringida - CEM [12], Algoritmo de Filtro Emparejado - MF [11, 10, 13], Proyector de Subespacio Ortogonal - OSP [14, 15], Algoritmo de Mapeo del Angulo Espectral - SAM [16, 17]. En el primer grupo de algoritmos, que modelan el fondo de manera estadística, se encuentran ACE, CEM y MF, el segundo grupo, que describe el fondo de manera geométrica, se encuentra el algoritmo OSP y SAM.

Los algoritmos de detección de objetivos usados en HSI en las dos últimas décadas han presentado mejoras, modificaciones y nuevas versiones [10, 11, 18]. Sin embargo, la tarea de encontrar un algoritmo de detección que cuente con un desempeño óptimo y que este libre de ajustes previos para diferentes escenarios y objetivos, aun es materia de investigación [19, 20]. La complejidad de los escenarios del mundo real y las dificultades para predecir la firma objetivo exacta *in situ*, incrementa la dificultad para encontrar un algoritmo óptimo, independiente a las áreas o escenarios donde serán evaluados los algoritmos. De acuerdo a lo anterior, la presente investigación realiza la identificación e implementación de algoritmos de detección de objetivos en HSI adquiridas por sensores remotos, sobre una zona del territorio colombiano. El objetivo es evaluar el desempeño de los algoritmos y realizar un proceso de transferencia tecnológica y científica a nivel nacional, mediante la apropiación, e investigación del uso y aplicación de HSI.

Para la evaluación de los cinco algoritmos de detección seleccionados (ACE, CEM, MF, SAM y OSP) y para el desarrollo de la presente investigación se trabajó con 20 HSI reales; 6 HSI sintéticas, 1 HSI sintética sin ruido y 5 con ruido (SNR = 10 dB, 15 dB, 20 dB, 25 dB, y 30 dB), se implantaron 5 objetivos sintéticos, creados a partir de la firma espectral pura de la alunita a diferentes grados de mezcla o fracción ($f = 10\%$, 20% , 30% , 40% , y 100%), y se extrajeron más de 115 firmas espectrales reales, de las cuales 11 fueron usadas como objetivo en la evaluación de los algoritmos, las cuales permitieron la caracterización de cinco cultivos agrícolas del nororiente colombiano (palma africana de aceite, caucho, pastos para ganadería, cítricos y caña de azúcar).

Colombia por su ubicación geográfica, posee características propias en sus escenarios, su localización en la zona ecuatorial determina la existencia de una gran variedad de ecosistemas, relieves geográficos y climas permitiendo una diversidad de cultivos agrícolas. Colombia se ubica en la zona tórrida o intertropical; permitiendo que cuente con la misma iluminación solar todo el año. Los rayos del Sol caen siempre con la misma verticalidad (ventaja para aplicaciones e investigaciones de SR), de modo que no hay estaciones, por lo que únicamente hay un invierno lluvioso y un verano seco. Estos sucesos hacen que Colombia posea condiciones geográficas y climáticas particulares, generando escenarios diversos, lo cual hace necesario evaluar el desempeño de los algoritmos seleccionados para diferentes tipos de escenarios, relieves, climas y cultivos agrícolas colombianos.

La presente investigación es una de las primeras en ser desarrolladas a partir de imágenes satelitales hiperespectrales, aplicadas a la agricultura colombiana, y es la primera en Colombia en usar HSI del sensor hiperespectral Hyperion para evaluar el desempeño de

algoritmos de detección a partir de objetivos agrícolas. Para tal fin, se solicitó al centro de vuelos espaciales Goddard de la NASA la adquisición de un conjunto de HSI a través del sensor hiperespectral Hyperion, ubicado en la plataforma satelital EO-1. Hyperion adquirió HSI sobre las cinco siguientes poblaciones: Puerto Wilches (PW), Río Negro-Lebrija (RL), Mogotes (MG) y Guadalupe (GL) ubicadas en el departamento de Santander, Colombia y San Alberto (SA), ubicada en el departamento del Cesar, Colombia.

Los resultados muestran que el algoritmo ACE presenta un mejor desempeño para diferentes HSI y objetivos de tipo agrícola, tanto en imágenes sintéticas como reales, seguido por los algoritmos CEM y MF que presentan un comportamiento aceptable con promedios de probabilidades de detección $P_D = 80\%$, en contraste con los algoritmos OSP y SAM que son capaces de detectar los objetivos con promedios de $P_D = 45\%$, sin embargo, el número de falsas alarmas (FA) son altas desmejorando sus desempeños. Finalmente, la presente investigación muestra que a partir de la transferencia tecnológica y científica del uso y aplicación de HSI y algoritmos de detección de objetivos, es posible su aplicación, uso, e investigación en cultivos agrícolas colombianos bajo variables y factores propios de la agricultura y geografía colombiana.

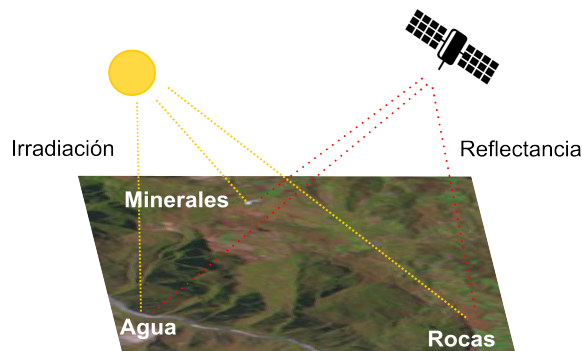
Este trabajo se encuentra organizado en siete capítulos y conclusiones. En el primer capítulo se presentan conceptos básicos de sensado remoto e imágenes hiperespectrales. En el segundo capítulo se introducen los conceptos de sensado remoto hiperespectral aplicado a la agricultura. En el tercer capítulo se presentan los algoritmos de detección usados en la presente investigación. El cuarto capítulo presenta las métricas usadas para la evaluación de desempeño de algoritmos de detección. El quinto capítulo muestra los resultados, inicialmente se presentan los conjuntos de datos reales y sintéticos, en seguida, muestra los resultados por cada área de estudio. En el capítulo seis se presentan las longitudes de onda de mayor importancia en el estudio espectral de la vegetación. En el capítulo siete se presenta la viabilidad técnica y científica para la creación de un prototipo de laboratorio de sensado remoto hiperespectral para el nororiente colombiano. Finalmente, el presente trabajo de investigación culmina con las conclusiones derivadas del proceso de investigación.

1 TEORÍA DE SENSADO REMOTO

1.1 SENSADO REMOTO

Sensado remoto (del inglés *remote sensing*) hace referencia a la técnica empleada para la adquisición de información sobre un objeto o escena sin hacer contacto físico con él [1], a través de sensores ubicados en plataformas aerotransportadas (aviones tripulado o no tripulados - UAV) o aeroespaciales (satélites) [21]. También existe el SR en campo o laboratorio, donde el sensor es dirigido al objeto o escena in situ [22]. La figura 1 ilustra los elementos de un sistema de SR, el principal elemento es el sensor que capta la energía electromagnética emitida desde una fuente de luz y que es reflejada por los elementos o materiales presentes en una escena. Mediante la implementación de algoritmos sobre los datos captados por los sensores se pueden detectar y/o clasificar objetos, materiales, vegetación, minerales, tipo de suelos y rocas debido a señales propagadas (por ejemplo, radiación electromagnética).

Figura 1: Sistema de SR, se conforma por una fuente de energía, una escena, la trayectoria de la señal desde su origen hasta su destino y el sensor.

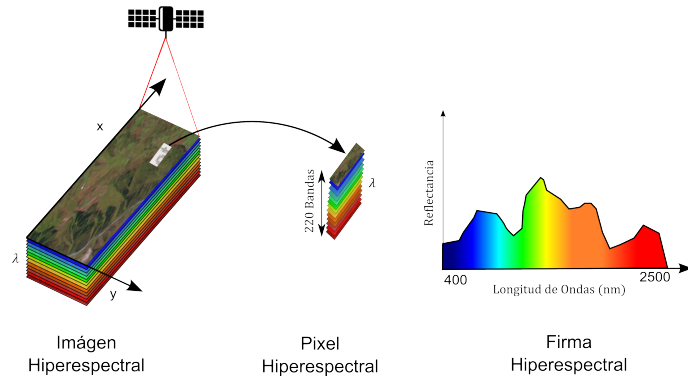


1.2 IMAGENES HIPERESPECTRALES

Las imágenes de alta resolución espectral se conocen como Imágenes Hiperespectrales (HSI). Un sensor hiperespectral capta imágenes espaciales bidimensionales (2D) en múltiples longitudes de onda del espectro electromagnético de forma contigua y angosta, creando un tercer eje con información espectral (λ). Cada imagen contiene información espectral en una longitud de onda determinada, conocida como banda espectral. Los sensores hiperespectrales miden la radiancia reflejada (reflectancia) en un gran número de longitudes de onda o bandas contiguas (>100 bandas), que representa su alta resolución espectral.

Al realizar el barrido sobre una superficie o escena, los sensores HSI detectan la energía solar reflejada por los diferentes materiales y miden la intensidad de la energía reflejada en diferentes longitudes de onda del espectro electromagnético (ver figura 2), creando las HSI, que son imágenes con información espacial (x, y) y espectral (λ), formando un hipercubo (3D). Cada píxel (x, y) contiene un espectro de alta resolución espectral llamado firma espectral.

Figura 2: Concepto de HSI y firma espectral. El sensor capta la energía solar reflejada por cada material presente en una escena espacial (x,y) , en diferentes longitudes de onda λ contiguas, creando un hipercubo. La firma espectral es la representación de un píxel (x,y) .



Las HSI en SR almacenan medidas espectrales de la región de reflectancia solar, entre los 400 nm y 2.500 nm, abarcando el espectro visible (VIS), el infrarrojo cercano (NIR) y la onda corta del infrarrojo (SWIR) [23]. La capacidad de los sensores HSI para adquirir mayor información espectral, se debe a que los materiales presentes en los diversos objetos o elementos en una escena tienen un comportamiento espectral diferente de acuerdo a su composición fisicoquímica [24]. Si se mide la radiación reflejada hacia el sensor en las diferentes longitudes de onda sobre una región espectral suficientemente amplia, la firma espectral resultante puede usarse para detectar y discriminar diferentes materiales en una escena [23] o caracterizar el comportamiento de un cultivo agrícola [25].

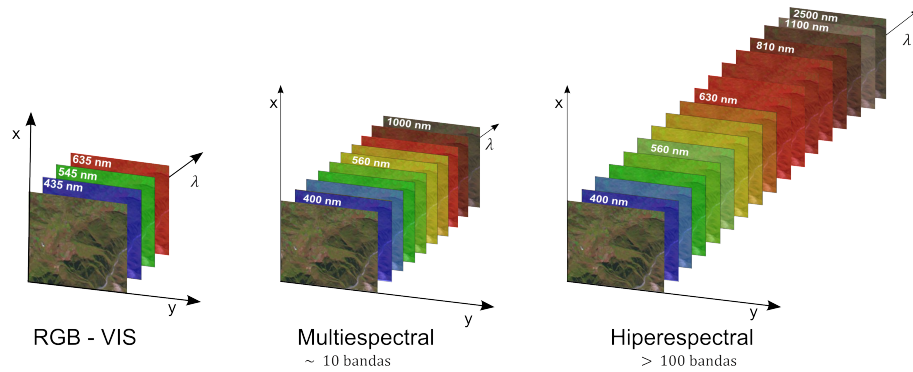


Figura 3: Clasificación de imágenes según su resolución espectral.

La precisión y calidad de los datos espectrales, y en específico la estimación de índices de vegetación espectrales son de gran uso en la agricultura, tienen una estrecha relación con la resolución espectral [26, 27] que es condicionada por la cantidad y ancho de las bandas espectrales usadas para medir la reflexión en diferentes longitudes de onda (firma espectral).

Por otro lado, las imágenes multispectrales captan información en solo decenas de bandas, que se caracterizan por ser discontinuas y de banda ancha (del inglés, *broadbands*). La principal ventaja de las HSI sobre las imágenes multispectrales es la gran cantidad de bandas angostas (del inglés, *narrowbands*) y contiguas que se pueden captar [28]. La figura 3 ilustra la diferencia de resolución espectral entre diferentes imágenes. El número

de bandas y el ancho de las mismas determinan la resolución espectral de una imagen adquirida por los sensores remotos.

1.3 SENSORES REMOTOS HIPERESPECTRALES

Los sensores remotos de HSI son instrumentos ópticos capaces de obtener datos a partir de la energía electromagnética reflejada un objeto a distancia en cientos de bandas espectrales [22]. Los sensores pueden clasificarse como pasivos, si su fuente de energía es el Sol y captan su reflejo, y activos los cuales emiten la energía desde la misma plataforma donde se encuentra el sensor, que después es medida por el sensor. Por otro lado, de acuerdo a plataforma los sensores se pueden categorizar en sensores aerotransportados (aviones), aeroespaciales (satélites) o terrestres (cámaras hiperespectrales de campo).

En la presente investigación, se usaron HSI captadas por el sensor Hyperion ubicado en la plataforma satelital EO-1 de la NASA. Hyperion es un sensor hiperespectral de alta resolución espectral, capaz de captar información espectral en 220 bandas, iniciando en la longitud de onda de los 355,59 nm hasta 2577,08 nm, con un ancho de banda de 10 nm y resolución espacial de 30 m [29, 30]. Actualmente, Hyperion está bajo el control del Servicio Geológico de Estados Unidos (USGS) y la NASA. Hyperion capta escenas de la Tierra principalmente para investigación científica solo a través de solicitudes específicas a través de la plataforma GeoBPMS [31].

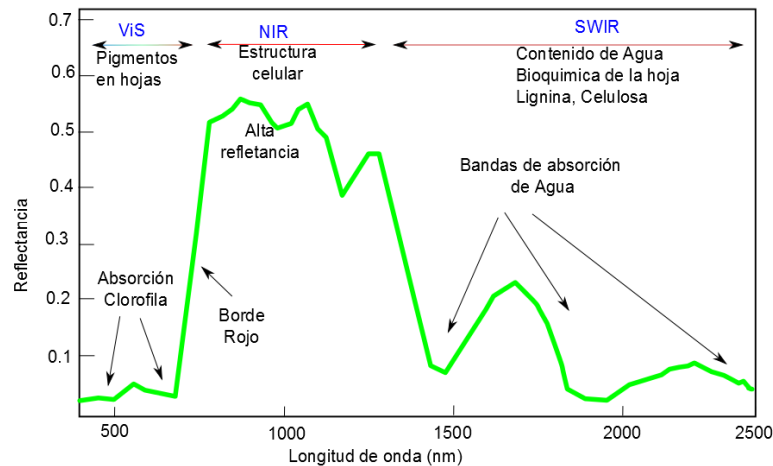
2 SENSADO REMOTO HIPERESPECTRAL APLICADO A LA AGRICULTURA

2.1 PROPIEDADES ESPECTRALES DE LA VEGETACIÓN

El comportamiento espectral de la vegetación y la cantidad de energía reflejada por una planta se puede medir a lo largo del espectro electromagnético, permitiendo su caracterización espectral. Este comportamiento espectral depende de la interacción con la radiación solar (400 nm – 2500 nm), y puede ser afectado por varios factores, tales como: nutrientes, agua, clima y su medio ambiente.

La figura 4 muestra las principales características espectrales de la vegetación, destacando las regiones espectrales o longitudes de onda de mayor importancia para el análisis y estudio de los cultivos agrícolas y vegetación en general [32]. La región espectral que corresponde al visible (VIS) en la vegetación se caracteriza por una baja reflectancia y transmitancia, debido a que los pigmentos y la clorofila absorbe la mayor parte de la energía en alrededor de $\lambda \sim 650$ nm (rojo) y $\lambda \sim 450$ nm (azul), esta absorción elimina estos colores de la cantidad de luz que se transmite y refleja, haciendo que el color visible predominante que llega a nuestros ojos sea el verde ($\lambda \sim 520$ nm) en esta longitud de onda se presenta el mayor pico de reflectancia en el VIS. La vegetación sana aparece como un color verde oscuro. La vegetación no saludable tendrá menos clorofila y por lo tanto generara tonos rojizos. Este aumento de la reflectancia roja, junto con el verde causa la apariencia amarilla general de las plantas no saludables.

Figura 4: Características espectrales de la vegetación. La firma espectral de la vegetación tiene un comportamiento muy específico, donde se destaca el gran contraste entre las regiones espectrales del VIS y el NIR.



Uno de los parámetros o regiones espectrales de importancia en la vegetación es el borde rojo (del inglés *red edge*), en el cual se presenta una fuerte pendiente entre la baja reflectancia en el VIS y la alta reflectancia del NIR, el borde rojo se ubica en $\lambda \sim 700-740$ nm. La segunda región espectral de importancia en la vegetación es el NIR ($\lambda \sim 740$ nm a

1040 nm). Debido a la alta reflectancia NIR, el NIR no es absorbido por ningún pigmento dentro de la planta, y tiene un pico de reflectancia que es sensible a la humedad presente en las plantas (940-1040 nm) y una reflectancia mínima alrededor de 982 nm [26].

2.2 SENSADO REMOTO HIPERESPECTRAL APLICADO A LA AGRICULTURA COLOMBIANA

Las HSI se han convertido en una técnica fundamental para la observación de la Tierra, y para satisfacer a nivel mundial la creciente demanda de mapeo de alta resolución espectral a gran escala en diferentes áreas, tales como recursos naturales, mineros, geológicos, ambiental, y principalmente en el sector de la agricultura.

Específicamente en Colombia, de acuerdo al informe de gestión del Instituto Geográfico Agustín Codazzi (IGAC) año 2012-2013 [33], el IGAC es la institución colombiana que más usa y aplica SR, principalmente en áreas de la cartografía, prevención y control de desastres, y el monitoreo meteorológico. Sin embargo, el mayor número de sensores remotos con los que se captan imágenes de Colombia son multiespectrales, a excepción de casos muy específicos, donde se han aplicado técnicas de SR hiperespectral a nivel comercial y privado en sus primeros pasos en el sector de exploración minera y de hidrocarburos. Empresas como AngloGold Ashanti (extractora de oro), con presencia en Colombia; junto a Specim (compañía de SR hiperespectral) desarrollaron, entre los años 2004 y 2008, una exploración minera y geológica de áreas de alta prospectividad de extracción de oro [34, 35], aplicando técnicas de SR hiperespectral. Por otro lado, la Agencia Nacional de Hidrocarburos (ANH) actualmente incluye en sus procesos de exploración el uso de HSI para estudiar áreas de alta prospectividad [36].

En el área de la academia e investigación científica, en los últimos 5 años, han surgido iniciativas para la investigación y aplicación de la técnica de SR en el país, específicamente se han desarrollado investigaciones de SR para la agricultura en Colombia, limitadas al orden multiespectral y mediciones *in situ* a través espectrorradiómetros. Investigaciones tales como: identificar los cultivos ilícitos usando imágenes de espectroscopia [37], estimación de índices de vegetación usando SR multiespectral aerotransportado aplicado en la caña de azúcar [38], en cultivos de café [39], cultivos de palma de aceite [40], y la estimación de la biomasa a partir de imágenes del sensor MODIS [41].

En SR hiperespectral a nivel de investigación se encuentran iniciativas en universidades, tales como: la implementación y creación de un centro regional de investigación e innovación en bioinformática y fotónica de la Universidad del Valle del Cauca, la creación de una librería de firmas espectrales colombianas por el IGAC [42], así como estudios en diferentes grupos de investigación de universidades colombianas como la Universidad Distrital [43], Universidad del Valle, Universidad de Manizales, y Universidad Industrial de Santander [44]. Instituciones que han mostrado su interés por la investigación y a aplicaciones de SR hiperespectrales en diferentes áreas (ambiental, forestal, geología y la agricultura). A nivel de centros de investigación especializados como Cenicaña, Cenipalma, Cenicafé, Centro Internacional de Agricultura Tropical (CIAT), y el IGAC, han logrado obtener un pequeño número de firmas espectrales propias de cultivos agrícolas colombianos apoyados en espectrorradiómetros a nivel de laboratorio y en campo, sin embargo, no están disponibles para su uso de manera similar como lo están librerías espectrales existentes como la ASTER [45].

Se evidencia que ha surgido el interés de apropiación, uso y aplicación de SR hiperespectral en Colombia. Las investigaciones en esta área continúan, tanto en términos de aplicaciones operativas, de generación de conocimiento, investigación y el desarrollo de técnicas y metodologías específicas en el uso de HSI aplicados a otros sectores. Principalmente el sector de la agricultura, donde Colombia posee condiciones geográficas y climáticas particulares y cultivos agrícolas específicos. Esta investigación es una de las primeras en ser desarrolladas a partir de imágenes satelitales hiperespectrales aplicadas a la agricultura colombiana.

3 ALGORITMOS DE DETECCIÓN DE OBJETIVOS EN IMÁGENES HIPERESPECTRALES

La tarea principal de los algoritmos de detección de objetivos en HSI es detectar un píxel deseado (objetivo) en una escena (imagen hiperespectral) con el menor número de falsas alarmas (FA).

Una HSI se puede expresar como un cubo de datos que contiene información de la radiancia espectral o reflectancia de la superficie, que pueden ser observada como una dispersión de puntos en un espacio Euclidiano K -dimensional, donde K es el número de bandas espectrales. Cada banda corresponde a un eje del espacio y todos los ejes son ortogonales entre sí. Dado que cada píxel en una HSI puede ser considerado como un vector $x = [x_1, x_2, \dots, x_k]^T$, que contiene un espectro puro (píxel completo o full-píxel) o varios espectros mezclados (píxel mezclados o subpíxel) de diferentes materiales presentes en una escena. Es posible medir la similitud entre dos espectros, o detectar un píxel deseado (objetivo) a través de la explotación de los datos contenidos en el espacio K -dimensional o HSI a través del uso de algoritmos de detección de objetivos. Dado que cada material se caracteriza por un espectro determinista único, su comportamiento espectral puede servir como huella digital espectral (firma espectral) y contribuir a su detección.

La base matemática de los algoritmos de detección se basa en el área de la estadística conocida como hipótesis binaria. Dado un espectro observado x , se pretende decidir entre una de las dos siguientes hipótesis [11, 10]:

$$H_0 : \text{Objetivo Ausente}$$

$$H_1 : \text{Objetivo Presente}$$

Si $p(x | H_0)$ y $p(x | H_1)$ son las funciones de densidad de probabilidad (PDF) de x bajo las dos hipótesis, esto se puede expresar como

$$\Lambda(x) = \frac{p(x | \text{Objetivo Presente})}{p(x | \text{Objetivo Ausente})}. \quad (1)$$

Si $\Lambda(x)$ excede un cierto umbral η , entonces la hipótesis de Objetivo Presente se da como verdadera. Un aspecto importante en los algoritmos de detección de objetivos es la selección del umbral η , esto debido a que si se selecciona un umbral η bajo, se aumenta la probabilidad de detección (P_D), sin embargo, si se maneja un umbral η alto, se obtienen bajas tasas de probabilidad de falsa alarma (P_{FA}). La relación existente entre P_D y P_{FA} determina el desempeño de los detectores o algoritmos, la métrica que mayor se usa en la comunidad científica para evaluar el desempeño de los algoritmos de detección son las curvas de caracterización operativa del receptor (ROC) [10], las curvas ROC dibujan la P_D del objetivo, frente a la P_{FA} como una función de todos los posibles valores del umbral η . Ahora, un detector óptimo es aquel que logra detectar los objetivos con una P_D alta y una P_{FA} muy baja.

Los algoritmos de detección de HSI deben superar dificultades como: (1) que con gran frecuencia el número de píxeles contenidos dentro de un hipercubo está sobre el orden de 10^5 con la limitante que la estimación de P_{FA} es menor a 10^{-4} por cada hipercubo; (2)

el número de objetivos de un tipo o clase en particular en una escena es frecuentemente muy pequeño y limita la precisión del desempeño de detección; (3) la verdad de terreno (del inglés, *ground truth*) es muy limitada para la totalidad de la escena y en la gran mayoría de investigaciones no se cuenta con ella, dificultando la validación de los resultados de desempeño de los algoritmos; (4) la existencia de píxeles mezclados (subpíxel) en la escena, dado que un píxel observado por el sensor puede recibir información del objetivo y del fondo; y (5) la gran variabilidad espectral, que se evidencia por el hecho de que un espectro único y fijo, teóricamente perfecto, no existe para ningún material dado. Los espectros observados a partir de muestras del mismo material no son idénticos, incluso en experimentos de laboratorio, debido a las variaciones en la superficie del material. La variabilidad es más pronunciada en aplicaciones de SR debido a las condiciones atmosféricas, ruido del sensor, la composición del material, la ubicación del sensor y la muestra, materiales circundantes, y otros factores [11, 13].

En la presente sección y de acuerdo al estado del arte de algoritmos de detección aplicados a HSI, se realiza una revisión de cinco de los algoritmos con mayores citas e importancia.

3.1 ALGORITMO DE FILTRO DE EMPAREJADO (MF)

El filtro de emparejado y sus diferentes variaciones han sido ampliamente usado en aplicaciones de detección de objetivos en HSI [11, 10, 13]. MF se basa en la prueba de las hipótesis binarias descrita en la ecuación (1). La PDF asociada a cada hipótesis es modelada como una distribución normal, y puede ser expresada por

$$H_0 : x \sim N(\mu_b, \mathbf{\Gamma}_b) \text{ Objetivo ausente} \quad (2)$$

$$H_1 : x \sim N(\mu_t, \mathbf{\Gamma}_t) \text{ Objetivo presente,} \quad (3)$$

Si se asume que las dos hipótesis tienen igual matriz de covariancia, $\mathbf{\Gamma}_b = \mathbf{\Gamma}_t = \mathbf{\Gamma}$, y partiendo que la selección del factor de normalización no afecta el desempeño del MF, este puede usar la siguiente ecuación

$$D_{MF}(x) = \frac{(d - \mu)^T \mathbf{\Gamma}^{-1} (x - \mu)}{(d - \mu)^T \mathbf{\Gamma}^{-1} (d - \mu)}, \quad (4)$$

donde d es el espectro objetivo, x es el espectro del píxel a evaluar (fondo o *background*), μ es el vector de que contienen la media del fondo y $\mathbf{\Gamma}$ es la matriz de covarianza.

3.2 ALGORITMO DE MINIMIZACIÓN DE ENERGÍA RESTRINGIDA (CEM)

El CEM está diseñado como un filtro de respuesta de impulso finito para pasar a través del objetivo deseado y reducir al mínimo la energía de salida resultante generada por el fondo con excepción de los objetivos deseados [12]. Una matriz de correlación o covarianza es utilizada para caracterizar la composición del fondo desconocido. Esta técnica maximiza la respuesta de la firma conocida y suprime la respuesta compuesta por información del fondo, esto permite detectar la firma objetivo conocida. El detector de CEM se expresa como

$$D_{CEM}(x) = \frac{d^T \mathbf{\Gamma}^{-1} x}{d^T \mathbf{\Gamma}^{-1} d} \quad (5)$$

donde d es el espectro objetivo, x es el espectro del píxel a evaluar, y $\mathbf{\Gamma}$ es la matriz de covarianza o correlación del fondo. CEM pertenece a los algoritmos basados en un enfoque estructurado donde se utiliza las estadísticas para describir el fondo, por esta razón se incluye en la presente investigación.

3.3 ALGORITMO ESTIMACIÓN DE COHERENCIA ADAPTATIVO (ACE)

El Estimador de Coherencia Adaptativo (ACE) puede ser extraído de la Tasa de Máxima Verosimilitud Generalizada (GLRT), partiendo de la asunción que se conoce la matriz de covarianza del fondo, ACE puede ser escrito como

$$D_{ACE}(x) = \frac{(d^T \mathbf{\Gamma}^{-1} x)^2}{(d^T \mathbf{\Gamma}^{-1} d)(x^T \mathbf{\Gamma}^{-1} x)}, \quad (6)$$

donde d es el espectro objetivo, x es el espectro del píxel a evaluar, y $\mathbf{\Gamma}$ es la matriz de covarianza del fondo. El ACE usa una forma diferente de estimar la estadística de detección y logra una mayor separación entre el objetivo y el fondo. Adicionalmente, ACE permite establecer con mayor facilidad un umbral P_{FA} constantes. Similar a CEM y MF, ACE ofrece un medio rápido para la detección de una firma objetivo específica basado en coincidencias frente a una firma de una librería espectral o una firma espectral de las misma HSI, Adicionalmente, ACE no requiere conocimiento de todos los espectros puros presentes dentro de una escena. En estudios previos el algoritmo ACE ofrece el mejor desempeño frente a los otros algoritmos [10], razón que lo hace viable para ser aplicado en la presente investigación.

3.4 ALGORITMO MAPEO DEL ÁNGULO ESPECTRAL (SAM)

El algoritmo SAM es usado para determinar que tan similar pueden ser dos espectros, a partir del computo del ángulo entre los dos vectores (vector objetivo y vector píxel a evaluar). SAM es rápido, fácil y su costo computacional es bajo, no requiere de ninguna información estadística de la escena [16, 17]. Cualquier píxel de la HSI o escena es tratado como un vector con dimensionalidad igual al número de bandas de la HSI y la similitud es calculada como el ángulo entre dos espectros

$$\cos\theta = \frac{d \bullet x}{|d| |x|}, \quad (7)$$

donde d es el espectro objetivo o vector deseado, x es el píxel a evaluar, y θ es el ángulo entre los dos. El algoritmo SAM también puede ser realizado por multiplicación de matrices y se puede expresar como

$$D_{SAM}(x) = \frac{d^T x}{(d^T d)^{1/2} (x^T x)^{1/2}}. \quad (8)$$

Los resultados de detección tendrán valores entre 0 y 1, los píxeles con mayor similitud serán valores cercanos a 0, debido a los ángulos pequeños entre los vectores en evaluación. SAM se incluye en la presente investigación como un algoritmo de detección de línea base por ser un algoritmo tradicional, fácil de usar, y que no requiere de información estadística de la escena, permitiendo ser evaluado contra algoritmos mas robustos.

3.5 ALGORITMO DE PROYECCIÓN SUBESPACIO ORTOGONAL (OSP)

El algoritmo OSP primero diseña un proyector de subespacio ortogonal para eliminar la respuesta de los píxeles no objetivos, a continuación, aplica un filtro de emparejado para buscar la coincidencia del objetivo deseado en los datos. OSP es eficiente y eficaz cuando la firma del objetivo es muy diferente con respecto al fondo. Cuando el ángulo espectral entre la firmas de la HSI y la firma objetivo es pequeña, la atenuación de la señal objetivo es dramática y el rendimiento de algoritmo OSP podría ser deficiente. El OSP se puede expresar como

$$D_{OSP}(x) = \frac{d^T P_U^\perp x}{d^T P_U^\perp d}, \quad (9)$$

donde d es el espectro objetivo o vector deseado, x es el píxel a evaluar, y $P_U^\perp = I_{L \times L} - UU^\#$ es el proyector de subespacio ortogonal. L es el número de bandas, U es la matriz de espectros no-objetivos, y $U^\# = (U^T U)^{-1} U^T$ es la pseudoinversa de U . El OSP es un algoritmo que usa un modelo de fondo estructurado para caracterizar la variabilidad espectral y es el primer algoritmo de filtro de emparejado geométrico con un enfoque diferentes a los anteriores algoritmos usados en la presente investigación.

4 MÉTRICAS DE DESEMPEÑO DE ALGORITMOS DE DETECCIÓN DE OBJETIVOS EN HSI

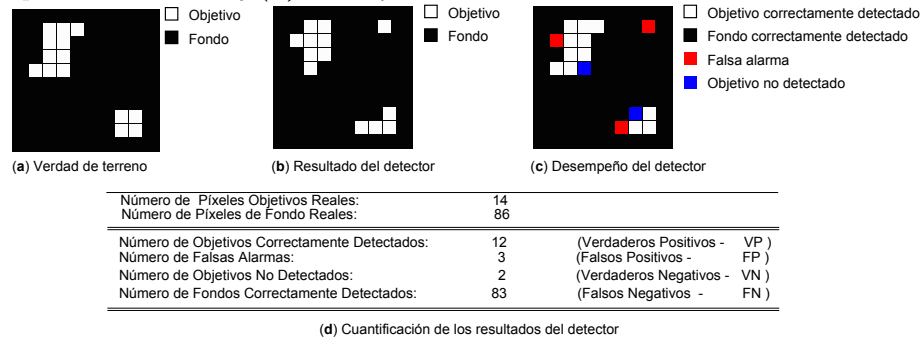
La tarea principal de cualquier métrica de desempeño de detección es determinar cómo un algoritmo detecta los objetivos deseados en función de detecciones correctas (verdaderos positivos, VP), falsas alarmas (falsos positivos, FP), pérdida de un objetivo (falso negativo, FN) o la correcta detección de los píxeles de fondo (verdadero negativo, VN). Para tal tarea, existen numerosas métricas para evaluar el desempeño de los algoritmos de detección de objetivos. Estas pueden ir desde métricas visuales hasta el mapeo o graficación de datos para análisis cuantitativos. Cada una de estas métricas tiene sus ventajas y desventajas en la presente sección se muestran las técnicas y métricas que determinan el rendimiento de detección de los algoritmos de objetivos en HSI.

4.1 MÉTRICA POR PÍXEL Y MATRICES DE CONFUSIÓN

La métrica basada por píxel es una métrica con una base sencilla y visual, sin embargo, de gran importancia y uso en todas las evaluaciones de desempeño de los algoritmos de detección. En esta métrica cada píxel puede ser clasificado como perteneciente a un objetivo y tratado como un objetivo individual, o el píxel puede ser clasificado como perteneciente a la clase fondo. Fondo o *background* en la terminología de detección de objetivos se refiere a todo píxel ubicado en una escena o imagen, que no es objetivo, por lo tanto, el resultado de un detector es un clasificador de solo dos clases (objetivo o fondo) o clasificador binario. Para que un objeto o material pueda ser perfectamente detectado como un objetivo, cada píxel del objeto debe clasificarse como objetivo.

La figura 5 es la representación gráfica de desempeño de un detector evaluando píxel por píxel. Para un total de 100 píxeles, en la figura 5.a se muestra la verdad de terreno donde el objetivo se conforma por 14 píxeles, la figura 5.b muestra el resultado del detector, y específicamente en la figura 5.c se ilustra de manera detalla la respuesta por cada posibilidad generada por el detector. El análisis cuantitativo de esta métrica es la entrada o *input* de otras métricas de tipo cuantitativo como la matriz de confusión. Una matriz de confusión muestra el número de detecciones correctas e incorrectas hechas por el detector en comparación con los resultados reales (verdad de terreno). La matriz es de $N \times N$, donde N es el número de clases resultado, para el caso específico de los detectores las clases solo son dos (objetivo o fondo) o expresado en términos de la respuesta del detector la respuesta puede ser verdadera o falsa. La figura 5. c y d. muestra con mayor detalle las posibles respuesta de un detector evaluado por píxel y la cuantificación de los resultados.

Figura 5: Métrica por píxel de objetivos y fondo. (a) verdad de terreno, (b) resultado del detector, (c) desempeño del detector, y (d) cuantificación de resultados del detector



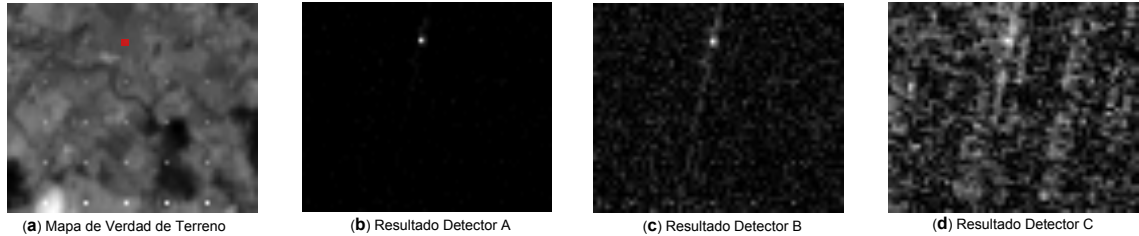
El mapa de verdad de terreno es esencial en la evaluación de las métricas de desempeño. Un mapa de verdad de terreno contiene la ubicación de todos los píxeles objetivos en una imagen. Además de la ubicación, el tipo de píxel objetivo también se puede observar. La figura 5.a muestra un ejemplo de un mapa de verdad donde se identifica la ubicación de los 14 píxeles objetivos mediante el color blanco y el fondo se identifica y etiqueta en color negro. Como se explicó en la introducción de la sección anterior (ver sección 3. Algoritmos de Detección de Objetivos), uno de los principales retos de los algoritmos es contar con mapas de verdad de terreno de las escenas que se captan, ya que la gran mayoría de escenas no cuenta con mapas de verdad de terreno. Sin embargo, los mapas de verdad de terreno se pueden construir con conocimiento de la áreas o escenas captadas y con imágenes de alta resolución espacial captadas desde plataformas aéreas. Específicamente, para la presente investigación donde se solicito por primera vez la captura de HSI para el nororiente colombiano, no se cuenta con mapas de verdad de terreno, sin embargo, para superar esta limitante, se desarrollaron estrategias durante el proceso de investigación como realizar cinco salidas de campo, una por cada área de estudio. Con el apoyo de un GPS e imágenes de otros satélites con mayor resolución espacial, como *GeoEye-1* que puede captar imágenes con una resolución que alcanza los 0,45 mt en blanco y negro, y en color a 1,65 mt. Adicionalmente, con el apoyo de un *drone* se adquirieron fotografías aéreas de algunas áreas de estudio con resoluciones de hasta 10 cm. Dichas estrategias permitieron en el desarrollo de la investigación conocer de manera detallada las áreas de estudio para validar las respuestas de detección de los algoritmos, así mismo, crear un mapa de verdad que sera tratado en la sección de resultados sobre un objetivo de plantas de palma africana en vivero bajo condiciones controladas.

4.2 MÉTRICA VISUAL

La métrica visual usa una representación gráfica en dos dimensiones (x, y) del resultado del algoritmo de detección. La representación muestra un mapa resultado de la detección en píxeles más brillantes (tonos blancos) que corresponden a posibles objetivos. Las respuestas mas brillantes se pueden comparar con el mapa de verdad de terreno para determinar si corresponde de manera correcta a un objetivo o por el contrario es una falsa alarma. El mapa de verdad de terreno puede ser superpuesto o visto a lo largo del lado del mapa resultado. El mapa de respuesta del detector también puede ser el resultado de un umbral η que permite diferenciar entre un objetivo probable o fondo. Los valores superiores al umbral η de decisión se detectan como píxeles objetivo, mientras que las respuestas por debajo del umbral son detectados como píxeles de fondo. Esto crea un mapa resultado binario

que luego se puede comparar con el mapa de verdad de terreno. La figura 6 muestra tres mapas de resultados de detección, donde los píxeles más brillantes corresponden a posibles objetivos. De igual manera, la primera gráfica muestra el mapa de verdad de terreno, donde se identifican la ubicación de diferentes objetivos, el objetivo identificado con el color rojo es el objetivo a detectar por los detectores A, B, y C, según la siguiente figura.

Figura 6: Imagen bidimensional (x, y) de resultados de algoritmos de detección de objetivos. (a) verdad de terreno, (b-d) resultados de los detectores A, B, y C.

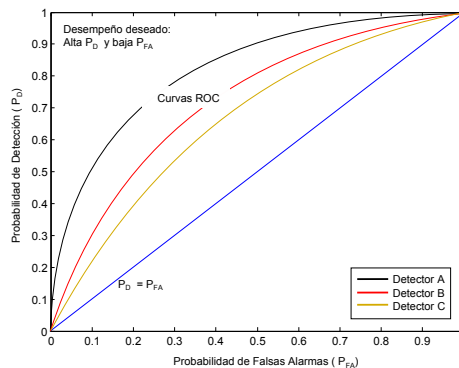


4.3 CURVAS ROC (RECEIVER OPERATOR CHARACTERISTICS)

La curva ROC describe la probabilidad de detección (P_D) frente a la probabilidad de falsa alarma (P_{FA}) como una función de todos los posibles valores de umbral η . Por lo tanto, las curvas ROC proporcionan los medios para evaluar el desempeño de un detector de forma independiente de la selección de un umbral. 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*). La FPR mide la cantidad de elementos negativos detectados incorrectamente como positivos, y la TPR la fracción de píxeles objetivos detectados correctamente como objetivos.

La figura 7 muestra de manera general e ilustrativa el concepto de las curvas ROC, como herramienta de medición de desempeño de los algoritmos de detección de objetivos. Donde el algoritmo con mejor desempeño es aquel que presente una mayor P_D con una muy baja P_{FA} .

Figura 7: Curvas ROC, métrica de desempeño que muestra la P_D frente a la P_{FA} de uno o varios detectores.



5 MÉTODO Y RESULTADOS

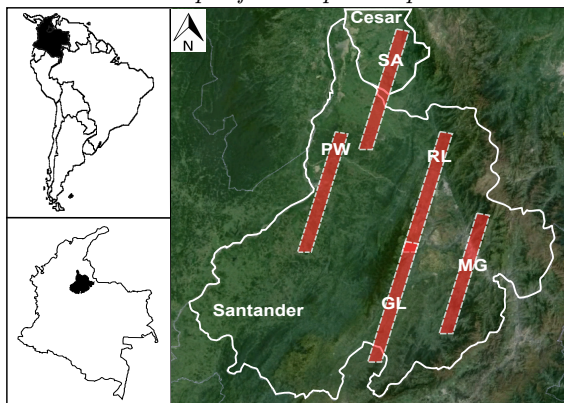
5.1 CONJUNTO DE DATOS DE IMÁGENES HIPERESPECTRALES REALES Y ÁREAS DE ESTUDIO

Para el desarrollo de la presente investigación se solicitó al centro de vuelos espaciales Goddard de la NASA la adquisición de un conjunto de HSI a través del sensor hiperspectral Hyperion, ubicado en la plataforma satelital EO-1. Hyperion captó escenas de cinco poblaciones del nororiente colombiano (ver figura 8). La selección de las áreas de estudio se realizaron previo análisis de las condiciones geográficas, climáticas y distribución de los principales cultivos en el nororiente colombiano, buscando que cumplieran con la restricción de poseer características geográficas y climáticas diferentes, permitiendo adquirir información espectral de varios cultivos agrícolas de importancia.

Específicamente, el departamento de Santander cumple en gran medida con esta condición. Santander se divide en dos sectores naturales: el andino o montaña hacia el oriente y el Valle del Magdalena Medio al occidente que es relativamente plano. Las dos regiones cubren al departamento de norte a sur y por mitades aproximadas. Gracias a esta geografía su mapa climático tiene una gran diversidad de pisos térmicos que van desde los 90 metros sobre el nivel del mar (m.s.n.m.) (región occidental) hasta los 4.000 m.s.n.m. (región oriental o andina); las temperaturas están en el rango de $9^{\circ}C$ y $33^{\circ}C$, generando una diversidad de cultivos agrícolas y condiciones diferentes para la áreas de estudio de la presente investigación.

Las coordenadas geográficas centradas en las áreas de estudio seleccionadas fueron suministradas a la NASA (ver tabla 1) para la captura de las HSI y corresponden a las poblaciones de Puerto Wilches (PW), Río Negro-Lebrija (RL), Mogotes (MG) y Guadalupe (GL) ubicadas en el departamento de Santander, Colombia y San Alberto (SA), ubicada en el departamento del Cesar, Colombia. Hyperion adquirió 20 imágenes entre febrero del 2014 y febrero de 2015. La tabla 1 resume las principales características de las áreas de estudio y los cultivos agrícolas principales en cada una de las áreas.

Figura 8: Áreas de estudio ubicadas en los departamentos de Santander y Cesar, Colombia. Las franjas rojas corresponden a las áreas específicas captadas por el sensor Hyperion.



Id.	Área	Coordenadas	Características Geográficas	Tipos de Cultivos
PW	Puerto Wilches, Santander, Colombia	7° 15' N, 73° 47' W	La zona se caracteriza por tener la mayor parte de sus superficies plana con un relieve llano, su clima es caliente y húmeda con una temperatura promedio de 33 ° C, tiene grandes ríos como el río Magdalena y el río Sogamoso. Su altura sobre el nivel del mar (msnm) es de 90 m.	El principal cultivo en PW es la palma de aceite, ubicado en la zona sur de PW, en el norte existen cultivos tradicionales como el maíz, plátano, y áreas con pastos para ganado
SA	San Alberto, Cesar, Colombia	7° 41' N, 73° 25' W	El área SA es muy similar a la zona de PW, sin embargo SA se ubica al sur del departamento del Cesar, Colombia y limita con el norte del departamento de Santander. La parte más baja del municipio se encuentra a 50 msnm y la parte de mayor altura se encuentra a 2600 msnm al extremo oeste.	El principal cultivo en SA es la palma de aceite
RL	Rio Negro y Lebrija, Santander	7° 14' N, 73° 10' W - 7° 00' N, 73° 14' W	Los dos municipios están ubicados en el centro-este del departamento de Santander. RL se caracteriza por poseer una topografía quebrada en su relieve, su clima es tropical que puede ir de templado a caliente, dependiendo de la altitud, con una temperatura promedio de 28 ° C.	Rio Negro se caracteriza por cultivos de cacao, café y cítricos. En Lebrija el principal cultivo es la piña.
MG	Mogotes, Santander, Colombia	6° 28' N, 72° 57' W	MG se encuentra en el sureste del departamento de Santander, Colombia. Con la excepción de la llanura donde se encuentra el área urbana y algunas riberas de los ríos y arroyos, el terreno es montañoso. Con una altura de 1746 MSL. La temperatura promedio de 18 ° C; en las partes más alta el clima es frío y al norte de MG es cálido. Para los meses de abril, mayo, junio y octubre se presenta lluvias fuertes y frecuentes.	Principal cultivo es la caña de azúcar. Sin embargo, hay presencia de cultivos de café y cacao en menor proporción
GL	Guadalupe, Santander, Colombia	6° 12' N, 73° 25' W	Guadalupe se encuentra en el sureste del departamento de Santander, Colombia. GL posee suelos francos y arcillosos con una fertilidad media a baja. Sus suelos se encuentran en zonas de bosque húmedo subtropical y bosque húmedo. La topografía es 40% ondulada, 30% plano y 30% con una altura media de 1.397 msnm. la temperatura media de 22 ° C	Debido a la calidad del suelo, GL no es apropiado para la agricultura extensiva, GL tiene pequeñas granjas donde existen varios cultivos pequeños, como el café, la yuca y la caña de azúcar. Sin embargo, cuenta con áreas cubiertas de sabana nativa para el ganado

Tabla 1: Áreas de estudio, cultivos y principales características.

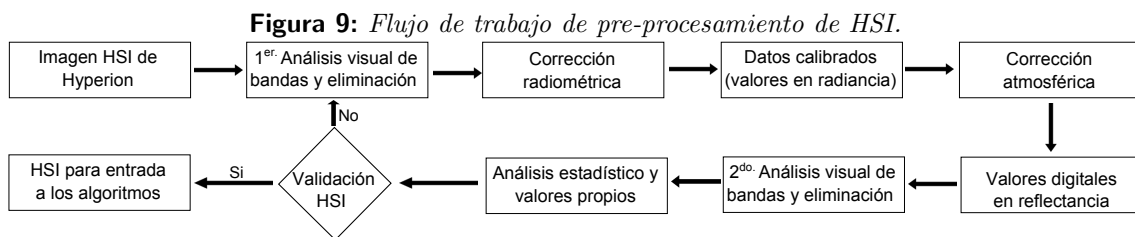
5.2 PREPROCESAMIENTO DE IMÁGENES HIPERESPECTRALES

Las HSI captadas presentan diferencias entre el valor de reflectancia registrado por el sensor y el valor real, debido a factores como el ángulo de azimut del sol, la altitud, las condiciones atmosféricas, la geometría de visión, entre otras características del sensor [46]. Por lo tanto, se requirió realizar actividades de pre procesamiento, tales como: correcciones radiométricas, atmosféricas y eliminación de bandas con baja relación señal/ruido (SNR). La corrección radiométrica elimina el ruido *striping* [47] causado por la no calibración de los detectores, y se manifiesta con la aparición de rayas verticales donde los valores digitales de una banda son sistemáticamente inferiores a los de las columnas adyacentes. Para el presente estudio, se aplicó la herramienta de calibración radiométrica del software ENVI.

Por otro lado, la corrección atmosférica es usada para compensar la pérdida de la señal, generada mientras viaja desde la fuente de iluminación hasta la escena y se refleja hacia el sensor. Durante la corrección atmosférica, los valores de radiancia son transformados en valores de reflectancia espectral.

Finalmente, con el objetivo de identificar y eliminar bandas espectrales redundantes y validar las bandas con baja SNR o bandas malas (que contienen pocos datos válidos), se realizó un análisis estadístico de las bandas en búsqueda de valores propios (del inglés *eigenvalues* [30]) bajos, que corresponden a bandas con poca información válida.

Para cada HSI seleccionada se aplico el flujo de trabajo de pre-procesamiento que se muestra en la figura 9, donde se ilustra la secuencia de actividades realizadas para obtener HSI que puedan ser usadas como entrada en los diferentes algoritmos de detección.

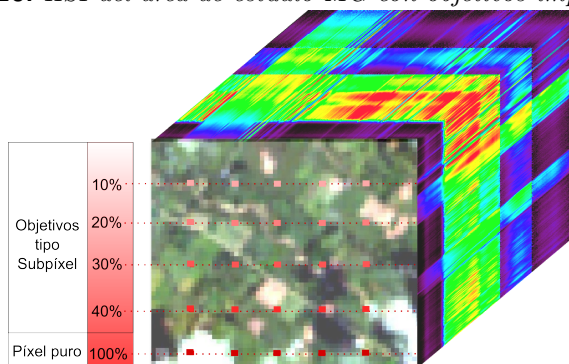


5.3 CONJUNTO DE DATOS SINTÉTICOS E IMPLANTACIÓN DE OBJETIVOS FRACCIONADOS

Con el objetivo de tener diferentes criterios de evaluación del desempeño de los algoritmos de detección en estudio, se crearon varias imágenes HSI sintéticas, a partir de la HSI captada para el área de estudio MG, la HSI real tiene 3341×931 píxeles, resolución espacial de 3 m y una resolución espectral de 242 bandas espectrales en el rango espectral de 355 nm hasta 2577 nm, después del preprocesamiento realizado se removieron bandas que presentaban bajo SNR generado por presencia de alta absorción de agua y bandas obsoletas o tradicionalmente conocidas en SR como bandas malas, finalmente se selecciono el sub-área de estudio para MG que corresponde a escena con presencia de cultivos de caña de azúcar obteniendo un cubo de datos de 52×62 con 170 bandas espectrales, con el cual se generaron los datos sintéticos y se implantaron objetivos dentro de la escena (ver figura 10).

Para la implantación de objetivos sintéticos en la HSI de MG se selecciono una firma espectral pura de un mineral no presente en la área de estudio MG (municipio de Mogotes, Santander). El mineral corresponde a la alunita y su firma espectral se extrajo de la librería ASTER Versión 2.0 [45]. Aster es un compendio de firmas espectrales de materiales naturales y artificiales, que contiene la librería espectral del Servicio Geológico de Estados Unidos (USGS), la librería espectral de la Universidad Johns Hopkins (JHU), y la librería del laboratorio de la NASA Jet Propulsion (JPL). En el orden de coherencia del experimento, la firma espectral de la alunita se réscalo para iguales rangos de reflectancia de la imagen y se seleccionaron las longitudes de onda acordes a las imágenes de Hyperion.

Figura 10: HSI del área de estudio MG con objetivos implantados.

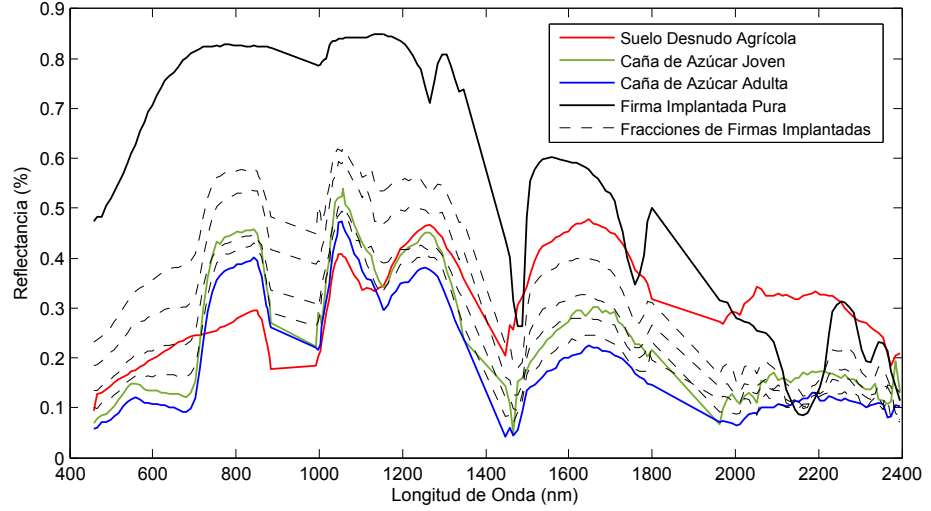


El método utilizado para la creación e implantación de objetivos parte de la base de crear una nuevo píxel objetivo d_{imp} conformado por de una fracción f de fondo o píxel existente y una fracción f de la firma espectral a implantar, el píxel sintético a implantar puede ser expresado como

$$d_{imp} = f \cdot d + (1 - f) \cdot x, \quad (10)$$

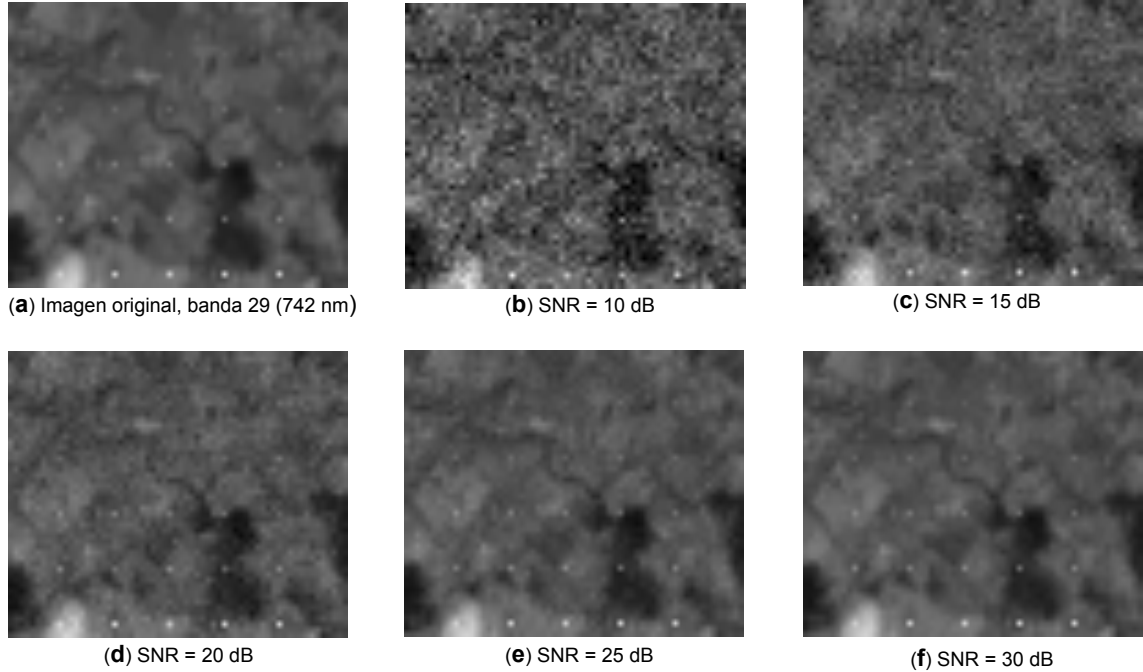
donde d es un vector $K \times 1$ que contiene la firma espectral pura de alunita, la cual se modifica su abundancia a partir una especifica fracción f , y es mezclada con la firma espectral del fondo x . Dentro de la HSI se implanta una grilla de objetivos de 5×5 con diferentes fracciones f implantadas de 0.1 (10%), 0.2 (20%), 0.3 (30%) y 0.4 (40%) para píxeles tipo sub-píxel y un ultimo conjunto de 5 objetivos puros. La figura 10 muestra la ubicación de cada píxel implantado y la fracción f implantada en cada objetivo, y la figura 11 muestra las firmas espectrales objetivo implantadas y las principales firmas dentro de la imagen HSI.

Figura 11: Firmas espectrales de fondo y objetivos implantados a diferentes niveles de mezcla.



Adicionalmente, se consideró efectos de ruido sobre la imagen y los objetivos implantados. En tanto, se adiciona ruido a las HSI con el fin de simular un comportamiento de los detectores semejante a condiciones reales de funcionamiento del sistema. Por ello, se adiciono ruido gaussiano blanco a la imagen sintética que contiene los objetivos implantados. Finalmente, se crearon cinco HSI sintéticas de prueba con SNR de 10 dB, 15 dB, 20 dB, 25 dB y 30 dB. El SNR corresponde a la tasa de la energía de la señal sobre la energía del ruido, matemáticamente se define como $SNR = \left(\frac{P_s}{P_n}\right)$, el SNR puede ser expresado en decibelios como $SNR_{dB} = 10 \log_{10}\left(\frac{P_s}{P_n}\right)$. La figura 12 muestra la banda 29 que corresponde a la longitud de onda $\lambda = 742nm$ del NIR, la banda se muestra a diferentes niveles de SNR.

Figura 12: Imágenes de banda espectral 29 que corresponde a la HSI sintética con los objetivos implantados, y diferentes niveles de SNR. (a) Banda espectral 29 sin SNR y con objetivos implantados, (b) Banda espectral 29 con SNR= 10 dB, (c) Banda espectral 29 con SNR= 15 dB, (d) Banda espectral 29 con SNR= 20 dB, (e) Banda espectral 29 con SNR= 25 dB, (f) Banda espectral 29 con SNR= 30 dB.



5.4 EVALUACIÓN DE DESEMPEÑO DE ALGORITMOS DE DETECCIÓN DE OBJETIVOS CON DATOS SINTÉTICOS

Para cada imagen sintética creada se evaluó el desempeño de los algoritmos de detección MF, CEM, ACE, SAM y OSP. Así mismo, se evaluaron los algoritmos con los objetivos sintéticos creados a partir de la firma espectral pura de la alunita a diferentes grados de mezcla (10 %, 20 %, 30 %, y 40 %), igualmente se evaluó el comportamiento ante la firma pura (100 %) de la alunita.

5.4.1 Experimentos y resultados con datos sintéticos sin ruido

La primera evaluación se realizó sobre la HSI sin ninguna clase de ruido, los algoritmos deben detectar la firma espectral mezclada con una fracción $f = 10\%$ de alunita y un 90 % de fondo, ubicada espacialmente en el píxel (10,30).

La figura 13 muestra una de las métricas visuales, donde se observa la ubicación del objetivo a detectar (verdad de terreno) y la respuesta de cada algoritmo (proyección bidimensional como respuesta de los detectores). Los algoritmos ACE, CEM y MF muestran el mejor desempeño. Para detallar y cuantificar con mayor precisión el desempeño de cada algoritmo, se evaluó el desempeño mediante la métrica de curvas de ROC (ver figura 14). El algoritmo ACE logro el mejor desempeño, alcanzando una P_D de 94 % con solo una FA, a diferencia del SAM que presento el menor desempeño con solo una P_D del 48 % obtuvo 124 FA.

Figura 13: Resultado de algoritmos con un objetivo implantado del 10% de la firma pura de la alunita en HSI de prueba sin ruido.

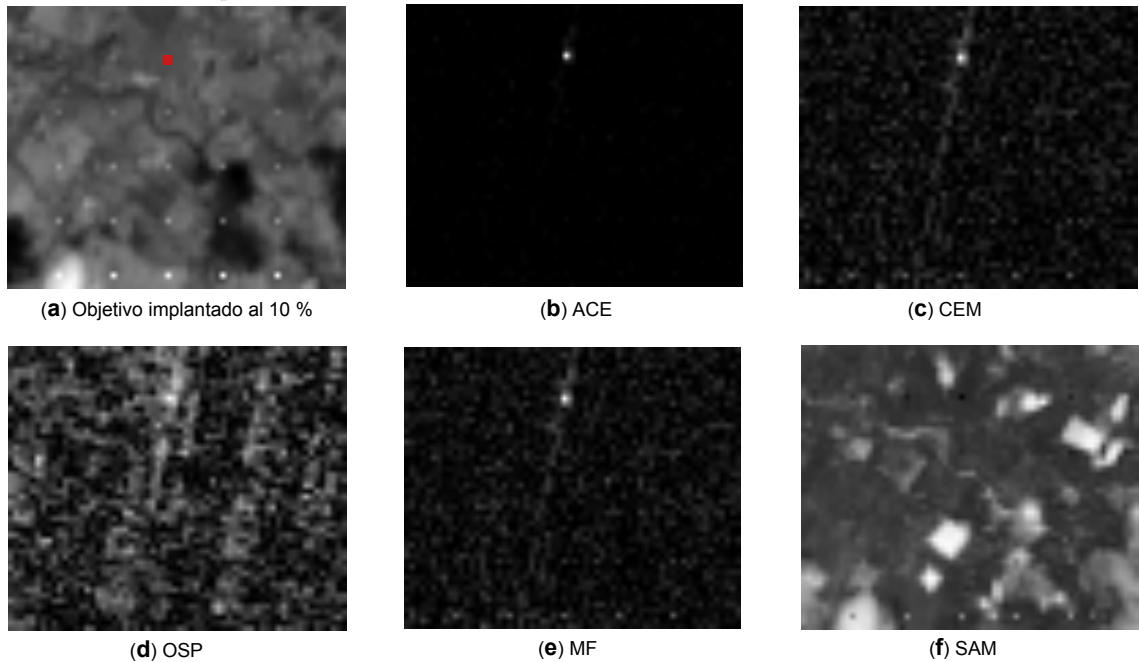
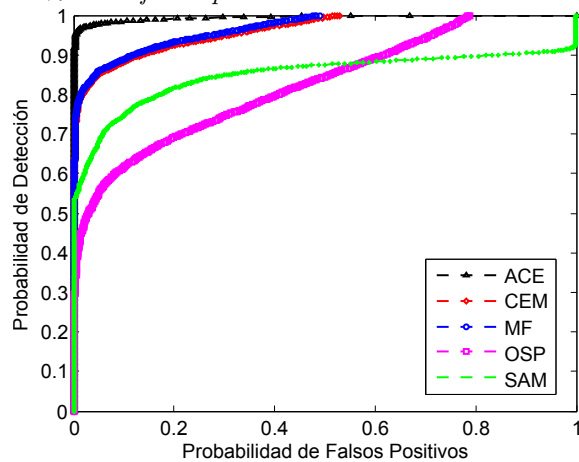
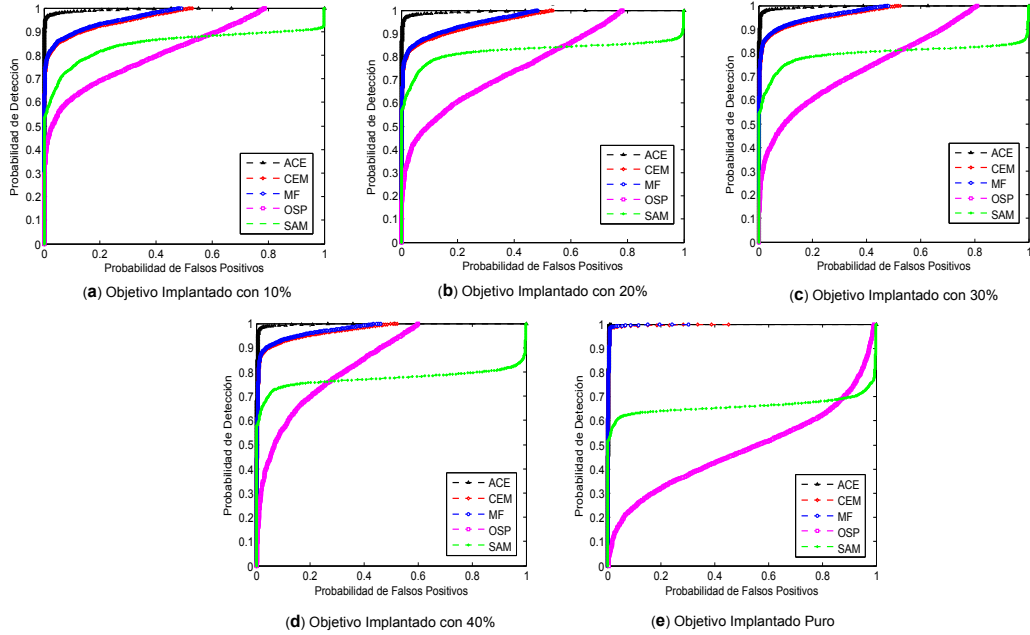


Figura 14: Curva ROC, desempeño de los algoritmos ACE, CEM, OSP, MF, y SAM con un objetivo implantado del 10% de la firma pura de la alunita dentro de la HSI de prueba sin ruido.



Así mismo, se evaluó el desempeño de todos los algoritmos para diferentes niveles de mezcla de la firma sintética, teniendo presente en el análisis objetivos de tipo sub-píxel para los objetivos con fracciones de $f = 10\%$, 20% , 30% y 40% . La figura 15 muestra las curvas ROC que grafican el desempeño de los diferentes algoritmos, donde se evidencia como los algoritmos ACE, CEM y MF incrementan su desempeño de P_D en la medida que la fracción f de la firma pura aumenta. Sin embargo, se observa que el comportamiento del algoritmos ACE es minimamente variante a los diferentes niveles de f . Adicionalmente, se observa un comportamiento atípico del detector OSP, ya que su rendimiento es inversamente proporcional en la medida que aumenta el porcentaje de f .

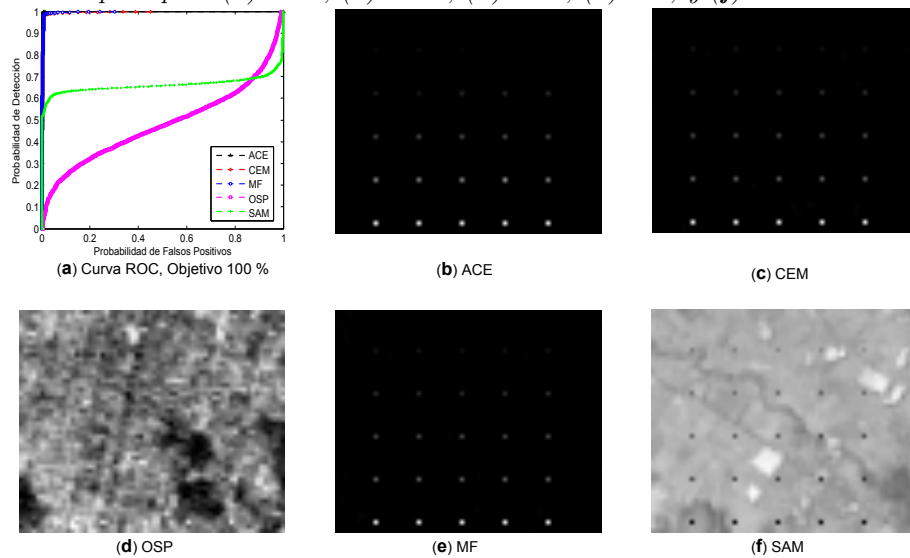
Figura 15: Curvas ROC desempeño de detección de objetivo implantado con diferentes fracciones de la firma pura de alunita dentro de HSI de prueba *sin ruido*. (a) $f = 10\%$, (b) $f = 20\%$, (c) $f = 30\%$, (d) $f = 40\%$, y (e) $f = 100\%$ firma pura.



Otro punto de importancia a entender en el comportamiento de los algoritmos, es cuando estos se enfrentan a objetivos tipo *full píxel* o píxel puro. Para evaluar este comportamiento se realizó una última prueba con la imagen sin ruido y con el objetivo de la firma espectral pura de la alunita implantada en la fila 50 de la HSI.

Se evidencia que los detectores ACE, CEM y MF muestran una muy buena respuesta frente a objetivos con firmas espectrales muy bien definidas (firmas puras). ACE, CEM y MF detectan la totalidad de los objetivos con P_D superior al 98%, ACE logra una P_D del 99.8% lo que corrobora la calidad de respuesta del detector observado en la figura 16.a.

Figura 16: Curvas ROC y respuesta de detección para *objetivos puros* en HSI sin ruido. (a) Curva ROC con píxel puro (b) ACE, (c) CEM, (d) OSP, (e) MF, y (f) SAM

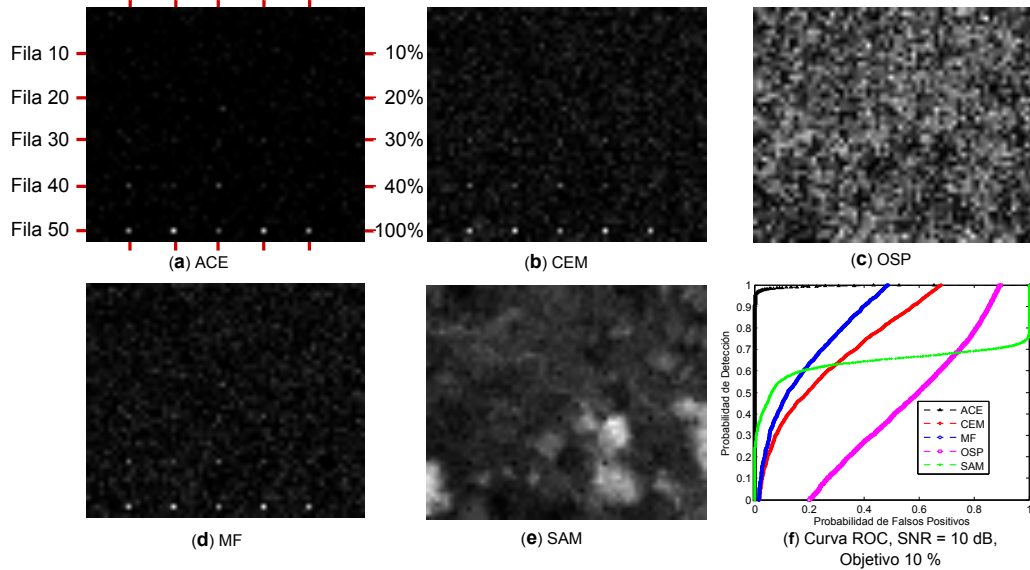


5.4.2 Experimentos y resultados con datos sintéticos con SNR=10 dB, SNR=15 dB, SNR=20 dB, SNR=25 dB y SNR=30 dB

Con el fin de analizar el desempeño y comportamiento de los cinco algoritmos seleccionados a diferentes niveles de ruido, se utilizó la HSI con los objetivos de detección implantados dentro de la misma y se adicionó ruido gaussiano blanco (10 dB, 15 dB, 20 dB, 25 dB y 30 dB), generando cinco HSI sintéticas.

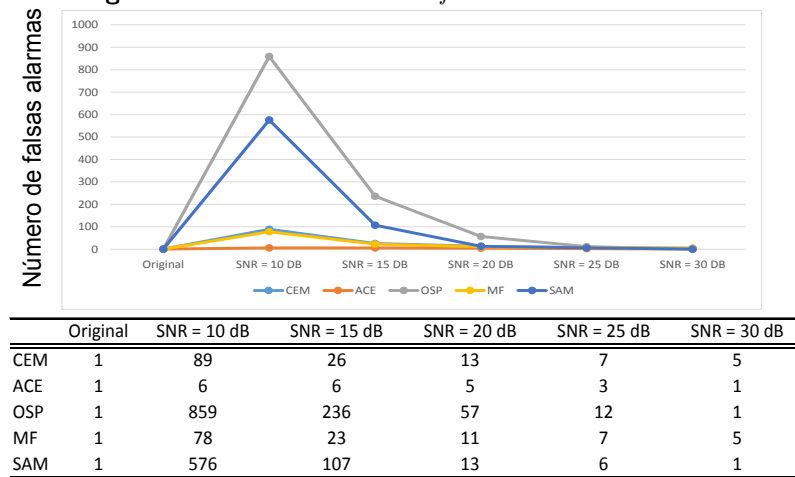
Para cada HSI sintética se corrieron los cinco algoritmos de detección, tanto para sub-píxeles como para píxeles puros (full-píxel). La figura 17 muestra el resultado de los algoritmos ejecutados sobre la imagen sintética con ruido SNR=10 dB, aquí se evalúa la capacidad de detectar un objetivo de tipo subpíxel formado por el 10% de alunita. Los resultados de los algoritmos muestran como ningún detector es capaz de detectar los 5 objetivos ubicados en la fila 10, dentro de una HSI con un SNR=10 dB, y el número de falsas alarmas se incrementa considerablemente. Sin embargo, con este nivel de ruido algunos logran detectar los píxeles implantados con firmas puras ubicados en la fila 50 de la HSI. La figura 17.a. muestra una malla donde los nodos indican la ubicación de los objetivos y las filas muestran el grado de mezcla del objetivo.

Figura 17: Resultado de respuesta de los detectores sobre HSI sintética con SNR=10 dB y con una firma mezclada con fracción $f = 10\%$ como objetivo de tipo subpíxel a detectar. (a) ACE, (b) CEM, (c) OSP, (d) MF, (e) SAM, y (f) Curvas ROC.



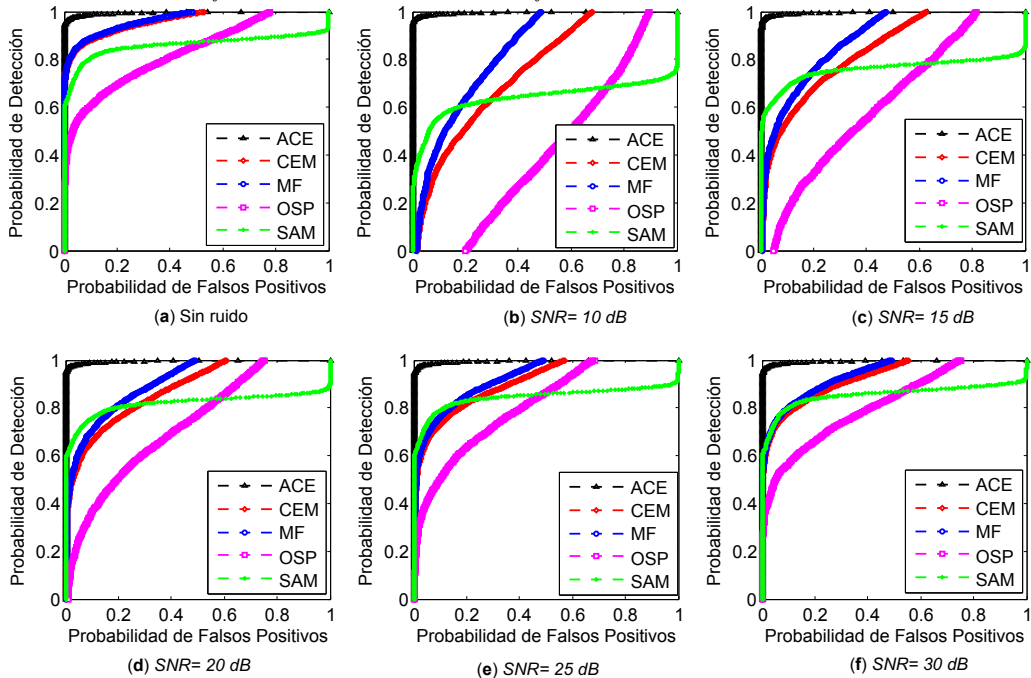
Así mismo, se evaluó tomando como criterio que cada algoritmo tuviese solo una falsa alarma en la imagen original como punto de partida, para realizar el análisis del comportamiento de número de falsas alarmas (FA) al ser adicionado diferentes niveles de ruido (ver figura ??). OSP y SAM presentaron el mayor número de FA con 859 FA y 576 FA respectivamente. Finalmente, todos los detectores se estabilizan respecto a un menor número de FA aceptables cuando se llegan a niveles de SNR=25 dB o superior.

Figura 18: Número de FA a diferentes niveles de ruido.



La figura 19 muestra seis curvas ROC, que permiten realizar un análisis general del resultado de todos los detectores a diferentes niveles de ruido, específicamente cuando la tarea es detectar un objetivo tipo subpíxel; los resultados arrojaron que los algoritmos CEM, MF, y OSP son drásticamente inferiores con SNR=10 dB o menores. Sin embargo, se observa que el algoritmo SAM presenta un mejor comportamiento, esto debido a que SAM no es sensible a modificaciones en las magnitud del vector que se ven afectadas por el ruido, por el contrario SAM al basarse en el calculo del ángulo entre píxeles solo puede ser afectado cuando la dirección de los vectores se altera.

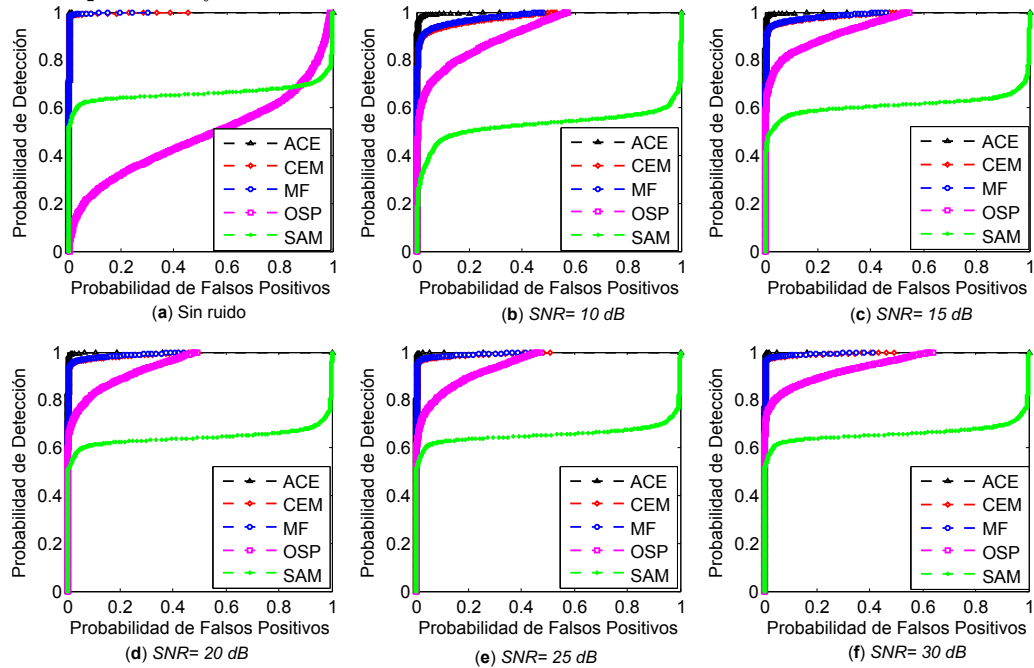
Figura 19: Curvas ROC de resultado de diferentes algoritmos al detectar un *objetivo tipo sub-píxel* creado con una firma mezclada al 10% a diferentes niveles de ruido.



Los resultados de desempeño de los diferentes algoritmos varían cuando se trata de objetivos tipo *full píxel* o píxel puro. La figura 20 muestra una mejor respuesta de los detectores a diferentes niveles de ruido inducido en las HSI cuando el objetivo a detectar es *full píxel*.

De igual manera, se observa que todos los algoritmos estabilizan su desempeño a niveles de ruido de 20 dB o superior, a diferencia cuando se trabaja con objetivos tipo sub-píxel que solo logran un desempeño aceptable con niveles de ruido superior a los 25 dB. Así mismo, el algoritmo CEM cuando se trata de objetivos tipo *full píxel* alcanza desempeños iguales al MF, algoritmo que siempre ha superado a CEM en las pruebas realizadas con datos sintéticos.

Figura 20: Curvas ROC de resultado de diferentes algoritmos al detectar un *objetivo de tipo píxel completo* a diferentes niveles de ruido



5.5 EVALUACIÓN DE DESEMPEÑO DE ALGORITMOS DE DETECCIÓN DE OBJETIVOS CON DATOS REALES

La principal tarea de la presente investigación es evaluar el uso y aplicación de las HSI a partir de la identificación e implementación de algoritmos de detección de objetivos en escenas o áreas agrícolas en Colombia, a partir de HSI. Para tal fin, se logro adquirir un número mayor a 20 HSI captadas por el sensor Hyperion desde el satélite EO-1 en las áreas de estudio seleccionadas (ver sección 5.1 Conjunto de datos de imágenes hiperespectrales reales y áreas de estudio).

La presente sección se abordará de manera individual y separada para cada área de estudio, con el objetivo de evaluar todos los aspectos y detalles surgidos en cada área y con cada HSI desde el plan de adquisición de las imágenes, las HSI, su préprocesamiento, y finalmente la aplicación de los algoritmos de detección.

5.5.1 Resultados de algoritmos de detección de objetivos en HSI del área de estudio de Puerto Wilches (PW)

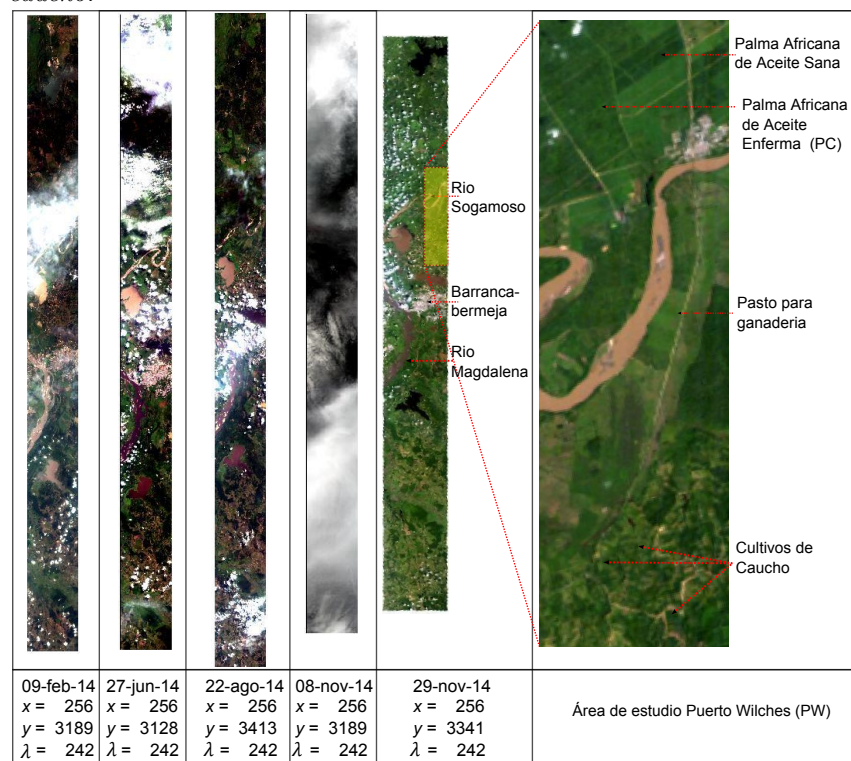
Puerto Wilches es un municipio ubicado al noroccidente del departamento de Santander. se caracteriza por poseer la mayor parte de su superficie plana. con un relieve de valles, su clima es templado y húmedo con una temperatura promedio de $33^{\circ}C$, por su superficie cruzan dos grandes ríos, el río Magdalena y el río Sogamoso. Su altura es de 90 msnm. El

principal cultivo en PW es la palma africana de aceite, ubicado en la zona norte de PW, en el sur existen cultivos tradicionales como el maíz, plátano, y áreas con pastos para ganado. Así mismo, se encuentran pequeñas plantaciones o áreas cultivadas con árboles de caucho.

El área de estudio de PW, es el área que posee mayor número de HSI adquiridas en el presente estudio. Se cuenta con siete HSI captadas entre el 9 de febrero del 2014 y el 29 de noviembre de 2014. La figura 21 muestra cinco HSI, dentro de las cuales se encuentra la imagen que finalmente se utilizó para el área de PW, y corresponde a la HSI captada el 29 de noviembre de 2014, por sus condiciones de muy baja nubosidad y sombras es la HSI más apropiada para el presente estudio.

De igual manera, se observa el problema de alta nubosidad en las diferentes escenas (ver figura. 21), siendo este uno de los retos a superar en la captura de las HSI. Debido a la resolución temporal (tasa de revisita) del satélite EO-1 sobre el territorio Colombiano (cada 16 días), la adquisición de escenas optimas es limitada. Colombia se caracteriza por ser un país con alta nubosidad, dificultando en cierto nivel estudios basados en SR.

Figura 21: HSI captadas por el sensor Hyperion para el área de estudio PW, en color verdadero (bandas 29=rojo, 20=verde, y 12=azul) y área de estudio específica a cultivos de palma africana de aceite y caucho.

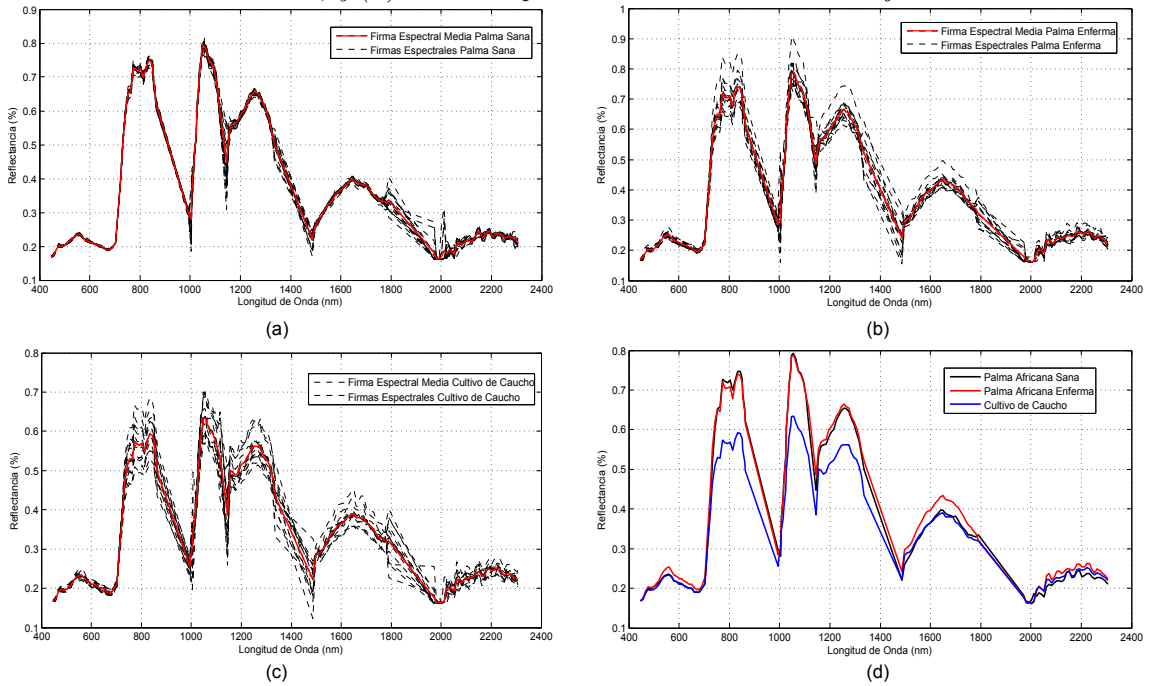


Para la extracción de las firmas espectrales a partir de HSI, es necesario superar el problema de la variabilidad espectral. Como resultado en la extracción de firmas espectrales para el área PW, se encontraron espectros correspondientes a los píxeles para el mismo tipo de superficie (ej. cultivo de palma africana de aceite) que presentan una variabilidad espectral inherente, impidiendo la caracterización homogénea de las palmas dentro de una misma área del cultivo. La figura 22 muestra la variabilidad espectral presente en el cultivo de palma africana, presentándose en mayor grado en los cultivos de palma africana enferma, debido a la irregularidad de los cultivos enfermos, donde se encuentran palmas muertas o

palmas enfermas afectadas en diferentes grados.

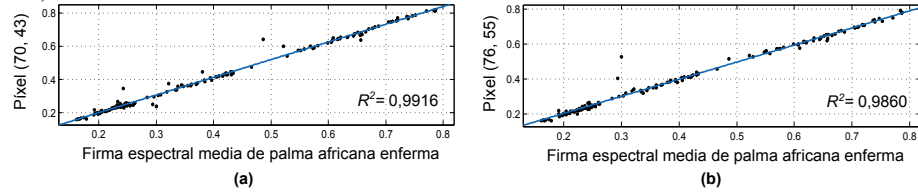
El área de estudio de PW presenta particularmente una gran afectación del cultivo de palma africana generada por la enfermedad denominada Pudrición de Cogollo (PC). La PC es considerada como un hongo. El hongo provoca que las hojas de la planta se sequen y adquieran un tono amarillento, acelerando el proceso de pudrición y muerte de la palma. En PW el área cultivada en palma africana es de aproximadamente 70.000 has de las cuales para el año 2013 más del 50 % (35.000 has) de los cultivos de palma poseen la enfermedad. Esta es la principal razón para extraer de esta área de estudio firmas espectrales como objetivos a detectar que correspondan a cultivos de palma africana sana y cultivos de palma africana enferma. Así mismo, en la parte inferior del área de estudio al sur de PW se encuentran plantaciones de caucho que serán también objeto de investigación.

Figura 22: Firmas espectral de cultivos característicos en área de estudio de PW. (a) Firma media de palma africana de aceite sana, (b) Firma media de palma africana de aceite enferma, (c) Firma media de cultivo de caucho, y (d) Firmas espectrales seleccionadas como objetivos



Se tomaron 15 firmas espectrales de la HSI por cada cultivo seleccionado o sub-área. A través de la estimación del coeficiente de determinación R^2 de cada firma como la variable dependiente respecto a la media del conjunto de firmas (ver figura 23), se retiraron las firmas que presentaron un menor R^2 , quedando un conjunto de 12 firmas con las cuales se generó la firma final de cada objetivo (ver figura 22.d).

Figura 23: Coeficientes de determinación R^2 de píxeles frente a la media de cultivo de palma africana enferma en PW. (a) R^2 de firma espectral del píxel (70,43), (b) R^2 de firma espectral del píxel (76,55)



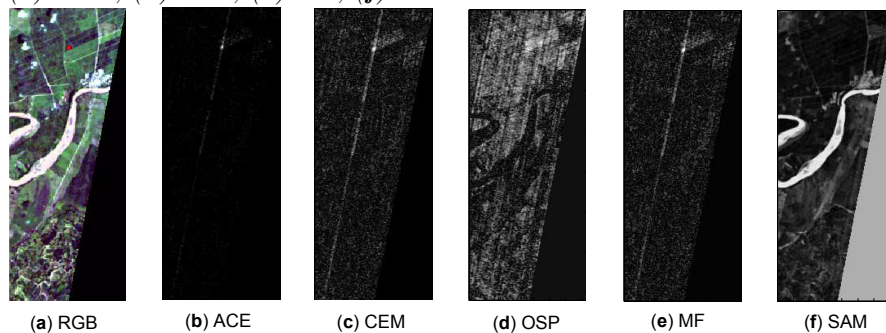
La figura 23 muestra los coeficientes de determinación de dos firmas espectrales tomadas para dos píxeles dentro de la HSI. Los píxeles corresponden a cultivos de palma africana enferma, como se comentó anteriormente la variabilidad presente en esta área del cultivo es alta, lo cual hace que el proceso de selección de una firma espectral u objetivo sea riguroso.

Con las tres firmas seleccionadas como objetivos se ejecutan los cinco algoritmos identificados (CEM, ACE, OSP, MF, y SAM) sobre la HSI preprocesada del área de estudio de PW. Finalmente, la HSI tiene una dimensión de 359×145 , y $\lambda = 149$ bandas espectrales.

5.5.1.1. Resultados de algoritmos para la detección de cultivos de palma africana sana como objetivo en PW

La figura 24 muestra los resultados de los diferentes algoritmos de detección para un objetivo caracterizado por la firma espectral de un cultivo de palma africana saludable. La ubicación del objetivo se identifica con un polígono rojo sobre la HSI en color verdadero (ver figura 24.a). Los cinco algoritmos de detección (ACE, CEM, OSP, MF y SAM) lograron detectar el objetivo. Sin embargo, el número de falsas alarmas generado por los diferentes algoritmos alteran su óptimo desempeño. Visualmente sobre la figura 24 se puede observar que el algoritmo ACE presenta el mejor comportamiento. Así mismo, los dos algoritmos con menor desempeño (SAM y OSP) aun teniendo un gran número de falsas alarmas, permiten mapear o visualizar el contorno de las áreas o parcelas con gran similitud espectral.

Figura 24: Resultados de los algoritmos en la detección de palma africana sana. (a) HSI en RGB, (b) ACE, (c) CEM, (d) OSP, (e) MF, (f) SAM.

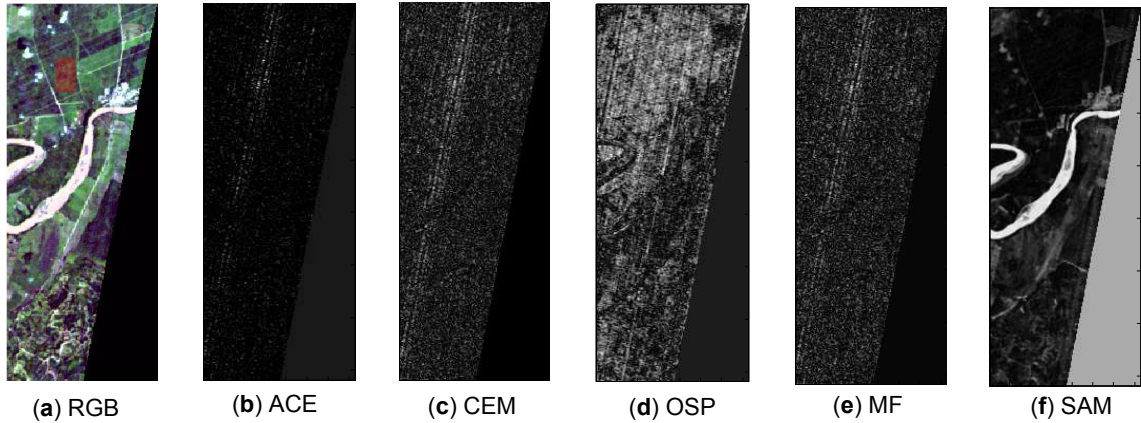


5.5.1.2. Resultados de algoritmos para la detección de cultivos de palma africana enferma como objetivo en PW

Para la detección de cultivos de palma africana enferma, todos los algoritmos redujeron su desempeño con relación a objetivos que tienen un comportamiento más homogéneo, la

irregularidad de las áreas cultivadas con presencia de palmas enfermas hace que la variabilidad espectral aumente, y por ende, el número de falsas alarmas. La figura 25 muestra el resultado de los diferentes algoritmos. Sin embargo, visualmente ningún detector muestra una área definida de detección del objetivo en su ubicación real, la ubicación del objetivo es señalada mediante un polígono rojo en la figura 25.a. Por lo tanto, es recomendable observar el desempeño de los algoritmos mediante métricas cuantitativas. Este análisis se realiza mediante las curvas ROC (ver figura 27), allí se observa que la probabilidad de detección de los algoritmos CEM, OSP y SAM cayó por debajo del 40% en su tarea de detectar las áreas con cultivos enfermos. Adicionalmente, las curvas ROC muestran que para este tipo de objetivo, solo el algoritmo ACE presenta un mejor comportamiento. Sin embargo, no es aceptable en términos del número de falsas alarmas (> 100) cuando ACE posee solo un 80% de seguridad en la detección de la firma objetivo.

Figura 25: Resultado de los algoritmos en la detección de palma africana enferma. (a) HSI en RGB, (b) ACE, (c) CEM, (d) OSP, (e) MF, (f) SAM.



5.5.1.3. Resultados de algoritmos para la detección de cultivos de caucho como objetivo en PW

Como ultimo objetivo a detectar utilizando en el área estudio de PW, se caracterizó la firma espectral de cultivos de caucho (ver figura 26). Los cultivos de caucho se ubican la parte inferior de la escena o HSI, y se identifica su ubicación mediante un polígono rojo en la figura 26.a. En la detección de este objetivo se volvió a presentar un mejor comportamiento de los detectores con respecto al anterior objetivo (cultivo de palma africana enferma). Las curva ROC (ver figura 27.c) para el cultivo de caucho muestra un comportamiento atípico en el algoritmo OSP, debido a la falta de valores de la imagen, causado por la manera como fue cubierta o barrida el área al momento de la captura por el sensor, debido a la órbita o trayectoria del satélite. El cultivo de caucho posee características homogéneas (alturas, estado del cultivo, follaje, etc), convirtiéndolo en un cultivo apto para ser detectado mediante el uso de algoritmos de detección en HSI, los algoritmos ACE y MF permitieron detectar con mayor detalles la presencia del cultivo. Desde el punto de vista espacial, se puede mostrar la capacidad de los algoritmos para detectar un objetivo. Específicamente en esta escena, el área de estudio posee una dimensión de 4.650 m \times 10.770 m, para un total de 46 millones de m^2 equivalente 4.685 hectáreas (has), y el área cubierta por el objetivo (cultivo de caucho) y que fue detectada es de tan solo 36 has.

Figura 26: Resultado de los algoritmos en la detección de cultivo de caucho. (a) HSI en RGB, (b) ACE, (c) CEM, (d) OSP, (e) MF, (f) SAM.

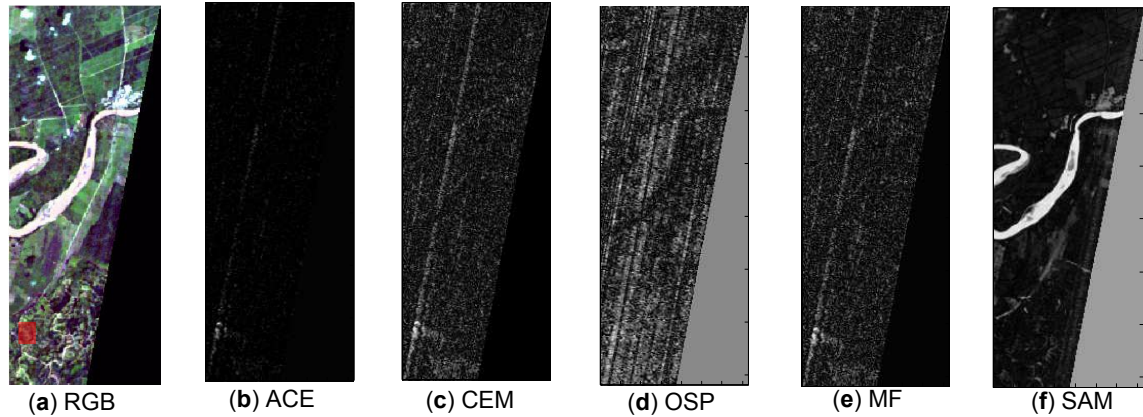
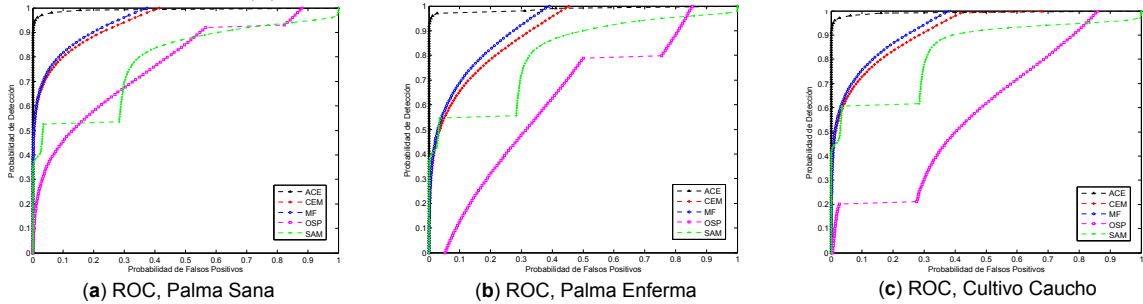


Figura 27: Curvas ROC de los algoritmos para los objetivos (a) Palma africana sana, (b) Palma africana enferma, y (c) Cultivo de Caucho.



Otro criterio de evaluación del desempeño de los detectores fue cuantificar el número de falsas alarmas a tasas de probabilidad de detección (P_D) del 80% y 90%. La tabla 2 muestra los resultados que validan el comportamiento observado en las imágenes de salida de los detectores y las curvas ROC. Se destaca de forma negativa el gran número de falsas alarmas generadas en el algoritmo OSP al intentar detectar cultivos de palma africana enferma, OSP con una probabilidad de detección del 90% alcanzo 4.654 falsas alarmas.

Tabla 2: Número de falsas alarmas frente a probabilidades de detección del 90% y 80% para los objetivos (a) Palma africana sana, (b) Palma africana enferma, (c) Cultivo de Caucho.

Algoritmos / PD	90%	80%
ACE	0	0
CEM	15	20
OSP	199	624
MF	15	19
SAM	N.A	N.A

(a) Número de FA, Palma Sana

Algoritmos / PD	90%	80%
ACE	0	0
CEM	71	166
OSP	4634	6997
MF	69	148
SAM	N.A	N.A

(b) Número de FA, Palma Enferma

Algoritmos / PD	90%	80%
ACE	0	0
CEM	36	14
OSP	632	1374
MF	31	64
SAM	N.A	N.A

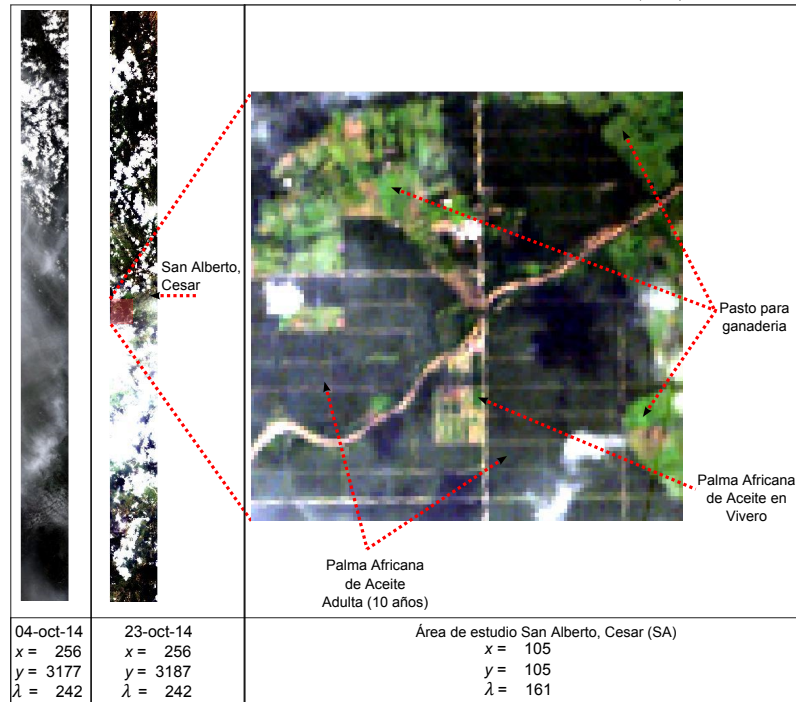
(c) Número de FA, Cultivo Caucho

5.5.2 Resultados de algoritmos de detección de objetivos en HSI del área de estudio de San Alberto (SA)

El municipio de San Alberto se ubica al sur del departamento del Cesar, Colombia, y limita con el norte del departamento de Santander. El área SA es muy similar a la zona de PW, el 80 % de SA posee valles de iguales características a los ubicados en PW. Así mismo, por su territorio circundan dos cuerpos de agua (río Lebrija y río San Alberto). Las condiciones geográficas de superficies planas y alta temperaturas ($33^{\circ}C$) son propicias para los cultivos agrícolas de palma africana de aceite y grandes superficies de pasturas para ganadería. A diferencia de PW, las plantaciones de palma africana en SA, son plantaciones saludables y los brotes de la enfermedad PC son mínimos.

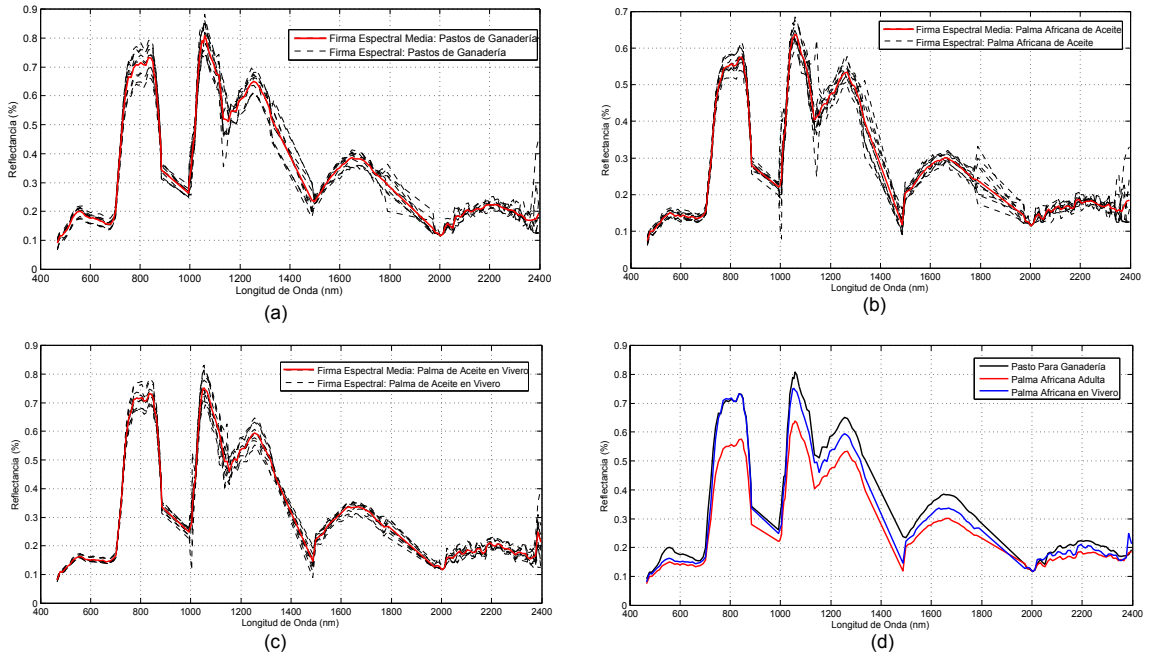
Para el área de estudio SA, Hyperion captó dos HSI en el mes de octubre de 2014, de igual modo como en PW, se presentaron dificultades por la alta nubosidad en los días de adquisición. Sin embargo, en la escena del día 23 de octubre fue posible captar el área de interés, donde se encuentran plantaciones de palma africana y grandes superficies de pastos para ganados (ver figura 28).

Figura 28: HSI del área de estudio de San Alberto (SA), Cesar.



El procedimiento para la extracción de las firmas espectrales objetivos del área de estudio de SA y que se usaron sobre los cinco algoritmos seleccionados, fue similar al usado en PW. Sin embargo, para la caracterización de la firma espectral de pastos dedicados a la ganadería fueron tomados píxeles en tres áreas separas por aproximadamente 2 Km entre sí, ver figura 28. Así mismo, para la caracterización de la firma espectral de plantas de palma africana de aceite en vivero, se tomo ventaja de la existencia del vivero de la plantación ubicado en la parte central de la escena, que corresponde a 10 píxeles de la HSI, permitiendo tomar como verdad de terreno los 10 píxeles. La figura 29 muestra las tres firmas seleccionadas como objetivo y que corresponden a pastos para ganadería, cultivo de palma africana adulta (10 años), y plantas de palma africana en vivero.

Figura 29: Firmas espectral de cultivos característicos en área de estudio de SA. (a) Firma media de pastos para ganadería, (b) Firma media de palma africana de aceite adulta (10 años), (c) Firma media de palma africana de aceite en vivero, y (d) Firmas espectrales seleccionadas como objetivos en SA.



Con las tres firmas seleccionadas como objetivos se ejecutan los cinco algoritmos identificados (CEM, ACE, OSP, MF, y SAM) sobre la HSI preprocesada del área de estudio de SA. La HSI usada para el área SA tiene una dimensión de 105×105 , y $\lambda = 161$ bandas espectrales. Las figuras 30, 31, y 32 muestran la respuesta de detección. Específicamente, la figura 30 muestra la detección de la firma espectral que caracterizan el objetivo de pastos dedicados a ganadería, distribuidos especialmente en la escena. Es posible visualizar y validar con la curvas ROC (ver figura 33.a) que los dos mejores algoritmos para este objetivo, fueron el ACE y MF. La figura 30.e que corresponde a la salida del detector MF, facilita visualmente la ubicación de las áreas de pastos para ganado que el algoritmo MF detecto, un análisis con mayor detalle y cuantitativo se realiza en la sección 5.5.2.1.

La respuesta de detección del objetivo que corresponde a palma africana adulta, se ubica en la figura 31 y en las curvas ROC graficadas en la 33.b. Donde se puede observar que todos los algoritmos disminuyen su rendimiento, aquí dos algoritmos (MF y CEM) que han mantenido un comportamiento estable decrecen con este objetivo, su P_D reportada en la curva ROC parte desde solo el 20 %, después de analizar la HSI en las bandas del infrarojo se encontró que la principal causa se debe a las sombras generadas por la nubes, que coinciden sobre gran parte del área con palma africana, generando aumento en la variabilidad espectral del cultivo a detectar.

Figura 30: Resultado de los algoritmos en la detección de cultivo o áreas dedicadas a pastos para ganado. (a) HSI en RGB, (b) ACE, (c) CEM, (d) OSP, (e) MF, (f) SAM.

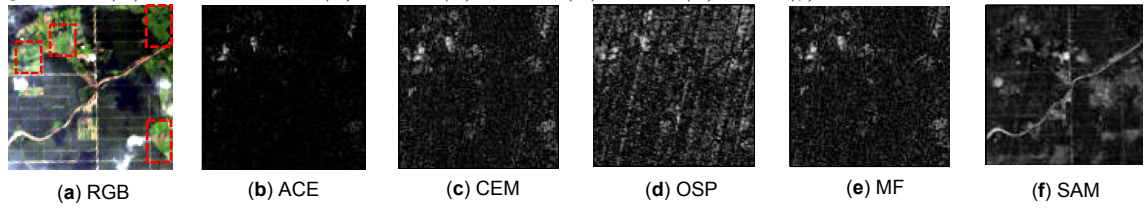


Figura 31: Resultado de los algoritmos en la detección de cultivos de palma africana adulta. (a) HSI en RGB, (b) ACE, (c) CEM, (d) OSP, (e) MF, (f) SAM.

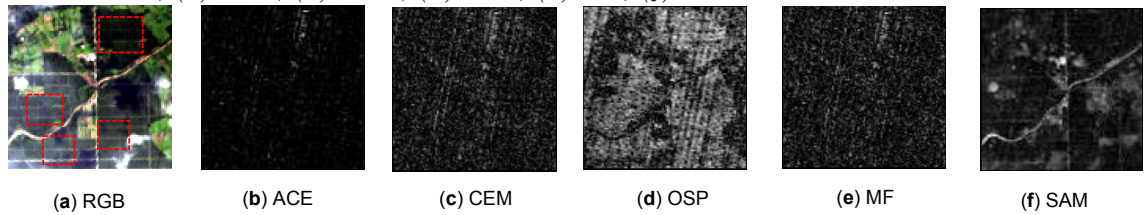


Figura 32: Resultado de los algoritmos en la detección de cultivos de palma africana en vivero. (a) HSI en RGB, (b) ACE, (c) CEM, (d) OSP, (e) MF, (f) SAM.

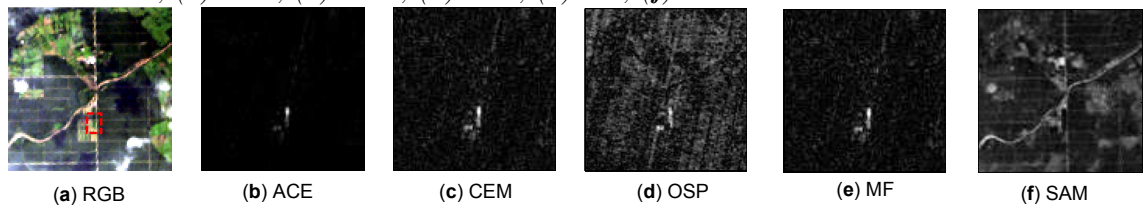
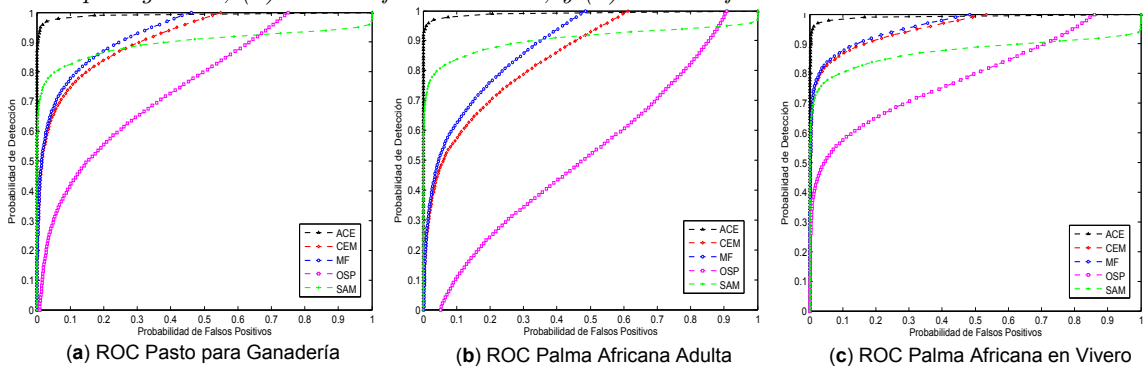


Figura 33: Curvas ROC de desempeño de los algoritmos en la detección de cultivos en SA. (a) Pasto para ganado, (b) Palma africana adulta, y (c) Palma africana en vivero.



5.5.2.1. Evaluación de resultados de algoritmos basados en métricas con umbrales para la detección de plantas de palma africana en vivero como objetivo en SA.

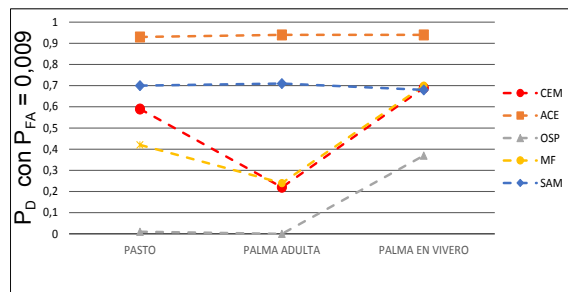
Para el análisis de los resultados de la firma espectral que caracteriza las plantas de palma africana en vivero, plantas que son controladas en condiciones optimas y con mayor

densidad (número de plantas por m^2) que las palmas adultas ubicadas sobre los cultivos productivos. Para la evaluación del desempeño de este objetivo, se aplicaran dos métricas que involucran el uso de umbrales, tanto para P_D , como para la P_{FA} . Para este experimento se creó un mapa de verdad de terreno tomando ventaja de la existencia del vivero de la plantación ubicado en la parte central de la escena, una área del vivero fue seleccionada como objetivo. El objetivo está formado por 10 píxeles de la HSI.

Existen dificultades cuando se intenta verificar con mayor detalle el desempeño de los algoritmos a través de la estimación de las curvas ROC. Las dificultades se deben al número de píxeles en un hipercubo tradicionalmente es del orden de 10^5 , lo que limita los métodos empíricos de estimación de la P_{FA} que es del orden de menos de 10^{-4} por cubo; adicionalmente, el número de píxeles objetivos en una escena es muy pequeño, lo que limita la precisión que se puede lograr en la estimación del rendimiento de detección las curvas ROC; para superar estas dificultades el número mínimo de píxeles N utilizados para estimar una probabilidad P debe ser $10/P$ o preferiblemente de $100/P$ [48].

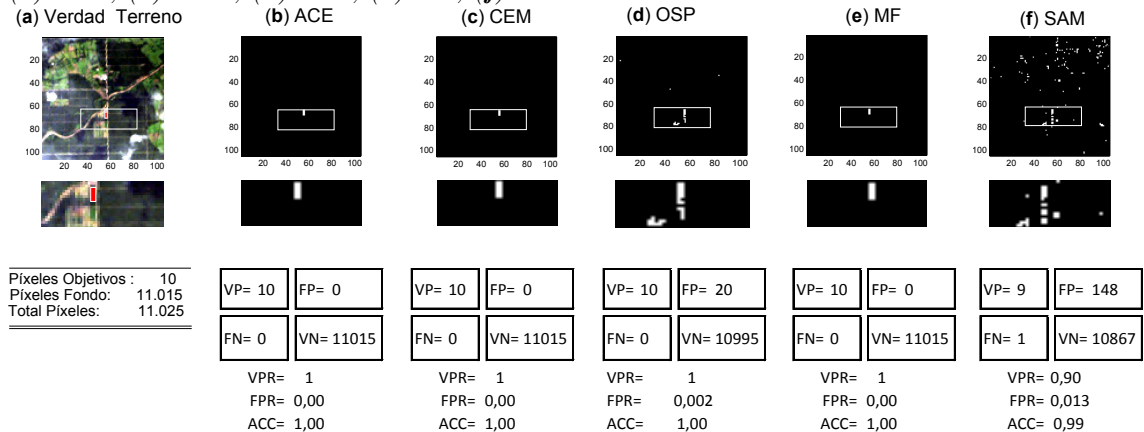
Por lo tanto, se utilizó una métrica que involucre una tasa de falsa alarma constante (TFAC) mediante la determinación de un umbral η para una $P_{FA} = 100/N$, donde N es el número de píxeles de la escena (11.025 píxeles), la figura 34 muestra la P_D de los tres objetivos con un umbral η con una $P_{FA} = 0,009$. La figura 34 permite validar desde esta otra métrica que el algoritmo ACE a umbrales de P_{FA} del orden de 10^{-3} mantiene el mejor desempeño. Así mismo, el bajo desempeño de los algoritmos CEM y MF cuando se trata del objetivo de palma africana adulta es notable, al perder aproximadamente 4 unidades porcentuales.

Figura 34: P_D de los algoritmos con una $P_{FA} = 0,009$ para los objetivos de pasto, palma africana adulta y palma africana en vivero.



Otro experimento realizado para analizar el comportamiento de los algoritmos, consistió en crear un umbral por cada detector tomando como base la menor respuesta de salida entre los píxeles que conforman el objetivo. De esta manera, se proyectaron las respuestas de cada detector en la figura 35.b-f. Así mismo, la primera imagen de la figura muestra la verdad de terreno y la estadística de número de píxeles objetivo, píxeles de fondo y el total de los mismos.

Figura 35: Desempeño de los algoritmos basados en umbrales sobre la respuesta de cada detector para el objetivo de palma africana en vivero y sus matrices de confusión. (a) verdad de terreno, (b) ACE, (c) CEM, (d) OSP, (e) MF, (f) SAM.

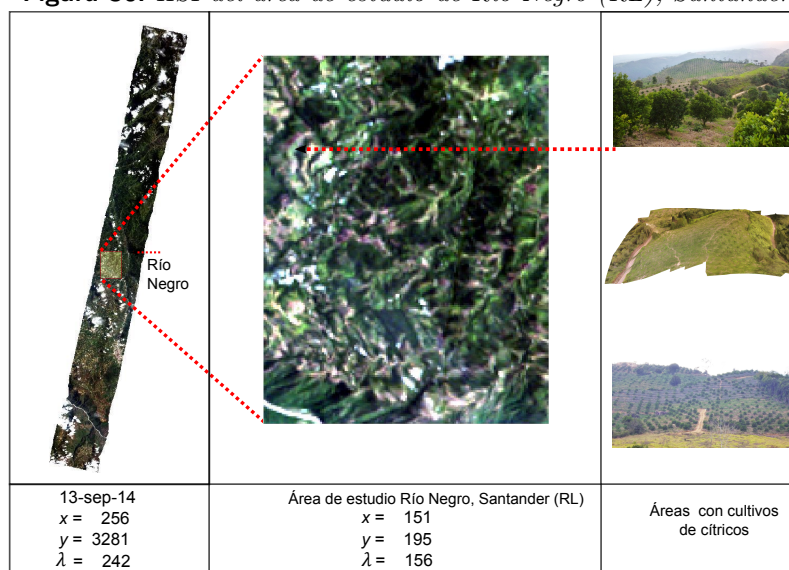


Como resultado de este experimento, se derivan otras métricas que permiten cuantificar la razón o tasa de verdaderos positivos (VPR) frente a la razón o tasa de falsos positivos (FPR), así como la exactitud (ACC). Para tal fin se crearon las matrices de confusión binaria para cada detector. La matriz de confusión está compuesta de cuatro categorías: los verdaderos positivos (VP) que son los objetivos correctamente detectados. Los falsos positivos (FP), también llamados falsas alarmas son las detecciones de objetivos cuando no lo son. Los verdaderos negativos (VN), detecciones negativas o de fondo detectadas correctamente como fondo. Y falsos negativos (FN) que son los píxeles objetivos que el algoritmo detecta de manera errada como fondo y los marca como fondo. La figura 35 muestra el resultado de cada algoritmo bajo diferentes métricas, desde las visuales, pasando por detecciones de objetivos con umbrales basados en las respuesta de los píxeles, la creación de las matrices de confusión binaria, y finalmente la estimación de las tasas de VPR, FPR y ACC. De manera general los resultados muestran como los algoritmos ACE, CEM y MF presentan un comportamiento optimo, en contraste con los algoritmos OSP y SAM que son capaces de detectar el objetivo, sin embargo, sus FP son altas desmejorando su rendimiento, estos resultados validan los evaluados desde las curvas ROC (ver figura 33) y los obtenidos con la métrica basada en umbrales para la P_{FA} (ver figura 34).

5.5.3 Resultados de algoritmos de detección de objetivos en HSI del área de estudio de Río Negro (RL)

El Municipio de Río Negro, es la primera área de estudio en la presente investigación en poseer una geografía montañosa, con una temperatura promedio de 28°. RL se caracteriza por su agricultura tradicional, con cultivos de menor extensión o superficie como cacao, café, y cítricos. RL se diferencia de PW y SA por su geografía, clima y cultivos industrializados de palma africana y pastos para ganadería extensiva, siendo cultivados en grandes superficies. La figura 36 muestra la HSI y el sub-área seleccionada, donde se observa el sistema montañoso presente.

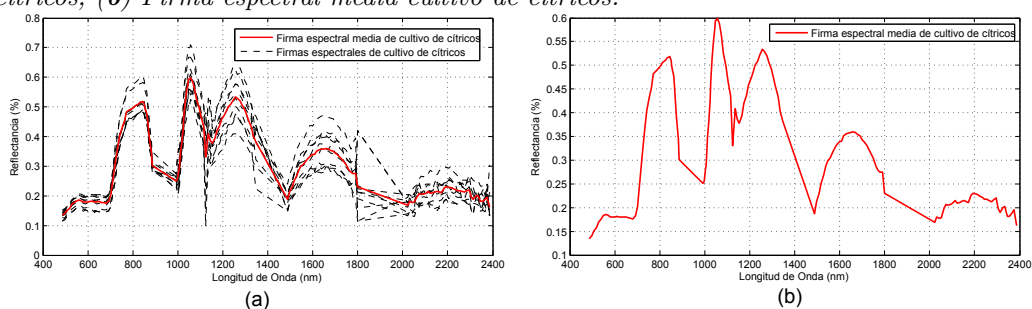
Figura 36: HSI del área de estudio de Río Negro (RL), Santander.



En RL se presentaron tres limitantes, la primera hace referencia al número de HSI, solo se logro adquirir una HSI, la segunda se debe al tipo de agricultura tradicional, con presencia de cultivos de menor tamaño en superficie cultivada, y la tercera limitante se debe a la sombras generadas por las montañas presentes en la escena, debido a la geografía del terreno y la hora de adquisición de la imagen (2:00 pm).

Como objetivo para el área de RL se seleccionó un superficie con un cultivo de cítricos que posee una superficie superior a $30m^2$ (tamaño de un píxel de Hyperion). La figura 37.a muestra las firmas espectrales de diferentes píxeles presentes en el área cultiva con cítricos, y la figura 37.b muestra la firma espectral media utilizada como objetivo para detectar el cultivo de cítricos.

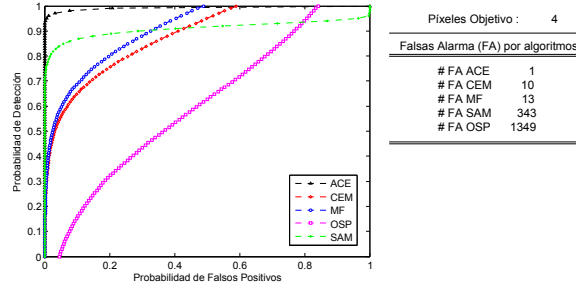
Figura 37: Firma espectral de cultivos de cítricos en RL, (a) diferentes firmas extraídas de cultivos de cítricos, (b) Firma espectral media cultivo de cítricos.



Para evaluar el desempeño de los algoritmos a la hora de detectar el cultivo de cítricos en el área de estudio RL, el objetivo se representó por 4 píxeles como la verdad de terreno, y para su evaluación se aplicaron dos métricas. La figura 38 muestra las curvas ROC, donde se observa como el algoritmo SAM, a pesar de mostrar desempeños bajos en otras áreas de estudio, para RL supera los algoritmos CEM y MF. Esto se debe a que SAM no se ve afectado por problemas de sombras o cambios en los niveles de iluminación de los píxeles. Desde otro punto de vista, se cuantificó el número de FA que presentan los algoritmos en el proceso de detectar los 4 píxeles objetivos, se observa que el algoritmo ACE presenta el mejor desempeño, logrando detectar el objetivo con solo una FA, así mismo se observa

que aunque SAM en la curva ROC muestra el segundo mejor desempeño, al momento de cuantificar el número de FA, pasa a la cuarta posición de desempeño, el efecto se explica al observar detalladamente el comportamiento de la curva ROC de SAM, esta logra detectar inicialmente los objetivos con 75% de P_D , sin embargo, se desplaza por el eje de P_{FA} sumando FA.

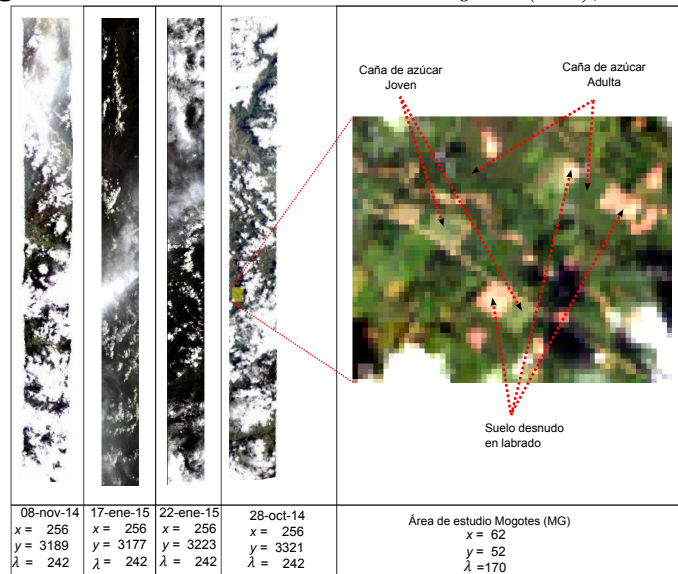
Figura 38: Curvas ROC de desempeño de algoritmos de detección de cultivo de cítricos como objetivo en RL, y el número de FA para en la detección de 4 píxeles.



5.5.4 Resultados de algoritmos de detección de objetivos en HSI del área de estudio de Mogotes (MG)

El área de estudio de MG es similar a RL, posee una geografía de montaña, sin embargo, MG tiene una mayor altura sobre el nivel del mar y su clima es más frío, con una temperatura promedio de 18°. La caña de azúcar es uno de los principales cultivos agrícolas en MG. Con el fin de corregir los problemas presentados en RL, respecto a sobras generadas por montañas y la falta de cultivos con superficies cultivadas en grandes extensiones, para el área de estudio de MG se seleccionó un valle donde se encuentran grandes superficies cultivadas con caña de azúcar en diferentes etapas de su desarrollo y crecimiento. La figura 39 muestra las diferentes HSI adquiridas, donde se visualiza la alta nubosidad presente, de igual manera se muestra el sub-área donde se identifican las ubicaciones de diferentes tipos de cultivos de caña de azúcar.

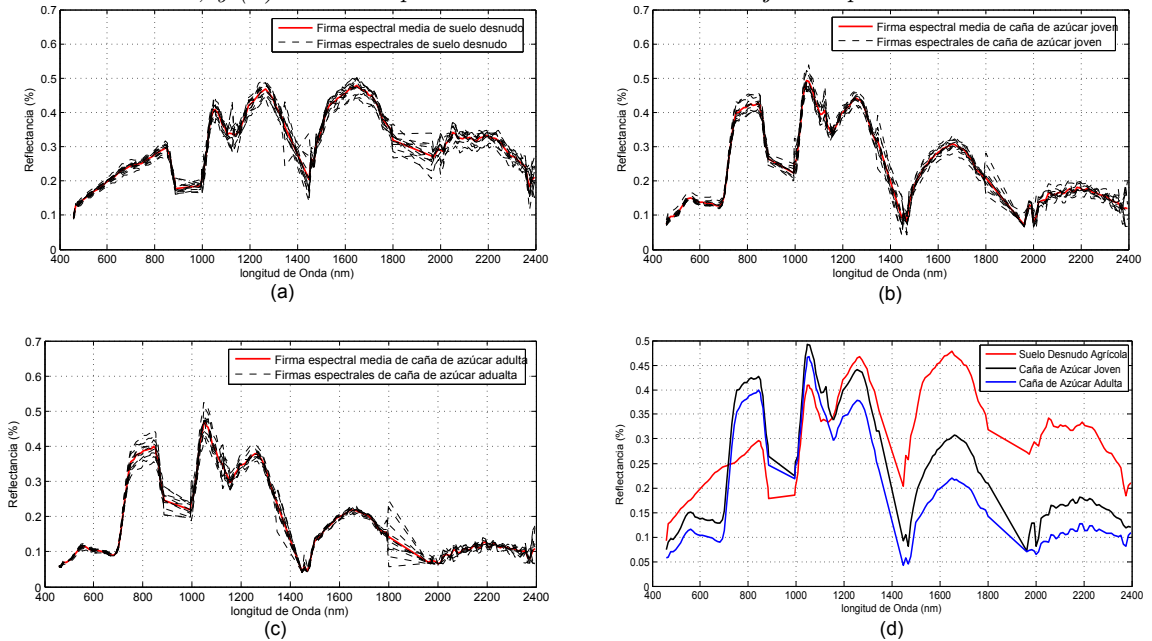
Figura 39: HSI del área de estudio de Mogotes (MG), Santander.



La selección de firmas espectrales en MG, tomó ventaja de la presencia de cultivos de caña de azúcar en sus diferentes niveles de desarrollo y crecimiento, permitiendo caracterizar

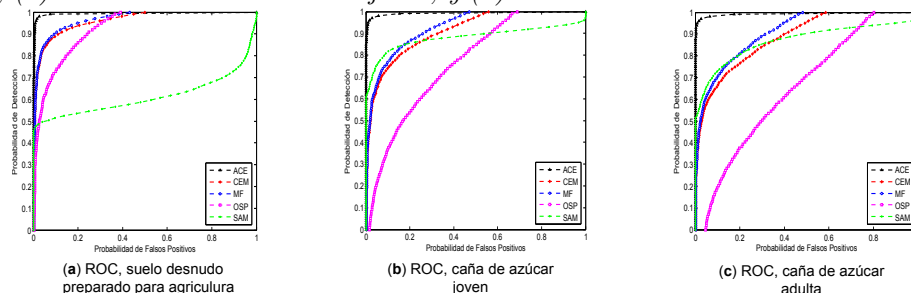
tres firmas espectrales (suelo desnudo en preparación para cultivar caña de azúcar, caña de azúcar joven (3 meses), y caña de azúcar adulta o madura (> 8 meses)). La figura 37 muestra las tres firmas espectrales seleccionadas como objetivos y su variación espectral. Se destaca en la figura 37.a la firma espectral del suelo desnudo por su forma y comportamiento espectral, debido a la alta reflectancia de las longitudes de onda del rojo por efecto del suelo.

Figura 40: Firmas espectral de cultivos característicos en área de estudio de MG. (a) Firma media de suelo desnudo, (b) Firma media de caña de azúcar joven, (c) Firma media de cultivo de caña de azúcar adulta, y (d) Firmas espectrales seleccionadas como objetivos para MG



Para comparar el desempeño de los algoritmos ACE, CEM, OSP, MF y SAM en el área de estudio de MG se realizaron tres pruebas, cada una, con una firma espectral caracterizada en la figura 40.d. El resultado mostró que los algoritmos tienen un mejor desempeño, incrementando la P_D en 30 puntos porcentuales, cuando se corren sobre escenas en territorios planos, con grandes áreas cultivadas de un mismo cultivo, condiciones como las anteriores permiten tener escenas homogéneas y libres de sombras, por ende se disminuye la variabilidad espectral presente en los objetivos. La figuras 41.a, b y c muestran las curvas ROC para los tres objetivos, El mejor desempeño para los algoritmos ACE, CEM, MF y OSP, se logró al detectar la firma espectral de suelo desnudo, principalmente por dos aspectos: el primero de ellos se debe a la baja variabilidad espectral presente en la firma espectral de suelo desnudo, y la segunda se genera porque el desempeño de estos detectores aumenta cuando el objetivo a detectar es fácilmente diferenciable del fondo, específicamente, el 75% del área de estudio de MG es vegetación y un 20% es suelo desnudo, el restante 5% corresponde a otros materiales u objetos (ríos y vías). Por lo tanto, la firma espectral de suelo desnudo tiene una gran diferencia del fondo (vegetación), debido a la alta reflectancia del rojo (ver 40.d).

Figura 41: Curvas ROC de desempeño de detección de los algoritmos para los objetivos (a) Suelo desnudo, (b) Cultivo de caña de azúcar joven, y (c) Cultivo de caña de azúcar adulta.

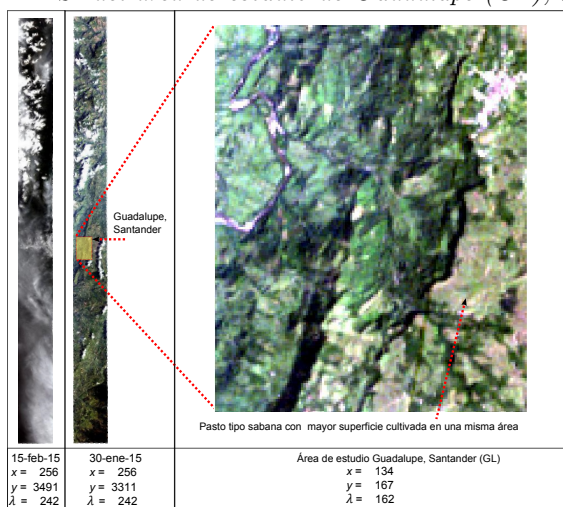


Adicionalmente, las figuras 41.b y c ilustran como los algoritmos MF y CEM, pierden desempeño con respecto a SAM, debido a la sensibilidad de los detectores cuando existen alta similitud entre los píxel a evaluar (ej. firmas espectrales de cultivos de caña de azúcar en diferentes etapas de crecimiento). Finalmente, el desempeño de los tres mejores algoritmos en orden descendente de desempeño son ACE, MF y CEM, debido a la forma en que modelan el fondo de manera estadística, para lograr una mayor separación entre fondo y objetivo.

5.5.5 Resultados de algoritmos de detección de objetivos en HSI del área de estudio de Guadalupe (GL)

La ultima área de estudio es Guadalupe (GL), GL se encuentra en el sureste del departamento de Santander, posee geografía montañosa, y algunas áreas onduladas. GL es una región agrícola, sin embargo, su agricultura es la más tradicional de las cinco áreas de estudio en la presente investigación. En GL existen cultivos en pequeñas fracciones o extensiones de tierra, tales como café, yuca, y caña de azúcar, debido a que la gran parte de los predios o tierras de GL son pequeños minifundios o fincas donde una familia puede tener en solo dos hectáreas mas de cinco cultivos diferentes (café, yuca, plátano, cítricos, huertas caseras, entre otros); las áreas con mayor superficie cultivada y destinada a fines agrícolas corresponde a pastos tipo sabana nativa, dedicada a la crianza de ganado. La figura 42 muestra las HSI de GL y el área de estudio seleccionada.

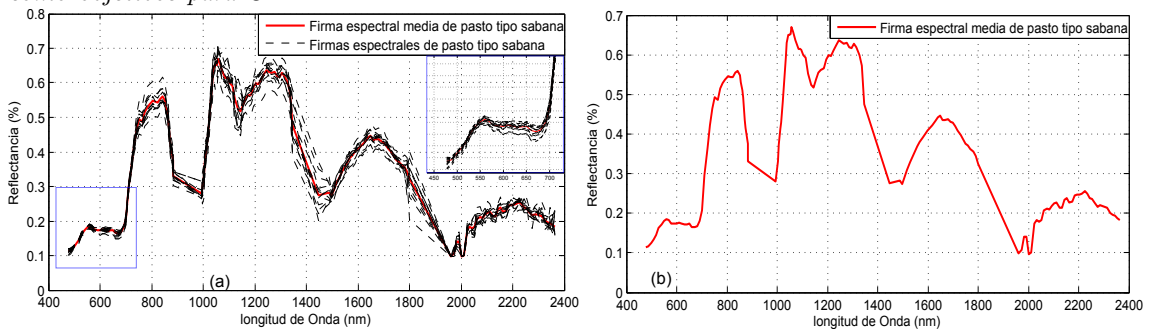
Figura 42: HSI del área de estudio de Guadalupe (GL), Santander.



En GL se caracterizó una sola firma espectral, debido al pequeño tamaño de los cultivos presentes en la zona, el cultivo con más extensión y que facilitó la extracción de varios

píxeles para caracterizar una firma espectral, fue una superficie cultivada con pastos tipo sabana dedicada a la crianza de ganado. Esta área permitió la extracción de más de 10 píxeles, logrando tener una firma media como objetivo. La figura 43.a ilustra la variabilidad espectral presente en el área cultivada con pasto, debido a la influencia realizada por el comportamiento espectral del suelo (fondo). Los pastos tipo sabana se identifican por su poca altura, generando una mezcla de los espectros del suelo y el pasto, este fenómeno se visualiza por el comportamiento espectral entre las longitudes de onda de 600 nm a 680 nm, donde la reflectancia del rojo aumenta por influencia del suelo, y no por el comportamiento espectral del pasto. La figura 43.b muestra la firma espectral media de pasto tipo sabana usada como objetivo.

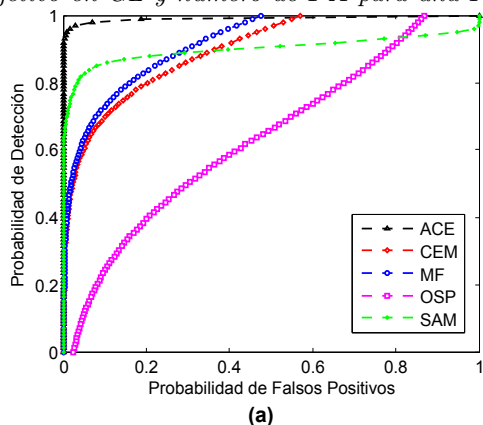
Figura 43: Firmas espectral de cultivos característicos en área de estudio de GL. (a) Firmas espectrales de pasto tipo sabana, (b) Firma espectral media de pasto tipo sabana seleccionadas como objetivos para GL



Para evaluar el desempeño de los algoritmos de detección en esta ultima área de estudio, se utilizaron dos métricas las curvas ROC y el número de FA con un umbral de $P_{FA} = 0,004$, donde el umbral de probabilidad se estimo como $100/N$, donde N es el número de píxeles de la HSI de GL (22378 píxeles). La figura 44.a ilustra las curvas ROC de desempeño de los cinco algoritmos a evaluar, los resultados muestran que el algoritmo ACE presentó el mejor desempeño, adicionalmente, al observar el comportamiento de los cinco algoritmos para todas las P_{FA} , el algoritmo MF y CEM tendrían el segundo y tercer mejor desempeño, respectivamente.

Sin embargo, se debe tener presente que la precisión de las curvas ROC se ven limitadas por el pequeño número de píxeles objetivo a detectar, que son del orden de 10^{-4} por cubo. Por lo tanto, la figura 44.b permite validar el desempeño de los cinco algoritmos mediante una tasa de falsa alarma constante (TFAC) con un umbral η para una $P_{FA} = 0,004$. los resultados de esta métrica muestran que el algoritmo ACE mantiene el mejor desempeño, sin embargo, el algoritmo SAM bajo $P_{FA} = 0,004$ tiene solo 14 FA superando a los algoritmos CEM, MF. El algoritmo OSP solo logra sus primeras detecciones a partir de $P_{FA} > 0,025$, teniendo el desempeño más bajo a la hora de detectar pasto de tipo sabana como objetivo en el área de estudio de GL.)

Figura 44: Curvas ROC de desempeño de algoritmos de detección de cultivo de pasto tipo sabana como objetivo en GL y número de FA para una $P_{FA} = 0,004$



$P_{FA} = 0,004$

Algoritmos	Número de FA
ACE	13
CEM	95
OSP	*
MF	94
SAM	14

* Para una $P_{FA} = 0,0255$
el número de FA es de 570

(b)

5.6 RESUMEN DE RESULTADOS

La principal tarea de la presente investigación es identificar e implementar algoritmos de detección de objetivos en HSI adquiridas por sensores remotos, aplicadas al sector de la agricultura Colombiana. Para su desarrollo, en este capítulo se evaluaron de cinco algoritmos seleccionados, estos algoritmos son: Estimador de coherencia adaptativo - ACE, Algoritmo de minimización de energía restringida - CEM, Algoritmo de filtro emparejado - MF, Proyector de subespacio ortogonal - OSP, Algoritmo de mapeo del ángulo espectral - SAM.

Este capítulo inicia realizando una introducción del conjunto de datos de HSI reales adquiridas y agrupadas en cinco áreas de estudio que corresponde a las poblaciones de Puerto Wilches (PW), Río Negro-Lebrija (RL), Mogotes (MG) y Guadalupe (GL) ubicadas en el departamento de Santander Colombia, y San Alberto (SA), ubicada en el departamento del Cesar Colombia. La selección de las áreas de estudio se realizó buscando que cumplieran con la restricción de poseer características geográficas y climáticas diferentes, permitiendo adquirir información espectral de varios cultivos agrícolas de importancia para el nororiente colombiano, tales como: palma africana de aceite, caucho, pastos para ganadería, cítricos y caña de azúcar. Así mismo, se crearon seis HSI sintéticas, una HSI sintética sin ruido y cinco con SNR de 10 dB, 15 dB, 20 dB, 25 dB y 30 dB. Se implantaron cinco objetivos sintéticos, creados a partir de la firma espectral pura de la alunita a diferentes grados de mezcla o fracción ($f = 10\%$, 20% , 30% , 40% , y 100%).

Para evaluar el desempeño de los cinco algoritmos, han sido diseñados diferentes experimentos, los primeros experimentos fueron realizados sobre el conjunto de datos sintéticos a diferentes niveles de SNR, junto con la implantación de firmas espectrales sintéticas con diferentes grados de mezcla o fracciones de la firma espectral pura de la alunita. El primer experimento evaluó el desempeño de los cinco algoritmos sobre una HSI sin ninguna clase de ruido, donde el objetivo a detectar es una firma espectral mezclada con una fracción $f = 10\%$ de alunita y un 90% de fondo, en este experimento los algoritmos ACE, CEM y MF muestran el mejor desempeño. El algoritmo ACE logró el mejor desempeño, alcanzando una P_D del 94% con solo una FA, a diferencia del SAM que presentó el menor desempeño, donde con una P_D del 48% y 124 FA.

Adicionalmente, los algoritmos se evaluaron con las HSI sintéticas a diferentes niveles de ruido. Las evaluaciones muestran que el desempeño de los algoritmos se afecta en la medida que el nivel de ruido aumenta. Sin embargo, su desempeño se estabiliza a niveles de ruido de 25 dB para objetivos de tipo sub-píxel y de 20 dB para objetivos de píxeles puros. El algoritmo con mejor desempeño; implementado con datos sintéticos variando los niveles de ruido y fracciones de abundancia de la firma; es el ACE. Sin embargo, el algoritmo SAM presenta un comportamiento minimamente variante a partir de 15 dB hacia arriba, esto se debe a que SAM no es sensible a la variabilidad espectral generada por ruido (sombras o variaciones de luminosidad) debido a que SAM depende solo de la forma espectral del objetivo y es independiente de las diferencias de magnitud entre el objetivo y los píxeles de fondo. Un último experimento con los datos sintéticos fue evaluar el comportamiento de los algoritmos con objetivos tipo *full píxel* o píxel puro, como resultado se observó que los detectores ACE, CEM y MF muestran un buen desempeño al detectar objetivos con firmas espectrales muy bien definidas (firmas puras), debido a la baja variabilidad espectral de las mismas. ACE, CEM y MF detectan la totalidad de los objetivos con P_D superior al 98 %, ACE logro una P_D del 99.8 %.

Los resultados de desempeño de los algoritmos de detección de objetivos ejecutados sobre el conjunto de datos de HSI reales, mostró que los cultivos con un comportamiento homogéneo (similitud en alturas, estado del cultivo, follaje, etc), como la palma africana de aceite para el área de estudio de PW y SA, y la caña de azúcar para MG, presentan una menor variabilidad espectral, por ende generan un mejor desempeño en todos los algoritmos. Así mismo, se observó que la geografía y las dimensiones de las áreas cultivadas, inciden directamente en el desempeño de los algoritmos. Los resultados mostraron que los algoritmos tienen un mejor desempeño, incrementando la P_D en 30 puntos porcentuales, cuando se corren sobre escenas en territorios planos, con grandes áreas cultivadas de un mismo cultivo, condiciones como las anteriores permiten tener escenas homogéneas y libres de sombras, disminuyendo la variabilidad espectral presente en los objetivos.

Finalmente, y de acuerdo a todas las evaluaciones realizadas en la presente investigación, el algoritmo ACE presentó el mejor desempeño para diferentes HSI y objetivos de tipo agrícola, tanto en HSI sintéticas como en HSI reales, seguido por los algoritmos CEM y MF que presentan un comportamiento aceptable con promedios de $P_D = 80\%$, en contraste con los algoritmos OSP y SAM que son capaces de detectar el objetivo, sin embargo, el número de FA son altas desmejorando sus desempeños.

6 LONGITUDES DE ONDA CON MAYOR IMPORTANCIA EN EL ESTUDIO ESPECTRAL DE LA VEGETACIÓN

El sensado remoto hiperespectral ha permitido nuevas formas de monitoreo del desarrollo y crecimiento de cultivos y la estimación de variables biofísicas y bioquímicas de las plantas. Existe una gran número de estudios centrados en la relación entre la concentración de pigmentos, nutrientes, estado de las plantas y las propiedades ópticas de las hojas (ver sección 3.1 Propiedades espectrales de la vegetación). De igual manera existen investigaciones que han estudiado este comportamiento a nivel de follaje (del inglés *canopy*) de diferentes cultivos [26, 49, 46, 25].

La base de la relación entre el desarrollo y crecimiento de las plantas y las propiedades ópticas se debe a que es posible medir el comportamiento espectral de la vegetación y la cantidad de energía reflejada por una planta a lo largo del espectro electromagnético ($400nm - 2500nm$), permitiendo la caracterización espectral a partir de su firma espectral (ver figura 4). Dado que una firma espectral puede ser considerada como un vector $x = [x_1, x_2, \dots, x_k]^T$, donde K corresponde al número de bandas espectrales y cada posición del vector a una respuesta espectral en una específica longitud de onda. Lo anterior permite la estimación de los índices de vegetación espectrales (V_{IS}).

Los (V_{IS}) se estiman a partir de combinaciones de dos o más bandas espectrales según las longitudes de onda (λ) y su ancho ($\Delta\lambda$). La mayoría de los (V_{IS}) comprende combinaciones lineales entre (λ) del NIR, (λ) del rojo en la región de visible (VIS), y (λ) del borde rojo. Las longitudes de onda situada en estas tres regiones poseen un alto contenido de información espectral de la vegetación [50]. La reflectancia espectral de la vegetación se caracteriza por una muy baja reflectancia en el espectro del rojo seguido por un aumento brusco de la reflectancia en (λ) entre $\sim 700 \pm 740nm$ (borde rojo), y con sus mayores valores de reflectancia en el NIR.

De esta manera, las imágenes captadas por sensores hiperespectrales permiten mejorar la estimación de índices de vegetación, gracias a la gran cantidad de bandas o longitudes de onda (λ) angosta y estrecha ($\Delta\lambda$) (*narrowbands*), en comparación con sistemas tradicionales de sensado remoto multiespectral o de bandas extensas y discontinuas (*broadbands*) [51]. Sin embargo, la alta dimensionalidad de los datos, específicamente el número de $\lambda > 100$ o 200 genera una incertidumbre entre cual o cuales son las (λ) que entregan mayor información para cuantificar, detectar, y clasificar características de los cultivo a través de la estimación de V_{IS} .

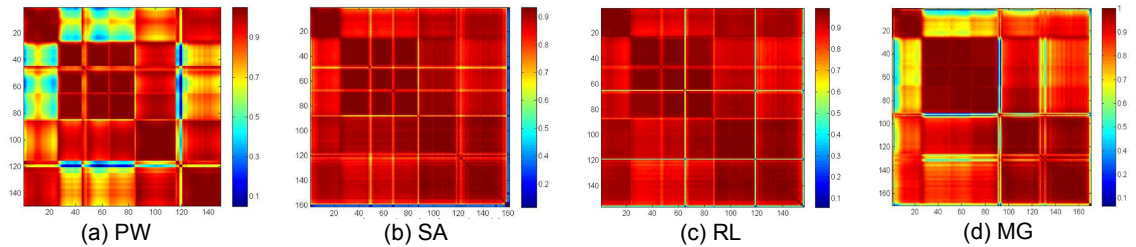
Con el objetivo de analizar cuales longitudes de onda o bandas espectrales son las más apropiadas para las HSI y firmas espectrales usadas en la presente investigación, se realizaran tres tareas: (1) Creación de matrices de correlación de cada HSI y análisis de los valores propios (*eigenvalues*) de cada HSI para determinar que bandas brindan mayor información, e identificar la existencia de bandas redundantes y bandas malas. (2) Análisis de la distribución espacial y espectral de los píxeles que corresponden a bandas del rojo, borde rojo, y NIR mediante gráficos de dispersión de (λ_i contra λ_j). (3) Finalmente, se validaran los resultados con respecto al estado del arte de longitudes de onda espectral

óptimas para el estudio de la vegetación.

6.1 ANÁLISIS DE LONGITUDES DE ONDA A PARTIR DE LAS MATRICES DE CORRELACIÓN Y LOS VALORES PROPIOS

La resolución espectral se define por las longitudes de onda (λ) y su ancho ($\Delta\lambda$), por ende es posible establecer el número de λ o bandas espectrales. En HSI el número de bandas se caracteriza por su cantidad (> 100). Sin embargo, cuando se requiere la estimación de V_{IS} esta es una limitante, por lo tanto, se deben aplicar técnicas que reducen la dimensionalidad con el objeto de reducir la carga computacional, eliminar ruido e información redundante en la HSI. El objetivo es obtener una representación mínima de la imagen que contenga la información indispensable para realizar el análisis sobre un sub-conjunto reducido de la imagen original. Adicionalmente, las técnicas de reducción dimensional permiten una mejora de la relación SNR en los datos a través de la eliminación de ruido.

Figura 45: Matrices de correlación de HSI de áreas de estudio de: (a) Puerto Wilches, (b) San Alberto, (c) Rionegro, y (d) Mogotes.



Específicamente en la tarea de encontrar las longitudes de onda o bandas espectrales más apropiadas para el análisis espectral de la vegetación, se aplicó la técnica análisis de matrices de correlación, que permite evidenciar la elevada correlación existente entre bandas consecutivas en una HSI, principalmente en cada región espectral, como se observa en la figura 45, los bloques con alta correlación ($r \geq 0,95$) corresponden al VIS, NIR y SWIR. Las figuras 45.a y 45.d muestran una baja correlación entre las regiones del VIS y el NIR para las áreas de estudio de PW y MG, debido a la presencia de superficies de la escena que no corresponden a cultivos o vegetación, tales como áreas urbanas, carreteras, ríos y suelo desnudo para el caso de MG.

Para cada región espectral del VIS, NIR y el borde rojo se seleccionaron las bandas o longitudes de onda que presentaban menor correlación dentro de su región espectral ($r < 0,95$). Adicionalmente, se tomó como criterio de selección las longitudes de onda con mayores y menores picos de reflectancia presente en las firmas espectrales resultantes de este estudio. La tabla 3 muestra la selección de 10 longitudes de onda.

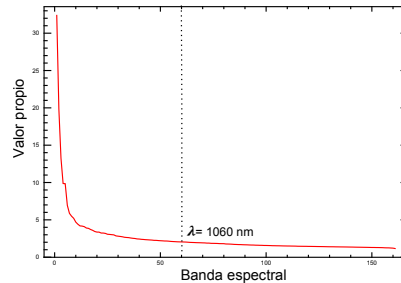
Tabla 3: Longitudes de onda seleccionadas por análisis de las matrices de correlación.

λ	VIS			Borde Rojo		NIR				
	Azul	Verde	Rojo	BR 1	BR 2	NR 1	NR 2	NR 3	NR 4	NR 5
# Banda	460 nm	560 nm	650 nm	680 nm	720 nm	750 nm	800 nm	850 nm	1050 nm	1060 nm
	1	10	19	22	26	29	34	39	59	61

El criterio para evaluar solo las regiones espectrales del VIS, NIR y el borde rojo se basó en el estudio de los valores propios de las HSI, donde todas coincidieron en el mismo com-

portamiento, las bandas con mayores valores propios poseen mayor información y bandas con valores propios inferiores a 2 son bandas con bajo SNR. La figura 46 muestra tal comportamiento, se puede observar que es posible determinar un umbral y este corresponde aproximadamente a la banda 60 en la longitud de onda $\lambda = 1060$ nm.

Figura 46: Comportamiento de los valores propios frente a cada banda espectral.

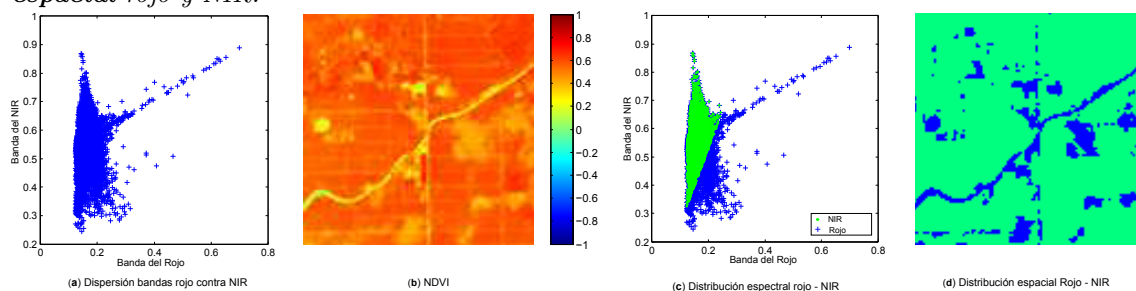


6.2 ANÁLISIS DE LA DISTRIBUCIÓN ESPACIAL Y ESPECTRAL DE LA VEGETACIÓN EN UNA HSI

En la aplicación de HSI para fines agrícolas o en el estudio de la vegetación, es importante determinar la distribución espacial de la vegetación dentro de una escena. Basado en un gráfico de dispersión de píxeles de diferentes bandas, y con la estimación de V_{IS} como el índice de diferenciación normalizada (NDVI) es posible realizar este tipo análisis de manera automática dentro de una HSI. Las anteriores técnicas también permiten determinar cual son las mejores longitudes de onda o bandas espectrales. A través de la estimación de los coeficientes de determinación r^2 , donde se busca encontrar las bandas con mejor comportamiento para estimar V_{IS} [25, 52, 53, 54], que posean una fuerte relación entre el resultado de los V_{IS} y la medición de variables biofísicas o bioquímicas realizadas tradicionalmente *in situ* con equipos, tales como: medidores de contenido de clorofila, índice de área foliar, contenido de agua, nutrientes, fertilizantes, entre otros.

Para el presente estudio no fue posible obtener medidas de las variables biofísicas o bioquímicas de las áreas de estudio, debido a que no se contaba con este tipo de equipos. Sin embargo, basado en las HSI captadas se analizó para dos áreas de estudio (PW y SA) la distribución espacial y espectral para las longitudes de onda centradas en el rojo ($\lambda = 650$ nm) y el NIR ($\lambda = 800$ nm) de acuerdo a la tabla 3. Adicionalmente, la distribución espacial se validó mediante visita de campo a las cinco áreas de estudio. La figura 47.a muestra la dispersión de los píxeles de la banda del rojo contra la banda del NIR para el área de estudio de SA, la figura 47.b ilustra el resultado del V_{IS} NDVI. El NDVI es el V_{IS} con mayores citaciones y uso en la literatura, fue desarrollado en [55] y se basa en el contraste entre la absorción máxima la longitud de onda del rojo, debido a pigmentos de clorofila y la reflexión máxima en el NIR causada por la estructura celular de la hoja. Los valores del NDVI se estiman entre -1 y 1, valores arriba de 0.4 corresponden a presencia de vegetación [56]. En esta figura se puede observar la presencia de un río y suelo desnudo en contraste a la área con presencia de vegetación.

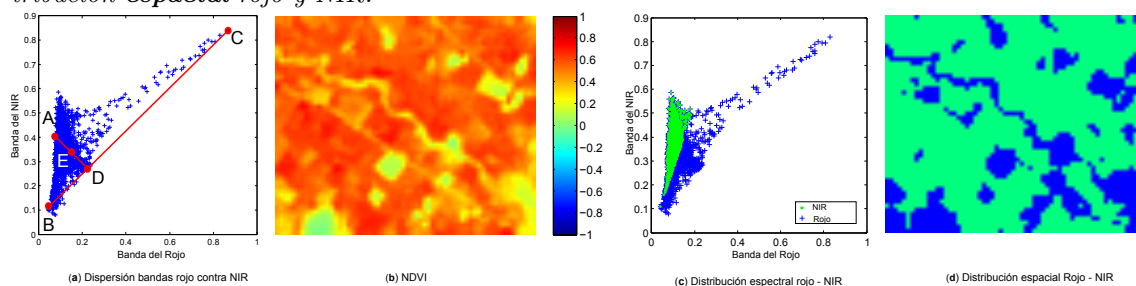
Figura 47: *Dispersión de espacial y espectral de (λ) del NIR y (λ) del rojo en HSI de SA. (a) dispersión de bandas del Rojo y NIR, (b) NDVI, (c) distribución espectral rojo y NIR, (d) distribución espacial rojo y NIR.*



La figura 47.c valida espectralmente el dominio que tiene la banda del NIR sobre el rojo en la escena, debido a alta reflexión del NIR generada por la vegetación, y la 47.d visualiza mediante la binarización del NDVI con valores superiores a 0.4 todas las áreas que corresponden a vegetación (color verde). El 88 % de los píxeles dentro de la HSI de SA corresponden a vegetación.

La figura 48 muestra las gráficas resultado de los análisis espacial y espectral de la HSI del área de estudio MG. La diferencia entre el área de SA y MG es el porcentaje de cobertura vegetal, para MG es de solo 75 %, a pesar que el área de MG al igual que SA está conformada por superficies agrícolas. Sin embargo, en MG hay presencia de cultivos transitorios o de periodo corto (4 a 6 meses) como la caña de azúcar, por lo tanto, la mayor parte de los polígonos azules en la figura 48.d corresponden a suelo desnudo de lotes de caña de azúcar. debido a lotes de caña de azúcar que se acaba de cortar o cosechar, a lotes en preparación para próximas siembras, o siembras tempranas.

Figura 48: *Dispersión de espacial y espectral de (λ) del NIR y (λ) del rojo en HSI de MG. (a) dispersión de bandas del Rojo y NIR, (b) NDVI, (c) distribución espectral rojo y NIR, (d) distribución espacial rojo y NIR.*



El gráfico de dispersión de las bandas del rojo y el NIR poseen una forma triangular, permitiendo no solo analizar la distribución espacial de la cobertura vegetal dentro de una escena. Gracias a su forma triangular es posible caracterizar en un espacio rojo-NIR la distribución de humedad del suelo o vegetación. [57], permitiendo la aplicación de V_{IS} basados en índices de línea de suelo, como el índice de sequía perpendicular (PDI) [58].

La figura 48.a muestra la distribución de características de la vegetación mediante los puntos A, B, C, D, y E así:

- Punto A: Cobertura vegetal total.

- Punto B: Suelo húmedo.
- Punto C: Suelos seco.
- Punto D: Suelo desnudo.
- Punto E: Cobertura vegetal parcial.
- Línea BC: Conocida como línea de suelo. En el espacio espectral, la zona húmeda generalmente se encuentra cerca al punto B y la zona de suelo seco se ubica en el punto C. Por lo tanto, la humedad del suelo disminuye a lo largo de la línea BC, permitiendo analizar la severidad de sequías.
- Línea AD: Representa el cambio de la cubierta vegetal, iniciando en el punto A (cobertura de vegetación total), pasando por el punto E (cobertura de vegetación parcial), terminando en el punto D (suelo desnudo).

6.3 ESTADO DEL ARTE DE LONGITUDES DE ONDA Y BANDAS ESPECTRALES OPTIMAS PARA EL ESTUDIO DE LA VEGETACIÓN

Con el objetivo de validar y analizar los resultados obtenidos en la presente investigación de manera global, se revisó el estado del arte con respecto a longitudes de onda usadas y referenciadas por científicos expertos en el área de SR hiperespectral. Este análisis ayuda a explicar mejor la variabilidad de características biofísico y bioquímicas de la vegetación [25, 59] que pueden ser estudiadas a mayor profundidad en longitudes de onda específicas y para cultivos específicos en trabajos futuros.

La tabla 4 muestra un resumen de las principales longitudes de onda agrupadas en las regiones espectrales del azul, verde, rojo, borde rojo, NIR y SWIR, y relacionadas con las longitudes de onda seleccionadas y descritas en la tabla 3.

Tabla 4: Resumen de longitudes de onda y bandas espectrales según estado del arte.

Región Espectral	Thenkabil et al 2002	Thenkabil et al 2002	Migliani et al 2008	Thenkabil et al 2013	* Presente estudio	Características principales
	λ (nm)	λ (nm)	λ (nm)	λ (nm)	λ (nm)	
Azul				405		Sensible a la pérdida de clorofila y senescencia y efectos de fondo del suelo.
	490	495	436	450	460	
Verde	520			515		Presenta el punto de reflectancia máxima en el VIS. La región del verde está fuertemente relacionada con la clorofila total de la planta y su relación con el nitrógeno.
		555		531		
	550			550	560	
	575		599	570		
Rojo			609			Sensible al contraste entre el cultivo y el suelo. Presenta el punto de máxima absorción de la clorofila. La absorción puede variar de manera significativa debido a los cambios en factores tales como la biomasa, LAI, fondo de suelo, tipos de cultivos, la estructura del dosel, nitrógeno, humedad, y el estrés en las plantas.
			620			
			640			
	660	655	650	650	650	
	675	675	671			
			681	687		
Borde rojo	700	705	742	705	680	Banda sensible a las variaciones temporales en el crecimiento del cultivo y sensible al estrés de la vegetación.
	720	735	752	720	720	
NIR	845	885	772	760	750	Región de reflectancia máxima del espectro en la vegetación entre 400 nm y 2500 nm. Permite clasificar ciertos tipos y / o etapas de crecimiento de la vegetación o cultivos. Existe una fuerte correlación con el contenido de biomasa, clorofila total, y el índice LAI. Entre las longitudes de onda aproximadamente de 1000 nm a 1100 nm. estas bandas son útiles para el cálculo de la humedad de los cultivos.
	905	915	782			
	920		793			
	975		803		800	
			854	855	850	
			874			
		985		970		
		1085	1003			
			1013			
			1023			
			1033			
			1043	1045	1050	
			1053		1060	
			1063			
		1083	1100			
SWIR		1135		1180		Bandas usadas para analizar contenidos de agua, lignina, celulosa, absorción de almidón, proteína, nitrógeno.
		1215				
		1245		1245		
		1285				
		1445		1450		
				1548		
				1620		
		1675		1650		
				1690		
		1725	1739	1760		
		2005	2002	1950		
		2035		2025		
				2133		
				1245		
			2173			
	2235		2205			
	2295		2295			
	2345					

7 VIABILIDAD TÉCNICA Y CIENTÍFICA PARA LA CREACIÓN DE UN PROTOTIPO DE LABORATORIO DE SENSADO REMOTO HIPERESPECTRAL PARA EL NORORIENTE COLOMBIANO

7.1 JUSTIFICACIÓN DE UN PROTOTIPO DE LABORATORIO DE SENSADO REMOTO HIPERESPECTRAL PARA EL NORORIENTE COLOMBIANO

El nororiente colombiano y específicamente Santander es una región con un alto potencial agrícola. Sin embargo, durante las últimas décadas se ha presentado un cambio en la agricultura y se han desplazado los productos de corto plazo, semestrales o anuales como el maíz, el sorgo, la soya y el arroz, cultivos también caracterizados por ser de pequeñas extensiones. Desplazamiento que se a dado especies perennes como el caucho, el cacao, el café, la palma africana, los forestales y otras especies que permiten tener una proyección de largo plazo y que su principal característica es ser cultivados en grandes superficies de tierras. Así mismo, se han presentado cambios en los ciclos climáticos cada vez con mayor intensidad (fenómenos del niño y la niña), que han generando en el nororiente colombiano problemas de sequías, inundaciones y ampliación de la frontera desértica, reduciendo las superficies agrícolas. Estos fenómenos climáticos han desencadenado otros riesgos naturales en la región, tales como: problemas de erosión, deslizamiento, e inundaciones causadas por los principales ríos (Magdalena, Sogamoso, Suarez, Chicamocha, Surata). Adicionalmente, se presentan geoamenazas como las tres fallas sísmicas que cruzan a Bucaramanga y su área Metropolitana, junto con la frecuencia histórica de los temblores generados por la presencia del segundo nido sísmico en el mundo ubicado en la Mesa de los Santos.

De acuerdo a los anteriores escenarios, de manera general el nororiente colombiano requiere tener herramientas, equipos y laboratorios que permitan mejorar las investigaciones y conocimientos de la dinámica de su territorio a nivel espacial y temporal, buscando un conocimiento integral de los fenómenos, su ubicación, magnitud e impacto en la agricultura y en la área de riesgo naturales.

El estudio espacial y temporal, tanto para en el área agrícola y de geoamenazas tradicionalmente se realiza con estudios en campo (*in situ*). Sin embargo, el estudio de grandes superficies requiere de recursos de tiempo y personal profesional en mayor número para el reconocimiento y evaluación total de áreas que contiene o pueden presentar afectaciones. Gracias a la evolución electrónica, óptica y el desarrollo de la técnica de SR, es posible obtener nuevos datos del espectro electromagnético, mejores, con mayor frecuencias temporales (resolución temporal), mayor resolución espacial, y principalmente cubriendo grandes superficies. Esto permite mejorar los estudios y análisis espaciales, temporales y sobre todo espectrales contribuyendo a la generación de nuevo conocimiento de la dinámica agrícola y geoamenazas del nororiente colombiano.

El SR permite captar imágenes en el visible (RGB), multiespectrales, hiperespectrales o térmicas, mediante sensores ópticos ubicado en plataformas aerotransportadas (aviones tripulados o no tripulados), plataformas aeroespaciales (satélites) o incluso en tierra a

través de sensores compactos y ligeros. SR está siendo aplicado a nivel internacional principalmente en agricultura y gestión del riesgo de desastres naturales. Específicamente, en la agricultura. Desde el plataformas aéreas es posible conocer el estado de desarrollo de los cultivos, compararlo con su situación en años anteriores, estimar las cosechas con meses de antelación, clasificar el usos de suelo, detectar cambios o realizar medidas de las superficies correspondientes a los distintos tipos de usos agrícolas. Ahora, en el área de gestión del riesgo de desastres naturales, el SR permite recopilación de información gráfica (fotografía y/o vídeo) antes, durante y después de un evento catastrófico, con el fin de proporcionar información rápida en situaciones de emergencia y desastre. Como parte de los objetivos de la presente investigación se busca presentar la viabilidad técnica y científica para la creación de un prototipo de laboratorio de sensado remoto hiperespectral para el nororiente colombiano. El prototipo debe permitir implementar un sistema de procesamiento y análisis a partir de la captura remota de imágenes en el VIS y NIR con equipos propios o a partir de bancos o bibliotecas de imagenes existentes, para el estudio y análisis espacial, temporal y espectral inicialmente en las áreas agrícolas y de gestión del riesgo de desastres naturales.

7.2 ESTRUCTURA DE UN LABORATORIO DE SENSADO REMOTO

El objetivo principal un laboratorio de SR hiperespectral, es desarrollar investigación básica y aplicada en el campo del SR o teledetección y más concretamente en el área de la espectro-radiometría o espectroscopia. Una de las actividades fundamentales de este tipo de laboratorios consiste en la obtención de medidas espectrales y el desarrollo de protocolos y técnicas de análisis que permitan ser analizadas, por si solas, o en combinación con información procedente de sensores aerotransportados y/o espaciales, para estimar parámetros biofísicos, bioquímicos, e indicadores relevantes en la vegetación. Un laboratorio de SR o centro de SR, centra sus operaciones en la captura, procesamiento y análisis de imágenes captados por diferentes sensores remotos (VIS, multiespectrales, HSI o térmicos), de acuerdo a su área de aplicación (agricultura, recursos naturales, geología, hidrológica, meteorología, etc.) el número de equipos y sensores puede variar. Alrededor del mundo existen grandes laboratorios o centros de SR que cuentan dentro de sus equipos desde satélites, plataformas receptoras (antenas) en tierra para recibir directamente datos y aviones tripulados de uso exclusivo para captura y procesamiento de imágenes de carácter científico. Se destacan las agencias aeroespaciales de la NASA (USA), la agencia espacial de Europa (ESA), el centro de sensado remoto nacional de China (NRSCC), y el centro de recepción, proceso, archivo y distribución de datos de observación de la Tierra (CREPAD) en España [60]. Los anteriores son centros de SR de carácter oficial o iniciativas a nivel de país. Desde la academia existen universidades que poseen sus propios laboratorios de SR, se mencionan como referencia algunas de las principales a nivel mundial como:

- Centro de RS de la universidad de Boston - USA.
- Centro de RS de la universidad de Arizona - USA.
- Laboratorio de RS (RSL) de la universidad de Zurich - Suiza.
- Laboratorio de RS (RSL) de la universidad politécnica de Cataluña - España.
- Laboratorio de RS para la agricultura de la universidad de Illinois - USA.
- Centro espacial y de investigación de sensado remoto (CSRSR) de la universidad central nacional - China

Un laboratorio o centro de SR debe contar con recursos, tales como: equipos de adquisición de imágenes e información espectral, hardware, software, personal técnico y científico, banco de imágenes, área de recepción, almacenamiento y tratamiento de la información, y un área de análisis e interpretación. En la sección siguiente se detallara que equipos básicos se requiere para el montaje inicial de un prototipo de laboratorio de sensado remoto.

7.3 PROTOTIPO DE LABORATORIO DE SR PLANTEADO

De acuerdo a la investigación realizada de los principales laboratorios y centros de SR a nivel mundial. Específicamente, referente a recursos y equipos base para la conformación de un laboratorio de SR, se realizó un estudio previo de empresas fabricantes y soluciones comerciales existentes de equipos para SR.

Dentro de las soluciones comerciales se destacan empresas líderes a nivel mundial como Headwall - USA (fabricante de sensores), Terra Remote Sensing - Chile (prestadora de servicios de SR e HSI), Specim - Finlandia (fabricante de sensores), Analytical Spectral Devices ASD - USA (fabricante de espectroradiómetros), SpecTir - USA (prestadora de servicios de SR), Resonon - USA (fabricante de sensores), y Hypspec - Noruega (fabricante de sensores). Estas compañías han logrado desarrollar sensores HSI de tamaños y pesos muy bajos con dimensiones de solo 11cm de alto por 7cm alto y 7cm de fondo, y un peso aproximado de 680 gramos, permitiendo abrir una nueva tendencia en el uso y aplicación de sensores HSI sobre plataformas aéreas no tripuladas (UAV).

Investigaciones demuestran que UAV's están siendo empleados como plataformas para ubicar diferentes sensores de manera satisfactoria, permitiendo su uso en el área de SR para diferentes aplicaciones. los UAV's son plataformas aéreas de manejo rápido, poseen la ventaja de tener bajos costos comparados con soluciones aéreas tripuladas, aeroespaciales o la compra de imágenes satelitales. Otra ventaja, es la existencia de plataformas UAV multi-sensor, que permite de manera fácil intercambiar de sensor y realizar planes de vuelos y captura de imágenes en diferentes rangos o bandas del espectro electromagnético, permitiendo extraer mayor información a analizar que la obtenida con UAV que son diseñados solo para un sensor, la mayoría de ellos en el rango visible. El uso de sensores VIS, multiespectrales, hiperespectrales, y térmicos sobre una única plataforma UAV permite abrir diferentes aéreas de investigación. La tabla 5 muestra las diferentes áreas de aplicación y actividades que se pueden desarrollar desde un laboratorio de SR.

Tabla 5: *Áreas de aplicación de prototipo de laboratorio de SR.*

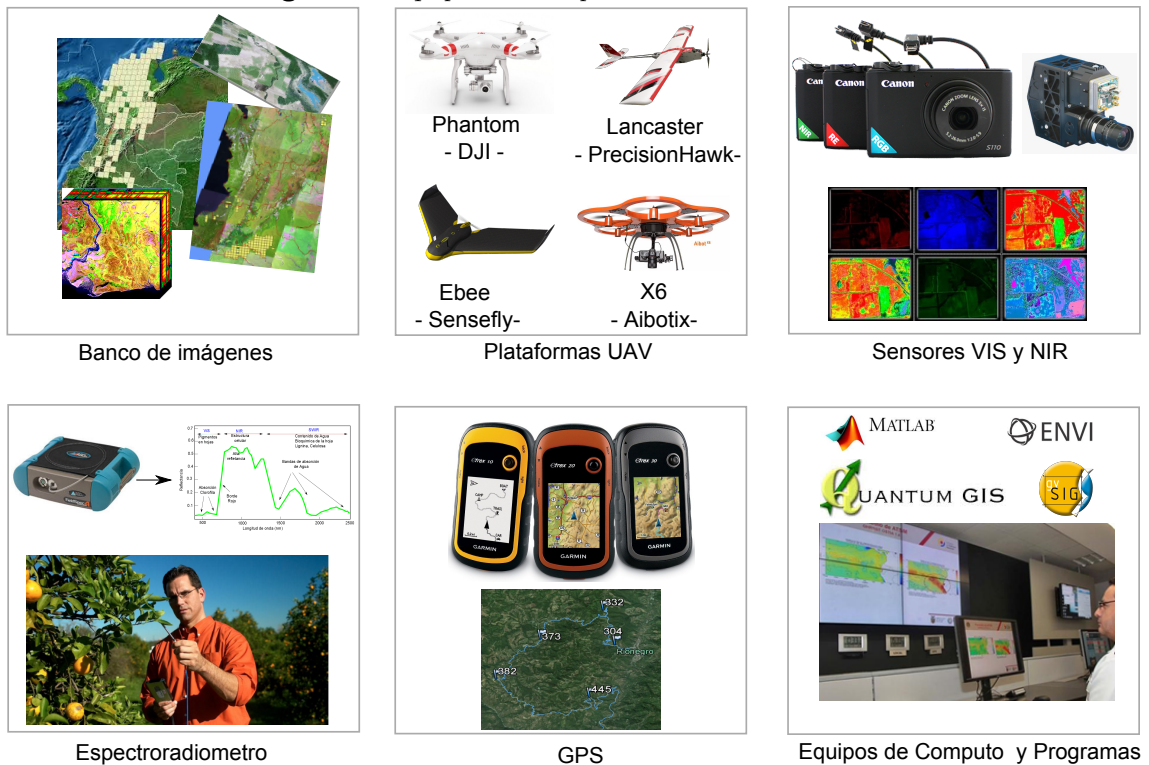
ÁREA	ACTIVIDAD
Agricultura	Identificación de cultivos y sanidad vegetal. Identificación de suelos y sus características. Estudios y monitoreos espaciales, temporales y espectrales. Monitoreo de rendimiento. Humedad del Suelo. Índices de vegetación.
Gestión del Riesgo	Planes de prevención de riesgo. Planes de respuesta a eventos críticos. Monitoreo de zonas de riesgos. Búsqueda y rescate.
Gestión del Agua	Mapeo de tierras con sistemas de riego. Planificación de cuentas. Monitoreo de ríos y quebradas. Contaminación en fuentes hídricas.
Clima y Medio Ambiente	Cartografía de la cubierta terrestre. Monitoreo de reservas forestales. Salud de Bosques. Índice de vegetación y biomasa.
Minería e Hidrocarburos	Monitoreo a derrames de hidrocarburos. Mineralogía. Monitoreo a oleoductos. Evaluación ambiental.

La oferta comercial de empresas fabricantes de UAV ha crecido en los últimos 5 años,

presentándose una gran variedad de opciones. Comercialmente, se destacan plataformas como el Aibot X6 de Aibotix en Alemania, plataforma Lancaster de la empresa Precisionhawk en USA, plataforma eBee de Sensefly en Suiza, y la plataforma Phantom de DJI en Alemania. Este crecimiento ha permitido la fabricación de UAV's con menor tamaño, peso y costo. Sin embargo, sus capacidades respecto a capacidad de carga, autonomía de vuelo y seguridad han aumentado. Permitiendo ser plataformas aptas dentro de un laboratorio de SR de alcance regional, como el prototipo que se plantea en la presente investigación.

La figura 49 muestra los equipos base para el prototipo de laboratorios de SR planteado en la presente investigación, donde se destacan diferentes plataforma aéreas de bajo costo para captura de imágenes aérea, un espectro-radiómetro de punto para uso en campo y laboratorio, un GPS, un sensor VIS, un sensor NIR, y una computadora de gama media con dos pantallas para procesamiento, análisis y visualización de la información captada.

Figura 49: Equipos básicos para un laboratorio de SR.



7.4 SELECCIÓN DE EQUIPOS BÁSICOS PARA EL PROTOTIPO DEL LABORATORIO DE SR PLANTEADO

Finalmente, con base a cotizaciones solicitadas a cada fabricante se realizó un estudio de costos y características técnicas de diferentes equipos, facilitando seleccionar las cantidades y equipos básicos que pueden permitir colocar en marcha el prototipo de laboratorio de SR planteado. La tabla 6 agrupa por diferentes clases de equipos por categorías, así mismo, muestra los precios, características y observaciones generadas en el análisis realizado por cada opción. La columna denominada base PLRS (prototipo de laboratorio de sensado remoto) asigna un asterisco (*) a los equipos seleccionados basado en su funcionalidad y capacidad técnica de realizar la captura de información de imágenes aérea e información

espectral *in situ* y laboratorio.

Tabla 6: Estudio comparativos de costos y características técnicas de equipos de prototipo de laboratorio de SR.

	Empresa	Equipo	Características	Valor	Observaciones	Base PLSR
CAMARAS HIPERESPECTRALES	Headwall	Camara HSI Nano Hyperspec	VIS/NIR de 400 a 1000nm	\$ 315.000.000	Es un sensor HSI pequeño y de bajo peso (0,5 kg). Sin embargo, su costo es muy elevado.	
	Specim	Camara HSI VIS/NIR	VIS/NIR de 400 a 1000nm	\$ 118.000.000	Diseñada para aviones tripulados y costo elevado	
	Resonon	Camara HSI PIKA XC VIS/NIR	VIS/NIR de 400 a 1000nm	\$ 80.000.000	Diseñada para uso en aviones, campo y laboratorio es robusta y grande no apta para UAV	
	HySpex	Camara HySpex VNIR-1024	VIS/NIR de 400 a 1000nm	\$ 252.000.000	Diseñada para uso en aviones y laboratorio, su costo es muy elevado	
PLATAFORMAS UAV	Aibotix	Drone X6, hexacóptero	Carga util 2 Kg, multisensor, autonomia 30 min	\$ 102.000.000	UAV gama alta, permite el uso de diferentes sensores, posee despegue vertical. Sin embargo, su costo es elevado.	
	PrecisionHawk	UAV Lancaster, Ala fija	Carga util 1,2 Kg, multisensor, autonomia 25 min	\$ 60.000.000	Ideal para cubrir grandes áreas en poco tiempo, es multisensor. Sin embargo requiere de un area abierta para el aterrizaje.	
	SenseFly	UAV eBee, Ala fija	Carga util 0,2 Kg, multisensor, autonomia 50 min	\$ 50.000.000	Cubre en un vuelo hasta 12 Km ² es multisensor, Sin embargo soporta sensores muy livianos	
	DJI	Drone Phantom, con 4 elises o motores	Carga util 0,2 Kg, multisensor, autonomia 25 min	\$ 4.760.000	Drone de bajo costo, estable y permite cambiar sensores y hasta montar un espectroradiometro	*
SENSORES ESPECTRALES	PrecisionHawk	Camara en el visible	Resolución de 1,3 cm por píxel	\$ 5.000.000	Modificada para su plataforma	
	PrecisionHawk	Camara multiespectral, 5 bandas	Bandas del azul, verde, rojo, borde rojo y NIR	\$ 18.000.000	Permite la captura de las 5 bandas en una sola toma.	
	SenseFly	Camara en el visible	Capta RGB, Resolución de 3,5 cm por píxel	\$ 800.000	Requiere realizar vuelo adicional si se requiere datos del NIR	
	SenseFly	Camara NIR	Capta las bandas verde (550 nm), roja (625 nm) y NIR (850 nm)	\$ 15.000.000	Requiere realizar vuelo adicional si se requiere datos del NIR, diseñada para su propia plataforma	
	Ocean Optics	Espectroradiometro MINI	VIS/NIR (350 am a 800 nm), peso 68 g.	\$ 6.000.000	Uso en campo, laboratorio y puede ser adaptado en drones	*
Ocean Optics	Espectroradiometro de campo	VIS/NIR (350 am a 1000 nm), peso 190 g.	\$ 40.000.000	Uso en campo y laboratorio.		
OTROS EQUIPOS	Dominio publico	Banco de imágenes satelitales	Imágenes RGB, multiespetrales e HSI	N/A	Aunque las imagene satelitales se pueden comprar a empresas comerciales, la NASA bajo el sitio web http://earthexplorer.usgs.gov/ permite el acceso y descarga de manera libre de imágenes actualizadas de mas de 10 satelites con diferentes sensores	*
		GPS	GPS submetrico portable	\$ 3.500.000	Permite la captura de las 5 bandas en una sola toma.	*
		Equipo de computo	Intel Core i7, Ram min 16 GB, Disco de 2 TR, dos pantallas	\$ 8.000.000	Equipo de procesamiento, analisis de imágenes y datos espectrales	*
	Libre	Programas de procesamiento de imagenes e informacion geografica		N/A	Existen programas o SIG de codigo abierto, que permiten realizar tareas profesionales	*

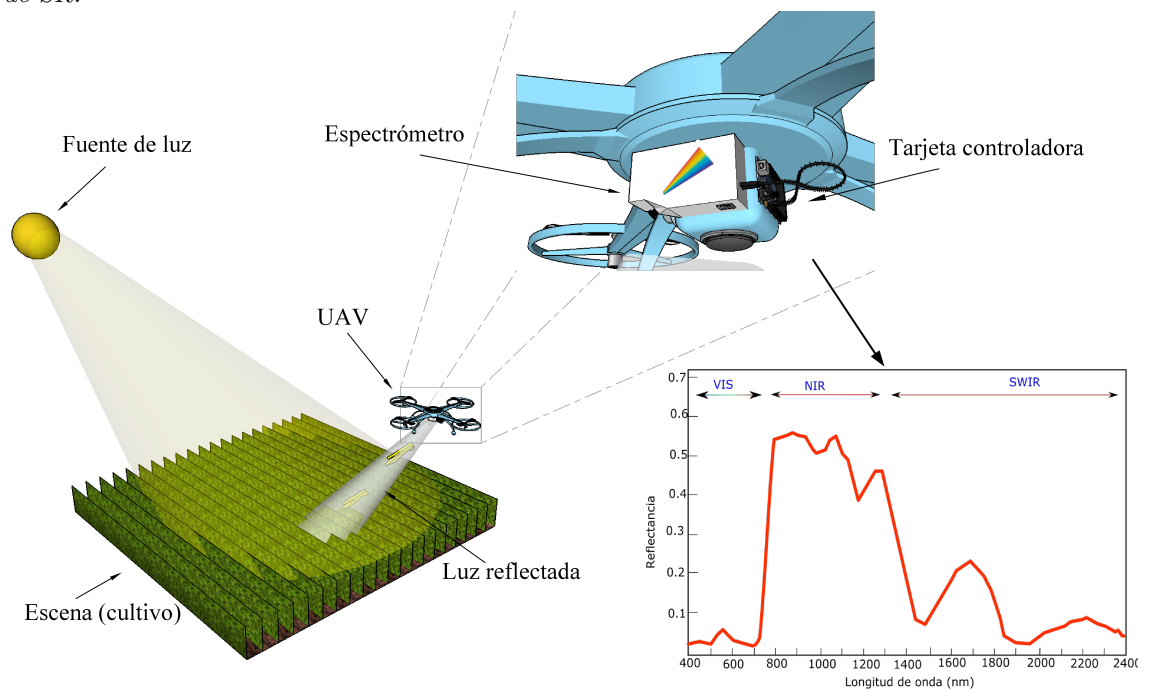
PLRS: Prototipo de laboratorio de SR, *: Equipos básicos seleccionados para el PLSR.

En el mercado actual, una cámara hiperespectral (Headwall, Specim, Hypex, Resonon) tienen un costo promedio de 150 millones de pesos, y un equipo UAV robusto (Precision-Hawk, Aibotix) con capacidad de carga para soportar sensores, cuesta alrededor de 70 millones de pesos; además, el software para el procesamiento y análisis de datos tiene un costo cercano a los 12 millones de pesos. En consecuencia, una solución completa, para la captura aérea y procesamiento de datos espectrales supera los 230 millones de pesos.

Como objetivo de la presente investigación, se estudia la viabilidad de proponer el desa-

rollo de un prototipo laboratorio de SR basado en un sistema de espectroscopia aérea y en tierra, y captura de imágenes aéreas utilizando equipos de bajo costo. Este sistema está orientado hacia el estudio de cultivos agrícolas, mediante la fusión de técnicas de espectroscopia, la estimación de índices de vegetación espectral y sensado remoto aéreo a través de plataformas UAV de bajo costo. Dado que el prototipo propuesto está orientado al análisis en cultivos, se plantea capturar información espectral mediante un espectrómetro en la región visible y infrarrojo cercano (350-800 nm) ubicado en una plataforma UAV, conectado a una tarjeta electrónica programable, que puede ser controlada a través de un teléfono inteligente, tableta u otro equipo de cómputo en tierra, mediante una señal wifi emitida por la antena ubicada en el UAV. Tanto el espectroradiómetro como la tarjeta controladora pueden ser programados para capturar y analizar la información espectral. Así mismo, el sistema puede captar imágenes aéreas en las bandas del VIS (RGB) e información espectral en tierra. Con las imágenes y datos espectrales captados en campo son procesados en el equipo de cómputo y junto al banco de imágenes disponibles de otros sensores es posible realizar análisis y estudios en el prototipo de laboratorio de SR propuesto. La figura 50 muestra un esquema del sistema de bajo costo propuesto el prototipo de laboratorio de SR planteado.

Figura 50: Componentes del sistema de SR de bajo costo propuesto para el prototipo de laboratorio de SR.



8 CONCLUSIONES, CONTRIBUCIONES Y TRABAJOS FUTUROS

8.1 CONCLUSIONES

- Es posible obtener un conjunto de datos de imágenes hiperespectrales reales sobre el territorio colombiano para introducción, apropiación, e investigación del uso y aplicación de HSI que se ha venido realizando a nivel internacional, facilitando el proceso de transferencia tecnológica. Permitiendo desarrollar y explotar aún más esta temática en nuevas investigaciones, procesos, aplicaciones, y áreas a nivel nacional.
- Las HSI permiten estudiar y analizar el comportamiento espectral de la vegetación. Específicamente, es posible su aplicación, uso, e investigación en cultivos agrícolas colombianos bajo variables y factores propios de la agricultura y geografía colombiana.
- Una selección de algoritmos de detección de objetivos en HSI fueron elegidos, los cuales abordan el problema de detección de diferentes maneras. El algoritmo de detección SAM que no requiere información estadística de los datos, fue elegido junto con otros dos diferentes grupos de algoritmos que abordan el problema de modelar la variabilidad espectral, principalmente del fondo de dos maneras diferentes. En el primer grupo de algoritmos, que modelan el fondo de manera estadística, se encuentran ACE, CEM y MF, el segundo grupo, que describe el fondo de manera geométrica, se encuentra el algoritmo OSP.
- Datos sintéticos y firmas espectrales implantadas permiten evaluar el desempeño de los algoritmos de detección de objetivos en HSI a diferentes niveles de ruido ($SNR = 10$ dB, 15 dB, 20 dB, 25 dB, y 30 dB) y a diferentes fracciones de firma espectral ($f = 10\%$, 20%, 30%, 40%, y 100%). Las evaluaciones muestran que el desempeño de los algoritmos se afectan en la medida que el nivel de ruido aumenta. Sin embargo, su desempeño se estabiliza a niveles de ruido de 25 dB para objetivos de tipo sub-píxel y de 20 dB para objetivos de píxeles puros. El algoritmo con mejor desempeño; implementado con datos sintéticos variando los niveles de ruido y fracciones de abundancia de la firma; es el ACE. Sin embargo, el algoritmo SAM presenta un comportamiento minimamente variante a partir de 15 dB hacia arriba, esto debido a que SAM no es sensible a la variabilidad espectral generada por ruido (sombras o variaciones de luminosidad) debido a que SAM depende solo de la forma espectral del objetivo y es independiente de las diferencias de magnitud entre el objetivo y los píxeles de fondo.
- Los cultivos con comportamiento homogéneo (similitud en alturas, estado del cultivo, follaje, etc) presentan menor variabilidad espectral, por ende generan un mejor desempeño en todos los algoritmos.
- Existen limitaciones cuando se intenta evaluar con mayor detalle el desempeño de los algoritmos a través de la estimación de las curvas ROC debido a que el número de píxeles en un hipercubo tradicionalmente es del orden de 10^5 , limitando la estimación

de la P_{FA} que es del orden de menos de 10^{-4} por cubo. Adicionalmente, el número de píxeles objetivo en una escena es muy pequeño. Por lo tanto, es necesario aplicar diferentes métricas de evaluación de desempeño basadas en el establecimiento de umbrales y en las matrices de confusión.

- ACE presenta un mejor desempeño para diferentes HSI y objetivos de tipo agrícola, tanto en imágenes sintéticas como reales, seguido por los algoritmos CEM y MF que presentan un comportamiento aceptable con promedios de $P_D = 80\%$, en contraste con los algoritmos OSP y SAM que son capaces de detectar el objetivo, sin embargo, el número de FA son altas desmejorando sus desempeños, estos resultados se concluyen de todas las evaluaciones realizadas en la presente investigación.
- Las HSI pueden estar conformadas por cientos de bandas espectrales, entregando un gran volumen de información del comportamiento espectral de los cultivos agrícolas. Sin embargo, es posible la estimación de índices de vegetación a partir de solo decenas de bandas espectrales en longitudes de onda específicas.
- El desarrollo de laboratorios de SR es viable a partir de equipos y sistemas de bajo costo, permitiendo realizar desde la captura, pre-procesamiento, análisis, visualización y archivo de información espacial, espectral y temporal a nivel regional.

8.2 CONTRIBUCIONES

- Un artículo en revista indexada u homologada A2, titulado “Applications and challenges of hyperspectral remote sensing in the colombian geology”. El artículo se encuentra publicado en la revista facultad de ingeniería de la Universidad Pedagógica y Tecnológica de Colombia - UPTC. Vol. 24, No. 40, pp. 17-29, Septiembre 2015.
- Presentación de avances y resultados en una conferencia científica internacional mediante una ponencia oral realizada en el III Congreso de percepción remota y sistemas de información espacial, adelantado dentro de la VII convención de agrimensura 2015 en la Habana, Cuba, titulada “Hyperspectral vegetation indices on agricultural crop in Colombia using remote sensing Of Hyperion imagery”, septiembre 2015.
- Presentación de la temática de SR hiperespectral aplicado a la agricultura mediante presentación de un poster en el primer simposio internacional en biotecnología y agroindustria - 2014 en Bucaramanga, Colombia, titulado “Sensado remoto hiperespectral aplicado a la detección y clasificación de tierras agrícolas en Santander, Colombia, a través de imágenes captadas por el sensor satelital Hyperion”. diciembre 2014.
- Un artículo en desarrollo para revista indexada u homologada A2, titulado “A comparative study of target detection algorithms in hyperspectral imagery applied agricultural crops in colombia”. El artículo se encuentra en desarrollo para ser presentado la revista Tecnura de la Universidad Distrital Francisco José de Caldas en el primer semestre del 2016.
- Participación en la creación, desarrollo y presentación de cinco propuestas de investigación en temáticas relacionadas con SR, tres en convocatorias externas (Colciencias) y dos en internas (VIE-UIS).

8.3 TRABAJO FUTURO

Las HSI usadas en SR presentan interesantes áreas de investigación futura, tales como: el desmezclado espectral (del inglés *unmixing*) y el desarrollo de nuevos algoritmos de clasificación y detección sobre imágenes captadas en el dominio comprimido (del inglés *compressive sensing*). A nivel regional, se pueden desarrollar y profundizar investigaciones y aplicaciones durante todo el proceso de crecimiento y desarrollo de cultivos específicos.

REFERENCIAS BIBLIOGRÁFICAS

- [1] John A. Richards. *Remote Sensing Digital Image Analysis*. Springer-Verlag Berlin Heidelberg, 5 edition, 2013.
- [2] USGS. GEOLOGICAL SURVEY. LandSat 8, 2014.
- [3] EO-1 Earth Observing 1. Earth Observing 1 (EO-1) / Sensor Hyperion, 2011.
- [4] Inc Digital Globe. Digital Globe, 2014.
- [5] Wing-Kin Ma, Jose M Bioucas-Dias, Jocelyn Chanussot, and Paul Gader. Signal and image processing in hyperspectral remote sensing [from the guest editors]. *Signal Processing Magazine, IEEE*, 31(1):22–23, 2014.
- [6] Minghua Zhang, Zhihao Qin, Xue Liu, and Susan L Ustin. Detection of stress in tomatoes induced by late blight disease in california, usa, using hyperspectral remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, 4(4):295–310, 2003.
- [7] Sayed M Arafat, Mohamed A Aboelghar, and Eslam F Ahmed. Crop discrimination using field hyper spectral remotely sensed data. 2013.
- [8] Lênio Soares Galvão, Antônio Roberto Formaggio, and Daniela Arnold Tisot. Discrimination of sugarcane varieties in southeastern brazil with eo-1 hyperion data. *Remote Sensing of Environment*, 94(4):523–534, 2005.
- [9] Camille CD Lelong, Jean-Michel Roger, Simon Brégand, Fabrice Dubertret, Mathieu Lanore, Nurul A Sitorus, Doni A Raharjo, and Jean-Pierre Caliman. Evaluation of oil-palm fungal disease infestation with canopy hyperspectral reflectance data. *Sensors*, 10(1):734–747, 2010.
- [10] Dimitris Manolakis, David Marden, and Gary A Shaw. Hyperspectral image processing for automatic target detection applications. *Lincoln Laboratory Journal*, 14(1):79–116, 2003.
- [11] Nasser M Nasrabadi. Hyperspectral target detection: An overview of current and future challenges. *Signal Processing Magazine, IEEE*, 31(1):34–44, 2014.
- [12] Chein-I Chang, Jih-Ming Liu, Bin-Chang Chieu, Hsuan Ren, Chuin-Mu Wang, Chien-Shun Lo, Pau-Choo Chung, Ching-Wen Yang, and Dye-Jyun Ma. Generalized constrained energy minimization approach to subpixel target detection for multispectral imagery. *Optical Engineering*, 39(5):1275–1281, 2000.
- [13] D. Manolakis, E. Truslow, M. Pieper, T. Cooley, and M. Brueggeman. Detection algorithms in hyperspectral imaging systems: An overview of practical algorithms. *Signal Processing Magazine, IEEE*, 31(1):24–33, Jan 2014.
- [14] Qian Du, Hsuan Ren, and Chein-I Chang. A comparative study for orthogonal subspace projection and constrained energy minimization. Technical report, DTIC Document, 2003.

- [15] Chein-I Chang. Orthogonal subspace projection (osp) revisited: a comprehensive study and analysis. *Geoscience and Remote Sensing, IEEE Transactions on*, 43(3):502–518, 2005.
- [16] FA Kruse, AB Lefkoff, JW Boardman, KB Heidebrecht, AT Shapiro, PJ Barloon, and AFH Goetz. The spectral image processing system (sips)—interactive visualization and analysis of imaging spectrometer data. *Remote sensing of environment*, 44(2):145–163, 1993.
- [17] Roberta H Yuhas, Alexander FH Goetz, and Joe W Boardman. Discrimination among semi-arid landscape endmembers using the spectral angle mapper (sam) algorithm. In *Summaries of the third annual JPL airborne geoscience workshop*, volume 1, pages 147–149. Pasadena, CA: JPL Publication, 1992.
- [18] Heesung Kwon and Nasser M Nasrabadi. A comparative analysis of kernel subspace target detectors for hyperspectral imagery. *EURASIP Journal on Applied Signal Processing*, 2007(1):193–193, 2007.
- [19] Yuval Cohen, Yitzhak August, Dan G Blumberg, and Stanley R Rotman. Evaluating subpixel target detection algorithms in hyperspectral imagery. *Journal of Electrical and Computer Engineering*, 2012:2, 2012.
- [20] D Manolakis, Ronald Lockwood, Thomas Cooley, and John Jacobson. Is there a best hyperspectral detection algorithm? In *SPIE Defense, Security, and Sensing*, pages 733402–733402. International Society for Optics and Photonics, 2009.
- [21] Adam C Watts, Vincent G Ambrosia, and Everett A Hinkley. Unmanned aircraft systems in remote sensing and scientific research: Classification and considerations of use. *Remote Sensing*, 4(6):1671–1692, 2012.
- [22] Robert A Schowengerdt. *Remote sensing: models and methods for image processing*. Academic press, 2006.
- [23] Gary A Shaw and Hsiao-hua K Burke. Spectral imaging for remote sensing. *Lincoln Laboratory Journal*, 14(1):3–28, 2003.
- [24] James B Campbell. *Introduction to remote sensing*. CRC Press, 2002.
- [25] Prasad S Thenkabail, Isabella Mariotto, Murali Krishna Gumma, Elizabeth M Middleton, David R Landis, and K Fred Huemmrich. Selection of hyperspectral narrowbands (hnbs) and composition of hyperspectral twoband vegetation indices (hvis) for biophysical characterization and discrimination of crop types using field reflectance and hyperion/eo-1 data. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 6(2):427–439, 2013.
- [26] Prasad S Thenkabail, Ronald B Smith, and Eddy De Pauw. Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote sensing of Environment*, 71(2):158–182, 2000.
- [27] PS Thenkabail, AD Ward, and JG Lyon. Landsat-5 thematic mapper models of soybean and corn crop characteristics. *Remote sensing*, 15(1):49–61, 1994.
- [28] Elhadi Adam, Onesimo Mutanga, and Denis Rugege. Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review. *Wetlands Ecology and Management*, 18(3):281–296, 2010.

- [29] Center for Earth Resource Observation EROS and Science. Hiperion Level 1G (L1GST) Product Output Files Data Format Control Book (DFCB). Technical report, Sioux Falls, South Dakota, 2006.
- [30] Michael K Griffin, Su M Hsu, Hh K Burke, SM Orloff, and CA Upham. Examples of eo-1 hiperion data analysis. Technical report, DTIC Document, 2005.
- [31] NASA-Stuart Frye. GeoBPMS, 2015.
- [32] John R Jensen and Kalmesh Lulla. Introductory digital image processing: a remote sensing perspective. 1987.
- [33] IGAC. Informe 2012-2013 Instituto Geografico Agustin Codazzi, IGAC. Technical report, Bogotá, Colombia, 2013.
- [34] AngloGold Ashanti. Anglogold Ashanti Colombia, 2015.
- [35] Ariolfo Camacho Velasco, César Augusto Vargas García, Fernando Antonio Rojas Morales, Sergio Fernando Castillo Castelblanco, and Henry Arguello Fuentes. Aplicaciones y retos del sensado remoto hiperespectral en la geología colombiana. *Facultad de Ingeniería*, 24(40):17–29, 2015.
- [36] ANH. ANH, Agencia Nacional de Hidrocarburos, 2015.
- [37] Ángel López, Yoseline Betzaida, et al. *Metodología para identificar cultivos de coca mediante análisis de parámetros red edge y espectroscopia de imágenes*. PhD thesis, Universidad Nacional de Colombia.
- [38] Edwin Erazo García, César Edwin; Herrera, Fabio Andrés; Mesa. Basic methodology to generate vegetation indices using airborne multispectral imagery applied in sugarcane crops. In *XVI Simposio Internacional SELPER 2014, La Geoinformación al Servicio de la Sociedad*, Medellín, Colombia, 2014.
- [39] Óscar Gonzalo Castillo Romero. Centro virtual de noticias de la educación, Ministerio de Educación Gobierno de Colombia, 2013.
- [40] Andrés Araque Torres, Leonardo ; Jiménez. Characterization of the Spectral Signature from Remote Sensors for the Management of Plant Health in the Oil Palm Crop. *PALMAS*, 30(3):63–79, 2009.
- [41] Alicia Anaya, Jesus A; Chuvieco Salinero , Emilio ; Palacios. Aerial biomass estimation in Colombia based on MODIS images. *Revista de Teledetección*, 30(1):5–22, 2008.
- [42] ONU-SPIDER. Oficina de Apoyo Regional de ONU-SPIDER, 2014.
- [43] Ivan Lizarazo. Vegetation condition assessment using proximal and remote sensors, 2013.
- [44] Ariolfo Camacho Velasco and Henry Arguello Fuentes. Sensado remoto hiperespectral aplicado a la detección y clasificación de tierras agrícolas en Santander, Colombia, a través de imágenes captadas por el sensor satelital Hyperion. In *Primer simposio internacional en biotecnología y agroindustria*, Bucaramanga, Colombia, 2014.
- [45] AM Baldrige, SJ Hook, CI Grove, and G Rivera. The aster spectral library version 2.0. *Remote Sensing of Environment*, 113(4):711–715, 2009.

- [46] Bisun Datt, Tim R McVicar, Tom G Van Niel, David LB Jupp, and Jay S Pearlman. Preprocessing eo-1 hyperion hyperspectral data to support the application of agricultural indexes. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(6):1246–1259, 2003.
- [47] Youcef SMARA, Zohra HAMADACHE, and Seloua CHOUAF. Preprocessing eo-1 hyperion hyperspectral data applied to forests ad vegetation classification. In *ForestSAT2014 Open Conference System*, 2014.
- [48] S Kay. *Fundamentals of statistical signal processing*, prentice hall, 1998.
- [49] Driss Haboudane, John R Miller, Elizabeth Pattey, Pablo J Zarco-Tejada, and Ian B Strachan. Hyperspectral vegetation indices and novel algorithms for predicting green lai of crop canopies: Modeling and validation in the context of precision agriculture. *Remote sensing of environment*, 90(3):337–352, 2004.
- [50] Davoud Ashourloo, Mohammad Reza Mobasheri, and Alfredo Huete. Evaluating the effect of different wheat rust disease symptoms on vegetation indices using hyperspectral measurements. *Remote Sensing*, 6(6):5107–5123, 2014.
- [51] Ariolfo Camacho Velasco, Cesar Augusto Vargas Garcia, and Henry Arguello Fuentes. Hyperspectral vegetation indices on agricultural crop in Colombia using remote sensing Of Hyperion imagery. In *III Congreso de percepción remota y sistemas de información espacial*, La Habana, Cuba, 2015.
- [52] Peng Gong, Ruiliang Pu, Greg S Biging, and Mirta Rosa Larrieu. Estimation of forest leaf area index using vegetation indices derived from hyperion hyperspectral data. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(6):1355–1362, 2003.
- [53] Niels Henrik Broge and E Leblanc. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote sensing of environment*, 76(2):156–172, 2001.
- [54] Ruiliang Pu, Peng Gong, Greg S Biging, and Mirta Rosa Larrieu. Extraction of red edge optical parameters from hyperion data for estimation of forest leaf area index. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(4):916–921, 2003.
- [55] Jhon W Rouse, Robert H Haas, John A Schell, Donald W Deering, and James C Harlan. *Monitoring the vernal advancement and retrogradation (greenwave effect) of natural vegetation*. Texas A & M University, Remote Sensing Center, 1974.
- [56] Chaoyang Wu, Zheng Niu, Quan Tang, and Wenjiang Huang. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *agricultural and forest meteorology*, 148(8):1230–1241, 2008.
- [57] Abduwasit Ghulam, Qiming Qin, and Zhiming Zhan. Designing of the perpendicular drought index. *Environmental Geology*, 52(6):1045–1052, 2007.
- [58] Qiming Qin, Chuan Jin, Ning Zhang, and Xuebin Yang. An two-dimensional spectral space based model for drought monitoring and its re-examination. In *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*, pages 3869–3872. IEEE, 2010.
- [59] Prasad S Thenkabail, John G Lyon, and Alfredo Huete. *Hyperspectral remote sensing of vegetation*. CRC Press, 2011.

- [60] CREPAD. Centro de Recepción, Proceso, Archivo y Distribución de datos de Observación de la Tierra: Programa CREPAD.

BIBLIOGRAFÍA

ADAM, E, Mutanga, O and Rugege, D. Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review. *Wetlands Ecology and Management*, Vol;.18 No 3. p 281-296, 2010.

ANAYA, A. *et al.* Aerial biomass estimation in Colombia based on MODIS images. *Revista de Teledetección*, Vol;. 30. p. 5-22, 2008.

ARAQUE, A and Jiménez L. Characterization of the Spectral Signature from Remote Sensors for the Management of Plant Health in the Oil Palm Crop. *PALMAS*, Vol 30 No 3, p. 63-79, 2009.

ASHOURLOO, D., Reza, M. and Alfredo Huete. Evaluating the effect of different wheat rust disease symptoms on vegetation indices using hyperspectral measurements. *Remote Sensing*, Vol;. 6. .p. 5107-5123, 2014.

CAMACHO, A. y Arguello, H. Sensado remoto hiperespectral aplicado a la detección y clasificación de tierras agrícolas en Santander, Colombia, a través de imágenes captadas por el sensor satelital Hyperion. Primer simposio internacional en biotecnología y agroindustria, Bucaramanga, Colombia, 2014.

CAMACHO, A. *et al.* . Aplicaciones y retos del sensado remoto hiperespectral en la geología colombiana. *Facultad de Ingeniería*, Vol;. 24 No 40. p. 17-29, 2015.

CAMACHO, A., Vargas, C. and Arguello, H. Hyperspectral vegetation indices on agricultural crop in Colombia using remote sensing Of Hyperion imagery. In III Congreso de percepción remota y sistemas de información espacial, La Habana, Cuba, 2015.

BALDRIDGE, A. *et al.* The aster spectral library version 2.0. *Remote Sensing of Environment*, Vol;. 113 No 4. p. 711-715, 2009.

CHANG, C. Orthogonal subspace projection (osp) revisited: a comprehensive study and analysis. *Geoscience and Remote Sensing, IEEE Transactions on*, Vol;.43 No 3, p. 502-518, 2005.

CHANG, C. *et al.* Generalized constrained energy minimization approach to subpixel target detection for multispectral imagery. *Optical Engineering*, Vol;. 39. No 5. p 1275-1281, 2000.

COHEN, Y. *et al.* Evaluating subpixel target detection algorithms in hyperspectral imagery. Journal of Electrical and Computer Engineering, Vol;. 2012 No 2, 2012.

DATT, B. *et al.* Preprocessing eo-1 hyperion hyperspectral data to support the application of agricultural. Geoscience and Remote Sensing, IEEE Transactions, Vol;. 41 No 6 p.1246-1259, 2003.

DU, Q., Ren, H and Chang, C. A comparative study for orthogonal subspace projection and constrained energy minimization. Technical report, DTIC Document,2003.

GONG, P. *et al.* Estimation of forest leaf area index using vegetation indices derived from hyperion hyperspectral data. Geoscience and Remote Sensing, IEEE Transactions on, Vol;. 41 No 6 p.1355-1362, 2003.

HABOUDANE, D. *et al.* Hyperspectral vegetation indices and novel algorithms for predicting green lai of crop canopies: Modeling and validation in the context of precision agriculture. Remote sensing of environment, Vol;. 90 No 3. p. 337-352, 2004.

HENRIK, N. and Leblanc, E. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. Remote sensing of environment, Vol;. 76 No 2. p. 156-172, 2001.

KRUSE, F. *et al.* The spectral image processing system (sips)|interactive visualization and analysis of imaging spectrometer data. Remote sensing of environment, Vol;.44 No 2. p. 145-163, 1993.

KWON, H and Nasrabadi, N. A comparative analysis of kernel subspace target detectors for hyperspectral imagery. EURASIP Journal on Applied Signal Processing, Vol;. 2007. p. 193-199, 2007.

LELONG, CAMILLE. *et al.* Evaluation of oilpalm fungal disease infestation with canopy hyperspectral reflectance data. Sensors, Vol;.10. P.734-747, 2010.

LÓPEZ, A. *et al.* Metodología para identificar cultivos de coca mediante análisis de parámetros red edge y espectroscopia de imágenes. PhD tesis, Universidad Nacional de Colombia.

MANOLAKIS, D. *et al.* Detection algorithms in hyperspectral imaging systems: An overview of practical algorithms. Signal Processing Magazine, IEEE, 31(1):24{33, Jan 2014.

MANOLAKIS, D. *et al.* Is there a best hyperspectral detection algorithms In SPIE Defense, Security, and Sensing, p. 733402-733402. International Society for Optics and Photonics, 2009.

MANOLAKIS, D., MARDEN, D and SHAW, G. Hyperspectral image processing for automatic target detection applications. Lincoln Laboratory Journal, Vol.;14 p.79-116, 2003.

MA, W. *et al.* Signal and image processing in hyperspectral remote sensing. Signal Processing Magazine, IEEE, Vol.; 31 , p.22-23, 2014.

NASRABADI, N. Hyperspectral target detection: An overview of current and future challenges. Signal Processing Magazine, IEEE, Vol.; 31, p.34-44, 2014.

RICHARDS, J.A. (2013) Remote Sensing Digital Image Analysis. Springer-Verlag Berlin Heidelberg, 5 edition.

SHAW, G. and Burke H. Spectral imaging for remote sensing. Lincoln Laboratory Journal, Vol;.14. p 3-28, 2003.

THENKABAIL, P. *et al.* Selection of hyperspectral narrowbands (hnbs) and composition of hyperspectral twoband vegetation indices (hvis) for biophysical characterization and discrimination of crop types using field reflectance and hyperion/eo-1 data. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, Vol;. 6 No 2. p. 427-439, 2013.

THENKABAIL, P., Smith, R and Pauw. E. Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. Remote sensing of Environment, Vol;. 71 No 2. P. 158-182, 2000.

WATTS, A. Ambrosia, V. and Hinkley, E. Unmanned aircraft systems in remote sensing and scientific research: Classification and considerations of use. Remote Sensing, Vol;.4 No 6. P.1671-1692, 2012.

YUHAS, R., Goetz, A and Boardman, J. Discrimination among semi-arid landscape endmembers using the spectral angle mapper (sam) algorithm. In Summaries of the third annual JPL airborne geoscience workshop, Vol;. 1, p. 147-149. 1992.

ZHANG, M., QIN, Z., LIU,X and USTIN, S. Detection of stress in tomatoes induced by late blight disease in california, usa, using hyperspectral remote sensing. International Journal of Applied Earth Observation and Geoinformation, Vol;. 4 No4, p.295-310, 2003.