



ANÁLISIS DE IMÁGENES DIGITALES PARA RECONOCIMIENTO DE
COBERTURA DE SUELOS

TESIS

QUE PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS DE LA COMPUTACIÓN

PRESENTA

JUAN CARLOS MORENO SANCHEZ

TUTOR ACADÉMICO

DR. ADRIAN TRUEBA ESPINOSA

TUTORES ADJUNTOS

DRA. ROSA MARIA VALDOVINOS ROSAS

DR. JESUS GUILLERMO FIGUEROA NAZUNO

TEXCOCO ESTADO DE MÉXICO

Enero de 2013



DEDICATORIAS

Agradezco a toda mi familia en especial a mi esposa y mamá por apoyarme escuchando mis ideas al realizar este trabajo.

Agradezco el apoyo de mi director de tesis así como el apoyo de mis profesores, considero que he aprendido cosas muy valiosas para mi vida laboral y personal.

Agradezco la comprensión de mis amigos así como su constante apoyo para presentar este trabajo.



AGRADECIMIENTOS

Me complace agradecer a través de este trabajo el apoyo brindado por CONACYT, COMECYT, UNIVERSIDAD AUTONOMA DEL ESTADO DE MEXICO, CENTRO UNIVERSITARIO TEXCOCO, al otorgarme una beca de manutención, una beca para escribir la tesis y facilitar mi estancia en España, con lo que fue posible realizar este trabajo y permitirme tomar la decisión de continuar mis estudios de Doctorado, con el fin de apoyar al Estado de México y mi país México con el uso e implementación de nuevas tecnologías en imágenes y tener información oportuna de esta temática.



Contenido

	pag.
Resumen	11
ABSTRAC.....	12
II Introducción	13
III Planteamiento del problema	15
IV Justificación.....	17
V Objetivo	18
5.1 Objetivos Específicos	18
VI Marco Teórico	19
6.1 Imagen digital	19
6.1.1 Imagen de satélite.....	19
6.1.2 Imágenes SAR.....	21
6.1.3 Imágenes LANDSAT	21
6.1.4 IKONOS.....	21
6.1.5 QUICKBIRD	22
6.2 Imágenes híper-espectrales	22
6.2.1 El escáner híper-espectral AHS	23
6.2.2 Características radiométricas y espectrales	24



6.2.3 Geometría de adquisición.....	24
6.3 Caracterización de imágenes	26
6.3.1 Histograma	27
6.3.2 Texturas.....	28
6.3.3 Umbral	32
6.3.4 Ventana	33
6.3.5 Índice Normalizado Diferencial de Vegetación.....	33
6.3.6 Píxeles	34
6.4 Segmentación.....	35
6.4.1 Detección de bordes	35
6.4.2 Umbralización.....	39
6.4.3 Segmentación orientada a regiones.....	41
6.4.4 Segmentación en imágenes híper-espectrales	42
6.5 Clasificación	45
6.5.1 Técnicas no estadísticas	45
6.5.2 Estadísticos clásicos	45
6.5.3 Algoritmos basados en inteligencia artificial.....	47
6.5.4 ANN	48
6.5.5 SVM	51
6.6 Filtros.....	55



6.6.1 Ejemplos de filtros	56
6.7 Trabajos relacionados	56
6.7.1 Clasificación con Máxima verisimilitud	57
6.7.2 Clasificación con índices.....	58
6.7.3 Clasificación con Transformaciones	59
6.7.4 Clasificación con ANN	59
6.7.5 Clasificación por Clustering.....	61
6.7.6 Clasificación con Umbrales	61
6.7.7 Clasificación con SVM	62
VII Metodología	63
7.1 Materiales.....	63
7.2 Caracterización de la imagen experimental	63
7.3 Desarrollo del experimento.....	64
7.3.1 Primera etapa ConjuntosF	64
7.3.2 Segunda etapa Conjuntos4Máquina.....	69
7.3.3 Tercera etapa Máquinas	70
7.3.4 Cuarta etapa DibujaRGB.....	74
VIII Resultados.....	76
8.1 Cuadros de precisión	76
8.2 Graficas	79



8.3 Parámetros SVM Sigmoid	86
8.4 Kappa, sensibilidad y especificidad	87
8.5 Imágenes	88
IX Discusión	93
X Conclusiones	97
XI Bibliografía	98



Índice de Figuras

	pag.
Figura 1 Ejemplo de un vector de pixeles híper-espectral.	23
Figura 2 Esquema de vuelo simultáneo AHS y AMDC.....	26
Figura 3 Ejemplo de umbralización de una imagen con ayuda de su histograma.....	27
Figura 4 Ejemplos de texturas y sus correspondientes texels.....	30
Figura 5 La ecuación de la línea para Hough.	39
Figura 6 Segmentación orientada a regiones.....	41
Figura 7 División y fusión de regiones.....	42
Figura 8 Diagrama de las partes de una SVM.....	51
Figura 9 Ejemplo grafico de un <i>kernel</i>	55
Figura 10 Diagrama de flujo súper-función <i>ConjuntosF</i>	66
Figura 11. Ejemplo de los 3 tipos de parámetros.	67
Figura 12 Ejemplo de imagen filtrada.	69
Figura 13 Diagrama de flujo de súper-función <i>Conjuntos4Maquina</i>	70
Figura 14 Resultado de SVM, asociados a los diferentes <i>kernel</i>	71
Figura 15 Diagrama de flujo de súper-función <i>Maquina</i>	73
Figura 16 Diagrama de flujo súper-función <i>DibujarRgb</i>	74
Figura 17 Diagrama general de la metodología.	75
Figura 18 Grafica de resultados generales de las 3 imágenes	81
Figura 19 Grafica de resultados filtro Cruz.....	82
Figura 20 Grafica de resultados filtro Cuadrado	83



Figura 21 Grafica de resultados filtro Estrella	84
Figura 22 Grafica radial de promedios	85
Figura 23 Grafica radial de cultivos.	86
Figura 24 Ejemplo de imagen IMG_63_1.....	89
Figura 25 Ejemplo IMG_63_1 clase Experimental Plot	89
Figura 26 Resultado IMG_63_1 con filtro Cruz 1 clase Avena.	90
Figura 27 Resultado IMG_63_1 con filtro Cruz 1 clase Cebolla.	90
Figura 28 Resultados IMG_63_1 con filtro Cruz 1 clase Sin Vegetación.	90
Figura 29 Resultado IMG_63_1 con filtro Cruz 1 clase Experimental Plot	91
Figura 30 Resultados IMG_63_1 con filtro Cuadrado 3.	91
Figura 31 Imagen IMG_63_1 filtro Cuadrado 3.	92



Índice de cuadros

	pag.
Cuadro 1 Clasificación de las 8 bandas espectrales.	20
Cuadro 2 Configuración espectral del AHS.	24
Cuadro 3 Kernels más comunes.	55
Cuadro 4. Precisiones de las diferentes clases con las dos pruebas	62
Cuadro 5 Lista de clases en la imagen IMG_63.....	63
Cuadro 6 Algoritmo para el filtro.....	68
Cuadro 7 Precisión de todos los filtros IMG_63_1	76
Cuadro 8 Precisión de todos los filtros IMG_63_2.....	77
Cuadro 9 Precisión de todos los filtros IMG_63_3.....	78
Cuadro 10 Pequeñas variaciones en precisiones	82
Cuadro 11 Parámetros k, θ para Estrella 5.	87
Cuadro 12 Estadísticos de imagen IMG_63_1.....	88



RESUMEN

Conocer los recursos naturales, cuantificar los desastres ecológicos y planear las campañas de reforestación. Necesitan de información actual pronta y precisa, esta información se puede obtener por técnicas manuales (censos, estudio de campo, inferencias estadísticas) o análisis de imágenes aéreas. El análisis de imágenes aéreas ha tenido grandes avances tecnológicos, lo cual permite colocar sensores remotos de mayor capacidad con el fin de obtener imágenes que brinde información detallada del área de estudio. Una tecnología desarrollada es la imagen híper-espectral. Las imágenes híper-espectrales permiten obtener más información de una fotografía área que otro tipo de imágenes, esta información se empleó para obtener la cobertura de suelo en la imagen. Esta cobertura de suelo es generada por 2 partes; la segmentación y la clasificación, para segmentar y clasificar imágenes existen técnicas como (ANN, Likelihood y SVM). En este trabajo, la segmentación se realiza de manera conjunta con la clasificación esto es posible al usar los pixeles de forma individual, con lo que se obtiene la clasificación y segmentación en un solo paso. Para clasificar se emplea una SVM con *kernel Sigmoid* la cual ofrece altas precisiones en imágenes terrestres y aéreas. Antes de emplear el SVM los pixeles son filtrados por 9 diferentes filtros creados a partir del concepto de Markov, buscando solo incluir a los pixeles más próximos al píxel de interés, estos 9 filtros están dado por 3 patrones (cruz, cuadrado y estrella) combinados con 3 magnitudes (1, 3, 5). Debido a que SVM es solo para clasificación binaria se ha introducido la técnica OAA para obtener multi-clasificación. Los resultados experimentales muestran que con la combinación de técnicas ha obtenido una precisión promedio de 90% para clasificar diferentes tipos de cultivos en imágenes híper-espectrales.



ABSTRAC

Knowing the natural resources, quantify ecological disasters and plan reforestation campaigns. Require prompt, accurate and current information, this information may be obtained: manual techniques (surveys, field study and statistical inferences) or analysis of aerial images. The analysis of aerial images has made great technological advancements, allowing remote sensors placed higher capacity in order to obtain images that provide detailed information of the study area. A technology developed is the hyperspectral image. The hyperspectral images allow more details of an aerial picture that other type of images, this information is used for land cover image. Ground cover is generated by two stages: the segmentation and classification, to segment and classify the images have various techniques (ANN, SVM and Likelihood). In this paper, the segmentation is performed in the classification because it is possible to use the pixels individually, so that the classification and segmentation is obtained in one step. SVM with Sigmoid kernel is used to classify because provides high accuracy in terrestrial and aerial imagery. Before using the SVM, the pixels are filtered by 9 different filters created from the concept of Markov, including the closest pixels to the pixel of interest, the combination of 3 patterns (cross, square and star) and 3 magnitudes (1, 3 , 5) obtained 9 filter. Because SVM is only for binary classification has been introduced OAA technique for multiclassification. Experimental results show that the combination of techniques has obtained an average accuracy of 90 % for classifying different types of crops in hyperspectral images.



II INTRODUCCIÓN

La reducción de costos en sensores remotos, han permitido que el área de fotografía aérea aumente la resolución de las imágenes, así como la sensibilidad de la lente para captar el espacio electromagnético y generar nuevas bandas en las imágenes llamadas híper-espectrales debido a la gran cantidad de bandas que manejan del espectro electromagnético. De la misma forma el desarrollo de nuevos procesadores y la reducción de costos en memoria RAM permiten que las fotografías híper-espectrales sean procesadas con diversas técnicas para conocer la cobertura del suelo.

La cobertura de suelo se define por un tipo de vegetación en un tiempo específico (Velázquez, Duran, Larrazábal, López, & Medina, 2010) las técnicas más comunes implican el uso de personas, muestreos de áreas y técnicas para generalización (Inventario Forestal 2010, 2013) esto consume una cantidad de recursos, personas, mapas, fotografías y tiempo para realizar un estudio sobre la cobertura de suelo en una zona específica. Es por ello que si al analizar una fotografía híper-espectral con una técnica de segmentación y clasificación se reduce el uso de recursos y se mejora la precisión debido a que una computadora es capaz de desempeñar trabajos monótonos sin equivocarse por la complejidad de estos.

El proceso de análisis de una fotografía híper-espectral generalmente contiene 2 etapas la segmentación y la clasificación (Gao, 2008). En la segmentación el objetivo es dividir la imagen en elementos más pequeños, siendo estos más característicos e identificables, estos son conocidos como grupos de píxeles o zonas. Una vez que se tiene estos grupos la clasificación se encarga de etiquetarlos para identificar a que clase pertenecen.



La segmentación en este trabajo parte de la idea de formar grupos con un solo píxel como integrante debido a que los pixeles en una imagen híper-espectral son un vector de valores de niveles digitales (ND) y contiene información suficiente para ser únicos e identificables, además que con esta idea todo el trabajo de cobertura de suelo es cubierto por la clasificación.

En el aparatado de clasificación se emplean Maquinas de Soporte Vectorial (SVM), con la técnica de multi-clasificación uno contra todos (OAA), en SVM se emplea un *kernel* de tipo tangente hiperbólica (*Sigmoid*).



III PLANTEAMIENTO DEL PROBLEMA

Cuando el ser humano analiza una imagen digital, tiene la capacidad de hacer un reconocimiento con base a la observación y la experiencia para con ello delimitar zonas que tengan características similares y de interés. Analizar pocas imágenes requiere de poco tiempo, sin embargo cuando el número de imágenes aumenta, a tal grado que rebasa sustancialmente el esfuerzo de las personas, no se cumpliría en tiempo y forma con el análisis, además, de que se corre el riesgo de cometer errores involuntariamente por cansancio, estrés y conocimientos inadecuados.

En este sentido se podría dejar que las computadoras hagan estas tareas ya que no se cansarían por el trabajo continuo que realizan. Pero surge la pregunta ¿una computadora como analizará la imagen?, ¿Con que experiencia contará para hacerlo correctamente?, con estas dos cuestiones fundamentamos nuestro problema; cómo enseñar a una computadora para que analice imágenes, aprenda de cada análisis para brindar información eficiente. Por otro lado, las imágenes que capturan la cobertura de suelo, tienen la virtud de poder registrar, la situación de un estatus de un momento determinado (tiempo), de un área específica de una región determinada (espacio), y con ello saber la situación en la que estaba y si en otro momento se captura el mismo lugar, se podría determinar qué cambios hubo y con ello determinar en un momento dado si hay que ejecutar una acción que permita dar una respuesta a una situación específica, por ejemplo cuando ocurre un desastre natural;

Hay que conocer con anterioridad como se encontraba esta zona en un periodo de tiempo relativamente corto antes del desastre natural, después evaluar la zona del desastre natural y con el análisis de esta información; implementar las acciones para atender adecuadamente la contingencia en un tiempo mínimo. En este ejemplo encontramos los siguientes problemas tener la información de la zona organizada y de fácil acceso, poder obtener la información de la zona afectada en un



tiempo mínimo para tomar decisiones pertinentes, evaluar la diferencia para dar los apoyos económicos lo cual conlleva meses.

Considerando los dos aspectos expuestos, se tiene por un lado computadoras con alta capacidad de procesamiento y en segunda instancia imágenes a las que hay que analizar en tiempo real. Por lo que el problema principal a resolver es la generación de una metodología que permita obtener información en tiempo real de las imágenes digitales obtenidas por satélites o medios aéreos.



IV JUSTIFICACIÓN

Al lograr proponer una metodología de análisis de imágenes híper-espectrales con SVM como clasificador trabajando en un tiempo real se podrían conocer la situación de grandes extensiones al momento de ocurrir un desastre natural y con ello poder atender las necesidades de una manera oportuna y objetiva. Aunado a ello se reducen costos por una buena planeación que se reflejara en una atención oportuna a la población. O bien para evitar otro tipo de accidentes colaterales a los desastres naturales.

Con un algoritmo adecuado de análisis de imágenes se podrían responder las siguientes preguntas; ¿Qué tipo de terreno es? ¿Qué tipo de cultivo posee el terreno? ¿Qué tipo de desastre ha afectado al terreno? Además de reducción de recursos humanos altamente calificados y con todo ello una reducción económica significativa y sin duda lo más importante una atención a la sociedad en la planificación de acciones a tomar por contar con información oportuna.

Permitir que los datos obtenidos brinden información de una forma sencilla y rápida a cualquier persona sin la necesidad de un experto para su interpretación. Además de que la información contenida años atrás conlleve a predecir fenómenos futuros.

Brindar a los gobiernos elementos para realizar una distribuir adecuada los recursos económicos y sociales para brindar soluciones a las comunidades en situaciones de desastres naturales.



V OBJETIVO

Reconocer la cobertura de suelos a partir de imágenes digitales aéreas mediante algoritmos de clasificación.

5.1 Objetivos Específicos

- Identificación de píxeles de forma manual para la clasificación.
- Identificación de filtros para adaptar la imagen.
- Clasificación de píxeles empleando un algoritmo SVM.
- Segmentación de píxeles empleando el algoritmo de clasificación SVM.
- Sistema generador de una imagen resultado para obtener la cobertura de suelo.



VI MARCO TEÓRICO

6.1 Imagen digital

Una imagen digital en la concepción más sencilla está definida como un solo píxel, de esta manera se representan un solo color, representado con un número en un determinado punto del gráfico (Ordoñez, 2005). Estos puntos o píxeles se agrupan en una matriz rectangular llamada *mapa de bits* (bitmap). Una imagen más completa es la RGB: las cuales se componen de 3 bandas de color: rojo (Red), verde (Green) y el azul (Blue). Cada imagen guarda una relación de 256 combinaciones posibles por cada uno de sus colores de esta manera se obtienen un total de 17.6 millones de posibles colores que tiene el gráfico (Sachs, 1999).

6.1.1 Imagen de satélite

En una imagen de satélite, se ve reflejado el espectro electromagnético (radiación electromagnética que absorbe o emite una sustancia), el cual refleja la luz que reflejan los objetos, Estas imágenes se catalogan como imágenes multi-espectrales. Una imagen multi-espectral es un modelo de la realidad construido con base en una respuesta espectral de los elementos que componen una escena y son captadas por detectores sensibles a un rango de longitudes de onda del espectro electromagnético (Salvatierra, 2004).

Las imágenes multi-espectrales operan en diversas bandas no existe una estandarización con respecto de donde inicia una banda y donde termina otra, cada imagen es diferente y depende del fabricante del dispositivo, a continuación se muestran las bandas de una imagen del satélite LANDSAT (ver Cuadro 1).



Cuadro 1 Clasificación de las 8 bandas espectrales.

Bandas	Micrones		Nombre
Banda 1	0.45	0.52	Azul
Banda 2	0.53	0.61	Verde
Banda 3	0.63	0.69	Rojo
Banda 4	0.75	0.90	Infrarrojo cercano
Banda 5	1.55	1.75	Infrarrojo medio
Banda 6	10.4	12.5	Infrarrojo térmico
Banda 7	2.09	2.35	Infrarrojo de corta onda
Banda 8	0.52	0.90	Pancromática

Ejemplo de una segmentación de bandas, hecha por el satélite LANDSAT 7 ETM+, se muestran 8 bandas espectrales en este cuadro.

La combinación de estas bandas producen varias composiciones de color u opciones de procesamiento, los investigadores han estandarizado varios compuestos de color (Garcia & Perez, 1999): RGB (Red Green Blue) 3-2-1 es el "color verdadero" este es llamado así porque es como el ojo humano percibe los objetos, la combinación de bandas 4-3-2 se utiliza para el análisis de los diferentes tipos de vegetación, la combinación de bandas 4-5-2 muestra las variaciones conteniendo la humedad, por ultimo las bandas 7-4-2 es la combinación de bandas utilizado por la NASA para el estudio demográfico, estilos de vida y comportamientos de los 880 millones de personas de las principales economías del mundo lo cual se denomina con el nombre de *mosaico global* (Bratasanu, 2012).

Las imágenes se clasifican considerando los sensores, la resolución y la técnica empleada para obtener cada imagen, es por ello que las imágenes a considerar son las siguientes: SAR, LANDSAT, IKONOS y QUICKBIRD a continuación se explica cada una de estas imagenes.



6.1.2 Imágenes SAR

El radar de apertura sintética SAR es un radar activo que emite la energía en un intervalo de frecuencias de microondas (unos centímetros) en un período pequeño de tiempo, estas imágenes son tomadas a una velocidad de desplazamiento del vehículo espacial (7,5 Km/s aproximadamente). El dispositivo SAR puede ser instalado a bordo de un avión o de un satélite, el preprocesamiento de imágenes obtiene tonos de gris de 100x100 Km con una resolución de unos 25 m ERS (*Earth Resources Technology*) o entre 12,5 m y 200 m por píxel también conocida como RADARSAT (<http://www.tdx.cat>).

6.1.3 Imágenes LANDSAT

La constelación de satélites LANDSAT (LAND=tierra y SAT=satélite), que inicialmente se llamaron ERTS (*Earth Resources Technology Satellites*), fue la primera misión de los Estados Unidos para el monitoreo de los recursos terrestres. La forman 7 satélites, hoy en día sólo están activos el 5 y 7. Estos satélites cuentan con un *Multispectral Scanner Sensor* (MSS), seguido de *Thematic Mapper* (TM) que tiene mayor sensibilidad radiométrica que su antecesor y por último, *Enhanced Thematic Mapper Plus* (ETM+) que entre sus mejoras técnicas destaca una banda espectral (pancromática) con resolución de 15 metros. Una imagen LANDSAT está compuesta por 7 u 8 bandas espectrales, con resoluciones que varían de 15, 30, 60 y 120m (<http://www.inegi.org.mx>, 2013).

6.1.4 IKONOS

El satélite IKONOS es el primer satélite de tipo comercial que posibilita la captación de imágenes con 1 metro de resolución espacial, sus imágenes son de 3 tipos (<http://www.aeroterra.com>, 2013): pancromática 1-metro de resolución, color multi-espectral 4-metros de resolución con bandas RGB e imágenes *pan-sharpened* 1-metro de resolución con bandas RGB e infrarroja.



6.1.5 QUICKBIRD

Estas imágenes tienen un ancho de 16.5 Km, tanto en la banda pancromática (0,61 metros de resolución) como en las multi-espectrales (azul, verde, rojo, infrarrojo cercano, con 2,44 metros de resolución) (<http://www.geosoluciones.cl>, 2013).

6.2 Imágenes híper-espectrales

La observación remota de un determinado objeto está basada en la captación, por parte de un instrumento de medida o sensor, de la radiación electromagnética proveniente de la interacción entre el objeto y la fuente de la radiación (Roger, 1999). La radiación electromagnética recibe varios nombres dependiendo de la longitud de onda que la caracteriza. Para medir la radiación emitida o reflejada por una determinada superficie es preciso cuantificar la cantidad de flujo energético que procede de la misma. Para ello se utiliza la medida de la radiancia, que depende de factores como la percepción de brillo, reflectancia y ángulos de observación (Hernández & Hernández, 2003).

Las técnicas de teledetección espectral se basan en; todos los materiales en el mundo real reflejan, absorben y emiten energía electromagnética de forma distinta en diferentes longitudes de onda (Rojas, 2009).

El valor asociado a cada píxel viene definido por un valor numérico denominado ND. El nombre se justifica por tratarse de un valor numérico, no visual, pero que puede fácilmente traducirse a una intensidad visual o nivel de gris mediante cualquier convertidor digital, analógico. Así, en un esquema puramente espacial, un píxel está constituido por un único valor discreto, mientras que, en un esquema espectral, un píxel consta de un conjunto de valores. Estos valores pueden ser entendidos como vectores N dimensionales, siendo N el número de bandas espectrales en las que el sensor mide (Hsieh & Landgrebe, 1998).

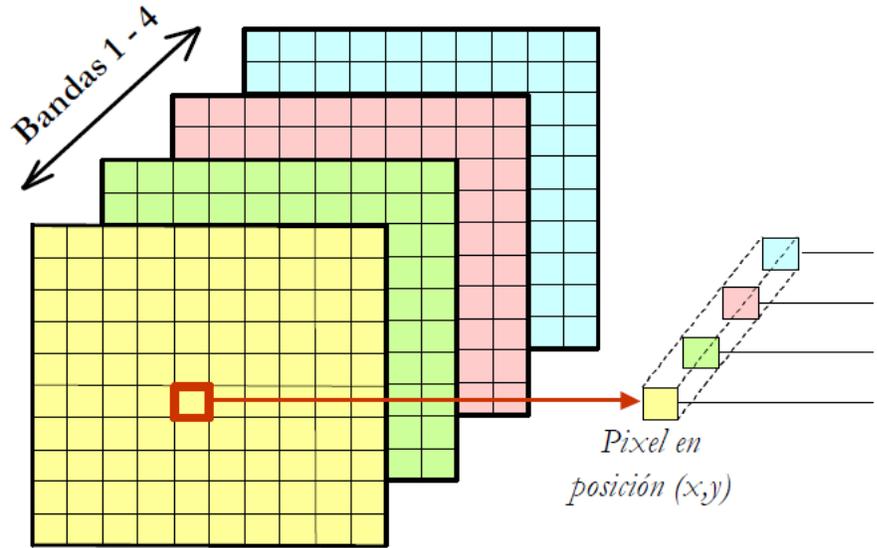


Figura 1 Ejemplo de un vector de píxeles híper-espectral.

En esta figura 1 se detalla el valor de un píxel en una imagen de 4 bandas. El hecho de tener N bandas nos permite conocer si existen características ocultas que no son posibles observar en una imagen RGB, es por ello que las imágenes híper-espectrales son la herramienta principal para desarrollar este trabajo.

Así mismo es importante conocer de una manera rápida el funcionamiento del escáner híper-espectral, así como los parámetros que permiten adaptar para poder obtener las imágenes híper-espectrales.

6.2.1 El escáner híper-espectral AHS

El espectrómetro de imagen Airborne Hyperspectral Scanner (AHS) es un escáner lineal aeroportado tipo *whiskbroom* de 80 bandas para el espectro electromagnético, desarrollado por la casa *Imaging Group ArgonST* bajo los requerimientos especificados por el Instituto Nacional de Técnica Aeroespacial (INTA). Las características técnicas del sensor se definen por sus



resoluciones (espacial, radiométrica, espectral y temporal) así como por el tipo de geometría de adquisición, que a continuación se detalla.

6.2.2 Características radiométricas y espectrales

El AHS registra datos en 12 bits dentro de las ventanas atmosféricas del visible al infrarrojo térmico y divide la radiación incidente en 5 espectrómetros o “puertos” individuales. Dispone de dos cuerpos negros como fuentes de referencia, controlables en temperatura en un rango de -15°C a $+25^{\circ}\text{C}$ con respecto a la temperatura de la cabeza sensora a continuación se muestra el Cuadro 2 la configuración de cada puerto.

Cuadro 2 Configuración espectral del AHS.

Canal	Limites espectrales Inferior Superior μm μm	Rango espectral	Ancho espectral μm	NER Est. @ 25 r.p.s $\text{w. cm}^{-2} \cdot \text{nm}^{-1}$ $\cdot \text{sr}^{-1}$	
1 a 20	0.43 1..030	VIS+NIR	0.030	2.0E-08	PORT 1
21	1.550 1.750	SWIR	0.20	1.0E-08	PORT 2A
22 a 65	1.994 2.540	SWIR	0.013	1.5E-08	PORT 2
66 a 69	3.30 5.40	MIR	.30	1.50	PORT 3
70 a 80	5.10	LWIR	.30	1.00	PORT 4

6.2.3 Geometría de adquisición

La modelización y caracterización de la geometría de adquisición en sensores de observación de la Tierra, es una tarea fundamental para el uso riguroso y efectivo de estos instrumentos en proyectos de ingeniería y de teledetección en los que se requiera un dato georreferenciado. Para algunos sistemas, como por ejemplo los escáneres lineales de adquisición por barrido de los que es un caso el AHS, la orientación directa se presenta como la única alternativa eficaz y operativa para el tratamiento de los datos.



Hay que determinar los parámetros que describen la geometría interna del sensor. Para el AHS, el FOV (*Field of View*) y el IFOV (*Instantaneous Field of View*) junto a la velocidad de barrido especifican su orientación interna y la cobertura y resoluciones espaciales a las que puede registrar un dato. La gama de velocidades de barrido del AHS es 6,25; 8,3; 12,5; 16,7; 25; 40 Hz, que para un FOV de 90° y un IFOV de 2.5 mrad permiten obtener resoluciones que van de 2.4 a 13.7 m en alturas típicas de vuelo (Casal, et al, 2009).

La distancia focal nominal del AHS son 6 pulgadas (15.24 cm). Para estimar la distancia focal efectiva y ajustar el parámetro teórico, se ha empleado la función matemática conocida para sensores *whiskbroom* y la relación de esta con la altura de vuelo y la distancia focal en una cámara matricial, aprovechando la capacidad que tiene el sistema de teledetección aeroportado del INTA para realizar vuelos en tándem con dos sensores (Müller, et al, 2002).

La distancia focal de AHS se puede expresar en función de su ángulo y arco (s) de barrido como (Rejas, et al, 2003):

$$s/2 = f_{AHS} * (FOV_{AHS}/2) \quad (1)$$

De la misma manera la cobertura espacial se deduce como:

$$Y_{AHS}/2 = (H - h) * \tan(FOV_{AHS}/2) \quad (2)$$

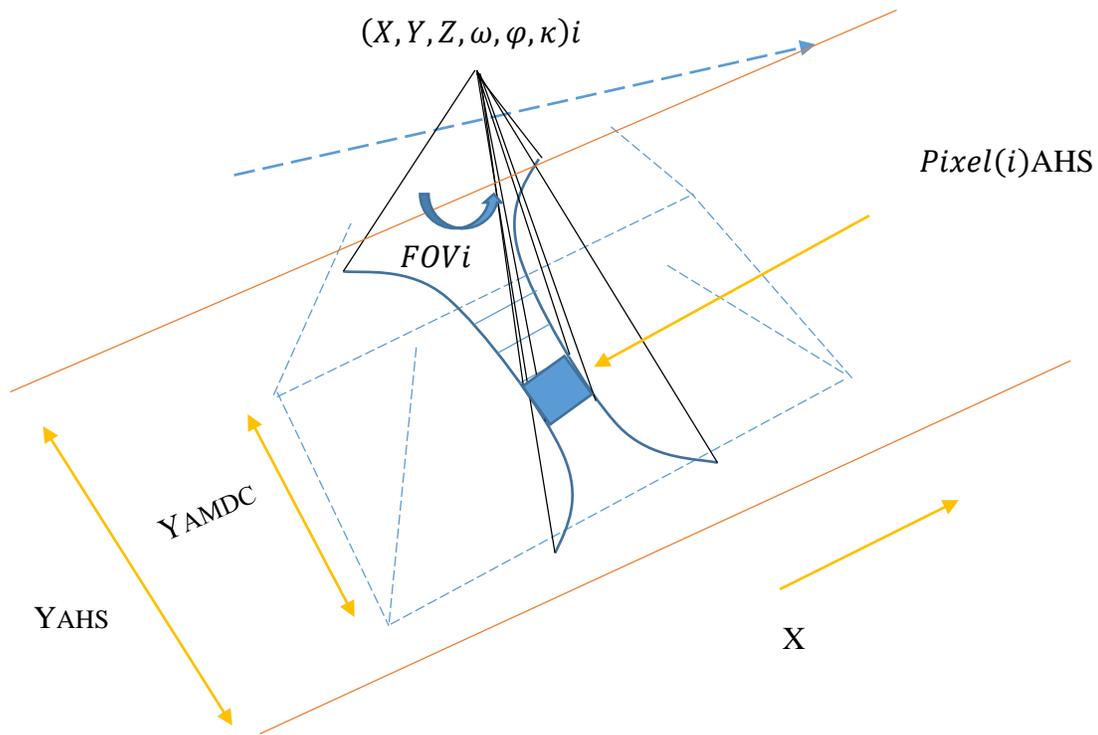


Figura 2 Esquema de vuelo simultáneo AHS y AMDC.

Para una misma altura de vuelo ($H-h$) se ha establecido la relación entre escalas para el AHS y la cámara digital AMDC, sensor cuyo detector es un CCD matricial de tamaño conocido, y consiguientemente se ha determinado la distancia focal efectiva del AHS que hemos adoptado en su orientación interna.

6.3 Caracterización de imágenes

Las imágenes están compuestas por píxeles, cada píxel contiene un vector de valores ND, con estos valores es posible identificar los objetos dentro de la imagen, debido a que los objetos son un conjunto de píxeles es necesario caracterizarlos bajo las siguientes condiciones: histogramas, texturas, umbrales, ventana, índices y los mismos píxeles son representativos siempre que se trate de una imagen híper-espectral.

6.3.1 Histograma

Un histograma representa cualquier número de cosas, ya que su único propósito es resumir gráficamente la distribución en un conjunto de una variable o más variables de datos (Chapelle, Haffner, & Vapnik, 1999). Por lo cual, un histograma de una imagen permite ver como se distribuye un determinado color en un grupo de píxeles. Con el histograma se observa gráficamente la distribución de cada textura, y de esta manera formular un esquema matemático que se adapte a esta distribución.

El histograma resulta una herramienta de gran interés también en tareas de análisis automático de imágenes, como las que se llevan a cabo en visión por computadora. En la Figura 3 se observa un ejemplo de un histograma, en este se observa 2 crestas en su gráfica, la primera se ve reflejado el objeto de interés (llave inglesa), mientras que en la segunda parte está el fondo del objeto. Para lograr esta separación la clave está en encontrar un nivel de gris o umbral, tal que todos los píxeles más oscuros (a su izquierda el histograma) correspondan a la herramienta y los más claros al fondo. Este nivel podría determinarse a partir del análisis del histograma (Atienza, 2011).

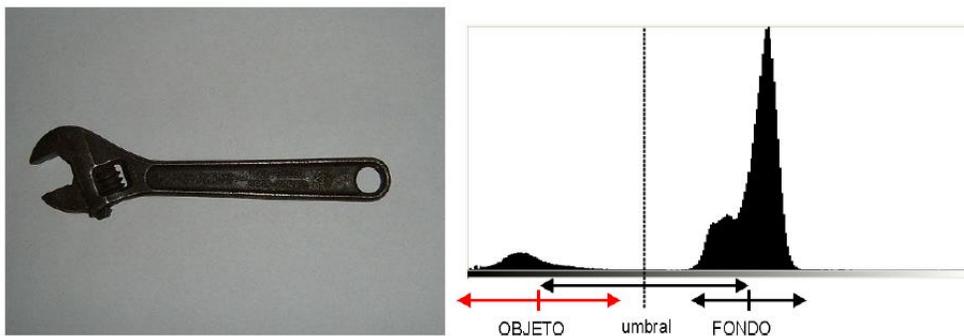


Figura 3 Ejemplo de umbralización de una imagen con ayuda de su histograma.

La figura muestra un típico histograma bimodal, donde hay dos grandes agrupaciones de niveles que aparecen con cierta frecuencia en la imagen, una correspondiente al objeto oscuro y



otra debida al fondo claro. Una vez determinadas estas dos agrupaciones, el nivel de gris equidistante entre ambas, resultaría un nivel de umbral adecuado.

El histograma de una imagen digital con L niveles de gris para el rango $[0, L-1]$ es una función discreta de la forma:

$$h(r_k) = \frac{n_k}{N} \quad (3)$$

Donde:

r_k es el k -ésimo nivel de gris

n_k es el número de píxeles en la imagen con el nivel de gris r_k

N es el número total de píxeles de la imagen

$k = 0, 1, 2, \dots, L-1$ niveles de gris

6.3.2 Texturas

La textura, en un concepto cotidiano es (Pajares & de la Cruz, 2002) : la capa exterior de un objeto que se percibe por el tacto o la vista, el concepto se modificó debido a la imagen digital ya que el realismo en esta ha aumentado de una forma exponencial. En este trabajo la textura se define como: la representación de un patrón espacial básico, cuya estructura es periódica o parcialmente periódica. Las primeras hablan de una distribución regular, como la tela o una pared de ladrillos, pero en este escrito trabajaremos con las segundas; donde el patrón se representa de forma aleatoria (Sonka, 2008).

Las texturas se describen de forma: estadística y sintáctica. La primera se describe con un vector, el cual se representa como un punto en un espacio de características multidimensionales, mientras que la segunda busca encontrar primitivas que constituyen la textura y las reglas de colaboración (La Serna, 2010).



Para caracterizar las texturas (Jiménez, 2008) comenta que existen básicamente tres formas de procesar la imagen y extraer su información: descriptores en frecuencia, descriptores estructurales y descriptores probabilísticos.

6.3.2.1 *Descriptores en Frecuencia*

Los modelos espectrales o de frecuencia consisten en obtener la transformada en frecuencia de la imagen y a partir de ésta, obtener ciertas características. Dichas características son más fáciles de obtener del espectro en coordenadas polares y sirven de base para la clasificación (Jaramillo & Branch, 2008).

Los descriptores en frecuencia se basan principalmente con el análisis del espectro de Fourier mediante la transformada del mismo nombre. El espectro de Fourier está especialmente indicado para describir la direccionalidad de patrones bidimensionales periódicos de una imagen, ya que estos patrones de textura son fácilmente distinguibles como concentraciones altas de energía por el espectro. Las características más importantes vienen dadas por (Miyer, Suarez, & González, 2011):

- Magnitud de “picos” prominentes en frecuencia: proporciona información sobre la direccionalidad de los patrones de textura.
- Localización de los “picos”: proporciona información sobre el periodo espacial fundamental de los patrones.
- Aplicar técnicas estadísticas a partes a-periódicas, una vez separada de la parte periódica mediante un filtro.

6.3.2.2 Descriptores Estructurales

.En cuanto a los descriptores estructurales, su desarrollo se basa en obtener patrones primitivos, texel (es la unidad mínima de una textura aplicada a una superficie), de la textura para generar una descripción de cómo se agrupan éstos para formar la textura en sí. Es decir, es una primitiva visual con propiedades invariantes que ocurre a diferentes posiciones, deformaciones y orientaciones en un área. Estas propiedades invariantes pueden ser: forma, tamaño, nivel de gris o color (Puig, Montseny, & Sobrevilla, 1999).

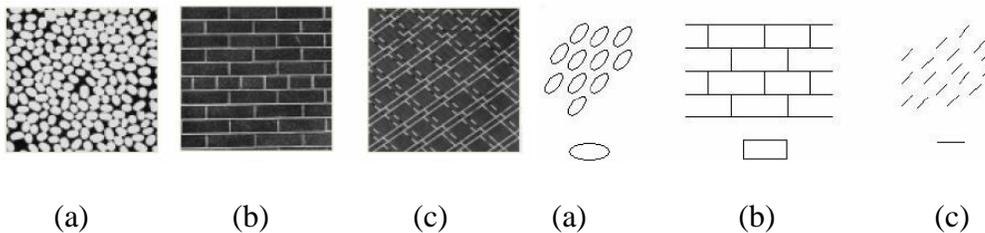


Figura 4 Ejemplos de texturas y sus correspondientes texels.

6.3.2.3 Descriptores Basados en Modelos

Estos métodos se basan en construir un modelo de la imagen, cuyos parámetros estimados sirven para describir y sintetizar una textura, ya que reúnen las características esenciales que caracterizan a la textura. Se consideran como tales las cadenas de Markov y los modelos fractales.

6.3.2.3.1 Cadenas de Markov

Los métodos que utilizan los campos aleatorios de Markov Gaussianos caracterizan la relación estadística entre un píxel y sus vecinos. Es decir, este método se basa en la hipótesis de que la intensidad de un píxel depende de las intensidades de los píxeles vecinos. El modelo estocástico resultante consta de un número de parámetros igual al tamaño de la máscara de la vecindad. Los parámetros se pueden estimar mediante un algoritmo de mínimos cuadrados sobre cada posición de la máscara en la imagen.



La máscara se forma tomando una serie de píxeles vecinos, siendo las más sencillas la de primer orden, formada por los cuatro píxeles directamente conectados a él, y la de segundo orden, formada por los ocho píxeles que lo rodean. Una máscara bastante habitual es la máscara simétrica.

6.3.2.3.2 *Fractales.*

El análisis de texturas basándose en fractales fue introducido por Pentland (Pentland, 1984) y muestra la correlación existente entre la dimensión fractal de una textura y su 'tosquedad'. La descripción fractal de texturas se basa en la determinación de la dimensión fractal.

La propiedad de auto-similaridad conlleva que la dimensión fractal de una imagen es independiente de la escala: un conjunto acotado A se dice que es auto-similar, si está constituido por la unión de N copias no solapadas y distintas de sí mismo, cada una de ellas escaladas con un factor r . La dimensión fractal correspondiente a dicha región viene dada por:

$$D = \frac{\log(Nr)}{\log(r^{-1})} \quad (4)$$

La dimensión fractal D da idea de la rugosidad de una superficie, de modo que cuanto mayor sea D más rugosa será la superficie.

6.3.2.4 *Descriptores Probabilísticos*

Por último, los descriptores probabilísticos son uno de los métodos más simples y se apoyan principalmente en estadísticos del histograma de niveles de gris de una imagen. Estos estadísticos pueden ser de Primer, Segundo o de Tercer orden (Ilea & Whelan, 2011):

Primer orden: las medidas texturales de primer orden son calculadas a partir de los valores de gris originales de la imagen y su frecuencia, como la media, varianza, desviación estándar. En estas medidas no se considera la relación entre los píxeles.



Segundo Orden: son las medidas que consideran la relación de co-ocurrencia entre grupos de dos píxeles de la imagen original y a una distancia dada.

Tercer y más Orden: se consideran las relaciones entre 3 o más píxeles. Si bien su cálculo es teóricamente posible, pero no se implementan porque requieren mucho tiempo de cálculo y su resultado es de difícil interpretación.

6.3.3 Umbral

Los umbrales son técnicas donde se busca obtener rangos que delimitan si un píxel pertenece a una clase o no pertenece a dicha clase, estas técnicas son usadas de forma binaria para separar objetos de interés del fondo de la imagen (Carcedo, 2004), aun así existen formas de usar rangos para clasificar más de 2 objetos.

Se divide la umbralización en 2 partes (Suárez, 2009) la umbralización basada exclusivamente en los niveles de luminosidad se denomina global, mientras para la umbralización local se utiliza información puntual de cada píxel, es decir, propiedades de un píxel que dependen de su localización en la imagen.

Existen dos métodos básicos para abordar el problema de la umbralización global (Nowak, Jurie, & Triggs, 2006):

- Basada en la búsqueda de mínimos.
- Basada en las técnicas de reconocimiento de patrones.

La umbralización basada en la búsqueda de mínimos trata de encontrar niveles de intensidad que umbralicen el histograma de una imagen. Por otro lado, la umbralización basada en las técnicas de reconocimiento de patrones consiste, precisamente, en buscar patrones o formas dentro del histograma.



6.3.4 Ventana

Esta técnica consiste empelando un vecindario de pixeles, el vecindario es delimitado por una figura geométrica, con el fin de aplicar un filtro, obtener texturas o valores estadísticos. Las figuras geométricas poseen un tamaño y buscan sobreponer una malla de estas figuras sobre la imagen, son empleados con el fin de reducir datos o usarlos como una pre segmentación, el objetivo de esta técnica consiste en elegir la figura adecuada así como el tamaño de la misma (Herrada-Mateo & Kober, 2002).

Las figuras más empleadas son simples matrices teniendo medidas de 3×3 , 5×5 y 7×7 , pero esto es variable y depende de la problemática que posee el caso de estudio, para este trabajo se emplean están magnitudes y figuras como cuadrados, estrellas y cruces.

6.3.5 Índice Normalizado Diferencial de Vegetación

En percepción remota existen un conjunto de parámetros numéricos que permiten estimar y evaluar el estado de salud de la vegetación con base a la medición de la radiación que las plantas emiten o reflejan. Estos parámetros permiten la construcción de índices, el índice más conocido y usado es el Índice Normalizado Diferencial de Vegetación (NDVI – Normalized Difference Vegetation Index). El NDVI se utiliza ampliamente para mejorar la discriminación de cubiertas vegetales, medir el vigor de las plantas y la producción de la biomasa (Al-Awadhi, 2007). El peculiar comportamiento radiométrico de la vegetación, relacionado con la actividad fotosintética y la estructura foliar de las plantas, permite determinar si la vegetación se encuentra sana o si tiene algún grado de estrés.

Los valores del NDVI están en función de la energía absorbida o reflejada por las plantas en diversas partes del espectro electromagnético. La respuesta espectral que tiene la vegetación



sana, muestra un claro contraste entre el espectro del visible, especialmente la banda roja, y el Infrarrojo Cercano (IRC).

Mientras la banda visible absorbe los pigmentos de la hoja la mayor parte de la energía que reciben, en el IRC, las paredes de las células de las hojas, que se encuentran llenas de agua, reflejan la mayor cantidad de energía.

El cálculo del NDVI implica el uso de una simple fórmula con dos bandas, una en la región visible y otra en la región IRC. SINAVEF utiliza imágenes del sensor AVHRR, incorporado en la serie de satélites NOAA, para la construcción del índice. Dentro de éstas imágenes, las bandas correspondientes a los canales 1 y 2 son las que se utilizan para obtener el NDVI (Lloyd, 1990).

$$NDVI = \frac{R_{IRC} - R_{R,Vis}}{R_{IRC} + R_{R,Vis}} \quad (5)$$

Donde R_{IRC} es la reflectancia espectral del canal infrarrojo cercano y $R_{R,Vis}$ representa la reflectancia del canal rojo del visible. Esta fórmula indica que existe una relación inversa entre el valor de reflectividad de estas bandas, por lo que es posible su uso para discriminación de cubiertas vegetales. Cada mapa de NDVI es un compuesto que se conforma con imágenes de 10 días, con el fin de poder eliminar en lo posible la cobertura de nubes, por lo que al mes se generan tres mapas con los valores de éste índice.

6.3.6 *Píxeles*

Al igual que las técnicas de histogramas, texturas o umbrales estas técnicas emplean los valores de los píxeles, pero al hablar de imágenes híper-espectrales ya no son un solo valor como en las imágenes de escalas de grises o 3 valores en imágenes RGB, sino que son un vector por lo cual contiene información que lo identifica y caracteriza como un objeto dentro de la imagen



(Rojas, 2009). Es por ello que en lugar de usar promedios se ha optado por usar los píxeles como vectores para poder introducirlos a las técnicas de segmentación o clasificación.

6.4 Segmentación

La mayoría de las imágenes están constituidas por regiones o zonas que tienen características homogéneas. Generalmente estas regiones están conformadas por un único objeto. La segmentación consta en separar la imagen en zonas homogéneas y disjuntas a partir de su contorno (Deng & Manjunath, 2001). Para identificar estos píxeles es necesario conocer la discontinuidad, similaridad y conectividad entre los píxeles, la discontinuidad es solo la propiedad que se tiene para que un píxel no pertenezca a cierta zona mientras que la similaridad es la propiedad que el píxel si pertenezca a la zona, mientras la conectividad se refiere al grupo de píxeles que delimitan la zona a segmentar (Ruíz, Morales, & Hernández, 2011).

6.4.1 Detección de bordes

Los métodos de extracción de bordes de una imagen, se basan en la diferencia que experimenta una característica en dos regiones adyacentes y que indican la existencia de un borde (Alberola-Lopez, Martin-Fernandez, & Ruiz-Alzola, 2004). A la vez los bordes pueden clasificarse por su anchura, ángulo de su pendiente de variación, y las coordenadas de su punto medio. En general, se identifican diferentes modelos de bordes o contornos: línea, tipo escalón, tipo rampa y tipo tejado. Las discontinuidades son detectadas usando derivadas de primer y segundo orden, para el caso de derivadas de primer orden se utiliza el operador gradiente, mientras que en derivadas de segundo orden se utiliza el operador Laplaciano (Palomino & Concha, 2009).

6.4.1.1 Derivada de primer orden, el operador gradiente

Una de las técnicas más comunes para detectar un borde de dos dimensiones (2D), es el uso de derivadas de primer orden mediante el operador gradiente. El gradiente de una imagen es una función $f(x, y)$ que se define como el vector bidimensional presentado en la ecuación 6.

$$G(F(x, y)) = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} dF/dx \\ dF/dy \end{bmatrix} \quad (6)$$

El operador gradiente G alcanza su máximo valor en la dirección en que la variación es máxima, por lo tanto puede definirse como la ecuación 7.

$$G(F(x, y)) = ([G_x(x, y)]^2 + [G_y(x, y)]^2)^{1/2} \quad (7)$$

Es práctica común aproximar el gradiente mediante los valores absolutos, con el fin de conseguir un procesamiento más rápido a costa de un error despreciable, y la expresión se aproxima a la ecuación 8.

$$G(F(x, y)) = [G_x(x, y)] + [G_y(x, y)] \quad (8)$$

A partir de las ecuaciones 7 y 8 se deduce que el cálculo del gradiente de una imagen se basa en la obtención de las derivadas parciales dF/dx y dF/dy en cada posición de pixel. Las derivadas de primer orden permiten encontrar lugares en una imagen en donde la intensidad que cambia rápidamente es mayor en magnitud que un umbral especificado. Existen varios estimadores derivativos como los de Sobel y Prewit.

6.4.1.2 Detección de bordes utilizando derivadas de segundo orden, el operador Laplaciano

El Laplaciano de una función bidimensional $f(x, y)$ es una derivada de segundo orden definida por la ecuación 9.

$$\Delta^2 = \frac{\delta^2}{\delta x^2} + \frac{\delta^2}{\delta y^2} \quad (9)$$

6.4.1.3 Enlazado de bordes y detección de límites.

Si bien hay varias técnicas que detectan discontinuidades de intensidad, y que deberían dar como resultado pixeles que estuvieran en la frontera entre un objeto y su fondo, en la práctica, este conjunto de pixeles rara vez caracterizan una frontera completamente debido al ruido, interrupciones en la frontera debido a la iluminación no uniforme, y otros efectos que introducen discontinuidades de intensidad *¹espurias. Por ello, los algoritmos de detección de bordes están seguidos por una unión y otros procedimientos de detección de frontera diseñados para reunir a los pixeles del borde en un conjunto con significado de fronteras de objetos.

6.4.1.3.1 Procesamiento local

Consiste en analizar un entorno de vecindad (por ejemplo, 3x3, 5x5) sobre todos los puntos (x, y) de una imagen en los que haya sido llevado a cabo un proceso de detección de borde, de forma que todos los puntos que presenten características similares se unen formando una frontera común. Se utilizan dos propiedades para determinar la similitud entre pixeles del borde: a) Magnitud del vector gradiente (valor del umbral para determinar el borde), y b) Dirección del gradiente.

La primera propiedad establece que un píxel de coordenadas (x, y) es similar a otro (x, y) dentro de su entorno de vecindad si la ecuación 10 se cumple.

¹ Espuria: es una relación matemática en la cual dos acontecimientos no tienen conexión lógica, aunque se puede implicar que la tienen debido a un tercer factor no considerado aún (llamado "factor de confusión" o "variable escondida")

$$G[F(x, y)] - G[F(x', y')] \leq T \quad (10)$$

Donde T es un umbral no negativo.

La segunda propiedad, es decir la dirección del gradiente (G), se puede establecer utilizando el ángulo del vector (θ) gradiente que viene dado por la ecuación 11.

$$\theta = \text{tg}^{-1} \left(\frac{G_x}{G_y} \right) \quad (11)$$

Donde θ representa el ángulo respecto al eje de x.

Así, un píxel del borde situado en (x, y) tiene un ángulo similar a (x, y) si se cumple la ecuación 12.

$$|\theta - \theta'| \leq A \quad (12)$$

Donde A es un umbral angular y θ ángulo del vector gradiente.

Un punto de la vecindad predefinida de (x, y) está enlazado con el píxel de (x, y) si se satisfacen los criterios de magnitud y dirección. Se repite para cada posición de la imagen. Un registro debe conservar los puntos enlazados a medida que el centro de la vecindad se desplaza de píxel a píxel.

6.4.1.3.2 *Procesamiento global usando la Transformada de Hough*

Originalmente fue diseñado para detectar líneas y curvas, utilizando ecuaciones analíticas conocidas de bordes de objetos, sin embargo con este método original, no siempre es posible encontrar ecuaciones analíticas que describan bordes. La transformada de Hough generalizada hace posible esto, es decir detectar bordes de objetos aun cuando las expresiones analíticas de bordes no es conocida.

La idea básica del método puede ser presentado desde un problema simple como la detección de líneas rectas, en este caso el objetivo de la transformada de Hough es encontrar puntos alineados que puedan existir en la imagen, es decir, puntos en la imagen que satisfagan la ecuación de la recta para distintos valores de ρ y θ . Para ello, de manera general, se siguen los siguientes pasos (Urrea & Ospina, 2004):

- La ecuación de la línea puede ser definida como (Figura 5-A)
- La transformada de Hough de esa línea es un punto para el plano (r, θ) (Figura 5-B)
- Para cada punto de la imagen se calcula su representación en el plano.
- Para reducir la búsqueda, se cuantifica el espacio de parámetros r y θ , dividiendo este en un espacio de grupos de posibles líneas.
- Los máximos locales en (r, θ) son las líneas que se desean encontrar.

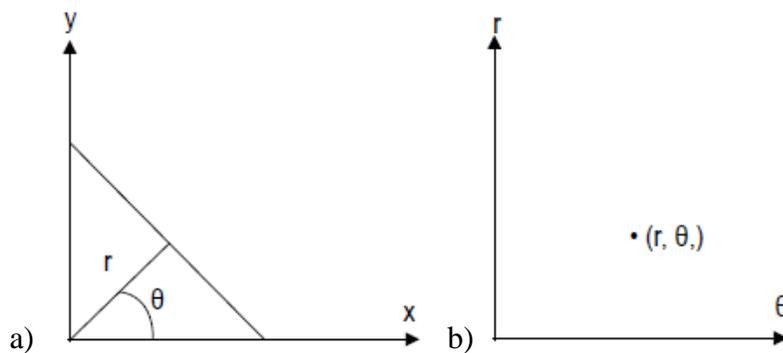


Figura 5 La ecuación de la línea para Hough.

a) Línea recta y los parámetros r y θ , y b) Punto en el plano (r, θ) .

6.4.2 Umbralización

La umbralización es uno de los métodos más importantes de la segmentación de imágenes. Se define el umbral como una función que convierte una imagen con diferentes tonalidades en una imagen en blanco y negro. Si la imagen original es $f(x, y)$, la imagen umbralizada $g(x, y)$ y se fija



un umbral U ($0 < U < 255$), la operación de umbralizado se define como la ecuación 13 (Ballard & Brown, 1982).

$$g(x, y) = 255 \quad \text{si } f(x, y) > \text{Umbral} \quad (13)$$

$$g(x, y) = 0 \quad \text{si } f(x, y) \leq \text{Umbral}$$

Se selecciona un umbral que permita agrupar los pixeles de una imagen pertenecientes a los diversos objetos de la misma imagen diferenciándolos del fondo. De esta manera la segmentación basada en el histograma se basa en la elección de uno o varios umbrales que permiten agrupar los puntos de la imagen en regiones de características similares en función de sus niveles de gris. El umbral de una imagen se define como la ecuación 14.

$$T = T [x, y, p(x, y), f(x, y)] \quad (14)$$

Donde $f(x, y)$ es la intensidad o nivel de gris del punto (x, y) y $p(x, y)$ representa alguna propiedad local medida en un entorno de vecindad de este punto. La imagen que se obtiene al llevar a cabo un procesamiento de umbral se define como la ecuación 15.

$$g(x, y) = 1 \quad \text{si } f(x, y) > T \quad (15)$$

$$\text{si } f(x, y) \leq T$$

De este modo los pixeles marcados con 1 corresponden a objetos, mientras que los pixeles marcados con 0 corresponden al fondo. En la ecuación 14 se ha supuesto que la intensidad de los objetos es mayor que la intensidad del fondo. Este proceso es conocido como *binarización*.

Hay varios tipos de umbrales, los más utilizados (Pal & Pal, 1993): a) Umbral global, cuando T depende solamente de $f(x, y)$, b) Umbral local, si T depende de $f(x, y)$ y de $p(x, y)$ y c) Umbral dinámico, si T depende de las coordenadas x e y , además de $f(x, y)$ y de $p(x, y)$.

6.4.3 Segmentación orientada a regiones

Se utilizan propiedades espaciales de una imagen para segmentarla por regiones, es decir la imagen es dividida en regiones conexas, en donde cada región tiene propiedades distintas que las diferencian unas de otras. En suma, se trata de extraer los objetos de una imagen, las que pueden ser procesadas de manera independiente (Beauchemin & Thomson, 1997).

6.4.3.1 Crecimiento de regiones.

Es una técnica muy utilizada, que consiste en el crecimiento de regiones agrupando píxeles adyacentes que presentan características o propiedades similares. Se parte de un conjunto de puntos “semilla” y se hace crecer las regiones, añadiendo a los puntos semilla los puntos vecinos que tengan propiedades similares, como intensidad, textura, color, etc. Por ejemplo, si la propiedad es intensidad, un criterio utilizado para incluir un píxel en una región, puede ser que la diferencia absoluta entre la intensidad del píxel y la intensidad de la semilla sea menor que un umbral, por ejemplo $T = 8$. Los puntos semillas pueden ser escogidos manualmente o de manera automática. En la Figura 6, se observa de manera gráfica el proceso de esta técnica (Gonzalez & Woods, 2002).



Figura 6 Segmentación orientada a regiones.

6.4.3.2 División y fusión de regiones.

Consiste en dividir inicialmente una imagen en un conjunto de regiones arbitrarias disjuntas, por ejemplo 64 divisiones, después, dependiendo del criterio de segmentación, regiones adyacentes son fusionadas si tienen propiedades similares como nivel de gris similares, o son

divididos si no comparten las mismas propiedades, como variaciones considerables de niveles de gris. Finalmente, la imagen queda segmentada en un conjunto de regiones homogéneas. La Figura 7, muestra una imagen segmentada en 1405 regiones utilizando este método (Rivera & Buri, 2012).

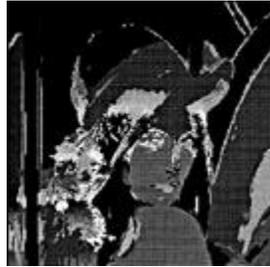


Figura 7 División y fusión de regiones

6.4.4 Segmentación en imágenes hiper-espectrales

En primer lugar, hemos de saber que el término *target* se aplica a objetos con características espectrales diferenciadas. Un conjunto de *targets* estaría por tanto formado por un conjunto de firmas espectrales (puras o no) suficientemente representativas de los diferentes objetos presentes en la imagen y con características espectrales distintas entre sí (Paz, 2009) unos de los algoritmos para realizar la segmentación de imágenes híper-espectrales son: ATGP y RX.

6.4.2.1 Automated Target Generation Process

Automated *Target* Generation Process o ATGP fue desarrollado inicialmente para encontrar firmas espectrales utilizando proyecciones ortogonales. Este algoritmo hace uso de un operador de proyección ortogonal que viene dado por la expresión mostrada en la ecuación 16.

$$P_U^\perp = I - U(U^t U)^{-1} U^t \quad (16)$$

Donde U es una matriz de firmas espectrales, U^t es la traspuesta de dicha matriz e I es la matriz identidad.

El algoritmo ATGP utiliza la proyección ortogonal de la ecuación 16 en cada iteración para encontrar una serie de píxeles o vectores de bandas ortogonales a partir de un píxel inicial que se

le pasa al algoritmo como valor ATGP y que por defecto suele ser el píxel más brillante. Este algoritmo realiza los siguientes pasos (Amorós, et al, 2005):

1. Calcular t_0 , el píxel más brillante de la imagen híper-espectral, utilizando la siguiente expresión que muestra la ecuación 17, donde $F(x, y)$ es el píxel (vector) en las coordenadas (x, y) de la imagen. Como puede comprobarse, el píxel más brillante es aquel que resulta en mayor valor al realizar el producto vectorial entre el vector asociado a dicho píxel y su transpuesto $F(x, y)^T$.

$$T_0 = \arg\{ \max_{(x,y)} F(x, y)^T \cdot F(x, y) \} \quad (17)$$

2. Aplicar un operador de proyección ortogonal que denotamos como P_{U^\perp} , basado en la ecuación 16, con $U = t_0$. Este operador se aplica a todos los píxeles de la imagen hiperespectral.
3. A continuación, el algoritmo encuentra un nuevo *target* denominado t_1 con el mayor valor en el espacio complementario $\langle t_0 \rangle^\perp$, ortogonal a t_0 , de la siguiente forma utilizando la ecuación 18. En otras palabras, el algoritmo busca el píxel con mayor ortogonalidad con respecto a t_0 .

$$T_1 = \arg\{ \max_{(x,y)} [P_{U^\perp}^\perp \cdot F(x, y)]^T \cdot [P_{U^\perp}^\perp \cdot F(x, y)] \} \quad (18)$$

4. El siguiente paso es modificar la matriz U añadiendo el nuevo *target* encontrado, es decir:
$$U = [t_0 \ t_1].$$
5. Seguidamente, el algoritmo encuentra un nuevo *target* denominado t_2 con el mayor valor en el espacio complementario $\langle t_0, t_1 \rangle^\perp$, ortogonal a t_0 y t_1 , utilizando la ecuación 19. Es preciso tener en cuenta que, a diferencia del paso 3 en el que $U = t_0$, en este punto el

proyector ortogonal se basa en una matriz $U = [t_0 \ t_1]$, por lo que el concepto de ortogonalidad es diferente.

$$T_2 = \arg \left\{ \max_{(x,y)} [P_U^\perp \cdot F(x,y)]^T \cdot [P_U^\perp \cdot F(x,y)] \right\} \quad (19)$$

6. El proceso se repite, de forma iterativa, para encontrar un tercer *target* t_3 , un cuarto *target* t_4 , hasta que se satisface una determinada condición de terminación para el algoritmo. La condición de terminación considerada en el presente trabajo es alcanzar un número de *targets* p determinado a priori como parámetro de entrada al algoritmo.

6.4.2.2 Algoritmo de detección de anomalías de Reed y Xiaoli (RX)

RX es un algoritmo de detección de anomalías propuesto por Reed y Xiaoli conocido como RXD, es usado en el proceso de detección de *targets*, este algoritmo encuentra los píxeles que son espectralmente distintos que sus píxeles vecinos. Podríamos resumir su funcionamiento en una serie de pasos (Yu & Reed, 1990):

1. Calcula el vector media m donde cada elemento del vector es la media de los valores de los píxel de cada banda espectral.
2. Calcula la matriz de covarianza espectral K utilizando el resultado del paso anterior.
3. Realiza, utilizando los datos obtenidos en los puntos anteriores, una operación de filtrado que viene dada por la conocida distancia de Mahalanobis a todos los píxeles de la imagen. El cálculo realizado viene dado por ecuación 20.

$$\delta(RXD)_{F(x,y)} = (F(x,y) - m)^T K^{-1} (F(x,y) - m) \quad (20)$$



4. Finalmente, basta con seleccionar los p (siendo p un valor escogido por el usuario) píxeles con mayor valor $\delta(RXD)_{F(x,y)}$ como conjunto de *targets* resultado del algoritmo RX.

6.5 Clasificación

Una vez que se dispone de un conjunto de clases y de sus valores espectrales característicos, se realiza el proceso de asignar un píxel a una determinada clase, estas técnicas han ido evolucionando con el paso del tiempo y se han englobado en tres grupos (Recio, 2010):

1. No estadísticos (mínima distancia, paralelepípedos)
2. Estadísticos clásicos (máxima probabilidad)
3. Algoritmos basados en inteligencia artificial (redes neuronales artificiales, árboles de decisión y SVM)

6.5.1 Técnicas no estadísticas

Estas técnicas buscan ajustar los datos en relación a una tabla de distancias, dicha tabla es complicada de obtener ya que compara a todos los datos con todos los datos, además de que es difícil poder obtener una distancia cuando los datos muestran múltiples valores como en el caso de las imágenes híper-espectrales, un ejemplo de estas técnicas es K-means (Pascual, Pla, & Sánchez, 2007).

6.5.2 Estadísticos clásicos

Técnicas robustas que han retomado fuerza gracias al incremento de capacidad en los equipos de cómputo, estas técnicas trabajan considerando que los píxeles son una muestra de reflectividades para cada clase y cada banda, con lo cual se calcula la media y la desviación típica de cada una de las clases, y al utilizar algún modelo de distribución de probabilidad se obtiene la

clasificación (Fernández, López , & Yubero, 2009). Un ejemplo de estas técnicas es la de máxima probabilidad, la cual asume que los datos siguen una función de distribución normal para asignar la probabilidad de que un píxel cualquiera pertenezca a cada una de las clases, el píxel se asigna de este modo a la clase a la que es más probable que pertenezca.

6.5.2.1 Máxima probabilidad

El método de máxima probabilidad es un algoritmo paramétrico que asume alguna distribución estadística particular para las clases consideradas. Es un algoritmo bastante utilizado utiliza un modelo probabilístico, comúnmente la distribución gaussiana para formular sus reglas de decisión en la categorización de los píxeles (Bernabé, 2010).

Para definir dicho algoritmo es necesario conocer la media y la matriz de covarianza que se podrán obtener de los datos de las áreas de entrenamiento, a continuación se enlistan los pasos del algoritmo (Gil & Muñiz, 1975):

1. Calcular la media y la matriz varianza-covarianza para cada clase. El tamaño de la matriz va a depender de las bandas que se utilicen, en nuestro caso utilizaremos las bandas R, G y B, además introduciremos el valor de transparencia, en total tendremos una matriz de dimensión 4x4.
2. Clasificar cada uno de los píxeles de la imagen final, de tal modo que tendremos que aplicar la ecuación 21:

$$g = -\log((x - y)' * (inv(z)) * (x - y)) \quad (21)$$

3. Siendo g , la probabilidad del píxel en una clase determinada. Z es la matriz de varianza-covarianza. X es el píxel a clasificar. Y Es la media RGB de la clase que se está calculando la probabilidad.



4. Obtener la mayor probabilidad de cada píxel calculado en el paso anterior, para asignar dicho píxel a la clase que mayor probabilidad tiene de pertenecer. Al final tendremos una matriz de tamaño la imagen final en donde cada píxel estará clasificado en aquella clase donde la probabilidad es máxima de pertenecer.

Desde el punto estadístico, el método de máxima probabilidad se considera bastante sólido, versátil y aplicable a la mayoría de tipos de datos.

6.5.3 Algoritmos basados en inteligencia artificial

Estas técnicas usan conceptos matemáticos tratando de imitar el complejo y abstracto pensamiento humano, mientras algunas de estas técnicas no buscan imitar el pensamiento humano sino solo ofrecer soluciones a problemas que implican una serie iterativa de pasos (Jain, Duin, & Jianchang, 2000). Un ejemplo son los árboles de decisión, su procesamiento supone un análisis exhaustivo de las respuestas espectrales de las clases y del conjunto de datos disponibles, consiste en ir haciendo preguntas a cada píxel cuya respuesta positiva o negativa conducirá a otra pregunta y así sucesivamente hasta obtener la clase a la que pertenece. Se basa en los mismos principios que los sistemas expertos (Andrew, Webb, & Copsey, 2011).

Otra de estas técnicas es las de redes neuronales artificiales (ANN) que, imitan a las redes neuronales reales en el desarrollo de tareas de aprendizaje. Una neurona artificial es un objeto lógico que recibe diversas entradas, hace una media ponderada de las mismas y produce una salida a partir de la aplicación de una función umbral a la media ponderada (Basogain, 2008). Por último la técnica SVM, la cual busca separar los datos mediante un hiper-plano de separación, para casos no lineales emplea un *kernel* (Esta se detalla en el capítulo 6.5.5). En algunas ocasiones estas técnicas se complementan con sistemas *Fuzzy* (Kaufmann, 1982), los sistemas *Fuzzy* buscan obtener valores continuos en lugar de valores discretos, de tal forma que el clasificador ya no



regresa un valor binario (pertenece o no pertenece) a una clase específica, sino un valor dado en porcentaje el cual indica el grado de pertenencia a dicha clase.

6.5.4 ANN

Las ANN están inspiradas en las redes neuronales biológicas del cerebro humano. Están constituidas por elementos que se comportan de forma similar a la neurona biológica en sus funciones más comunes. Estos elementos están organizados de una forma parecida a la que presenta el cerebro humano. Las ANN están en el margen de "parecerse" al cerebro, presentan una serie de características propias del cerebro. Por ejemplo las ANN aprenden de la experiencia, generalizan de ejemplos previos a ejemplos nuevos y abstraen las características principales de una serie de datos (Basogain, 2008).

Aprender: adquirir el conocimiento de una cosa por medio del estudio, ejercicio o experiencia. Las ANN pueden cambiar su comportamiento en función del entorno. Se les muestra un conjunto de entradas y ellas mismas se ajustan para producir unas salidas consistentes.

Generalizar: extender o ampliar una respuesta. Las ANN generalizan automáticamente debido a su propia estructura y naturaleza. Estas redes pueden ofrecer, dentro de un margen, respuestas correctas a entradas que presentan pequeñas variaciones debido a los efectos de ruido o distorsión.

Abstraer: aislar mentalmente o considerar por separado las cualidades de un objeto. Algunas ANN son capaces de abstraer la esencia de un conjunto de entradas que aparentemente no presentan aspectos comunes o relativos.

Las neuronas artificiales son descritas por tres tipos de funciones matemáticas que determinan el tipo de comportamiento de la red. Estas funciones dependen del modelo de neurona,



de especificaciones del problema y en general son consideradas otra variable a ajustar cuando se está diseñando la solución a un problema (Del Rio, 2006).

6.5.4.1 Funciones Base.

Cada neurona suministra un valor y_k a su salida. Este valor se propaga a través de la red de conexiones unidireccionales hacia las demás neuronas de la red. Asociada a cada conexión hay un peso sináptico w_{ij} , que determina el efecto de la j -ésima neurona sobre la i -ésima neurona. Las entradas a la i -ésima neurona que provienen de las otras neuronas son acumuladas junto con el umbral externo, para obtener el valor de salida. Una función base es utilizada para determinar matemáticamente cómo se combinan las entradas y los pesos para obtener la llamada “entrada neta” (net input) (Isasi & Galván, 2004).

6.5.4.1.1 Función Base Lineal (LBF).

Para un estudio analítico, las redes de conexión son matemáticamente representadas por la función de base $u(w, x)$, donde w es la matriz de pesos, y x el vector de entrada. La función de base tiene dos formas típicas: Función Lineal de Base (LBF) es una función de tipo hiper-plano, o sea una función de primer orden. El valor de red es una combinación lineal de las entradas, como se muestra en la ecuación siguiente (Specht, 1991):

$$f_i(w, x) = \sum_{j=1}^n w_{ij} x_j \quad (22)$$

6.5.4.1.2 Función Base Radial (RBF).

Función de base Radial es una función de tipo hiper-esférica. Esto implica una función de base de segundo orden no lineal. El valor de entrada representa la distancia a un determinado patrón de referencia. La función de segundo orden se puede extender a otra más general llamada función de base elíptica (Jain, Jianchang, & Mohiuddin, 1996).

$$f_i(w, x) = \sqrt{\sum_{j=1}^n (x_j - w_{ij})^2} \quad (23)$$

6.5.4.2 Funciones de Activación.

Las funciones de activación más comúnmente utilizadas en las neuronas individuales son la función lineal, la función logística y la función hiperbólica (Schalkoff, 1997).

Función umbral:

$$\begin{aligned} \varphi(x) &= 0 \text{ para } x < a \\ \varphi(x) &= 1 \text{ para } x > a \end{aligned}$$

Función lineal:

$$y(x) = x$$

Función lineal por partes:

$$\begin{aligned} \varphi(x) &= 0 \text{ para } x < -a \\ \varphi(x) &= x \text{ para } -a < x < a \\ \varphi(x) &= 1 \text{ para } x > a \end{aligned}$$

6.5.4.3 Función de salida.

Esta función proporciona la salida global de la neurona. La neurona, en una red neuronal, envía generalmente señales a otras neuronas, o hacia el mundo exterior. La señal que la neurona envía es determinada por la función de salida. Ésta a su vez entrega una señal que depende del estado de activación de la unidad. En general la función de salida es la identidad, o sea, la red coloca a la salida el resultado de la función de activación. Sin embargo, no hay argumentos teóricos que impongan esto, por lo que se puede pensar en fijar una función específica para la salida de la neurona. En el caso de la llamada “Máquina de Boltzmann”, la función de salida es estocástica, por lo que la salida es probabilística (Koushal & Gour, 2012).

En general se tiene que: $\lambda(\varphi(v)) = \varphi(v)$

6.5.5 SVM

El SVM es una técnica de clasificación que permite separar solo 2 clases a la vez, esta forma de clasificar es llamada de tipo binario. El trabajo de SVM es clasificar buscando el mejor hiperplano de separación que pueda existir entre las clases, el hiperplano se obtiene de maximizar la distancia que existe entre una clase y otra clase. De esta forma se define que SVM busca el margen más grande que existe entre 2 clases sin que existan puntos intermedios dentro de este margen, el margen se considera como la distancia máxima de los dos puntos que corren en paralelo al hiperplano, estos puntos son llamados vectores de soporte (Cristianni & Shawe-Taylor, 2000).

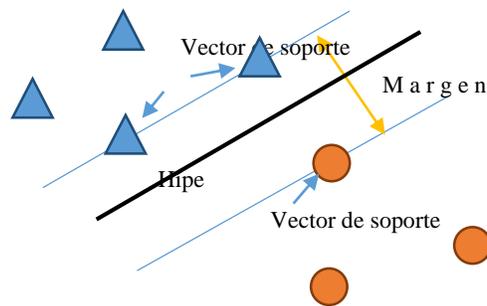


Figura 8 Diagrama de las partes de una SVM.

Los triángulos representan la primera clase, los cuadrados representan la segunda clase, el hiper-plano de separación se localiza al centro y los vectores de soporte se encuentran en paralelo al hiper-plano y los vectores de soporte crean el margen de separación.

El SVM es una técnica de clasificación supervisada, que debe de entrenar con un conjunto de puntos (vectores) x_i , junto con la clase y_i . Los puntos x_i contienen d dimensiones, cada vector está definido dentro de $x_i \in \mathbb{R}^d$, y su clase está definida por: $y_i = \pm 1$. Con esta información se tiene la siguiente ecuación del hiper-plano (Muller, et al, 2001).

$$f(x) = (w \cdot x) + b \quad (24)$$

Donde $w \in \mathbb{R}^d$ además de que w, x y b pertenecen a números reales.

Para obtener el mejor hiper-plano de separación es necesario encontrar a w y b , esto se logra minimizando a $\|w\|$ de tal manera que para todos los puntos (x_i, y_i) .

$$f(x) = y_i(w \cdot x_i) + b \geq 1 \quad (25)$$

Los vectores de soporte x_i son aquellos para los que cumplan con la siguiente ecuación:

$$y_i(w \cdot x_i) + b = 1 \quad (26)$$

Por conveniencia matemática, el problema se puede generalizar al minimizar a $\frac{w \cdot w}{2}$. Este es un problema de programación cuadrática. La solución óptima w, b permite la clasificación de un vector z de la siguiente manera:

$$\text{Clase}(z) = \text{signo}((w \cdot z) + b) \quad (27)$$

Donde $z \in \mathbb{R}^d$ y la clase depende del signo resultante de esta ecuación.

Para resolver este problema cuadrático es computacionalmente más simple resolverlo de manera dual. Para obtener la dualidad se emplea el positivo de Lagrange y el multiplicador α_i , para multiplicarlo a la ecuación 28 y restar la función objetivo.

$$L_P = \frac{1}{2}(w \cdot w) - \sum_i \alpha_i (y_i((w \cdot x_i) + b) - 1) \quad (28)$$

Para buscar un punto estacionario L_P sobre w, b . Se establece el gradiente de L_P igual a 0, para obtener:

$$w = \sum_i \alpha_i y_i x_i \quad (29)$$

$$0 = \sum_i \alpha_i y_i$$

Gradiente $L_P = 0$ en problema por dualidad.

Sustituyendo dentro de L_P , se obtiene la dual L_D :

$$L_D = \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (30)$$

Despeje en L_P para obtener L_D .

Cuando se maximiza sobre $\alpha \geq 0$. En general, cualquier α_i que es 0 es el máximo. Las α_i diferentes a 0 dentro del problema dual define el hiper-plano lo que da a w como la suma de $\alpha_i y_i x_i$.

De esta manera los valores donde x_i corresponden a una $\alpha_i \neq 0$ son los vectores de soporte.

Al derivar L_D con respecto a $\alpha_i \neq 0$, es 0 un grado óptimo. De esta manera se tiene como resultado:

$$y_i((w \cdot x_i) + b) - 1 = 0 \quad (31)$$

Derivada de L_D con respecto a $\alpha_i \neq 0$.

En particular, esto le da el valor de b a la solución, mediante la adopción de cualquier i con un $\alpha_i \neq 0$.

6.5.6.1 SVM para multi-clasificación.

Para realizar la multi-clasificación es posible a pensar que SVM es una técnica binaria, esto se puede solucionar empleando la técnica OAA. La formulación inicial del método de OAA requiere la unanimidad entre todos los SVMs: un punto de datos sería clasificado bajo una cierta clase, si y sólo si, SVM de esa clase lo aceptó y SVM de todas las demás clases lo rechazan. Mientras la precisión para la agrupación de clase es estrecha, este método deja regiones del espacio de características indecisas donde más de una clase acepta o rechaza todas las clases (Milgram, Cheriet, & Sabourin, 2004).

6.5.6.2 Kernel una solución para problemas no linealmente separables.

Algunos problemas de clasificación binarios no tienen un hiper-plano simple como un criterio de separación útil, para esos problemas, hay una variante del enfoque matemático que conserva casi toda la sencillez de un SVM. Esta variante es llamada *kernel* que nos permite transformar los datos en un espacio de función F que por lo general, tiene una gran dimensión, permitiendo de esta manera que los datos sean separables (Dumitru, 2012).

Para que los datos sean transformados a un espacio F , es considerada una función de mapeo ϕ esta opera desde un espacio de entrada X , convirtiendo los datos en una función de espacio Hilber (Baudat & Anouar, 2001).

$$\phi: X \rightarrow F \quad (32)$$

$$x \rightarrow \phi(x)$$

Se supone que m es el número de muestras, con el cual se define la matriz del *kernel* k basada en un producto punto:

$$K = (k_{ij})_{1 \leq i \leq m, 1 \leq j \leq m} \quad (33)$$

Donde $k_{ij} = \phi^t(x_i)\phi(x_j)$

La $m \times m$ K matriz con elementos $K_{ij} = k(x_i, x_j)$ es llamado matriz *kernel* o matriz Gram.

La matriz *kernel* es simétrica y positiva semi-definida por lo tanto para todo $a_i \in \mathbb{R}, i = 1, \dots, m$, con esto se tiene:

$$\sum_{i,j=1}^m a_i a_j K_{ij} \geq 0 \quad (34)$$

De manera gráfica el *kernel* se encarga de adaptar los datos para permitir una clasificación lineal y simple (ver Figura 9).

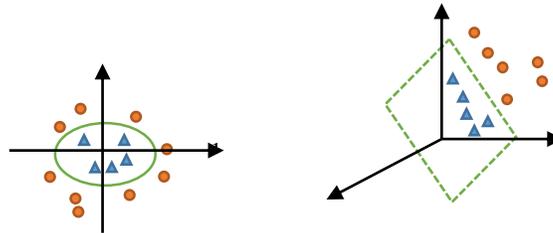


Figura 9 Ejemplo grafico de un *kernel*.

En la figura 9. Se ve un ejemplo grafico de como un *kernel* genera un nuevo espacio F para permitir una clasificación por un hiper-plano, separando los triángulos de los círculos.

Existe una gran variedad de *Kernel* que se emplean en las SVM, a continuación el Cuadro 3 muestra algunos ejemplos de Kernels (Ayat, Cheriet, & Suen, 202).

Cuadro 3 Kernels más comunes.

Nombre	Ecuación
Lineal	$k_0(x, y) = (x \cdot y)$
Polinomio	$k_1(x, y) = ((x \cdot y) + \theta)^d$
Gaussiano	$k_2(x, y) = e^{\left(-\frac{\ x-y\ ^2}{\gamma}\right)}$
Sigmoid	$k_3(x, y) = \tanh(k(x \cdot y) + \theta)$

6.6 Filtros

Al momento de trabajar con imágenes es necesario que estas sean filtradas para permitir un suavizado, con el objeto de normalizar los datos, y con ello lograr que las técnicas de clasificación y segmentación mejoren.



6.6.1 Ejemplos de filtros

Uno de los filtros más usados es el filtro Gaussiano que se explica en (Camapum & Fisher, 2001) y (Sarría, Recio, & Ruiz, 2003) donde demuestran que al usar filtros antes de clasificar los píxeles (individuales o agrupados en texturas) se mejora la clasificación de los píxeles y por ende de la imagen.

Otro filtro usado es el de Gabor explicado en (Ali, Georgsson, & Hellstrom, 2008) y (Yang & Newsam, 2010) en dichos trabajos se demuestra que los resultados de clasificación mejoran considerablemente, y son equiparables a los obtenidos con el filtro Gaussiano antes referido.

De acuerdo a (Kim & Hong, 2008), Para trabajar con un filtro Gaussiano empleando texturas es necesario dividir la imagen en ventanas, señalado también por (Chehade, Boureau, Vidal, & Zerubia, 2009), con lo cual se obtiene una nueva imagen la cual será procesada para obtener los parámetros de textura.

6.7 Trabajos relacionados

La necesidad de conocer los recursos naturales de una región, empleando mecanismos que reduzcan el tiempo de respuesta de resultados a bajos costos económicos es un factor esencial que se debe considerar para desarrollar nuevas técnicas para conocer la cobertura de suelo en grandes extensiones.

Para interpretar la cobertura de suelo de una imagen digital se han empleado diferentes técnicas de clasificación. Sin embargo, a medida que se tienen avances tecnológicos en la captura de imágenes, las técnicas para clasificarlas han evolucionado significativamente.

Usando la teoría de segmentación de imágenes para clasificar una imagen (Trejo & Hernández, 1996) proponen usar una imagen satelital de 7 bandas (Landast TM), para clasificar 14



categorías de cobertura de suelo en: bosques templados, bosques de encinos, bosques secundados, selvas subcaducifolias, matorrales, selva baja caducifolia, selva baja secundaria, agricultura-temporal, pastizal, agricultura, agricultura de riego, sin vegetación, uso urbano y cuerpos de agua. Ellos buscan patrones similares a uno propuesto en una ventana con un patrón cuadrado obteniendo los valores del espectro en cada ventana, después se ejecuta el algoritmo clustering de máxima verisimilitud expuesto en (Myung, 2003) obteniendo como resultado una precisión de un 70% al clasificar 14 distintas clases en esta imagen. Este resultado no es muy aceptable.

6.7.1 Clasificación con Máxima verisimilitud

También con imágenes de radar (ERS y RADARSAT) (Freitas, Lelio, & Dutra, 2000) realiza un estudio para obtener la cobertura de suelo de 6 categorías: tierras con pasturas, tierras abandonadas dedicadas al pastoreo, tierras con reforestación de 3 a 8 años, tierras con reforestación de 9 a 13 años, tierras con reforestación de 13 años en adelante y bosque con vegetación primaria. Logrando una precisión de 80% en las primeras imágenes y un 70% en las segundas, emplean una ventana con movimiento y estadísticos de texturas, usando el clasificador clustering de máxima verisimilitud (maximum likelihood). Mejorando en algunos casos la precisión.

En otro trabajo con una intensidad similar (Simard, Sasan, & DeGrandi, 2000) tratan de clasificar: Pasto de sabana. Zona urbana. Bosque abierto. Bosque deforestado. Árboles de sabana Jungla bosque muy tupido con lianas y enredaderas y dosel de árboles, los autores experimentaron con 2 clasificadores: clustering de máxima verosimilitud y árbol de decisión, sin lograr obtener grandes diferencias entre ellos, solo que el árbol de decisiones mejor los datos espaciales, más no las precisiones.

Hasta el momento las técnicas solo generan imágenes con resultados planos, por esto (Sanofer & R.De, 2010) propone el uso de un mapa 2d, los datos de entrenamiento son obtenidos



por un escalador, el cual irá dando información de lo que clasifica incluyendo la altura donde se localiza, estos datos se dividen en regiones dependiendo la altura y el tipo de terreno, después estos datos son clasificados con la técnica de máxima verisimilitud, obteniendo un promedio de un 80% y teniendo como resultado un mapa 2d, este autor comenta que clasifica tipos de copa de árboles sin especificar que especies.

6.7.2 Clasificación con índices

Con el objetivo clasificar los bosques de coníferas de las especies *Picea crassifolia*, *Sabina przewalskii* y pastos (Qisheng, 2009) propone usar el índice de biomasa aérea, considerando datos de texturas y del espectro de imágenes adquiridas por de satélite LiRDA. El logro una precisión de 87.54%.

Trabajando con el mismo índice de biomasa aérea (Haipeng & Kazuo, 2010), busca emplear la técnica de k distribuciones descritas en (Jakeman, 1980), con el objetivo de separar las siguientes cinco clases de árboles: coníferas en general, la especie particular de coníferas Larch, Pino rojo y Abeto, logrando una precisión promedio de 85%.

Con el fin de buscar separar áreas en una imagen (Lin, Wang, & Zha, 2010) trata de clasificar las siguientes clases: terraza urbanizada, nubes, arboles caducifolio, bosque latifolio, bosque mixto de árboles caducifolio, suelo desnudo, bosque de coníferas de hoja perenne, bosque mixto de árboles de hoja perenne, prados, sombras y agua. Para este trabajo se usó los índices normalizados de vegetación, el índice normalizado de agua, obtuvieron un promedio de 92% y con un índice Kappa de .89.



6.7.3 Clasificación con Transformaciones

Con el fin de poder automatizar la clasificación (Wang & Boesch, 2007) propone una técnica para segmentar imágenes, empleando la segmentación con base en la media J (JSEG) referirse a (Deng & Manjunath, 2001), se clasificaron bosque, cultivos, zonas urbanas y agua, empleando texturas y la transformada de Wavelet, logrando precisiones desde 85% hasta el 95%. Se tienen buenas precisiones, pero es conveniente resaltar que se clasificaron las clases muy contrastante a diferencia de los trabajos donde se trata de clasificar árboles que a la vista son clases sumamente parecidas.

Empleando un Frame Wavelet (D De Grandi, Lucas, & Kropacek, 2009) agrega el concepto de retro inspección tomando como base el análisis discriminatorio lineal de Fisher, logrando separar algunos tipos de árboles, pero encontrando el problema al identificar clases que no son linealmente separables.

Retomando la transformada de Wavelet (Ghasemi, Sahebi, & Mohammadzadeh, 2013) introduce el índice de biomasa aérea (AGB) descrita en (Hyypää, et al, 2000), y los índices de vegetación, en imágenes de satélite SPOT y LandSat con la tecnología de mapeo térmico avanzado de alta resolución radiométrico (AVHRR), logrando una precisión del 75%, al clasificar los arboles: beech (*Fagus sylvatica*) hornbeam (*Carpinus betulus*) alder (*Alnus glutinosa*) oak (*Quercus castaneafolia*), maple (*Acer velotonia*) ironwood (*Parotia pérsica*).

6.7.4 Clasificación con ANN

Con el fin de no solo usar los valores de los pixeles, (Pugh & Waxman, 2006), emplean una red neuronal artificial difusa (*Fuzzy ARTMAP neural classifier*) ver (Garpenter, Grossberg, Markuzon, Reynolds, & Rosen, 1992), introduciendo también el valor de la elevación en la imagen, logrando una precisión desde 73% hasta 97% al separar varios tipos de árboles entre ellos pino rojo



y pino blanco. En estos resultados se observa que las redes neuronales podrían jugar un papel muy importante en la clasificación de vegetal, dadas las altas precisiones del 97 %.

En otro trabajo (Dengru & Linders, 1999) evalúan la precisión al clasificar áreas quemadas de bosque, áreas de bosque talado y áreas pobladas por copas de árboles con vegetación como lianas y enredaderas) en imágenes obtenidas de un sensor de Radar de Apertura Sintética (SAR) (acrónimo en inglés, Synthetic Aperture Radar) en la banda C, el experimento consta en 3 etapas: selección de ventana de estudio, estadísticos y de clasificador. En la primera etapa la ventana se obtienen por una prueba de movimiento y una prueba disjunta. Los estadísticos se dividen en estadísticos de textura (promedio, desviación estándar, texturas) y estadísticos de espectro (entropía) y una combinación de estos. En la última etapa se emplean 2 clasificadores análisis discriminatorios paso a paso y una cuantificación de vector de aprendizaje por cuantificación vectorial (LVQ), (siglas en inglés Linear-Vector Quantization). Obteniendo precisiones arriba del 73% hasta el 78%, con la combinación de pruebas por movimiento, estadísticos de textura y el análisis discriminatorio paso a paso. Las precisiones son mejoradas con respecto al anterior método.

No solo es necesario probar las técnicas de clasificación sino que (Trias-Sanz, 2006) genera una nueva textura basada en el dominio de Fourier con el fin de obtener un gradiente y un sentido para realizar la clasificación entre viñedos, bosques, terrenos de cultivo y pomares, logrando precisiones del 81% al 95%. Aparte de transformar los datos al dominio de Fourier, es posible realizar otro tipo de transformaciones como la transformación de Wavelet (Shaker & Zou, 2008) con la cual realizan un ventana de 3x3x3 (tomando los valores de los pixeles como un cubo), después aplicando uno de estos tres filtro: pasa altas, pasa bajas o Gaussiano, después obteniendo



los estadísticos de promedio y desviación estándar, introduciendo estos datos a una red neuronal artificial (ANN).

6.7.5 Clasificación por Clustering

Considerando que las técnicas de clustering consumen una gran de recursos computacionales y tratando de disminuir este inconveniente (Sgrenzaroli, Baraldi, Eva, & De Grandi, 2002) emplean el algoritmo de clustering adaptativo de Pappas (MPAC) ver (Pappas, 1992), el cual emplea un filtro pasa bajas, después agrupa tomando en cuenta los archivos aleatorios de Markov (RFM) véase (Smits & Dellepiane, 1997), para clasificar Bosque, áreas deforestadas, áreas reforestadas, agua, nubes, sombras de nubes y Bosque tropical de sabana, obteniendo resultados obtenidos de 85% de precisión mejoran a los presentados por la técnica de máxima verisimilitud, además emplea menor cantidad de recursos computacionales para su ejecución.

6.7.6 Clasificación con Umbrales

Empleando imágenes de satélite IKONOS (Kawamura, Tsujino, & Tsujiko, 2004), realiza la clasificación de las siguientes especies: conífera cryptomeria, ciprés hinoki, pino rojo, áreas de pino muerto, karamatsu, roble, Castaño y Bambú. El experimento uso dos variantes en la primera se usaron valores del espectro y la segunda se utilizaron además de los valores espectrales se añadieron los valores derivados de las texturas de la imagen RGB. En este trabajo es importante destacar que se tienen clases diversas con texturas muy parecidas por lo que los resultados resultan ser muy importantes por las altas precisiones presentadas. Por lo que se pondrán todos los resultados obtenidos en el Cuadro 4.



Cuadro 4. Precisiones de las diferentes clases con las dos pruebas

Clase	Precisión con valores espectrales	Precisión con valores espectrales más valores derivados de las características de texturas
conífera cryptomeria	97.9	96.8
ciprés hinoki	99.7	100
pino rojo	28.8	30.6
áreas de pino muerto	37.4	47.4
karamatsu	79.9	90.5
Roble	67.1	31
Castaño	25.4	16.2
Bambú	96	84

Se observa que las precisiones mejoran en algunas clase con los puros valores espectrales y en otros casos sumando los valores derivados de las características de las texturas.

6.7.7 Clasificación con SVM

Con la ayuda de una máquina de soporte vectorial y empelando la textura de Haralick, promedio y desviación estándar (Benelcadi, et al, 2012) logra una precisión de 50%. Con las siguientes especies vegetales: bosque virgen, bosque templado caducifolio, bambú plantaciones, bosque de melaleuca, pasto de la sabana y manglares. Cabe mencionar que se usaron imágenes infrarrojas por lo que solo tienen una banda con datos útiles, esto hace que la precisión sea muy pobre.



VII METODOLOGÍA

7.1 Materiales.

Para obtener la cobertura de suelo en este trabajo se emplea una imagen híper-espectral la cual contiene 63 bandas, 512 pixeles de ancho y 512 pixeles de largo, la imagen fue obtenida mediante la tecnología híper-espectral AHS-ATM (Rejas, et al , 2003). Esta imagen será conocida como IMG_63, dentro de la imagen existen 15 clases, representadas por tierras de cultivos, zonas boscosas y tierras de labranza, la lista se observa en el Cuadro 5.

Cuadro 5 Lista de clases en la imagen IMG_63

Sin Vegetación	Avena
Experimental Plot	Cebolla
Maíz	Reforestación
Barbecho	Girasol
Fruta	Vid
Ajo	Masa de agua
Invernadero de Ajo	Trigo
Pasto	

Con fines de validación experimental, además de la imagen híper-espectral, se tiene una imagen RGB llamada resRGB en la cual se detallan los límites de cada una de las clases que existen, así como las zonas que no son de interés para el estudio.

La computadora empleada para realizar la investigación cuenta con las siguientes características: procesador AMD FX 6300 Black Edition a 3.5 GHz, 16 Gb en RAM y un disco duro de 2 TB.

7.2 Caracterización de la imagen experimental

De acuerdo al estudio realizado por (Srivastava, Asati, & Bhattacharya, 2010) se demostró que al trabar con pixeles de forma individual generan resultados favorables al combinar técnicas



Fuzzy con los pixeles de manera individual los resultados no se presentan con precisión u otro estadístico, pero se comenta que es posible usar técnicas de pixeles e implementarlas en equipos de cómputo. Aunado a esto, se considera que al trabajar con imágenes AHS-ATM contiene suficiente información al contar con 63 bandas, de esta manera es posible clasificar empleado SVM.

Debido a las características del equipo de cómputo y en virtud de que la imagen completa no pudo ser procesada por falta de memoria y desbordamiento de pila se decidió dividir la imagen IMG_63 en 3 partes con las siguientes dimensiones 172x512 pixeles, 172x512 pixeles y 168x512 pixeles. Cada nueva imagen será nombrada con un guion bajo y el número que corresponda: IMG_63_1, IMG_63_2 y IMG_63_3.

7.3 Desarrollo del experimento

Las tres imágenes pasan por 4 etapas o súper-funciones: *ConjuntosF*, *Conjuntos4Máquina*, *Máquina* y *DibujaRGB* para realizar el proceso de cobertura de suelo.

7.3.1 Primera etapa ConjuntosF.

En detalle, en esta etapa la imagen es caracterizada píxel por píxel, eliminando los pixeles que no son de interés, creando una meta-matriz (Es una matriz que almacena solo información referente a otras matrices) de nombre *etiquetaExisteResRgb*. La meta-matriz almacene el número de clases, los colores de cada clase, el número de pixeles que presenta cada clase y la cantidad de conjuntos que se crearán para cada clase. Del mismo modo se genera una índice-matriz (indexa otra matriz) de nombre *matClaseResRgb* en la cual se guardan las coordenadas de cada píxel, la clase y el conjunto al que pertenece cada pixel. Una vez que se tiene el índice matriz de todos los pixeles, se inicia el proceso de filtrado, proceso por el cual la imagen con todas las bandas será



transformada para permitir un suavizado y ajuste de sus fronteras. El filtrado se explica más adelante.

La Figura 10, muestra los pasos para realizar la súper-función *ConjuntosF*, es importante mencionar que las 2 tareas imagen caracterizada (*resRgb*) e imagen hiper-espectral (*mRgb*), son independiente y no es necesario terminar una tarea para iniciar otra, además que la parte de los filtros es exclusiva de *mRgb* y se realiza banda por banda.

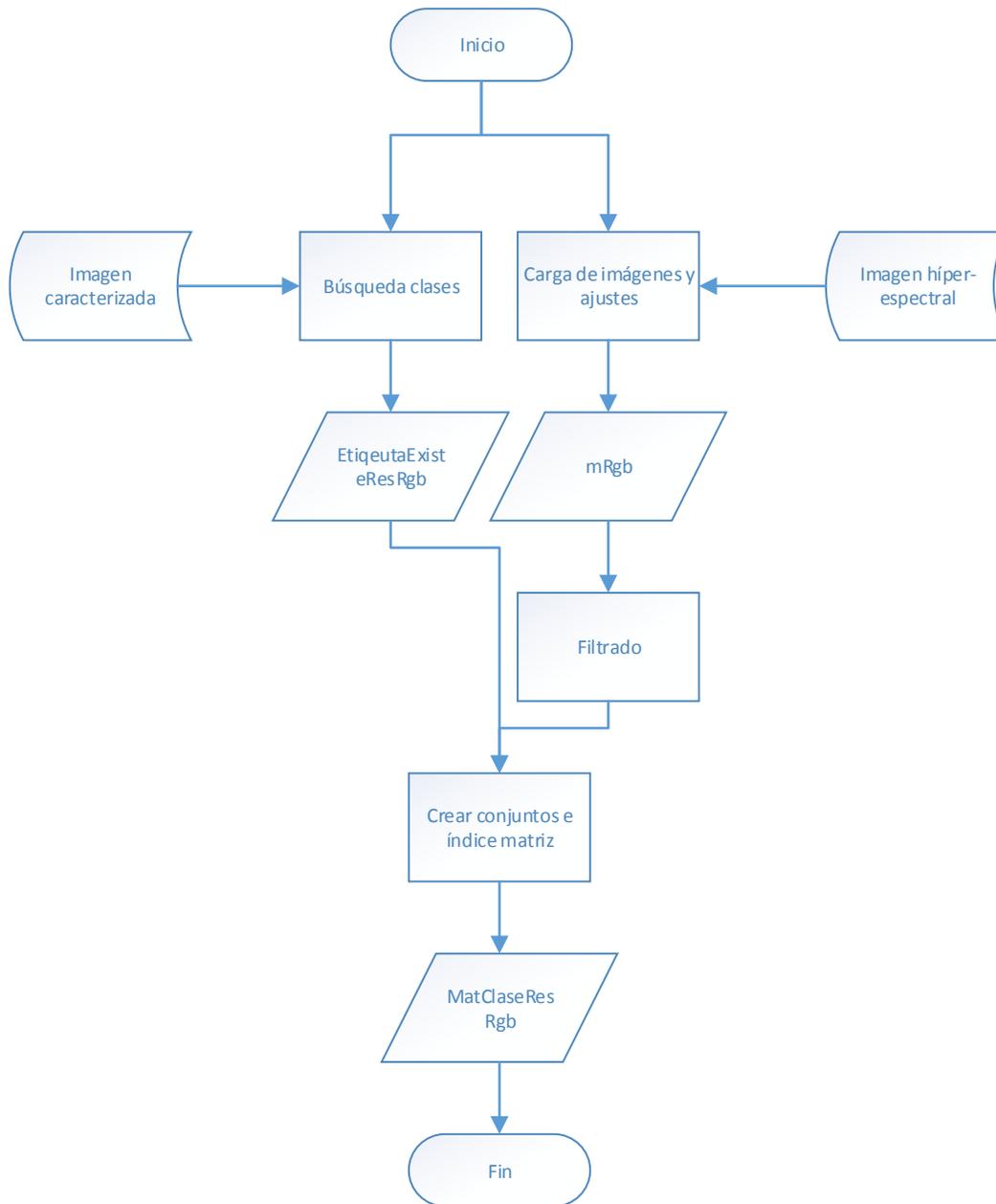


Figura 10 Diagrama de flujo súper-función *ConjuntosF*

7.3.1.1 Definición del filtro.

En esta investigación, en vez de usar ventanas se emplean 3 topologías de distribución de píxeles: cuadrado, cruz y estrella de acuerdo a lo propuesto por los parámetros de Markov referidos en (Petrou & Garcia Sevilla, 2006), con la Fórmula 35.

$$s = a(g_l + g_r) + b(g_t + g_b) \quad (35)$$

Donde g_l , g_r , g_t y g_b son los píxeles que se ubican a partir del píxel central, el de arriba el de debajo de la izquierda y de la derecha, a y b son los parámetros de Markov, en cada patrón se maneja una magnitud, que varía entre 1, 3 y 5. Se muestra un ejemplo en Figura 11.

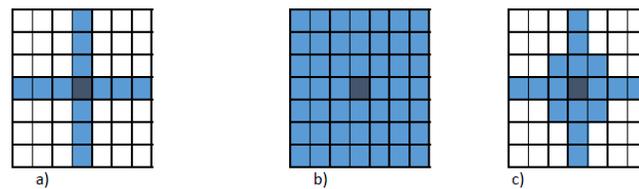


Figura 11. Ejemplo de los 3 tipos de parámetros.

- a) muestra una de tipo cruz con magnitud 3.
- b) muestra una de tipo cuadrado con magnitud 3.
- c) muestra una de tipo estrella con magnitud 3. Los coloreados son los considerados.

Los datos de entrada a la técnica de clasificación no son obtenidos por las texturas sino que cada píxel sufre una transformación, con base en un promedio de su vecindario, este valor está dado por la fórmula 33. Esto permite evitar el uso de técnicas de segmentación, y emplear el valor de cada píxel como descriptor de una clase.

$$X_{new} = \frac{\sum_{i=1}^{tF} \frac{X_{old} + X_i}{2}}{tF} \quad (36)$$

Donde

tF= total de pixeles que conforman la figura sin incluir el píxel de interés.

X_{old} = píxel de interés.

X_{new} = nuevo valor del píxel de interés.

X_i = pixeles vecinos al píxel de interés sin contar al píxel de interés.

7.3.1.2 Algoritmo.

Para implementar la ecuación 36 se ha implementado el siguiente algoritmo (ver Cuadro

6).

Cuadro 6 Algoritmo para el filtro.

```
Para cada i a la ancho de la imagen
  Para cada j a lo lardo de la imagen
    Para cada k por cada banda
      Hacer
        Obtener valor de pixel;
          Hacer por cada posición de figura
            Busca pixel hasta encontrar magnitud
              Por cada pixel
                pixelnew = pixelold + pixelFigura / 2;
                contador incrementa;
              fin
            fin
          fin
        pixelnew = pixelnew / contador;
      fin
    fin
  fin
```

Este algoritmo recorre la imagen por ancho, largo y profundidad, en cada una de estos ciclos busca todos los pixeles que conforman el vecindario del patrón elegido del filtro, este patrón se ve

afectado por una magnitud de tal forma, que si el patrón es una cruz de magnitud 1 solo recorre un pixel arriba, un pixel a la derecha, un pixel abajo y un pixel a la izquierda. Pero si se elige el mismo patrón pero con magnitud 3 entonces se recorren 3 pixeles a la derecha 3 pixeles arriba y así sucesivamente hasta terminar con la figura.

Un ejemplo en una imagen RGB muestra que los filtros caen dentro de filtros de suavizado esto se observa en la Figura 12.

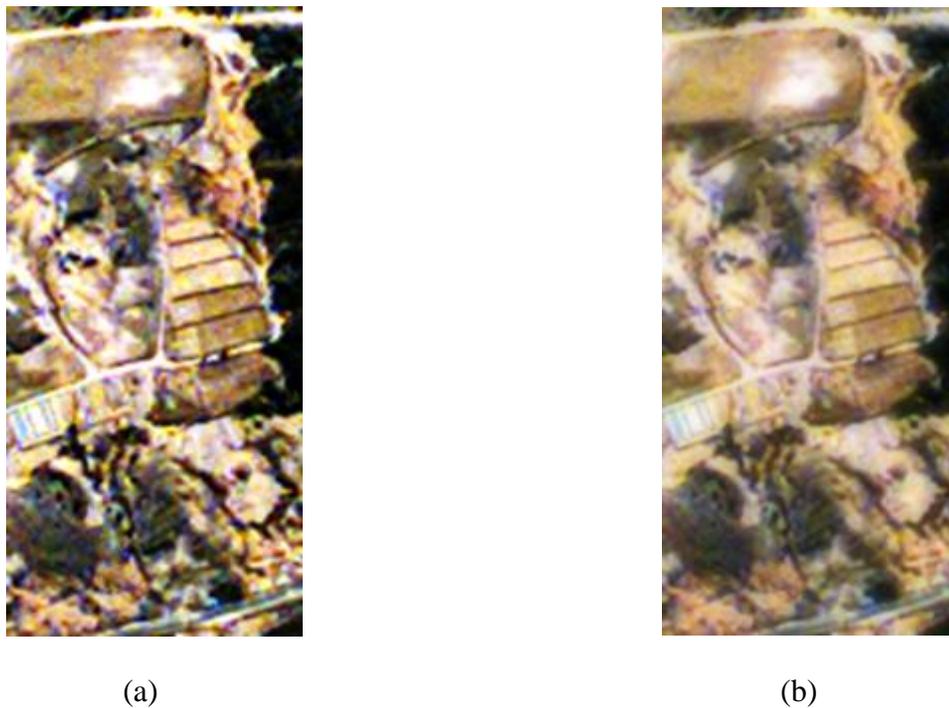


Figura 12 Ejemplo de imagen filtrada.
a) Imagen original.
b) filtrada por un filtro cruz magnitud 1.

7.3.2 Segunda etapa Conjuntos Máquina

La imagen IMG_63_FPT es dividida en 20 conjuntos por cada clase que tenga la imagen, estos conjuntos son generados de forma aleatoria, el algoritmo de selección de pixeles distribuye

los pixeles en subconjuntos con 20 pixeles como mínimo y en el caso que no tenga la clase suficientes pixeles crea menos subconjuntos (10 o 15).

Cada clase proporciona el 20 % de los subconjuntos de forma aleatoria para agregarlos a un súper-conjunto de entrenamiento y el 80 % restante es agregado a un súper-conjunto de prueba (estos 2 conjuntos albergan todos los pixeles de la imagen a analizar) (Rojas, 2009), si la clase solo cuenta con 1 subconjunto este se agrega al súper-conjunto de entrenamiento, una vez que se tiene el súper-conjunto de entrenamiento este es desordenado para evitar que la técnica de clasificación se quede en un mínimo local (ver Figura 13).

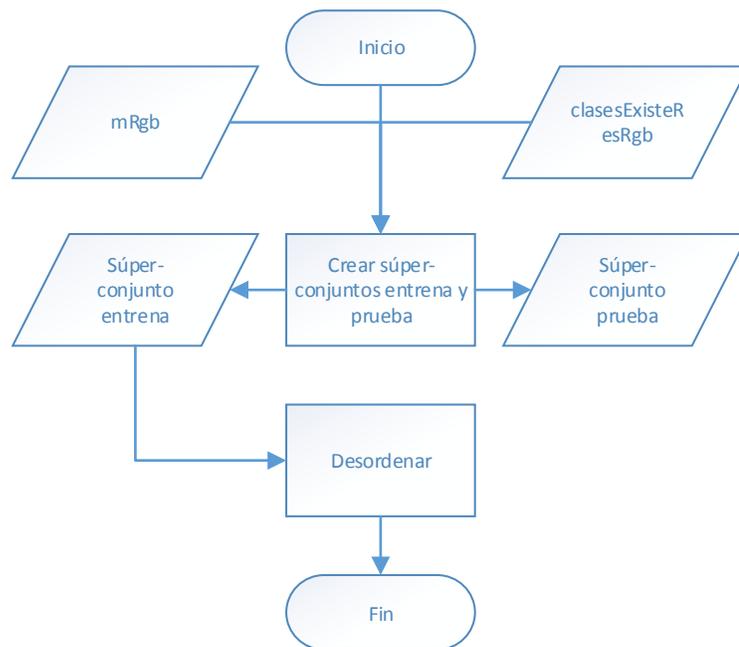


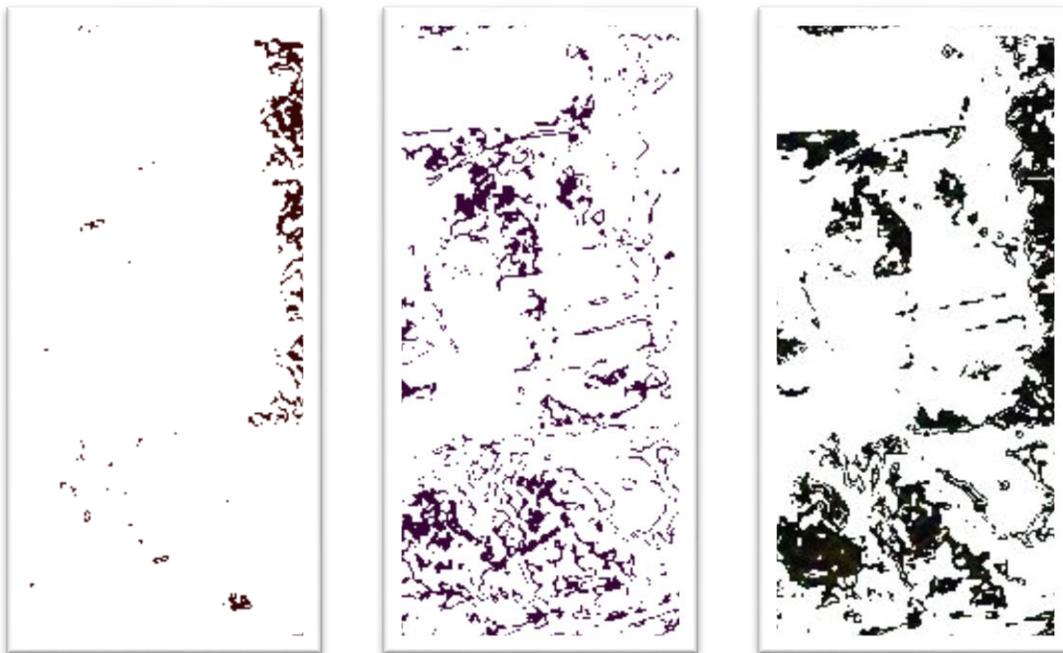
Figura 13 Diagrama de flujo de súper-función *Conjuntos4Maquina*

7.3.3 Tercera etapa Máquinas

El súper conjunto de entrenamiento se utiliza para entrenar a la SVM, para generar la clasificación de los pixeles en las diferentes clases que componen la imagen.

7.3.3.1 Clasificación de la imagen

Para realizar esta investigación se empleó el *kernel* Sigmoid, debido a que el comportamiento que tiene es muy constante a diferencia de otros kernels, para determinar cuál es el *kernel* más adecuado se realizó una prueba en un fotograma de avión de 3 bandas RGB, se buscó separar el bosque de matorrales y suelo desnudo (ver Figura 11) .Para este experimento se desarrolló una aplicación donde se selecciona un píxel que representa la clase deseada, a partir de ese píxel se seleccionan los pixeles más próximos en forma de cruz de magnitud 1 (ya descrito arriba). Después se hace la clasificación con una SVM: asociada los *kernel*: cuadrático, poligonal y Sigmoid los resultados se muestra gráficamente en la Figura 14.



a) Kernel cuadrático.

b) Kernel Poligonal.

c) Kernel Sigmoid.

Figura 14 Resultado de SVM, asociados a los diferentes *kernel*



Como se ve en las imágenes el *kernel* Sigmoid se considera como el mejor, por tal razón se eligió para este trabajo.

7.3.3.2 Procedimiento para realizar la clasificación

El súper-conjunto de entrenamiento es introducido a la técnica de clasificación, en este caso la técnica de clasificación es un SVM con un *kernel* dado por una perceptron multicapa de acuerdo a (Rauber & Berns, 2011), el algoritmo *Maquina* busca los parámetros óptimos del *kernel*, estos se dan por una combinación que va desde 0 hasta 1 y desde 0 hasta -1 incrementando estos dos parámetros en 0.1, por lo cual el algoritmo *Maquina* realiza 100 SVM por cada clase, empleando la técnica uno contra todos SVM (OAA) como lo recomiendan en (Chehade, Boureau, Vidal, & Zerubia, 2009). Por cada clase se tiene como resultado un vector de soporte, este grupo de vectores de soporte son ocupados para etiquetar un mapa de clases, el cual, pasa por el algoritmo de *DibujaRGB* para colorear la imagen y obtener la precisión.

Todo este procedimiento desde la selección de pixeles, el algoritmo *Máquina* y la obtención de la precisión se repite 20 veces, por cada combinación de filtro de prueba y magnitud ya definidos. Con lo cual se hacen 180 ejecuciones, agregando 20 ejecuciones más donde los datos que no han pasado por ningún filtro, haciendo en total 200 ejecuciones tardando un tiempo de 30 días.

El criterio de selección de los parámetros óptimos para SVM está definido en la ecuación 37.

$$pre = \frac{NA}{NA+NE} \quad (37)$$

Donde pre = precisión.
NA = número de aciertos.
NE= número de errores.

El diagrama de flujo de la Figura 15 se observa que la súper-función *Maquina* contiene 3 ciclos iterativos; el primer respecto a las clases que existen en la imagen, el segundo respecto a la búsqueda de parámetros y por ultimo un ciclo en relación a la magnitud de SVM (El valor asignado por el clasificador, para diferenciar entre clases).

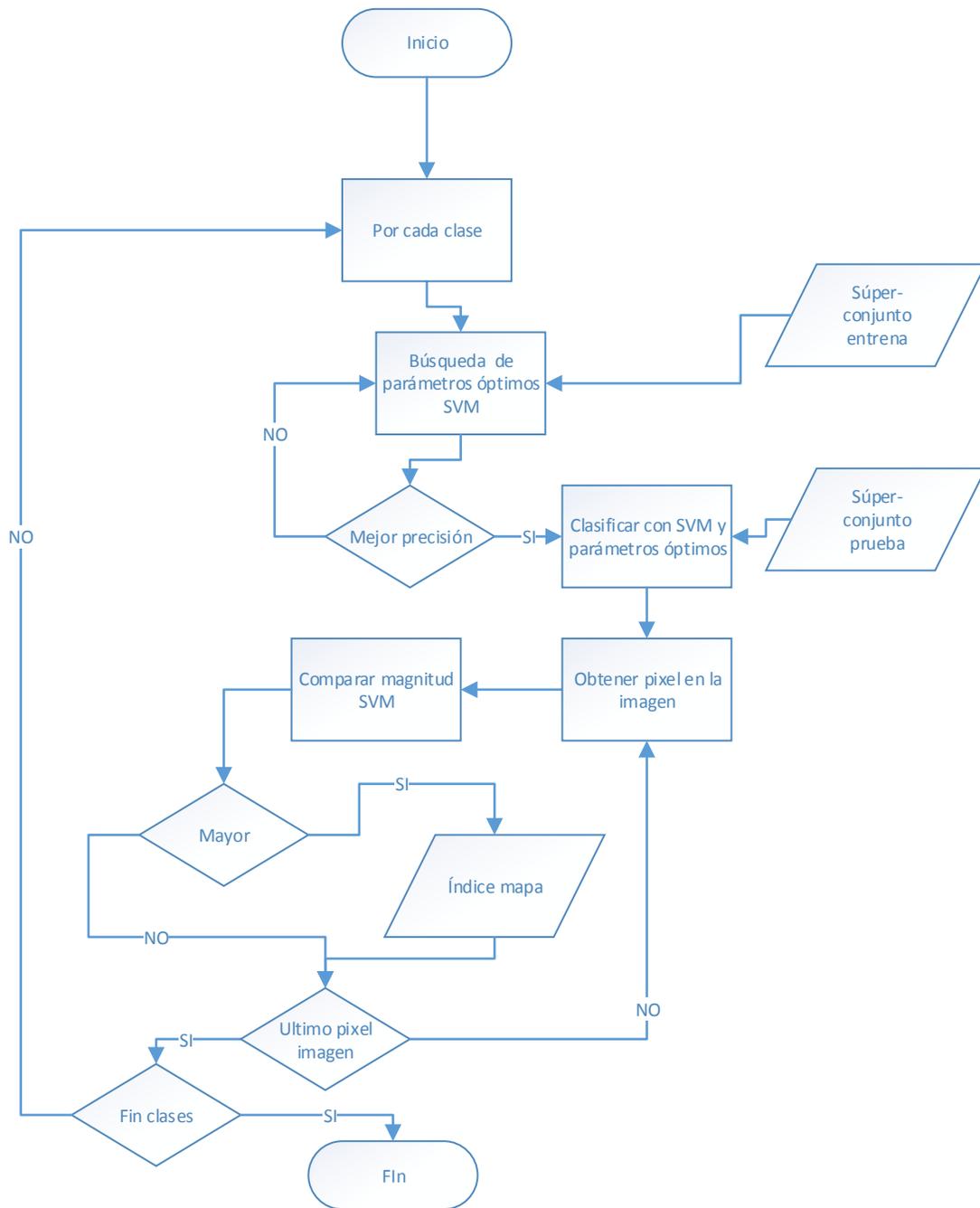


Figura 15 Diagrama de flujo de súper-función *Maquina*

7.3.4 Cuarta etapa *DibujaRGB*

En esta etapa el índice de mapa es convertido en una imagen RGB para que se pueda visualizar los resultados de la clasificación, esta opción permite crear tantas imágenes como clases contenidas en la IMG_63 inicial, en cada imagen aparecen 3 colores, en color negro todos los datos que fueron despreciados por que no fueron caracterizados en la imagen resRgB o son pixeles usados para el entrenamiento, en color gris todo lo que es considerado ajeno a la clase de interés y en un color diferente (este color esta dado como el identificador de la clase de interés, descrito en la imagen resRgB) los anteriores colores se ilumina la clase de interés (ver Figura16).

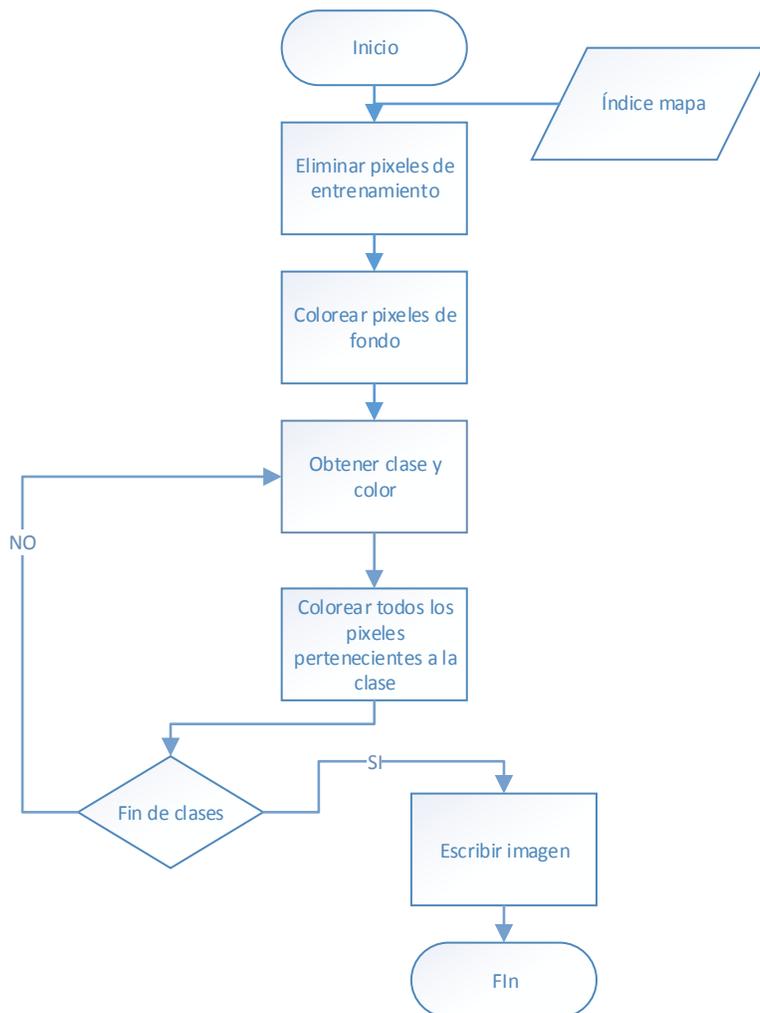


Figura 16 Diagrama de flujo súper-función *DibujarRgb*

En la figura 17 se muestra un diagrama general de los pasos a seguir durante el desarrollo del experimento, en este diagrama la imagen de n bandas es filtrada y almacenada en la matriz mR_{gb} , al mismo tiempo se obtiene la información de la imagen $resR_{gb}$, una vez terminados estos 2 pasos se crea los súper-conjuntos entrena y prueba. Los súper-conjuntos son utilizados en el SVM para obtener los parámetros que maximicen la precisión, una vez encontrados se almacena los valores resultados del SVM para ser comparados con las otras SVM de las clases restantes, con esto se obtiene una matriz de índice $newMatResR_{gb}$ para colorear la imagen final y obtener la cobertura de suelo.

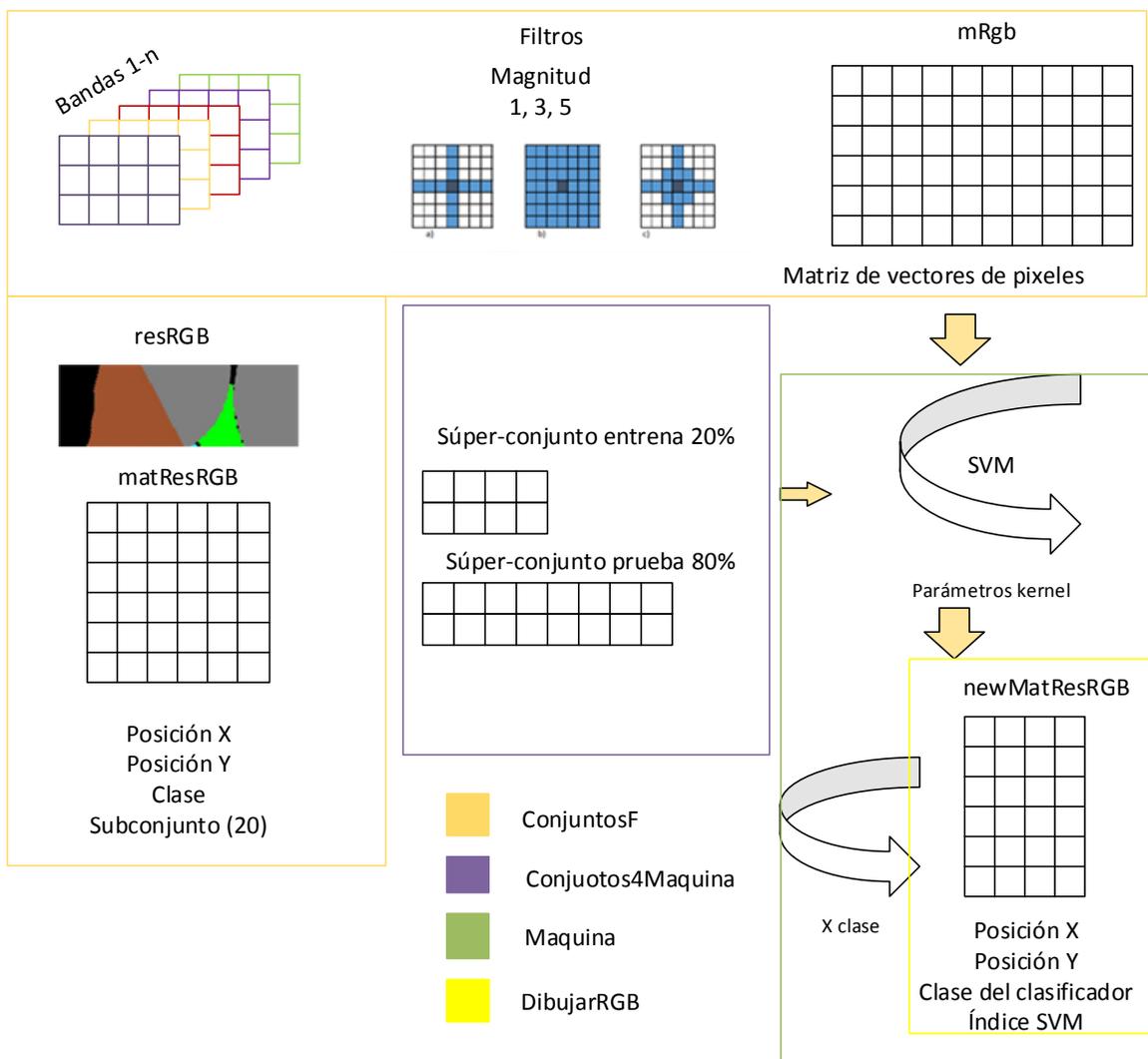


Figura 17 Diagrama general de la metodología.

VIII RESULTADOS

8.1 Cuadros de precisión

A continuación se muestran los resultados de las precisiones obtenidas en cada imagen, dentro de estos cuadros, se introducen todas las precisiones con relación al filtro y la clase. El filtro se compone por el nombre del filtro (Cruz, Estrella y Cuadrado), seguido de un número el cual indica su magnitud, al momento de no usar un filtro “Sin filtro” no contiene ninguna magnitud. Además de la precisión en los cuadros se anexa el promedio, desviación estándar y la varianza de estas.

Cuadro 7 Precisión de todos los filtros IMG_63_1

Filtro \ Clase	Avena	Cebolla	Experimental Plot	Sin Vegetación	Promedio	Des. Estándar	Varianza
Cruz 1	0.92409	0.71034	0.99938	0.93044	0.89106	0.12521	0.01568
Cruz 3	0.92322	0.71416	0.99936	0.92896	0.89143	0.12314	0.01516
Cruz 5	0.92445	0.70633	0.99780	0.92025	0.88721	0.12573	0.01581
Cuadrado 1	0.92398	0.70984	0.99935	0.93074	0.89098	0.12546	0.01574
Cuadrado 3	0.91914	0.71334	0.99936	0.93330	0.89128	0.12368	0.01530
Cuadrado 5	0.92527	0.70901	0.99936	0.92572	0.88984	0.12548	0.01575
Estrella 1	0.91610	0.71134	0.99935	0.92898	0.88894	0.12392	0.01536
Estrella 3	0.92315	0.71074	0.99936	0.93299	0.89156	0.12521	0.01568
Estrella 5	0.92301	0.71792	0.99936	0.92898	0.89232	0.12133	0.01472
Sin Filtro	0.91771	0.71269	0.99935	0.92953	0.88982	0.12346	0.01524
Promedio	0.92201	0.71157	0.99920	0.92899			
Des Estándar	0.00317	0.00317	0.00049	0.00376			
Varianza	0.00001	0.00001	0.00000	0.00001			



Cuadro 8 Precisión de todos los filtros IMG_63_2

Filtro \ Clase	Ajo	Avena	Barbecho	Cebolla	Experimental Plot	Fruta
Cruz 1	0.99249	0.88634	0.92539	0.90255	0.85994	0.83741
Cruz 3	0.99249	0.90177	0.92076	0.90074	0.86671	0.83860
Cruz 5	0.99193	0.89912	0.92128	0.90255	0.81137	0.83811
Cuadrado 1	0.84875	0.90334	0.92025	0.90253	0.86045	0.83801
Cuadrado 3	0.99236	0.89676	0.92066	0.90256	0.86673	0.83810
Cuadrado 5	0.99251	0.90311	0.92388	0.90256	0.85873	0.82131
Estrella 1	0.99251	0.90268	0.91720	0.90256	0.79732	0.83815
Estrella 3	0.99251	0.88334	0.92164	0.90255	0.86671	0.83809
Estrella 5	0.99249	0.90317	0.92562	0.90255	0.86423	0.83818
Sin Filtro	0.98514	0.90334	0.92080	0.90272	0.86142	0.83879
Promedio	0.97732	0.89830	0.92175	0.90239	0.85136	0.83647
Des Estándar	0.04523	0.00745	0.00256	0.00058	0.02517	0.00534
Varianza	0.00205	0.00006	0.00001	0.00000	0.00063	0.00003

Filtro \ Clase	Girasoles	Invernadero de ajo	Masa de agua	Reforestación	Sin Vegetación	Trigo
Cruz 1	0.98928	0.99989	0.97773	0.94499	0.75629	0.94038
Cruz 3	0.98931	0.99991	0.97780	0.93415	0.76404	0.94038
Cruz 5	0.98928	0.99989	0.97772	0.92016	0.75381	0.94038
Cuadrado 1	0.98926	0.99989	0.97795	0.93954	0.76400	0.94040
Cuadrado 3	0.98931	0.99992	0.97772	0.94216	0.75983	0.94038
Cuadrado 5	0.98930	0.99992	0.97786	0.90761	0.76432	0.91540
Estrella 1	0.98928	0.99989	0.97772	0.93649	0.76398	0.93556
Estrella 3	0.98928	0.99989	0.97773	0.91732	0.59143	0.94040
Estrella 5	0.98928	0.99989	0.97773	0.91179	0.76515	0.91837
Sin Filtro	0.98930	0.99987	0.97773	0.95229	0.76399	0.94038
Promedio	0.98929	0.99989	0.97777	0.93065	0.74468	0.93520
Des Estándar	0.00002	0.00002	0.00008	0.01530	0.05399	0.00980
Varianza	0.00000	0.00000	0.00000	0.00023	0.00291	0.00010



Filtro \ Clase	Promedio	Des. Estándar	Varianza
Cruz 1	0.91772	0.07351	0.00540
Cruz 3	0.91889	0.07063	0.00499
Cruz 5	0.91213	0.07759	0.00602
Cuadrado 1	0.90703	0.06979	0.00487
Cuadrado 3	0.91887	0.07179	0.00515
Cuadrado 5	0.91304	0.07244	0.00525
Estrella 1	0.91278	0.07767	0.00603
Estrella 3	0.90174	0.11093	0.01230
Estrella 5	0.91570	0.07008	0.00491
Sin Filtro	0.91965	0.07081	0.00501
Promedio			
Des Estándar			
Varianza			

Por razones de espacio el cuadro se dividió en 3 partes.

Cuadro 9 Precisión de todos los filtros IMG_63_3

Filtro \ Clase	Ajo	Barbecho	Experimental Plot	Girasoles	Invernadero de ajo	Maíz
Cruz 1	0.99885	0.98039	0.99400	0.92927	0.99591	0.67656
Cruz 3	0.99886	0.98039	0.99400	0.92907	0.99593	0.66854
Cruz 5	0.99886	0.98038	0.99401	0.92550	0.99591	0.66434
Cuadrado 1	0.99885	0.98039	0.99401	0.92930	0.99591	0.66946
Cuadrado 3	0.99886	0.98038	0.99400	0.92859	0.99591	0.67257
Cuadrado 5	0.99885	0.97591	0.99400	0.92920	0.99591	0.66124
Estrella 1	0.99886	0.98038	0.99400	0.92927	0.99591	0.67030
Estrella 3	0.99883	0.98039	0.99401	0.92924	0.99593	0.67740
Estrella 5	0.99886	0.98039	0.99400	0.92924	0.99593	0.66983
Sin Filtro	0.99886	0.98039	0.99400	0.92948	0.99591	0.67682
Promedio	0.99886	0.97994	0.99400	0.92881	0.99592	0.67071
Des Estándar	0.00001	0.00142	0.00001	0.00119	0.00001	0.00535
Varianza	0.00000	0.00000	0.00000	0.00000	0.00000	0.00003



Filtro \ Clase	Masa de agua	Pasto	Reforestación	Sin Vegetación	Vid
Cruz 1	0.99458	0.90005	0.94440	0.85484	0.87186
Cruz 3	0.99430	0.90007	0.92538	0.85481	0.85748
Cruz 5	0.99458	0.84782	0.95413	0.83739	0.84948
Cuadrado 1	0.98719	0.90002	0.94350	0.85403	0.85859
Cuadrado 3	0.99460	0.88146	0.92261	0.85484	0.83972
Cuadrado 5	0.99460	0.89910	0.94611	0.85483	0.82826
Estrella 1	0.97119	0.90007	0.94929	0.85213	0.87124
Estrella 3	0.99457	0.90168	0.96622	0.85485	0.83297
Estrella 5	0.99460	0.88775	0.95829	0.85480	0.89328
Sin Filtro	0.99460	0.90005	0.94587	0.85508	0.86578
Promedio	0.99148	0.89181	0.94558	0.85276	0.85686
Des Estándar	0.00750	0.01683	0.01340	0.00547	0.01994
Varianza	0.00006	0.00028	0.00018	0.00003	0.00040

Filtro \ Clase	Promedio	Des. Estándar	Varianza
Cruz 1	0.92188	0.09671	0.00935
Cruz 3	0.91808	0.09924	0.00985
Cruz 5	0.91295	0.10473	0.01097
Cuadrado 1	0.91920	0.09877	0.00975
Cuadrado 3	0.91487	0.09993	0.00999
Cuadrado 5	0.91618	0.10335	0.01068
Estrella 1	0.91933	0.09716	0.00944
Estrella 3	0.92055	0.09995	0.00999
Estrella 5	0.92336	0.09833	0.00967
Sin Filtro	0.92153	0.09700	0.00941
Promedio			
Des Estándar			
Varianza			

El cuadro se dividió en 3 partes por el espacio reducido.

8.2 Graficas

En los Cuadro 7, 8 y 9 se observa que la varianza general de la precisión en las tres imágenes son: 0.01158, 0.00147 y 0.00077 respectivamente, con esto se identifica que la imagen IMG_63_3 no existe diferencia alguna entre elegir algún filtro o no elegir filtro, y no existe diferencia al



clasificar cualquier tipo de cultivo. Esto no indica que se tendrá siempre la mejor precisión, sino que esta precisión será constante en todos los experimentos.

En la imagen IMG_63_3, se encuentra el cultivo de Maíz, este presenta una precisión promedio de 67% siendo este un resultado desfavorable para el experimento, mientras que las otras clases logran precisiones favorables y aun en esta imagen la varianza general es de 0.0007 lo cual indica que el comportamiento del SVM es constante en toda la imagen.

En la gráfica de la Figura 13 contiene los valores de las clases y los promedios obtenidos en cada clase, el apartado de las clases contiene un número, el cual representa a la imagen que pertenece dicha clase ejemplo: Ajo 2 (La primera parte indica el tipo de cultivo, mientras que el número indica a que imagen en este caso a la imagen IMG_63_2).

En la Figura 18 se observa que en la clase Cebolla de la IMG_63_1, Maíz IMG_63_3 y Sin Vegetación IMG_63_2 han tenido precisiones desfavorables; en las clases de Cebolla y Sin Vegetación este fenómeno se encuentra asociado al proceso de división de la imagen, por el cual el fragmento de imagen no posee los pixeles frontera que caracterizan a este tipo de clase. Mientras que la clase Maíz, no se da este fenómeno porque la clase solo existe en la imagen IMG_63_3, lo cual quiere indicar que los parámetros del SVM no fueron los adecuados para clasificar este cultivo.

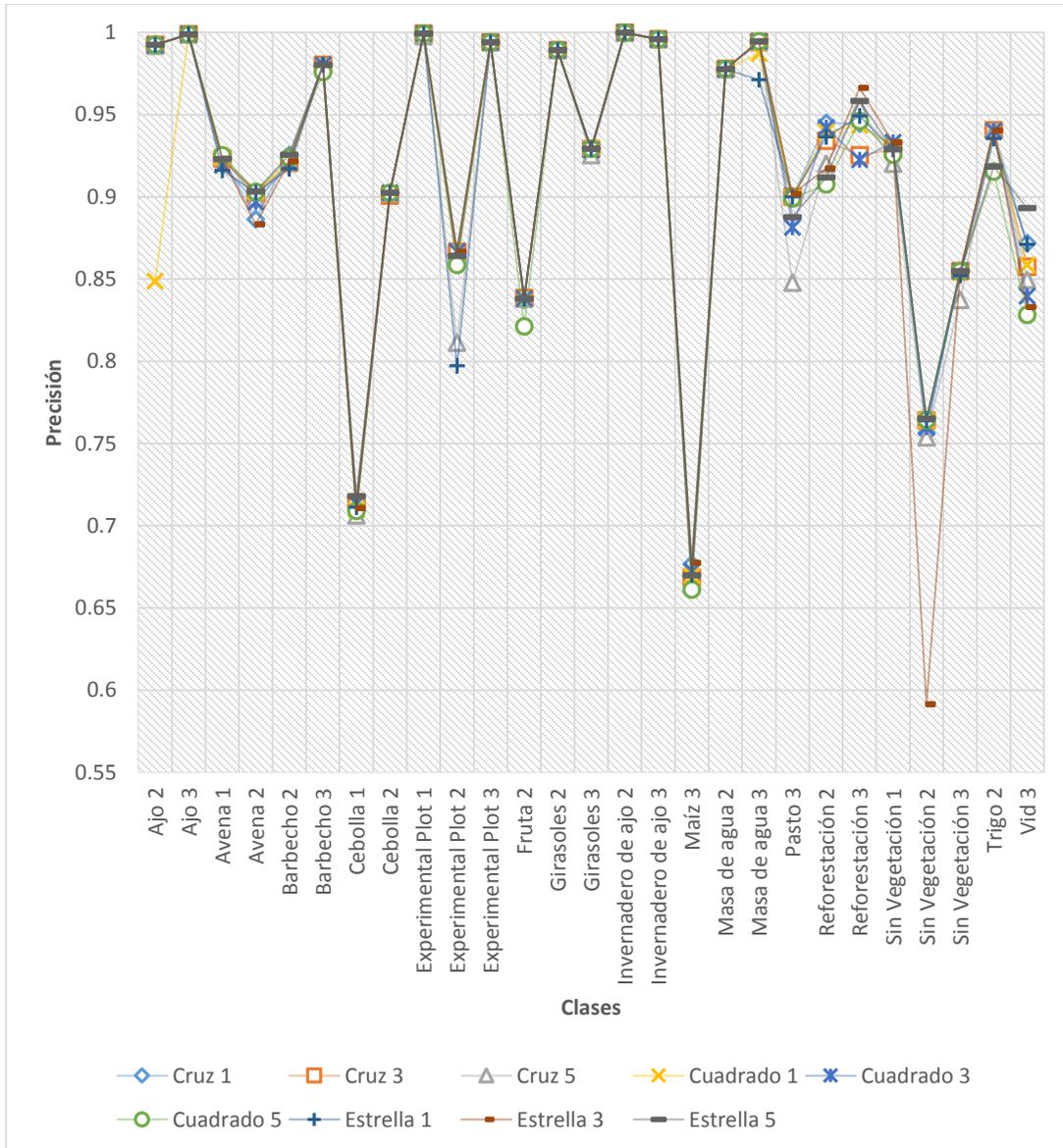


Figura 18 Grafica de resultados generales de las 3 imágenes

En las gráficas de las Figuras 19, 20 y 21, muestran los filtros agrupados por el tipo de filtro, donde se puede comprobar que no hay una gran variación con respecto de la precisión en cada cultivo y que solo se presentan pequeñas variaciones (ver Cuadro 10):

Cuadro 10 Pequeñas variaciones en precisiones

Filtro	Clases con pequeñas variación
Cruz	Avena 1, Experimental plot 2, pasto 3, reforestación 2 y reforestación 3.
Cuadrado	Ajo 2, pasto 3, reforestación 2 y reforestacion3.
Estrella	Avena 2, experimental plot 2, pasto 3, reforestación 2, reforestación 3, sin vegetación 2 y vid 3.

En las clases de: Pasto 3, Reforestación 2 y Reforestación 3 este fenómeno se repite en los tres tipos de filtros, esto se debe a que los pixeles son mezclas (Paz, 2009) lo cual indica que hay más de una sola clase dentro de cada pixel, por lo que el SVM genera resultados variables en cuanto a precisión.

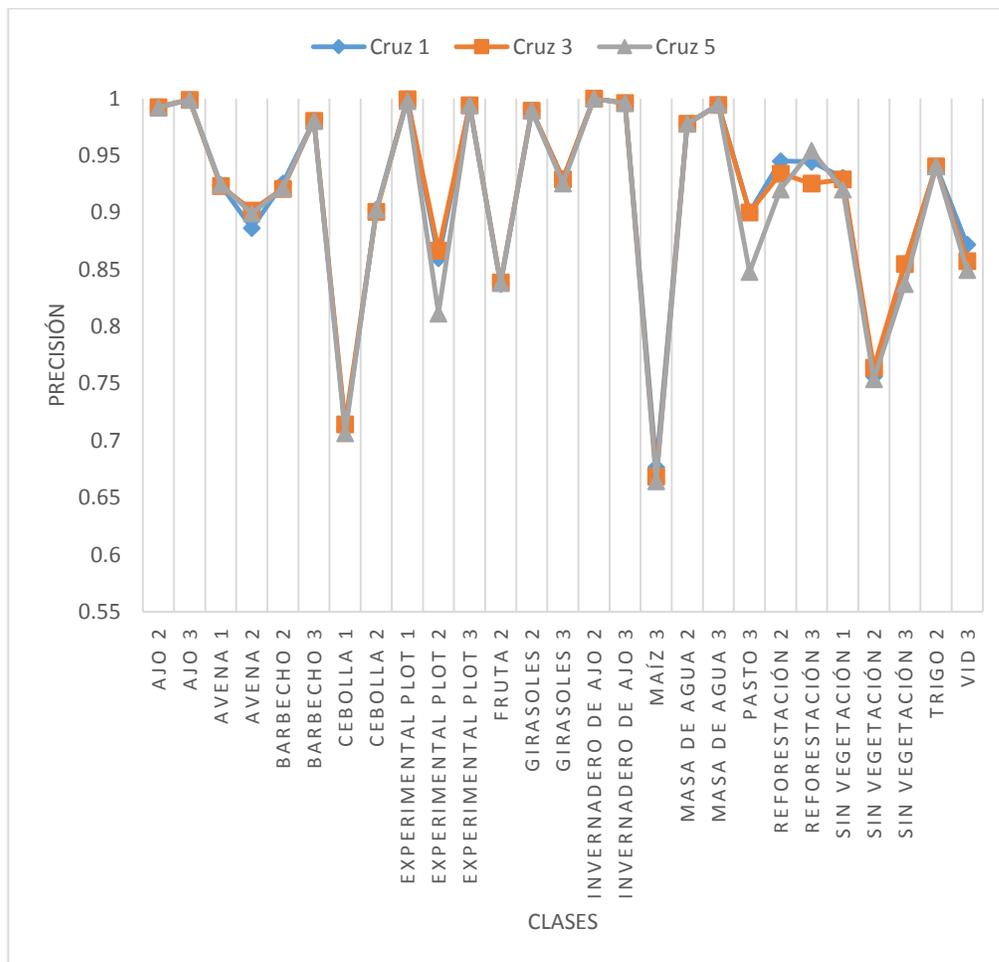


Figura 19 Grafica de resultados filtro Cruz

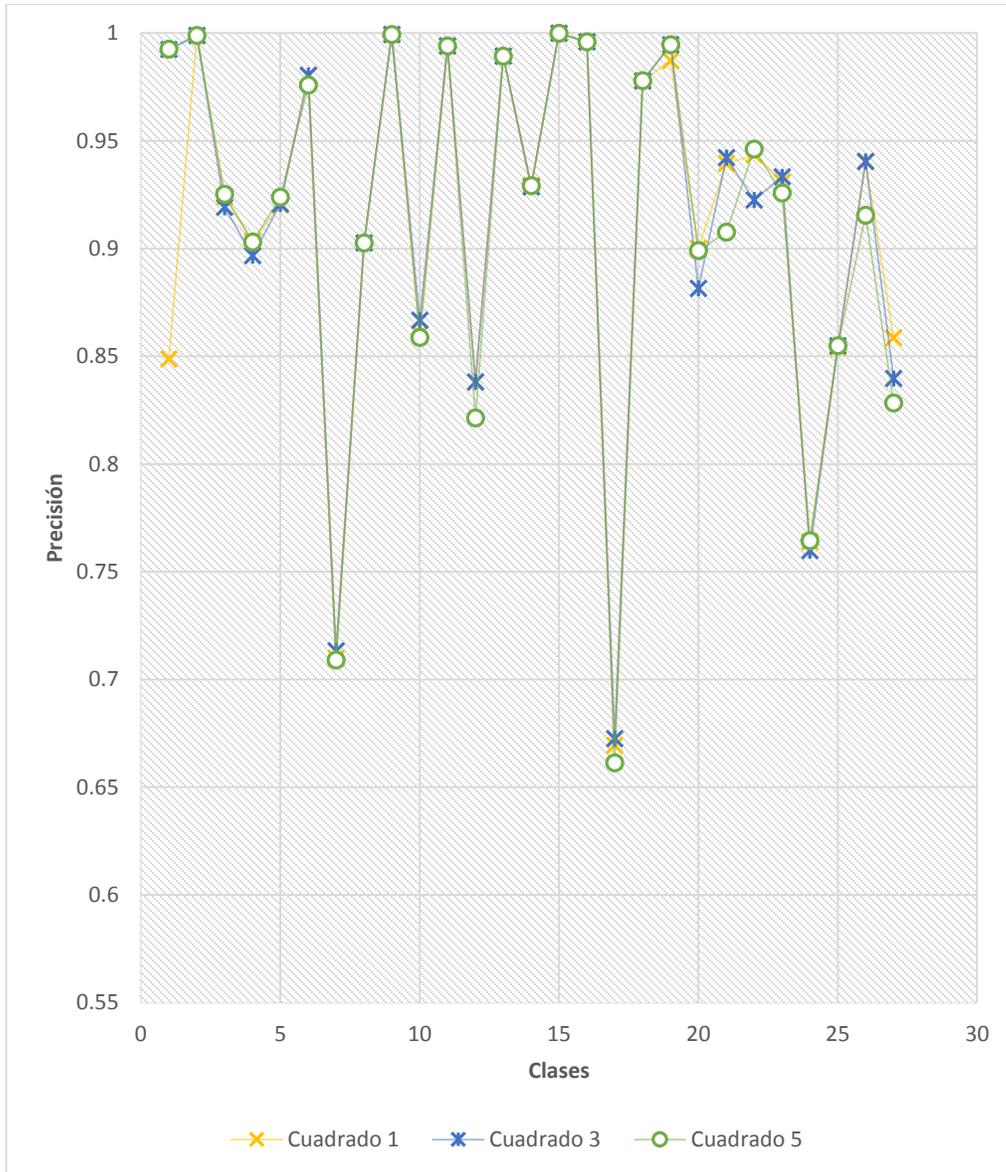


Figura 20 Grafica de resultados filtro Cuadrado

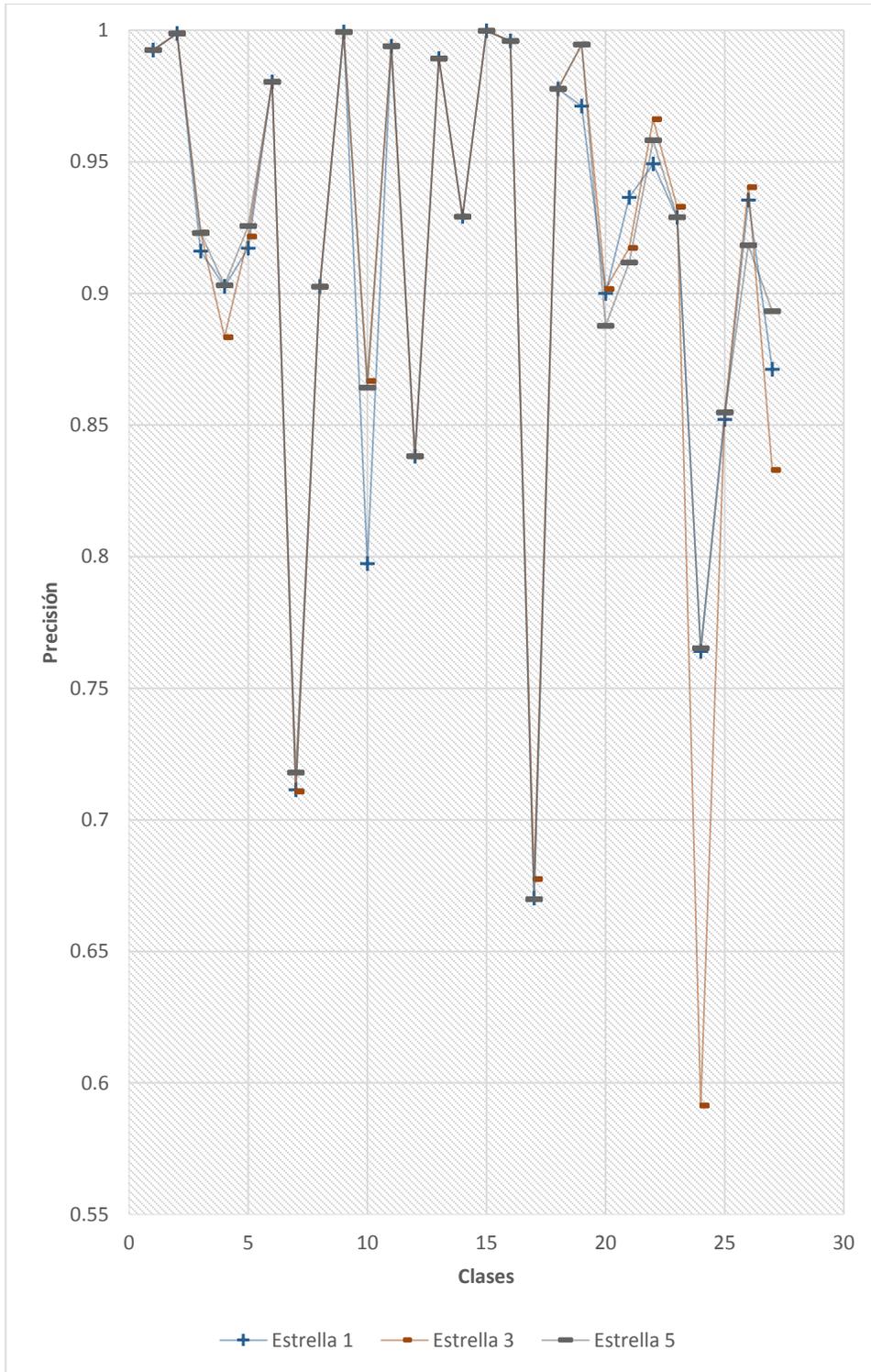


Figura 21 Grafica de resultados filtro Estrella

Una forma generalizada de ver las Figuras 19, 20 y 21, se encuentra en la Figura 22, la cual contiene una gráfica en forma radial, esta grafica contiene al centro las presiones más bajas, las presiones altas se alejan del centro del círculo y al contorno del semicírculo final se encuentran las etiquetas de las clases del experimento.

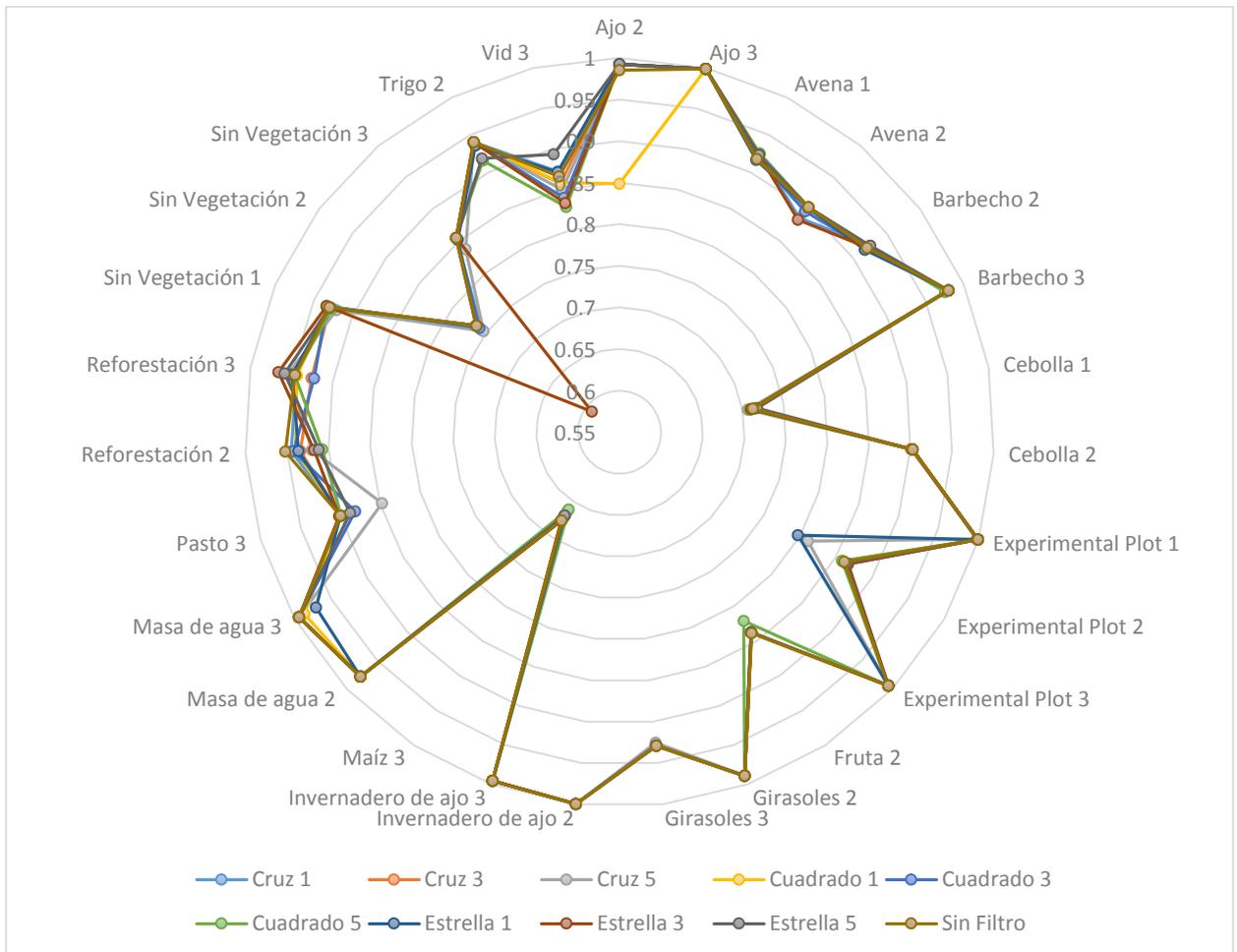


Figura 22 Grafica radial de promedios

Por ultimo en la Figura 23 contiene una gráfica de forma radial, para visualizar que el comportamiento de todos los filtros es constante en relación a una sola clase y aunque existen variaciones estas son mínimas.

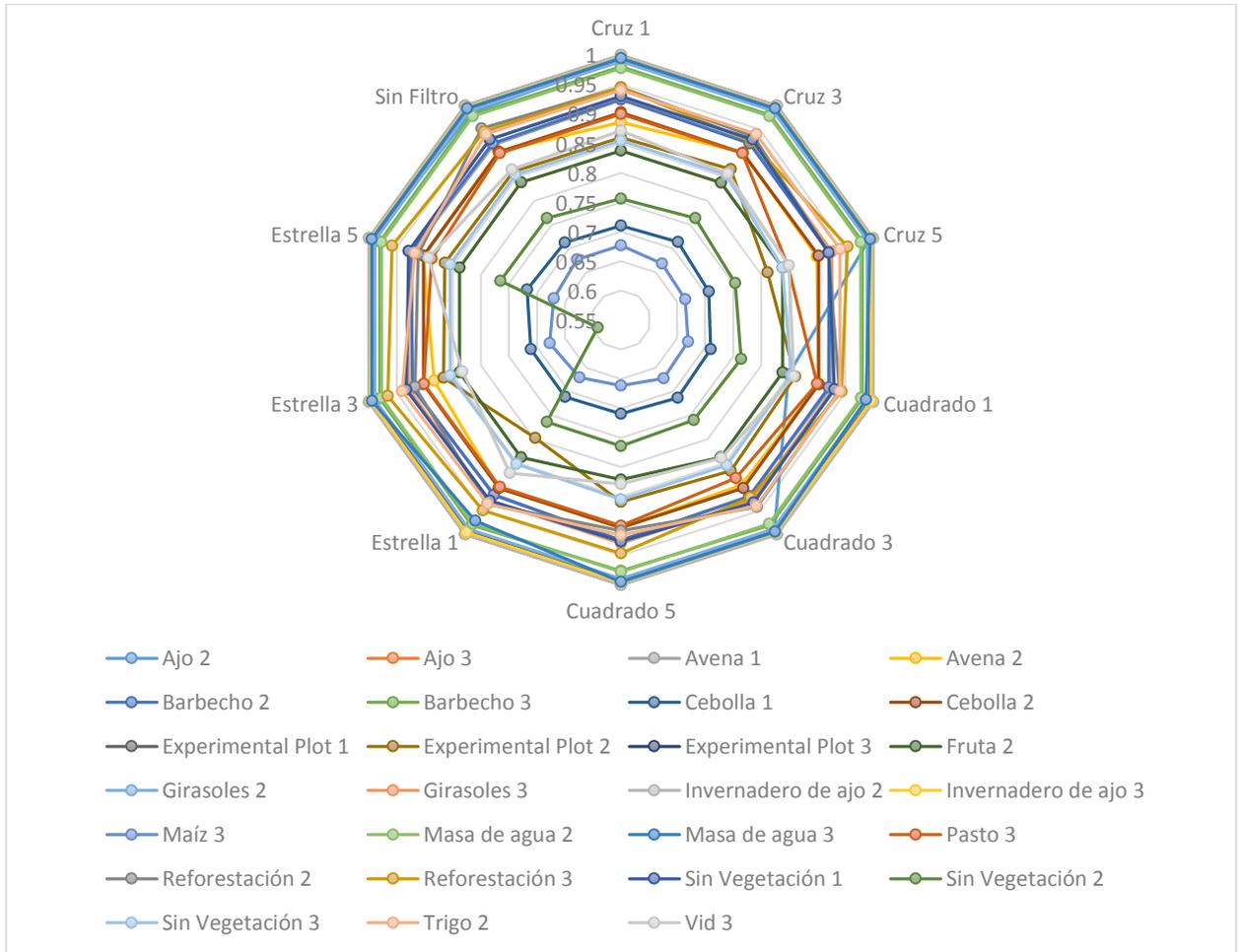


Figura 23 Grafica radial de cultivos.

8.3 Parámetros SVM Sigmoid

El SVM con *kernel Sigmoid* necesita de 2 parámetros k, θ el cuadro 11 se contabilizó la cantidad de apariciones k, θ en pares con el filtro Estrella magnitud 5.



Cuadro 11 Parámetros k, θ para Estrella 5.

Numero de repeticiones	Parámetros k, θ
7	0.1,-0.1
1	0.1,-0.5
1	0.1,-0.6
2	0.1,-0.8
1	0.1,-1
1	0.2,-0.1
1	0.2,-0.4
1	0.2,-0.6
1	0.2,-0.9
1	0.2,-1
1	0.4,-0.3
1	0.4,-0.4
2	0.4,-0.9
1	0.5,-0.3
1	0.5,-1
1	0.6,-0.1
1	0.7,-0.2
1	0.8,-0.1
1	1,-0.4

8.4 Kappa, sensibilidad y especificidad

Empleando otros estadísticos aparte de la precisión, a continuación en la cuadro 12, se muestra los datos de la imagen IMG_63_1, con: índice Kappa, sensibilidad y especificidad.



Cuadro 12 Estadísticos de imagen IMG_63_1.

Filtros	Avena 1 Numero de pixeles: 24954			Cebolla 1 Numero de pixeles: 44143			Experimental Plot 1 Numero de pixeles: 48			Sin Vegetación 1 Numero de pixeles: 5288		
	K	S	E	K	S	E	K	S	E	K	S	E
Cruz 1	0.5361	0.8756	0.6458	0.4266	0.6378	0.7851	0.0016	0.9995	0.0014	0.0000	0.8996	0.0461
Cruz 3	0.5274	0.8740	0.6388	0.4146	0.6280	0.7831	0.0002	0.9995	0.0007	0.1509	0.9553	0.1590
Cruz 5	0.5308	0.8746	0.6415	0.4130	0.6269	0.7826	0.0000	0.9993	0.0000	0.1507	0.9549	0.1592
Cuadrado 1	0.5208	0.8730	0.6333	0.4120	0.6252	0.7835	0.0000	0.9993	0.0000	0.0000	0.9031	0.0474
Cuadrado 3	0.5296	0.8751	0.6398	0.4241	0.6369	0.7835	0.0007	0.9994	0.0010	0.1861	0.9603	0.1805
Cuadrado 5	0.5350	0.8768	0.6434	0.4308	0.6370	0.7902	0.0004	0.9994	0.0008	0.0000	0.8978	0.0439
Estrella 1	0.5222	0.8755	0.6321	0.4074	0.6233	0.7807	0.0000	0.9993	0.0002	0.0000	0.9002	0.0483
Estrella 3	0.5279	0.8725	0.6409	0.3788	0.6077	0.7681	0.0004	0.9994	0.0008	0.0000	0.8998	0.0464
Estrella 5	0.5429	0.8767	0.6515	0.4297	0.6385	0.7874	0.0010	0.9994	0.0011	0.1154	0.9484	0.1390
Sin filtro	0.5326	0.8761	0.6417	0.4040	0.6233	0.7772	0.0000	0.9992	0.0005	0.0000	0.9012	0.0487

Cuadro de estadísticos de la imagen IMG_63_1, donde K = kappa, S= sensibilidad y E = especificidad

En esta figura se observa que el estadístico de especificidad, presenta los valores mínimos, cuando la clase disminuye el número de pixeles, situación similar al índice Kappa. Con esto es posible generalizar que el desbalance de las clases provoca que el SVM se comporte de manera errática, pero aun así sigue siendo constante con cada filtro que se emplea.

8.5 Imágenes

Las 3 imágenes que se generan al dividir la imagen IMG_63, producen un desbalance de pixeles, de tal manera que el SVM genera un sesgo, esto ocasiona una precisión favorable cerca del 100%, pero no logra identificar ningún objeto al momento de clasificar la imagen, es decir todos los pixeles no pertenecen a esta clase.

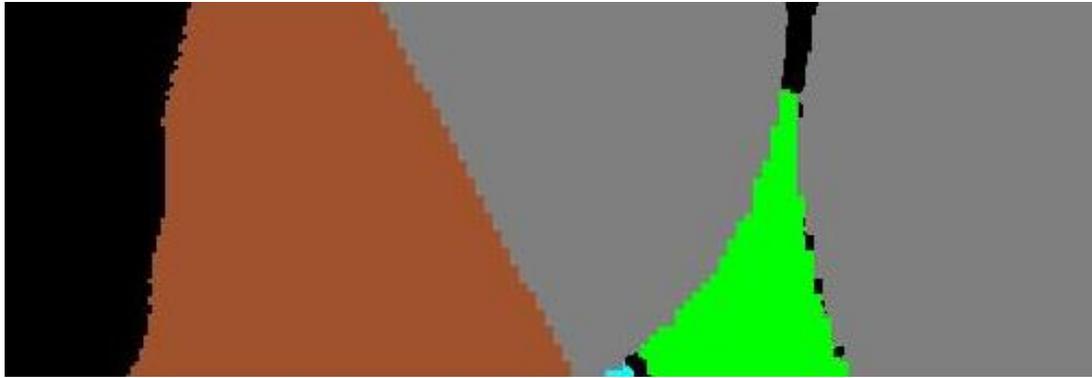


Figura 24 Ejemplo de imagen IMG_63_1.

Imagen donde se muestran 5 clases a continuación se enlistan las clases en orden de aparición de izquierda a derecha: primera clase datos sin interés o no clasificado, segunda clase “Avena”, tercera clase con 2 semicírculos “Cebolla”, cuarta clase entre las clases de cebolla se encuentra la clase de “Sin Vegetación”, y por último al centro en la parte de abajo se localiza la clase “Experimental Plot” de color azul.

Retomando la parte del sesgo de SVM, en la figura 25 se observa un ejemplo claro del desbalance de clases al momento de clasificar.

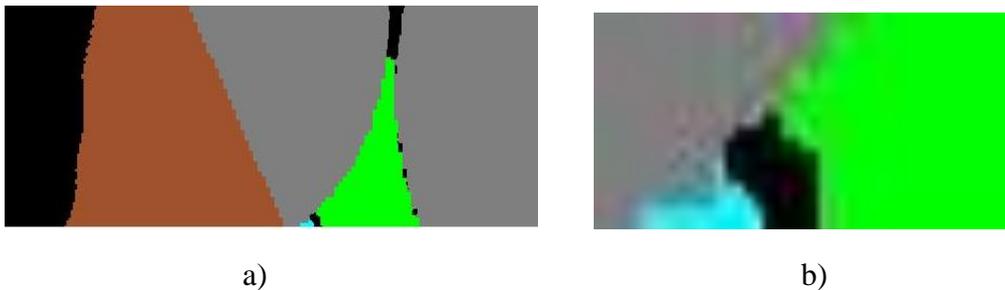
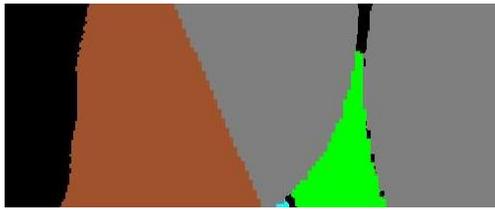


Figura 25 Ejemplo IMG_63_1 clase Experimental Plot

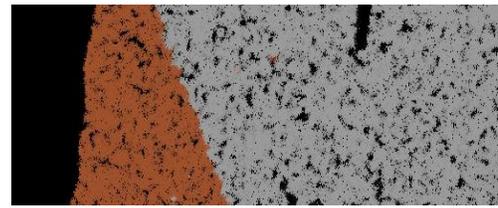
- a) Imagen de referencia o caracterizada de forma manual.
- b) Ampliación de la imagen a) donde se muestra la pequeña clase de Experimental Plot.

Continuando con el análisis, al introducir la imagen IMG_63_1, con un filtro Cruz 1, se obtiene como resultado 4 imágenes, cada imagen obtenida corresponde a las clases: Avena, Cebolla, Sin Vegetación y Experimental Plot. En las figuras 26, 27, 28 y 29 muestran gráficamente el resultado de la clasificación.

En la figura 29 muestra que al tener pocos pixeles el SVM ha producido un sesgo considerando que ningún pixel pertenece a esta clase “Experimental plot”.



a)

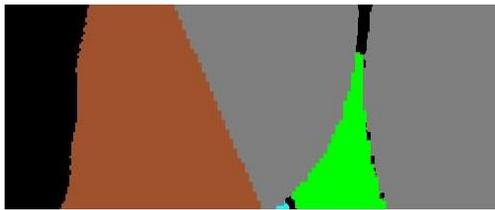


b)

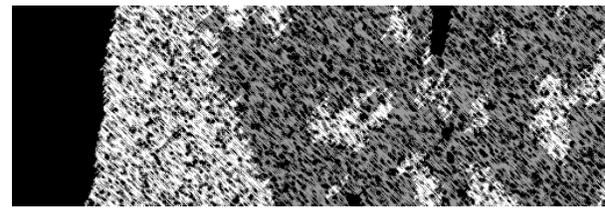
Figura 26 Resultado IMG_63_1 con filtro Cruz 1 clase Avena.

a) Imagen de referencia o caracterizada de forma manual.

b) Imagen resultado del SVM con filtro cruz magnitud 1 el color oscuro indica la clase Avena y el color claro indica todo lo que no es avena, el color negro indica lo que fue utilizado para entrenar y no considerado en la imagen.



a)

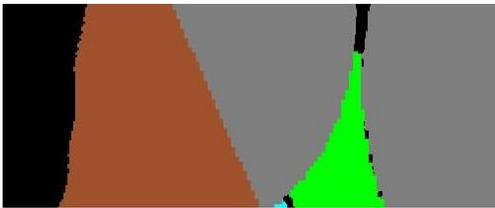


b)

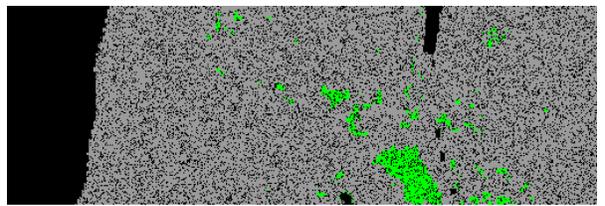
Figura 27 Resultado IMG_63_1 con filtro Cruz 1 clase Cebolla.

a) Imagen de referencia o caracterizada de forma manual.

b) Imagen resultado del SVM con filtro cruz magnitud 1 el color oscuro indica la clase Cebolla y el color claro indica todo lo que no es avena, el color negro indica lo que fue utilizado para entrenar y no considerado en la imagen.



a)



b)

Figura 28 Resultados IMG_63_1 con filtro Cruz 1 clase Sin Vegetación.

a) Imagen de referencia o caracterizada de forma manual.

b) Imagen resultado del SVM con filtro cruz magnitud 1 el color oscuro indica la clase Sin Vegetación y el color claro indica todo lo que no es avena, el color negro indica lo que fue utilizado para entrenar y no considerado en la imagen.

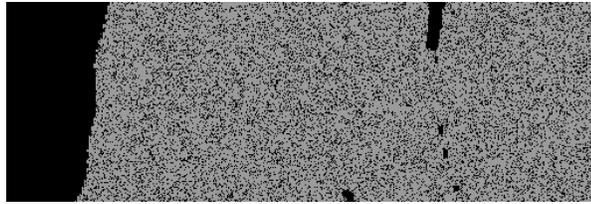


Figura 29 Resultado IMG_63_1 con filtro Cruz 1 clase Experimental Plot

La imagen resultado no presenta ningún elemento perteneciente a esta clase.

En la misma imagen pero con el filtro Cuadrado 1 se muestran los resultados obtenidos de las 4 clases en la Figura 30

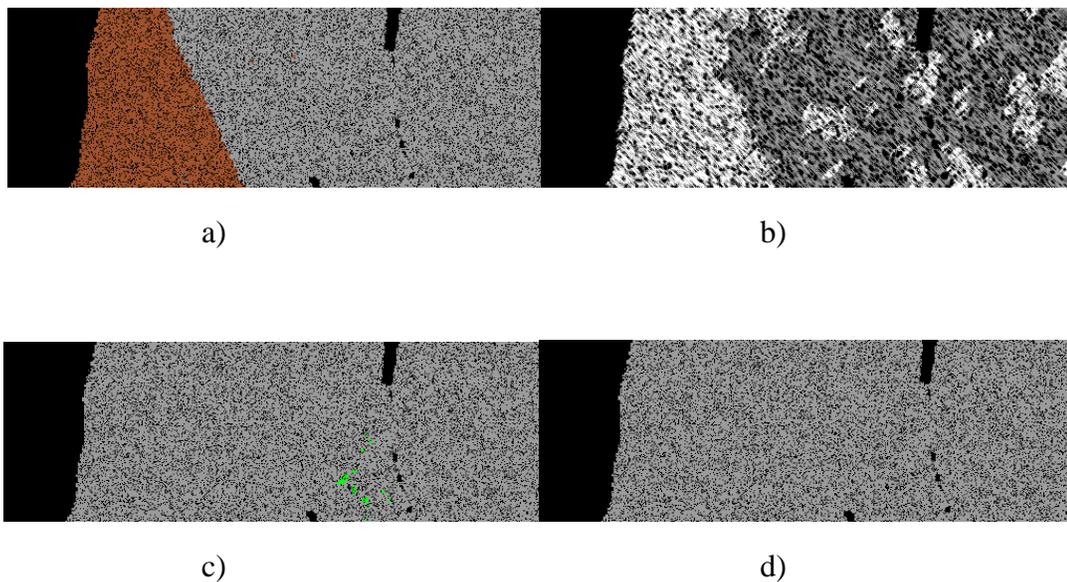


Figura 30 Resultados IMG_63_1 con filtro Cuadrado 3.

- a) Imagen resultado del SVM con filtro cuadrado magnitud 3 el color obscuro indica la clase Avena y el color claro indica todo lo que no es avena, el color negro indica lo que fue utilizado para entrenar y no considerado en la imagen.
- b) Imagen resultado del SVM con filtro cuadrado magnitud 3 el color obscuro indica la clase Cebolla y el color claro indica todo lo que no es avena, el color negro indica lo que fue utilizado para entrenar y no considerado en la imagen.
- c) Imagen resultado del SVM con filtro cuadrado magnitud 3 el color obscuro indica la clase Sin Vegetación y el color claro indica todo lo que no es avena, el color negro indica lo que fue utilizado para entrenar y no considerado en la imagen.
- d) Imagen resultado del SVM con filtro cuadrado magnitud 3 el color obscuro indica la clase Experimental Plot y el color claro indica todo lo que no es avena, el color negro indica lo que fue utilizado para entrenar y no considerado en la imagen.

Por último presentamos la misma imagen pero ahora con un filtro de estrella, magnitud 5

(ver figura 31).

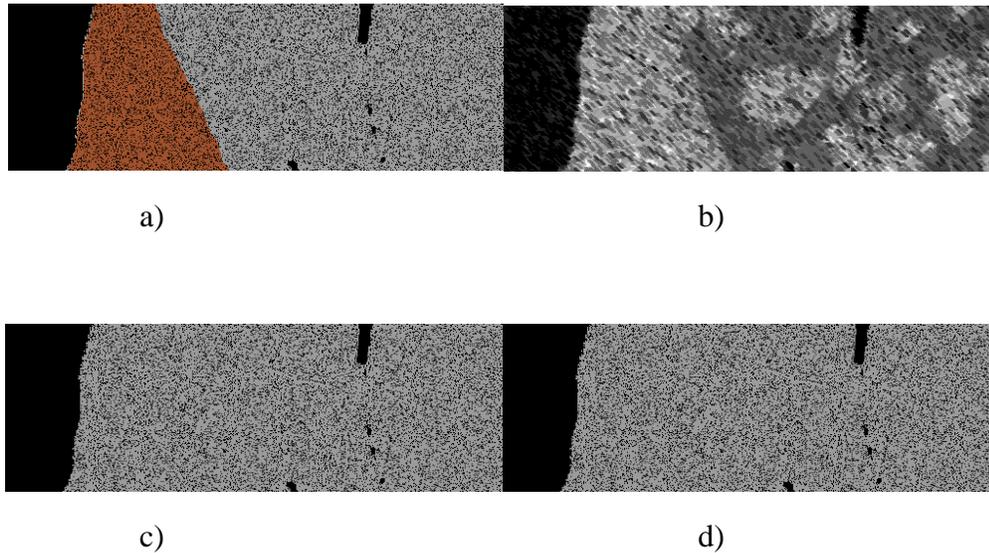


Figura 31 Imagen IMG_63_1 filtro Cuadrado 3.

- a) Imagen resultado del SVM con filtro estrella magnitud 5 el color obscuro indica la clase Avena y el color claro indica todo lo que no es avena, el color negro indica lo que fue utilizado para entrenar y no considerado en la imagen.
- b) Imagen resultado del SVM con filtro estrella magnitud 5 el color obscuro indica la clase Cebolla y el color claro indica todo lo que no es avena, el color negro indica lo que fue utilizado para entrenar y no considerado en la imagen.
- c) Imagen resultado del SVM con filtro estrella magnitud 5 el color obscuro indica la clase Sin Vegetación y el color claro indica todo lo que no es avena, el color negro indica lo que fue utilizado para entrenar y no considerado en la imagen.
- d) Imagen resultado del SVM con filtro estrella magnitud 5 el color obscuro indica la clase Experimental Plot y el color claro indica todo lo que no es avena, el color negro indica lo que fue utilizado para entrenar y no considerado en la imagen.



IX DISCUSIÓN

Tomando como referencia los resultados del Cuadro 8, hay que comentar que en general las precisiones son buenas, pero, es conveniente descartar la clasificación de 2 clases, las obtenidas en la clase “Invernadero con ajo” en todos los tipos de filtro, ya que las precisiones se aproximan a 1, sin embargo, hay que aclarar que solo se clasificaron 8 píxeles, lo que hace que estos resultados sean casi perfectos (≈ 1), esta clasificación debe de tomarse con todas las reservas para aseverar que la precisión es excelente por la cantidad de píxeles empleados. En contraste, para la clasificación “sin vegetación” es la más pobremente clasificada con cualquier filtro, es de destacar que es la que tiene más píxeles de todas las clases, esto es digno de tomarse en cuenta ya que baja sustancialmente la precisión en comparación a las otras clases. Con ello podríamos deducir que la SVM no en todos los casos tiene un buen desempeño, esto posiblemente se deba a lo que comenta (Aisen, 2006) al referirse a que la SVM crean un espacio en el hiper plano donde se alojan datos sin ser clasificados correctamente. Para aseverar esto habría que hacer más pruebas.

Otro aspecto que es digno de destacar es que en todas las clases existe una buena precisión, lo que hace pensar que los datos que proporciona una imagen hiper-espectral son mejores que los que se tienen con otro tipo de imagen, esto también se puede reafirmar si se observa que las clasificaciones de la clase “sin filtro” tienen el mismo comportamiento que las otras, incluso llegando a obtener mejores clasificaciones que algunos filtros, véase el cuadro 4, donde se puede ver claramente como los valores de la desviación estándar que se obtuvieron son muy pequeños, esto ya lo comentan (Rojas, 2009) y (Paz, 2009), esto es importante porque el filtrado consume mucho tiempo de uso de computadora.



Sin embargo si se requiere saber cuál es el mejor filtro aunque la diferencia sea por centésimas, el análisis demuestra que el no uso de filtro (sin filtro) es que obtuvo mejores precisiones en la mayoría de clases.

Las imágenes AHS presentan una resolución de 6 metros por pixel (Rejas, et al, 2003) a diferencia de imágenes como las de IKONOS que presentan una resolución de 1 metro por pixel (<http://www.aeroterra.com>) o las imágenes de QUICKBIRD que poseen una resolución de .6 metros por pixel (<http://www.satimagingcorp.es>).

La imagen IMG_63 fue un reto para ser procesada debido a que se contaba con 52,429 vectores de pixeles para entrenar y la SVM debía crear para el peor de los casos una matriz *kernel* de (52,429 x 52429) lo cual produce un consumo de memoria de 21.99107142432 Gigabytes sin contar con los recursos que consumen los demás procedimientos. Es por esto que la imagen se dividió en 3 partes para que cada matriz *kernel* solo necesitara de un máximo de 2.48196760352 Gigabytes (Boettcher, 2007).

$$8(n + 1)^2 \quad (38)$$

Donde n es el número de vectores que se introduce al SVM y el resultado está expresado en bytes.

Con respecto al uso de memoria RAM en la investigación se proponen 2 mejoras, la primera la adquisición de más memoria RAM y la segunda generar un algoritmo para elegir los puntos de clasificación, buscando un ajuste en estos para que los pixeles de entrenamiento sean más descriptivos por clase, además de permitir remplazos y usar lo mínimo de memoria RAM.

El equipo de cómputo que se ha empleado en este trabajo es de uso común, con esto se demuestra que hoy en día, con esta tecnología se pueden realizar trabajos científicos y que aun con



un presupuesto limitado es posible obtener resultados como los de (Trejo & Hernández, 1996), que emplearon un servidor dedicado al trabajo.

Al realizar la cobertura del suelo se han realizado varios trabajos como (Trejo & Hernández, 1996) (Kawamura, Tsujino, & Tsujiko, 2004) (Pugh & Waxman, 2006), en sus trabajos se emplea el mismo concepto, crear “Windows” (filtros cuadrados), para realizar la clasificación, pero en este trabajo se ha inferido con que el uso de pixeles, Filtros y SVM, se logra una buena precisión, además los resultados sugieren que el uso de “ventanas” no es la mejor opción para filtrar una imagen. Con la metodología propuesta al mismo tiempo se obtiene una segmentación cuando cada clase es coloreada, sin necesidad de realizar procedimientos alternos.

Para presentar resultados en una imagen coloreada que permite identificar cuáles son las zonas conflictivas para el clasificador, la técnica OAA-SVM presenta un inconveniente, debido a que más del 25% de los pixeles fueron clasificados por diferentes máquinas indicando que estos pixeles pertenecían a varias clases como lo comentan (Aisen, 2006), OAA-SVM genera regiones de espacio de características indeciso.

El *kernel* Sigmoid, presenta un buen desempeño, sin necesidad de realizar una mayor profundidad en la búsqueda de los parámetros k y θ , con esto se obtiene que el mejor parámetro para Sigmoid está dado por [0.1 y -0.1].

Los filtros no presentan una tendencia lineal, y es poco ortodoxo establecer un filtro para cada tipo de clase que se desea clasificar, con esto en mente se considera poder trabajar sin ningún filtro, pero si es necesario aumentar la precisión sin considerar el costo computacional que este proceso conlleva, se recomienda el uso del filtro Estrella de magnitud 5, en este experimento el tiempo promedio para obtener este filtro las imágenes IMG_63 fue de 5 horas por cada imagen.



La investigación permitió generar un marco de trabajo para plantear otros experimentos para clasificar la cobertura de suelo, con el fin de mejorar estas clasificaciones, al proponer filtros paralelos o introducir nuevos filtros, así mismo, probar con otras técnicas de SVM para multi-clasificación como; uno contra uno (one against one OAO) de acuerdo a (Milgram, Cheriet, & Sabourin, 2004), o con SPIDER SVM descrito en (Weston, Elisseeff, BakIr, & Sinz, 2004) entre otras técnicas.

La técnica OAA-SVM, no es adecuada para etiquetar imágenes con respecto a la cobertura de suelo, debido a que la cobertura de suelo contiene una gran cantidad de etiquetas, y genera un espacio de características indeciso al momento de etiquetar, quedando un píxel clasificado en varias clases o en ninguna.



X CONCLUSIONES

Los pixeles derivados de imágenes híper-espectrales, tienen características deseables para clasificar la cobertura vegetal del suelo al tener una firma espectral única y diferenciable de las demás firmas.

Las precisiones obtenidas en un promedio de 92% sugieren que las SVM con imágenes híper-espectrales y usando pixeles como elemento de clasificación son adecuados para clasificar cobertura vegetal de suelo.

Se recomienda repetir el experimento en más imágenes para determinar si el uso de filtros puede ser excluido y mejorar los tiempos de uso de la computadora.

La técnica OAA-SVM, no es adecuada para etiquetar imágenes con respecto a la cobertura de suelo, debido a que la cobertura de suelo contiene una gran cantidad de etiquetas, y genera un espacio de características indeciso al momento de etiquetar, quedando un pixel clasificado en varias clases o en ninguna.

Se requiere repetir este experimento con otras imágenes para aseverar que los resultados obtenidos son generalizables, además de agregar un nuevo módulo que permita balancear los pixeles al entrenamiento así como trabajar con toda la imagen.



XI BIBLIOGRAFÍA

- Aisen, B. (2006). A Comparison of Multiclass SVM Methods. *Neural Networks, IEEE Transactions on*, 13, 415-425. doi:10.1109/72.991427
- Al-Awadhi, T. (2007). Monitoring and Modeling Urban Expansion Using GIS & RS: Case Study from Muscat, Oman. *Urban Remote Sensing Joint Event*, 1,5, 11-13.
- Alberola-Lopez, C., Martin-Fernandez, M., & Ruiz-Alzola, J. (2004). A methodology for evaluation of boundary detection algorithms on medical images. *IEEE Transactions on Medical Imaging*, 23(5), 658-660.
- Ali, W., Georgsson, F., & Hellstrom, T. (2008). Visual Tree Detection for Autonomous Navigation in Forest Environment. *Intelligent Vehicles Symposium, IEEE*, 560-565.
- Andrew, R., Webb, K., & Copsey, D. (2011). *Statistical Pattern Recognition*. Wiley.
- Atienza, V. (2011). *El histograma de una imagen digital*. Valencia, España: Universitat Politecnica de Valencia.
- Ayat, N. E., Cheriet, M., & Suen, C. Y. (2002). KMOD - a two-parameter SVM kernel for pattern recognition. *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, 3, 331-334.
- Ballard, D. H., & Brown, C. M. (1982). *Computer Vision*. NJ: Prentice-Hall.
- Basogain, X. (2008). *Redes Neuronales Artificiales y sus Aplicaciones*. Bilbao: Campus Birtualeko azterketen.
- Baudat, G., & Anouar, F. (2001). Kernel-based Methods and Function Approximation. *International Joint Conference on Neural Networks*, 1244 - 1249.



- Beauchemin, M., & Thomson, K. B. (1997). The evaluation of segmentation results and the overlapping are matrix. *Remote Sensing*, 18, 389-3899.
- Bernabé, S. (2010). *Desarrollo de un sistema automático para búsqueda por contenido en imágenes de satélite*. España: Universidad de Extremadura.
- Boettcher, P. (2007). <http://stackoverflow.com>. Retrieved 07 12, 2013, from <http://stackoverflow.com/questions/15994222/out-of-memory-using-svmtrain-in-matlab>
- Bratasanu, D. . (2012). Spectral band discovery for advancing multispectral satellite image visual analysis. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 5, 207-224. doi:10.1109/JSTARS.2011.2169772
- Camapum, J. F., & Fisher, M. H. (2001). Multiscale Color And Texture Invariants For Image Recognition. *Image Processing, 2001. Proceedings. 2001 International Conference on*, 1, 862-865. doi:10.1109/ICIP.2001.959182
- Carcedo, A. (2004). *Programa de segmentación de regiones en imágenes médicas en MATLAB*. Puebla, México: Universidad de las Américas Puebla.
- Chapelle, O., Haffner, P., & Vapnik, V. N. (1999). Support vector machines for histogram-based image classification. *Neural Networks, IEEE Transactions on*, 10(5), 1055-1064. Retrieved from <http://www.earthboundlight.com/phototips/histogram.html>
- Cehade, N. H., Boureau, J.-G., Vidal, C., & Zerubia, J. (2009). Multi-Class SVM For Forestry Classification. *Image Processing (ICIP), 2009 16th IEEE International Conference on*, 1673-1676. doi:10.1109/ICIP.2009.5413395



Cristianni, N., & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods A Review*. Cambridge: Cambridge University Press.

D De Grandi, G., Lucas, R. M., & Kropacek, J. (2009). Analysis by Wavelet Frames of Spatial Statistics in SAR Data for Characterizing Structural Properties of Forests. *Geoscience and Remote Sensing, IEEE Transactions on*, 47, 494-507. doi:10.1109/TGRS.2008.2006183

Del Rio, B. M. (2006). *Redes Neuronales y Sistemas Difusos*. Alfaomega.

Deng, Y., & Manjunath, B. (2001). Unsupervised Segmentation of Color-Texture Regions in Images and Video. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23, 800-810. doi:10.1109/34.946985

Dengru, W., & Linders, J. (1999). A New Texture Approach to Discrimination of Forest Clearcut, Canopy, and Burned Area Using Airborne C-Band SAR. *Geoscience and Remote Sensing, IEEE Transactions on*, 37, 555-563. doi:10.1109/36.739113

Dumitru, C. (2012). Using SVM for Classification. *AUDCE*, 207-222.

Fernández, E., López, L., & Yubero, F. (2009). *Clasificación de texturas naturales mediante técnicas de visión por computador en imágenes aéreas*. Biblioteca Universidad Complutense.

Freitas, C., Lelio, W.-J., & Dutra, L. V. (2000). Assessment of Operational Radar Satellite for Monitoring Land Cover Change in Amazonia. *Geoscience and Remote Sensing Symposium, 2002. IGARSS '02. 2002 IEEE International*, 4, 2063-2065.

Gao, Y. (2008). *Comparación de Distintos Métodos de Clasificación Digital de Imágenes de Satélite*. Universidad Nacional Autónoma de México.



- Garcia, M. P., & Perez, M. E. (1999). Estudio mediante imágenes TM del entorno de Villacañas. *Revista de Teledetección, 11*, 2-8.
- Garpenter, G. A., Grossberg, S., Markuzon, N., Reynolds, J., & Rosen, D. B. (1992). Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. *Neural Networks, IEEE Transactions on, 3*, 698-713. doi:10.1109/72.159059
- Ghasemi, N., Sahebi, M. R., & Mohammadzadeh, A. (2013). Biomass Estimation of a Temperate Deciduous Forest Using Wavelet Analysis. *Geoscience and Remote Sensing, IEEE Transactions on, 51*, 765-776. doi:10.1109/TGRS.2012.2205260
- Gil, A., & Muñoz, M. (1975). *Aplicación Del Método De Máxima Verosimilitud Al Análisis Cuantal*. España: Ministerio de Agricultura. Instituto de Inveatigaciones Agrarias.
- Gonzalez, R., & Woods, R. (2002). *Digital Imagen Processing*. Upper Saddle River, New Jersey, USA: Prentice Hall.
- Haipeng, W., & Kazuo, O. (2010). A Simple Moment Method of Forest Biomass Estimation From Non-Gaussian Texture Information by High-Resolution Polarimetric SAR. *Geoscience and Remote Sensing Letters, IEEE, 7*, 811-815. doi:10.1109/LGRS.2010.2047839
- Hernández, J. F., & Hernández, L. (2003). Aplicación De Técnicas De Teledetección Al Estudio De La Turbidez De Las Aguas Litorales En La Costa Oriental De Gran Canaria Mediante Un Modelo Teórico. *Vegueta: Anuario de la Facultad de Geografía e Historia(7)*, 230-240.
- Herrada-Mateo, L., & Kober, V. (2002). *Mejoramiento de contraste local de imágenes con vecindarios espacialmente adaptivos*. México: Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE) .



Hsieh, P.-F., & Landgrebe, D. (1998). *Classification Of High Dimensional Data*. INDIANA , USA:
The Trustees of Princeton University.

<http://www.aeroterra.com>. (2013). Retrieved 11 23, 2013, from <http://www.aeroterra.com/p-ikonos.htm>

<http://www.geosoluciones.cl>. (2013, 11 25). Retrieved from
<http://www.geosoluciones.cl/quickbird/>

<http://www.inegi.org.mx>. (2013). Retrieved from
<http://www.inegi.org.mx/geo/contenidos/imgpercepcion/imgsatelite/landsat.aspx>

<http://www.tdx.cat>. (n.d.). Retrieved 11 27, 2013, from
http://www.tdx.cat/bitstream/handle/10803/6380/2_3_Satelites_con_SAR.pdf;jsessionid=20D413910A28602B04A626F76D52D3D9.tdx2?sequence=10

<http://www.um.es>. (2013, 11 29). Retrieved from
http://www.um.es/geograf/sigmur/temariohtml/node74_mn.html

Ilea, D. E., & Whelan, P. F. (2011). Image segmentation based on the integration of colour–texture descriptors—A review. *Pattern Recognition*, 44(10-11), 2479-2501.

Inventario Forestal 2010. (2013, 11 25). Retrieved from <http://portal2.edomex.gob.mx>:
http://portal2.edomex.gob.mx/probosque/publicaciones/inventario_forestal/groups/public/documents/edomex_archivo/probosque_pdf_inventario10_1.pdf

Isasi, P., & Galván, I. M. (2004). *Redes de neuronas Artificiales Un enfoque practico*. Madrid:
Pearson Education.



- Jain, A. K., Duin, R. P., & Jianchang, M. (2000). Statistical Pattern Recognition: a review. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 4-37.
- Jakeman, E. (1980). On the statistics of K-distributed noise. *Journal of Physics A: Mathematical and General*, 13, 31.
- Jaramillo, G. E., & Branch, J. W. (2008). DORIS: Sistema para la Recuperación de Imágenes de Piezas Mecánicas y de Automoción Utilizando Descriptores de Textura. *Revista Avances en Sistemas e Informática*.
- Jiménez, M. D. (2008). *Extracción De Características De Textura Basada En La Transformada Wavelet Discreta*. Sevilla, España.: Biblioteca de Ingeniería de la Universidad de Sevilla.
- Kaufmann, A. (1982). *Introduccion a la teoria de conjuntos borrosos*. Mexico: Continental.
- Kawamura, M., Tsujino, K., & Tsujiko, Y. (2004). Characteristic Analysis of High Resolution satellite Imagery for Forest Species Discrimination. *Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International*, 4, 2358-2361.
- Kim, C., & Hong, S.-H. (2008). The Characterization of a Forest Cover Through Shape and Texture Parameters from Quickbird Imagery. *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International*, 3, III - 692-III - 695.
- Koushal, K., & Gour, S. M. (2012). Advanced Applications of Neural Networks and Artificial Intelligence: A Review. *International Journal of Information Technology and Computer Science*, 4(6), 57-68.



- La Serna, P. N. (2010). Procesamiento Digital de textura: Técnicas utilizadas en aplicaciones actuales de CBIR. *Revista de Investigación de Sistemas e Informática Facultad de Ingeniería de Sistemas e Informática Universidad Nacional Mayor de San Marcos*, 57-64.
- Lin, W., Wang, Q., & Zha, S. (2010). *Construction and Application of Characteristic Bands of typical Land Cover based on Spectrum-photometric Method*. Wuhan, China: National Science Foundation of China.
- Lloyd, D. (1990). A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. *International Journal of Remote Sensing*, 11(12), 2269-2279.
- Milgram, J., Cheriet, M., & Sabourin, R. (2004). “One Against One” or “One Against All”: Which One is Better for Handwriting Recognition with SVMs? *In 10th International Workshop on Frontiers in Handwriting Recognition*.
- Miyer, A., Suarez, S. A., & González, M. (2011). Método Adaptivo de Descripción de Textura Utilizando el Patrón Espectrum y la Morfología Matemática. *Revista Colombiana de Tecnologías de Avanzada*, 1(17), 62-66.
- Myung, I. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, 47, 90–100.
- Nowak, E., Jurie, F., & Triggs, B. (2006). Sampling Strategies for Bag-of-Features Image Classification. In A. Leonardis, H. Bischof, & A. Pinz, *Computer Vision – ECCV 2006* (pp. 490-503). Springer Berlin Heidelberg.
- Ordoñez, C. A. (2005). Formatos de Imagen Digital. *Revista Digital Universitaria*, 5(7), 2-10.



Pajares, G., & de la Cruz, J. M. (2002). Clasificación de texturas naturales mediante k-Means.

Revista Electrónica de Visión por Computador., 6, 1-18.

Pal, N. R., & Pal, S. K. (1993). A review on image segmentation techniques. *Pattern Recognition*,

26(9), 1277 - 1294.

Palomino, N. L., & Concha, U. R. (2009). Técnicas de Segmentación en Procesamiento Digital de

Imágenes. *Revista de Ingeniería de Sistemas e Informática vol. 6, N.º 2, Julio - Diciembre.*

Pappas, T. N. (1992). An adaptive clustering algorithm for image segmentation. *Pattern*

Recognition Letters, 20, 57-68.

Pascual, D., Pla, F., & Sánchez, S. (2007). *Algoritmos de agrupamiento*. Santiago de Cuba, Cuba:

Departamento de Computación Universidad de Oriente.

Paz, A. F. (2009). *Procesamiento paralelo de imágenes hiperespectrales: Sistemas multi-*

procesador frente a tarjetas gráficas programables (GPUs). España: Universidad de Extremadura.

Pentland, A. P. (1984). Fractal-Based Description of Natural Scenes. *Pattern Analysis and Machine*

Intelligence, IEEE Transactions on, PAMI-6(6), 661-674.

Petrou, M., & Garcia Sevilla, P. (2006). *Image Processing Dealing Texture*. England: Jhon Wiley

& Sons Ltd.

Pugh, M. L., & Waxman, A. M. (2006). Classification of Spectrally-Similar Land Cover Using

Multi-Spectral Neural Image Fusion and the Fuzzy ARTMAP Neural Classifier.

Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on, 1808-1811. doi:10.1109/IGARSS.2006.467



- Puig, D., Montseny, E., & Sobrevilla, P. (1999). Aplicación de Operadores Estructurales a la Segmentación de Imágenes Aéreas. *V Congress of the Spanish Association for Fuzzy Technologies and Logic*. España.
- Qisheng, H. (2009). *Estimation of coniferous forest above-ground biomass using LiDAR and SPOT-5 data*. Nanjing, China: National Natural Science Foundation of China.
- Rauber, T. W., & Berns, K. (2011). Kernel Multilayer Perceptron. *Graphics, Patterns and Images (Sibgrapi), 2011 24th SIBGRAPI Conference on*, 337-343.
- Rivera, J. M., & Buri, A. D. (2012). *Segmentación de Imágenes de Placas Vehiculares Usando Técnica de Crecimiento de Regiones*. España: Escuela Superior Politecnica de Litoral.
- Roger, N. (1999). Spectroscopy of Rocks and Minerals, and Principles of Spectroscopy. *U.S. Geological Survey*. Retrieved from <http://speclab.cr.usgs.gov>.
- Rojas, M. (2009). *Caracterización de imágenes hiperspectrales utilizando Support Vector Machines y técnicas de extracción de características*. España.
- Ruíz, M. E., Morales, M., & Hernández, Y. (2011). Una estrategia de segmentación de imágenes digitales de huellas dactilares latentes. *Revista de Ingeniería Eléctrica, Electrónica y Computación*.
- Sachs, J. (1999). *Digital Image Basics*. Digital Light & Color.
- Salvatierra, C. H. (2004). *Fundamentos De Procesamiento Digital De Imágenes*. Mendoza: Departamento de Electrónica de la Universidad Técnica Federico Santa María.



Sanofer, J., & R.Deer, p. (2010). Detection of Individual Tree Crowns in High Spatial Resolution using 2D Walking Ant Histogram. *Proceedings of the International Conference on Communication and Computational Intelligence*.

Sarría, F., Recio, A., & Ruiz, J. (2003). *Análisis De Imágenes Mediante Texturas: Aplicación A La Clasificación De Unidades De Vegetación*. Valencia, España.: Departamento de Ingeniería Cartográfica, Geodesia y Fotogrametría.

Schalkoff, R. J. (1997). *Artificial Neural Networks*. McGraw-Hill Higher Education.

Sgrenzaroli, M., Baraldi, A., Eva, H., & De Grandi, G. (2002). Contextual Clustering for Image Labeling: An Application to Degraded Forest Assessment in Landsat TM Images of the Brazilian Amazon. *Geoscience and Remote Sensing, IEEE Transactions on*, 40, 1833-1848.

Shaker, K., & Zou, B. (2008). Analysis and Classification of Remote Sensing, by using Wavelet Transform and Neural Network. *Computer Science and Software Engineering, 2008 International Conference on*, 4, 963-969.

Simard, M., Sasan, S., & DeGrandi, G. (2000). Comparison of a Decision Tree and Maximum likelihood Classifiers: Application to SAR image of tropical forest. *Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International*, 5, 2129-2130.

Smits, P., & Dellepiane, S. (1997). Synthetic aperture radar image segmentation by a detail preserving Markov random field approach. *Geoscience and Remote Sensing, IEEE Transactions on*, 35, 844-857.

Sonka, M. H. (2008). *Image Processing, Analysis, and Machine Vision*. Thomson.



- Specht, D. F. (1991). A general regression neural network. *Neural Networks, IEEE Transactions on*, 2(6), 568-576.
- Srivastava, A., Asati, A., & Bhattacharya, M. (2010). Fast Hybrid Rough-Set Theoretic Fuzzy Clustering Technique with application to Multispectral Image Segmentation. *Proceedings of the First International Conference on Intelligent Interactive Technologies and Multimedia*, 126-129.
- Suárez, Á. (2009). *Análisis De Métodos De Procesamiento De Imágenes Estereoscópicas Forestales*. Madrid, España.: Facultad de Informática Universidad Complutense de Madrid.
- Trejo, I., & Hernández, J. (1996). *Identificación De La Selva Baja Caducifolia En El Estado De Morelos, México, Mediante Imágenes De Satélite*. Distrito Federal, México: Universidad Nacional Autónoma de México.
- Trias-Sanz, R. (2006). A Texture Orientation Estimator for Discriminating Between Forests, Orchards, Vineyards, and Tilled Fields. *Geoscience and Remote Sensing, IEEE Transactions on*, 44, 2755-2760.
- Urrea, J. P., & Ospina, E. (2004). Implementación de la Transformada de Hough para la Detección de Líneas para un Sistema de Visión de Bajo Nivel. *Scientia Et Technica*.
- Velázquez, A., Duran, E., Larrazábal, A., López, F., & Medina, C. (2010). *La cobertura vegetal y los cambios de uso del suelo*. México: INECC.
- Wang, Z., & Boesch, R. (2007). Color- and Texture-Based Image Segmentation for Improved Forest Delineation. *Geoscience and Remote Sensing, IEEE Transactions on*, 3055-3062.



Weston, J., Elisseeff, A., BakIr, G., & Sinz, F. (2004). *<http://people.kyb.tuebingen.mpg.de>*.

Retrieved 10 13, 2013, from <http://people.kyb.tuebingen.mpg.de/spider/>

Yang, Y., & Newsam, S. (2010). Bag-Of-Visual-Words and Spatial Extensions for Land-Use Classification. *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 270-279. doi:10.1145/1869790.1869829

Yu, X., & Reed, I. (1990). *Adaptive multiple-band cfar detection of an optical pattern with unknown spectral distribution*. *IEEE Transactions on Acoustics*, 10, 1760-1770.