



Trabajo Fin de Máster

Visualización de imágenes multi- e hiperespectrales usando operadores de agregación

Sara Iglesias Rey

Máster Universitario en Sistemas de Información Geográfica y
Teledetección

Curso 2017-2018

Propuesta de Trabajo Fin de Máster

Título en español: Visualización de imágenes multi- e hiperspectrales usando operadores de agregación
English title: Visualization of multi- and hyperspectral images using aggregation operators
Autora: Sara Iglesias Rey, Universidad Pública de Navarra
Directores: Carlos López Molina, Universidad Pública de Navarra;
Breve resumen del trabajo: En este trabajo se estudiarán alternativas de visualización para conjuntos de datos correspondientes a imágenes multi- e hiperspectrales. En particular, se propondrá la utilización de operadores de agregación, concretamente funciones OWA (<i>Ordered Weighted Averaging</i>). La implementación práctica del estimador propuesto requiere la selección de un conjunto de parámetros de pesos. Para esto último se sugerirán distintos vectores de ponderación y se realizará un estudio de su idoneidad.
Otras observaciones: Se realizarán estudios de visualización, evaluando el comportamiento de los operadores OWA con los métodos existentes. También se realizará una aplicación a datos reales, lo cual permitirá visualizar escenas multi- e hiperspectrales en distintas regiones de observación consideradas.

Agradecimientos

El presente trabajo no podría haber sido realizado sin los conocimientos adquiridos en el Máster Universitario en Sistemas de Información Geográfica y Teledetección, por ello dar las gracias al conjunto del profesorado.

En especial a mi tutor Carlos López Molina, sin cuya ayuda tampoco hubiese sido posible.

Índice general

Resumen	IX
Preludio	XI
1. Introducción	1
1.1. El papel de la visualización en la ciencia de datos	1
1.2. Imágenes multi- e hiperespectrales	2
1.3. Visualización de imágenes	3
1.4. Objetivo del trabajo	4
2. Preliminares	5
2.1. ¿Qué es una imagen?	5
2.2. Funciones de agregación	8
2.2.1. Conceptos básicos	8
2.3. Clasificación y propiedades generales	10
2.3.1. Clases principales	10
2.3.2. Propiedades principales	10
3. Estudio matemático del problema	15
3.1. La importancia del color	16
3.1.1. El sistema de color estándar	16
3.1.2. Espacio de color RGB	17
3.1.3. Otros espacios de color	18
3.2. Objetivos de diseño de Jacobson-Gupta	20
3.3. Evaluación crítica de los objetivos	22
3.4. Alternativas para la fusión	23
4. Propuesta	27
4.1. Operadores OWA	27
4.2. Ejemplos prácticos	33
5. Experimentos	35
6. Conclusiones y líneas futuras	41

Resumen

Resumen en español

El crecimiento de las imágenes multi- e hiperspectrales ha sido claro en los últimos años, y nada hace pensar que su uso decaerá a corto plazo. Sin embargo, sufre un problema debido a la alta dimensionalidad de la información almacenada en cada píxel: los problemas para su visualización en monitores estándar (tri-estímulo). Por ello, en el presente trabajo se ha propuesto un método de fusión de los datos mediante operadores de agregación, empleando para ello los operadores OWA (*Ordered Weighted Averaging*) para obtener una representación inicial de la información.

El comportamiento del proceso propuesto se evaluó y comparó con el método clásico de fusión de la información. En particular, se han estudiado distintos vectores de pesos w_i , mostrando la gran influencia que estos tienen sobre la visualización. Finalmente, se realizó una aplicación de los operadores de agregación escogidos a una serie de casos prácticos de conjuntos de datos de imágenes multi- e hiperspectrales obtenidos de las Universidad de País Vasco.

Se ha comprobado la utilidad y mejora de este método de visualización, respecto a otras propuestas anteriores, cumpliendo los objetivos de diseño propuestos presentados a lo largo de este estudio.

Palabras clave: Visualización, Imágenes multispectrales, Imágenes hiperspectrales; Operadores de agregación, Teledetección.

English abstract

There has been a boost in the use of multi- and hyperspectral imaging in the past years, and this trend is unlikely to change in a short term. However, there is a problem endemic to that imaging systems, due to the high dimensionality of the data stored at each pixel: the problems in the visualization in standard (tri-stimulus) screens.

The behaviour of the proposed process was studied and compared to the classic method of fusion of the data. In particular, different weighted vector w_i were studied, showing the huge influence that they have over the visualization. Finally, the aggregation operators were applied to a series of practical examples of multi- and hyperspectral images datasets, obtained from the University of País Vaco.

The utility and the improvement of this visualization method were verify, regarding other previous proposals, accomplishing the designs goals proposed along this work.

Keywords: Visualization, Multispectral images, Hyperspectral images, Aggregation operators, Remote sensing.

Preludio

La *teledetección* se define como la ciencia y el arte de obtener información sobre un objeto, área o fenómeno a través del análisis de la información obtenida a través de un instrumento que no se encuentra en contacto con el objeto, área o fenómeno bajo investigación [1]. Esta información es recolectada, de modo que pueda ser analizada o interpretada a posteriori. Los datos obtenidos mediante teledetección pueden ser adquiridos en múltiples formatos, incluyendo variaciones en distribuciones de fuerzas, distribuciones de ondas acústicas o distribuciones de energía electromagnética. En este trabajo nos centraremos en sensores de energía electromagnética, instalados sobre plataformas aéreas o espaciales para asistir en tareas de creación de mapas o de monitorización de recursos terrestres. Estos sensores adquieren información según la capacidad de emisión y reflectancia de energía electromagnética de una característica terrestre; estos datos serán analizados para proveer información sobre los recursos bajo investigación.

Las imágenes multi- e hiperspectrales serán presentadas en el Capítulo 1, junto con la problemática que conlleva su visualización. Una vez expuestos los obstáculos y las inconveniencias a los que debemos enfrentarnos en esta tarea, presentaremos los objetivos de este trabajo que intentaremos llevar a cabo.

Si bien la idea general de la propuesta resulta bastante intuitiva, debemos presentar con mayor profundidad las nociones de qué es una imagen desde un punto de vista más técnico, así como realizar un estudio y análisis de las herramientas matemáticas usadas en la proposición para la visualización.

Tras presentar de forma general la problemática a la que nos enfrentamos en este trabajo, estudiamos desde un punto de vista más técnico y matemático la propuesta en el Capítulo 3. Para ello, analizaremos los distintos objetivos de diseño que debemos cumplir, así como distintas alternativas de fusión existentes previas a este trabajo.

En el Capítulo 4 se presentarán las técnicas para la fusión de la información y poder realizar la visualización mediante operadores de agregación OWA. Se realizará una comparativa con métodos clásicos de fusión de imágenes, así como un balance entre distintos vectores de pesos usados.

Para comprobar el funcionamiento de los operadores OWA propuestos como método de visualización, se realizarán aplicaciones sobre diversas escenas, cuyos resultados se presentarán en el Capítulo 5.

Finalmente, en el último Capítulo 6 se incluyen las conclusiones y líneas futuras del trabajo de fin de máster.

Capítulo 1

Introducción

1.1. El papel de la visualización en la ciencia de datos

El avance de la tecnología digital ha pasado por diferentes fases en los últimos 50 años. En la actualidad, nos encontramos en un punto en el que la cantidad de información a captar y almacenar se ha disparado. El motivo de este hecho es la confluencia de muchas tecnologías diferentes, incluyendo la miniaturización de los sensores digitales, la reducción en peso y tamaño de las baterías y el abaratamiento de los sistemas de almacenamiento. No existe ningún campo de conocimiento que haya quedado ajeno a esta tendencia global, y así la tecnología basada en imagen también produce cantidades crecientes de información.

Dentro de las tecnologías de imagen, existen diferentes paradigmas que se han extendido más allá de la imagen convencional, todas posibilitadas por las nuevas tecnologías. De hecho, en diferentes foros se ha discutido su denominación como *imagen*. Véase, por ejemplo, el caso de las tecnologías avanzadas médicas. Un escaner PET (*Positron emission tomography*) genera representaciones 3D de la presencia de fluorodesoxiglucosa en la zona escaneada. ¿Debe esto considerarse una imagen, en la medida en que nada de lo medido es visible? Otro caso, si cabe, más enrevesado es el de las resonancias magnéticas (MRI). Siendo el resultado original de la inducción de campos magnéticos un tensor que modela el movimiento relativo de partículas de agua, ¿en qué momento puede considerarse una imagen? La clave de las consideraciones de estos campos de estudio como pertenecientes a la imagen médica es, precisamente, su capacidad de ser visualizados en un momento dado. Esta visualización no es directa, o sencilla, pero finalmente se alcanza para que los datos puedan ser revisados por los agentes interesados.

Este trabajo trata de una tecnología que, de alguna manera, sufre con los mismos condicionantes y críticas que la imagen médica avanzadas: la tecnología Multi- e Hiperespectral (MHE). Esta tecnología combina las tres patas sobre las cuales los datos de una imagen pueden dispararse. Primero, tenemos imágenes en alta resolución, incluso para sensores *in-lab* relativamente baratos. Segundo, tenemos una gran cantidad de valores por píxel, en cualquier caso superior a los 3 canales que puede visualizar un monitor estándar. Tercero, puede generar series temporales, sobre todo en caso de imágenes en laboratorio o en caso de combinar diferentes tomas de teledetección en el tiempo.

La tecnología MHE ha desarrollado una gran cantidad de literatura. De hecho, ha tomado protagonismo en muchos campos científicos diferentes, y se ha mostrado como una tecnología altamente práctica. Sin embargo, sufre de un problema básico que está muy poco estudiado. Quizá, precisamente, porque no se liga a ningún campo de conocimiento, sino que es genérico y propio de la tecnología: la visualización de imágenes MHE. Dada una *imagen MHE*, véase un cubo con profundidad $N > 3$, ¿existe algún método estándar de representación? Lo cierto es que no, o al menos diferentes autores tienden

a generar la solución *ad-hoc* que más les conviene en cada trabajo.

Así pues, podría parecer que la visualización es innecesaria, especialmente con el advenimiento de tecnologías que producen más y más datos. Sin embargo, esto no es aceptable. Gran parte de los avances, especialmente en adopción de nuevas tecnologías, se dan por el hecho de poder ser demostrados. La visualización de datos es una disciplina *per se*, no simplemente un añadido opcional a la ciencia de datos. Si usamos sistemas que manejan datos sin posibilidad de visualización tendremos sistemas inteligentes sin capacidad de ser auditados. Es decir, tendremos sistemas que toman decisiones (e.g. presencia de ciertos compuestos, ocupación de ciertas regiones,...) sin capacidad de mostrar por qué lo hacen. En el caso que nos ocupa, véase la tecnología MHE, necesitamos visualizar las imágenes para poder comprobar su contenido o verificar los diferentes procesamientos que se hacen sobre ella (e.g. registro o rectificación, en caso de imágenes cenitales). De la misma manera, e independientemente del tratamiento que vayamos a darle a la información a nivel de píxel, debemos poder representar estadios intermedios de la imagen en el procesado, lo cuál probablemente implique más de tres canales. Este trabajo se centra, por lo tanto, en la visualización de la información producida por tecnologías MHE.

1.2. Imágenes multi- e hiperespectrales

Con la aparición de nuevas cámaras fuera del visible, la teledetección sufrió un auge espectacular que modificó la forma de obtener información terrestre en todos los sentidos [2]. En el siglo XX, debido a la Primera Guerra Mundial, se desarrollaron las técnicas de obtención de fotografías aéreas como un activo militar de gran importancia; las aeronaves se centraron en un intenso desarrollo de la tecnología de reconocimiento aéreo. Al mismo tiempo se fomentaron sistemas y técnicas de imagen para la medición y la identificación de características en el suelo; la fotografía aérea y panorámica tuvo una gran aplicación para usos civiles de topografía, cartografía y creación de mapas. Durante esta época aparecieron las primeras medidas aéreas en el infrarrojo térmico, es decir, fuera de la región visible del espectro electromagnético. Este interés en técnicas de interpretación y medida continuó tras la guerra y fueron aplicadas a otras áreas fuera del campo militar, como en agricultura y silvicultura. En el periodo alrededor de la Segunda Guerra Mundial, se obtuvieron técnicas más sofisticadas en la interpretación de fotografías aéreas. Tras desarrollos significativos en la tecnología, se obtuvieron medidas en la región del infrarrojo y de las microondas, y finalmente de todo el espectro electromagnético.

La fotografía aérea a color-infrarroja y la digital a color natural, pueden ser consideradas como conjuntos de datos multispectrales de tres bandas [3]. Estas pueden ser convertidas a bandas del rojo, verde y azul o del verde, rojo e infrarrojo cercano de información digital. Aunque en muchas aplicaciones estos conjuntos de datos multispectrales de tres bandas son suficientes, existen múltiples casos en los que bandas espectrales localizadas en posiciones más óptimas del espectro electromagnético, aportarían una mayor información, siendo así de mayor utilidad para nuestra aplicación concreta. Afortunadamente, los ingenieros ópticos han desarrollado detectores sensibles a cientos de bandas en el espectro electromagnético.

La teledetección multispectral se define como la colección de energía reflejada, emitida o dispersada por un objeto o área de interés en múltiples bandas o regiones del espectro electromagnético. La teledetección hiperespectral conlleva la colección de datos en cientos de bandas. La mayoría de los sistemas de teledetección multi- e hiperespectrales recogen la información en formato digital.

Este tipo de imágenes tuvieron un gran impacto con la aparición de las fotografías aéreas satelitales y de los UAVs (*Unmanned Aerial Vehicle*) o drone, ya que estos datos multi- e hiperespectrales proveen un mejor entendimiento de nuestro planeta, desde uso y cubiertas del suelo, desarrollo urbano, hasta procesos terrestres. En los últimos años existe un auge de la cantidad de satélites orbitando alrededor de nuestro planeta, que va a la par de la mejora de las tecnologías espaciales, llegando a un número de

aproximadamente 3000 satélites. La tecnología drone posee grandes cualidades y amplias aplicaciones, mediante la medición y captura de datos en regiones del espectro fuera del visible. Los drones han ejercido una gran influencia en el sector de la topografía, ya que aporta grandes ventajas respecto a las técnicas tradicionales de topografía y cartografía logrando una mayor eficiencia y calidad.

1.3. Visualización de imágenes

Las imágenes aéreas y espaciales contienen un registro detallado de características del suelo en el momento de la adquisición de los datos [1]. Un intérprete de imágenes examina sistemáticamente los datos, usando material de apoyo como mapas y registros de observaciones de campo. Basándose en este estudio, se realiza una interpretación de la naturaleza física de los objetos y fenómenos que aparecen en las imágenes. Las interpretaciones pueden conllevar distintos niveles de complejidad, desde el simple reconocimiento de objetos en la superficie terrestre, hasta la derivación de información concerniente a las complejas interacciones entre características terrestres superficiales y subterráneas. El grado de éxito en la interpretación de la imagen depende del entrenamiento y experiencia del intérprete, de la naturaleza de los objetos o fenómenos a interpretar y de la calidad de las imágenes utilizadas.

Las imágenes son visualizadas en tres dimensiones que pueden ser representadas como canales RGB (rojo, verde y azul) de un dispositivo triestímulo estándar. La visión humana funciona de un modo similar; el espectro visible continuo que penetra en cada punto del ojo, es convertido en un color percibido tricomponente tras ser capturado por los conos L, M y S, que se corresponden aproximadamente con los canales RGB. Los humanos pueden interpretar rápidamente las imágenes a color que conllevan una rápida comprensión de la información.

Las imágenes multi- e hiperespectrales contienen muchas más bandas que las que pueden ser mostradas en una representación triestímula [4]. Estas últimas incluyen los monitores estándar y pantallas de cristal líquido, así como cualquier representación que tenga tres canales imagen, comunmente rojo, verde y azul. La interacción humana con imágenes multi- e hiperespectrales es vital, tanto para dirigir métodos computarizados como para validar análisis automatizados o para realizar decisiones e interpretaciones correctas. Debido a esto surge la cuestión de cuál sería la mejor representación de imágenes multi- e hiperespectrales, de forma que permita la mejor interacción humana con ellas y los resultados computados.

Toda representación a 3 bandas de una imagen originalmente de N -bandas, conlleva necesariamente una pérdida de información para $N > 3$. La información espacial es importante para ciertas tareas de clasificación y detección, pero las técnicas de reducción de dimensionalidad más populares no tienen esta en cuenta. Para un esquema de visualización concreto, un valor particular de color puede representar diferentes vectores N -dimensionales. La visión humana sufre el mismo problema, es decir, múltiples espectros del visible distintos conllevan las mismas estimulaciones L, M y S del cono fotorreceptor. Este espectro metamérico se diferencia usualmente por contexto, forma y otras pistas visuales.

Por ello, dado que todas las representaciones de imágenes multi- e hiperespectrales conllevan una pérdida de información, debemos hallar un método que nos permita juzgar que visualización es correcta. El mejor criterio es específico para cada aplicación particular y requiere una alta especialización y pruebas humanas para observar el grado de adecuación de cierta visualización a la tarea.

Sin embargo existen objetivos independientes a las aplicaciones que pueden ser usados para diseñar y evaluar visualizaciones efectivas. Estos objetivos se basan en acelerar y facilitar la interpretación humana de las escenas. Los objetivos de diseño para la representación de imágenes multi- e hiperespectrales, no presentan la misma importancia en todas las aplicaciones, así como no será posible lograr simultáneamente todos en muchos de los casos. Sin embargo, cada uno de estos objetivos, incrementará la transmisión efectiva de información.

1.4. Objetivo del trabajo

El objetivo de este trabajo se centra en hallar un método de visualización para imágenes multi- e hiperespectrales (imágenes de N -bandas con $N > 3$) con la menor pérdida de información posible. Nos centraremos en un procedimiento que no requiera una especialización para cada aplicación particular, si no que ofrezca un método genérico de representación. Para ello perseguiremos una serie de objetivos presentados en la sección 3.2, que nos proporcionarán unos requisitos generales de diseño para la consecución de nuestro objetivo.

En [4] y [5] se desarrollan distintos métodos de fusión de imágenes multi- e hiperespectrales. Se proponen distintos objetivos y soluciones para la visualización triestímula de imágenes de N -bandas, centrándose en la creación de representaciones consistentes de datos multi- e hiperespectrales que faciliten la compresión y el análisis de escenas, y que puedan ser usadas de forma conjunta con visualizaciones específicas para cada tarea. Este es un proceso de fusión de información por lo que conlleva de forma inherente una pérdida de información. Un conjunto de envolturas espectrales fijas ponderadas [4], dan lugar a visualizaciones que cumplen los objetivos de diseño expuestos en las sección 3.2 y son útiles para un amplio rango de tareas. Las tres bandas representadas R , G y B , serán integraciones lineales fijas de la imagen multi- o hiperespectral original, ponderada por tres envolturas espectrales distintas. Esto es similar a como funciona la visión humana fotópica (a la luz del día), donde tres tipos distintos de conos tienen diferentes envolturas espectrales de sensibilidad. Cada cono absorbe los fotones entrantes con una probabilidad que se corresponde con la sensibilidad del cono a esa longitud de onda. En este método de visualización, la sensibilidad probabilística ha sido reemplazada por una ponderación determinística.

Del mismo modo que en el método expuesto anteriormente, para reducir la dimensionalidad de un conjunto discreto de imágenes multivariantes para su visualización, es común proyectar el conjunto completo de imágenes multivariantes en tres funciones base [5]. Estas funciones representarán la probabilidad sensitiva de los conos para absorber los fotones entrantes. Sean r, g, b las funciones de base discretas, donde $r[n]$ denota el n -ésimo componente de la función de base r y $n = 1, \dots, N$. La respuesta en cada canal de color es una proyección lineal de los datos en cada función base, es decir, realizamos una aplicación entre el vector de datos normalizado x y unos componentes escalares de color R, G y B de la forma

$$R = r^T x \quad G = g^T x \quad B = b^T x.$$

Los componentes R, G y B pueden ser escalados y cuantificados para la visualización. Las funciones bases r, g y b pueden adaptarse a los datos, como ocurriría en un análisis de componentes principales (3.4).

Existen múltiples métodos de fusión de información que encaran el problema desde distintas perspectivas [6][7]. Nos centraremos en extender este trabajo de reducción de la dimensionalidad de la información, ampliando los objetivos de diseño planteados desde una perspectiva matemática. Como método de fusión de imágenes multi- e hiperespectrales, estudiaremos los operadores de agregación para poder llevar a cabo una visualización con la menor pérdida de información posible. Esta herramienta matemática nos proporciona una relación de preferencia colectiva, permitiéndonos de este modo resolver el problema de decisión de manera eficiente y flexible.

Capítulo 2

Preliminares

2.1. ¿Qué es una imagen?

Una imagen es una representación visual de una percepción de un objeto real o imaginario. Las imágenes digitales son matrices de números, es decir, una imagen se representa de forma lógica como una matriz de filas y columnas [8]. Estas matrices de datos que conforman las imágenes se incluyen en la clase general de información raster, lo cual conlleva que el valor informativo individual de un punto no se asocia explícitamente con una localización particular en el suelo. La ubicación de cada valor de información (o elemento de la imagen denominado píxel) viene implícito por su posición en la matriz; por consiguiente, si conocemos las coordenadas UTM de la celda superior izquierda de la matriz o raster y el espaciado entre celdas en metros, podremos calcular la posición de cualquier celda del raster. Los valores de los números almacenados en los elementos de la matriz se encuentran dentro de un rango específico, usualmente entre 0 y 255, que se corresponde con el rango de luminosidad del color asociado con la matriz imagen. El valor 0 indica una falta de los colores asociados (rojo, verde y azul) y el valor 255 es el nivel más brillante para el color representado. Los dos valores en los extremos del rango (0 y 255) pueden ser usados con otras finalidades.

La matriz de los valores de cada píxel se almacena en un área especial de la memoria del ordenador denominada *memoria gráfica*. La memoria gráfica se encuentra localizada normalmente en la tarjeta gráfica y no forma parte de la memoria de acceso aleatorio del ordenador (RAM).

En una imagen en escala de grises, tan solo se necesita almacenar una matriz de números para conservar los valores de los píxeles, cada uno de los cuales puede tomar un valor en los 256 niveles de brillo, comenzando desde el 0 (negro), pasando por el 127 (tono medio de grises) y terminando en el 255 (blanco). Una imagen en escala de grises tiene tan sólo un componente por píxel, el nivel de gris; por el contrario, una imagen a color almacena tres componentes por cada posición de píxel, siendo estos los niveles de los colores primarios que conforman la luz, rojo, verde y azul.

Una imagen a color, como ya hemos expuesto, es creada usando tres matrices raster que contienen los valores píxel que representan los niveles de los tres colores primarios de la luz. Los niveles del 0 al 255 representan el rango de cada color primario desde el 0 (negro) hasta el 255 (máxima intensidad de rojo, verde o azul, es decir RGB). Diferentes combinaciones de rojo, verde y azul producen los colores del espectro, como demostró Sir Isaac Newton con su famoso experimento del prisma. Los colores primarios de la luz son aditivos, es decir, juntando rojo y verde obtenemos amarillo; por el contrario, los colores usados en impresión son sustractivos, por ello las impresoras de inyección usan tinta en cian, magenta y amarillo. Presentamos en la tabla 2.1 algunos ejemplos de colores generados mediante la adición de diferentes proporciones de rojo, verde y azul. Remarcar que combinaciones RGB en las

Intesidad de rojo	Intensidad de verde	Intensidad de azul	Color
255	255	0	Amarillo
0	255	255	Cian
255	0	255	Magenta
127	0	0	Rojo medio
127	127	127	Gris medio
0	0	0	Negro
255	255	255	Blanco
241	0	171	Violeta
255	155	50	Naranja

Tabla 2.1: Combinaciones de los colores primarios asumiendo representación de 8-bits, escala 0-255.

cuales los niveles de los tres colores sean iguales, dan lugar a tonos de gris.

No todas las imágenes obtenidas mediante un sistema de teledetección ,presentan valores de píxel en el rango de 0-255. Por ejemplo, los datos obtenidos mediante AVHRR usan el rango 0-1023 y las bandas térmicas de las imágenes ASTER se miden en la escala 0-4095. Los valores mínimos y máximos de los datos ASTER tienen asociados usos específicos; por ejemplo los valores 0 y 4095 se usan para indicar *malos datos* y *píxeles saturados* respectivamente.

Los píxeles almacenados en las celdas que conforman la imagen digital (los valores píxel o intensidades píxel) se representan electrónicamente como una serie de dígitos binarios (en base dos) que pueden ser interpretados como interruptores de encendido y apagado. En base dos los números decimales 0, 1, 2, 3..., se escriben como 0, 1, 10, 11... donde cada columna de la izquierda representa sucesivamente una potencia mayor de dos, en lugar de una potencia de diez como en el sistema decimal. Si hacemos uso de ocho dígitos binarios para registrar el valor almacenado en cada píxel, en ese caso 0 y 255 se escriben como 00000000 y 11111111. Por lo tanto, un total de ocho dígitos binarios (bits) son necesarios para representar los 256 números en el rango 0-255. El rango de las intensidades píxel se denomina el rango dinámico de la imagen. Ya que se necesitan ocho bits para representar el rango de cada uno de los tres colores primarios, la imagen resultante se denomina *imagen de 24 bits*. Existen otros métodos de representar información píxel de una imagen; por ejemplo, una imagen de 10 bits unibanda provee 1024 niveles de gris, mientras que una imagen de 16 bits puede representar tanto datos positivos como negativos, o tan solo positivos, dependiendo de si usamos la representación con o sin signo. Los datos reales se utilizan comunmente para almacenar valores físicos en lugar de resultados cuantificados; por ejemplo, la amplitud y fase de una imagen SAR se almacena como un par de números reales de 32 bits.

Es importante observar que el número de bits por píxel en la memoria gráfica está fijado en 8 por el hardware. Sin embargo las imágenes obtenidas mediante teledetección se proporcionan en 10, 12, 16 y 32 bits, así como en la forma de 8 bits entero, 32 bits o incluso la representación 64 bits real. Los números reales tienen puntos decimales, mientras los enteros no se encuentran fraccionados más allá de la unidad. La apariencia en la pantalla de la imagen se encuentra afectada por el método elegido para transformar la imagen de entrada a 8 bits. Un método consiste en tomar el rango numérico de los datos para cada uno de los tres colores y realizar una aplicación lineal entre los valores de entrada

y de salida, usando la siguiente relación

$$\text{valor de salida} = \frac{(\text{valor de entrada} - \text{valor de entrada}_{\min})}{(\text{valor de entrada}_{\max} - \text{valor de entrada}_{\min})} \cdot 255, \quad (2.1)$$

donde el *valor de salida* toma valores entre 0 y 255, *valor de entrada_{max}* y *valor de entrada_{min}* son los valores máximo y mínimo respectivamente en la imagen de entrada y el *valor de entrada* es el valor píxel de la imagen a convertir. Los valores extremos pueden tener un efecto substancial en el rango de los datos de entrada y conllevar una falta de contraste en los resultados obtenidos mediante este método.

Un método alternativo de transformación de un imagen de entrada cuyo rango dinámico es mayor que 8 bits al rango 0-255, se denomina igualación. Los valores píxel de entrada se agrupan en 256 conjuntos mediante la mezcla de los 1024 niveles en 256 clases, las cuales contienen aproximadamente el mismo número de píxeles. Más que rangos iguales, las 256 clases de salida tienen frecuencias iguales. En comparación con la aproximación lineal, la igualación produce una imagen de salida con un mayor contraste. Sin embargo, ya que varios valores de entrada se transforman en el mismo valor de salida, la transformación es no reversible y también no lineal, ya que los pasos entre clases de salida individuales contiguas no se corresponden con rangos iguales en los valores de entrada.

De forma general, existen tres tipos distintos de imágenes que pueden ser almacenadas en una memoria gráfica y visualizadas en una pantalla: imágenes a color, imágenes en escala de grises e imágenes etiquetadas o clasificadas. Las imágenes a color pueden ser de tres clases. La primera denominada de color natural, está formada por tres componentes representando los colores del mundo real, las bandas del rojo, verde y azul del visible, en las reservas de memoria del rojo, verde y azul de la memoria gráfica. Las imágenes a color natural son como fotografías a color ordinarias, nos muestran los colores como los veríamos. Si las tres bandas escogidas para la visualización no representan el rojo, verde y azul como los veríamos en la realidad, el resultado es una imagen en falso color; por ejemplo, se pueden almacenar bandas del infrarrojo cercano como tonalidades de rojo y variaciones del rojo y el azul como variaciones del verde y azul. Tres bandas cualesquiera pueden ser almacenadas en las reservas de memoria para rojo, verde y azul. El tercer tipo de imágenes a color se denomina imagen a pseudocolor, ya que los datos ocupan una única reserva de memoria en lugar de tres. Esto implica que los valores píxel de una imagen de pseudocolor se encuentra en el rango entre 0 y 255. Estos 256 niveles están asociados con colores mediante una tabla de consulta. En esta representación, un único valor en una posición de píxel dada es transformado mediante conversores de digital a analógico a través de la tabla de consulta.

El segundo tipo de imágenes que pueden ser almacenadas en la memoria gráfica y visualizadas en pantalla son denominadas imágenes en escala de grises. Al igual que las imágenes a pseudocolor, las imágenes en escala de grises tienen tan solo un único valor de entrada, representando una única banda o canal.

Una imagen etiquetada o clasificada está compuesta por píxeles cuyos valores representan una equiqueta que indica una propiedad de algún tipo. La etiqueta por sí misma no presenta ningún significado numérico. Existen métodos de clasificación de imágenes que permiten identificar a cada píxel de la imagen dentro de una categoría específica, como un tipo específico de cubierta del suelo. Estas categorías están descritas por etiquetas como '1', '2', '3', etc. que indican 'agua', 'bosque arbolado' o 'suelo descubierto'. Para representar este tipo de imágenes en una pantalla, la imagen etiquetada es almacenada en la memoria gráfica y los tres conversores de digital a analógico (rojo, verde y azul) son programados para asignar valores RGB a los píxeles individuales unibanda, como en el caso de la imagen pseudocolor.

2.2. Funciones de agregación

En esta sección se recogen las principales nociones y resultados necesarios sobre las funciones de agregación, así como su clasificación y propiedades generales [9].

2.2.1. Conceptos básicos

Una función matemática es una regla que toma un valor de entrada, llamado argumento, y produce un valor de salida. Cada valor de entrada tiene un valor de salida único asociado a él. Una función se denota típicamente de la forma $y = f(x)$, donde x es el argumento e y es el valor de salida. El argumento x puede ser un vector, es decir una tupla de longitud n : $\mathbf{x} = (x_1, x_2, \dots, x_n)$, donde x_1, x_2, \dots, x_n son los denominados componentes de \mathbf{x} .

Las funciones de agregación son funciones con propiedades especiales. Nos centraremos en funciones de agregación que toman argumentos reales dentro del intervalo cerrado $[0, 1]$ y produce un valor real en $[0, 1]$. Usualmente denotamos esta transformación como $f : [0, 1]^n \rightarrow [0, 1]$ para funciones que toman argumentos con n componentes.

La finalidad de las funciones de agregación es combinar valores de entrada típicamente interpretados como grados de pertenencia en conjuntos difusos, grados de preferencia o apoyo a una hipótesis entre otros.

El valor de entrada 0 es interpretado como no pertenencia, no preferencia, no evidencia, no satisfacción, etc. y de forma natural, un proceso de agregación de n ceros debe producir el valor de salida 0. De forma similar el valor 1 es interpretado como pertenencia total, preferencia o evidencia más sólida, etc., y un proceso de agregación de unos debe producir de forma natural el valor 1. Esto implica una propiedad fundamental de las funciones de agregación, la conservación de los límites

$$f(0, 0, \dots, 0) = 0 \text{ y } f(1, 1, \dots, 1) = 1. \quad (2.2)$$

La segunda propiedad fundamental es la condición de monotonía. Consideramos la agregación de dos valores de entrada \mathbf{x} y \mathbf{y} , de forma que $x_1 < y_1$ y $x_j = y_j$ para todo $j = 2, \dots, n$. v. g., $\mathbf{x} = (a, b, b, b)$, $\mathbf{y} = (c, b, b, b)$, $a < c$. Pensemos en el argumento j -ésimo de f como un grado de preferencia respecto al j -ésimo criterio, y \mathbf{x} e \mathbf{y} como vectores representando dos alternativas A y B . Por consiguiente, respecto al primer criterio B presenta preferencia sobre A ; respecto al resto de criterios ambas alternativas presentan una preferencia equitativa. Por ello no es razonable preferir A frente a B . El número de criterios no es importante, por lo que la monotonía se mantiene no solo para el primero si no para cualquier argumento x_i .

Matemáticamente podemos expresar la monotonía no decreciente en todos los argumentos como

$$x_i \leq y_i \text{ para todo } i \in \{1, \dots, n\} \text{ implica } f(x_1, \dots, x_n) \leq f(y_1, \dots, y_n). \quad (2.3)$$

Usaremos con frecuencia el desigualdad vectorial $\mathbf{x} \leq \mathbf{y}$, que representa que cada componente de \mathbf{x} no es mayor que su correspondiente componente en \mathbf{y} . De este modo la monotonía no decreciente puede ser expresada como $\mathbf{x} \leq \mathbf{y}$ implica $f(\mathbf{x}) \leq f(\mathbf{y})$. La condición 2.3 es equivalente a la condición de que cada función invariante $f_{\mathbf{x}}(t) = f(\mathbf{x})$ con $t = x_i$ y el resto de componentes de \mathbf{x} estando fijos, es monótona no decreciente en t .

La monotonía en todos los argumentos y la preservación de los límites son las dos propiedades fundamentales que caracterizan las funciones de agregación generales. Si una de estas propiedades falla, no podemos considerar la función f como una función de agregación, ya que produciría valores de salida inconsistentes en procesos de decisión.

Definición 2.1. Una función de agregación es una función de $n > 1$ argumentos que proyecta el cubo unitario n -dimensional en el intervalo unitario $f : [0, 1]^n \rightarrow [0, 1]$, con las propiedades

1. $f(0, 0, \dots, 0) = 0$ y $f(1, 1, \dots, 1) = 1$
2. $\mathbf{x} \leq \mathbf{y}$ implica $f(\mathbf{x}) \leq f(\mathbf{y})$ para todo $\mathbf{x}, \mathbf{y} \in [0, 1]^n$.

A menudo debemos considerar en el mismo marco de trabajo valores de entrada de distintos tamaños en la agregación. En algunas aplicaciones, los vectores de entrada pueden presentar un número variable de componentes, v. g., algunos valores pueden faltar. En estudios teóricos, es apropiado en un elevado número de casos considerar una familia de funciones de $n = 2, 3, \dots$ argumentos con la misma propiedad subyacente. La siguiente construcción matemática de una *función de agregación extendida* nos permite definir y trabajar con esta clase de familias con cualquier número de argumentos.

Definición 2.2. Una función de agregación extendida es una aplicación

$$F : \bigcup_{n \in \{1, 2, \dots\}} [0, 1]^n \rightarrow [0, 1],$$

de forma que la restricción de esta aplicación al dominio $[0, 1]^n$ para un n fijo es una función de agregación n -aria f_n con la restricción $F(x) = x$ para $n = 1$.

Por consiguiente, en términos más sencillos, una función de agregación extendida es una familia de 2, 3, ... funciones de agregación variables, con la restricción $F(x) = x$ para el caso especial de $n = 1$. Usaremos la notación f_n cuando deseemos enfatizar que una función de agregación tiene n argumentos. En general, dos miembros de una de estas familias con distintos tamaños de entrada m y n no necesitan estar relacionados; sin embargo, observamos que en los casos más interesantes presentan una relación y, en algunos casos, pueden ser calculados usando una única fórmula genérica.

En la siguiente sección presentaremos algunas propiedades generales de funciones de agregación y funciones de agregación extendidas. En general una propiedad se mantiene para una función de agregación extendida F si y sólo si se mantiene para cada miembro de la familia f_n .

Mostramos a continuación a modo de ejemplo funciones de agregación que, debido a que son válidas para $n > 1$ actúan como funciones de agregación extendidas:

- *Media Aritmética:* $f_n(\mathbf{x}) = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$.
- *Media Geométrica:* $f_n(\mathbf{x}) = \sqrt[n]{x_1 x_2 \dots x_n}$.
- *Media Armónica:* $f_n(\mathbf{x}) = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}}$.
- *Mínimo:* $\text{mín}(\mathbf{x}) = \text{mín}\{x_1, \dots, x_n\}$.
- *Máximo:* $\text{máx}(\mathbf{x}) = \text{máx}\{x_1, \dots, x_n\}$.
- *Producto:* $f_n(\mathbf{x}) = x_1 x_2 \dots x_n = \prod_{i=1}^n x_i$.
- *Suma Limitada:* $f_n(\mathbf{x}) = \text{mín}\{1, \sum_{i=1}^n x_i\}$.

2.3. Clasificación y propiedades generales

2.3.1. Clases principales

Existen diversas semánticas de funciones de agregación, a partir de las cuales se definen las principales clases. En algunos casos se desea llevar a cabo un promedio entre los distintos valores de entrada; en otros casos las funciones de agregación modelizan conectivas lógicas, conjunción o disyunción, de forma que los valores de entrada se refuerzan entre ellos; a veces el comportamiento de las funciones de agregación dependen de los valores de entrada.

Las cuatro clases principales de funciones de agregación son: *Promedio*, *Conjunción*, *Disyunción* y *Mixto*.

Definición 2.3. Una función de agregación f presenta un comportamiento promedio si para cada \mathbf{x} está limitada por

$$\text{mín}(\mathbf{x}) \leq f(\mathbf{x}) \leq \text{máx}(\mathbf{x}).$$

Definición 2.4. Una función de agregación f presenta un comportamiento conjuntivo si para cada \mathbf{x} está limitada por

$$f(\mathbf{x}) \leq \text{mín}(\mathbf{x}) = \text{mín}(x_1, x_2, \dots, x_n)$$

Definición 2.5. Una función de agregación f presenta un comportamiento disjuntivo si para cada \mathbf{x} está limitada por

$$f(\mathbf{x}) \geq \text{máx}(\mathbf{x}) = \text{máx}(x_1, x_2, \dots, x_n)$$

Definición 2.6. Una función de agregación f presenta un comportamiento mixto si no pertenece a ninguna de las clases anteriores, es decir, si exhibe diferentes tipos de comportamientos en distintas partes del dominio.

2.3.2. Propiedades principales

Definición 2.7. Una función de agregación f se denomina idempotente si para cada valor de entrada $\mathbf{x} = (t, t, \dots, t) = t$, $t \in [0, 1]$ la salida es $f(t, t, \dots, t) = t$.

Debido a la monotonía de f , la idempotencia es equivalente a un promedio.

Las funciones de agregación *mínimo* y *máximo* son las únicas funciones que son al mismo tiempo conjuntivas (disjuntivas) y promedio, y por tanto idempotentes.

Ejemplo 2.8. La media aritmética es una función de agregación promedio (idempotente)

$$f(\mathbf{x}) = \frac{1}{n}(x_1 + x_2 + \dots + x_n).$$

Ejemplo 2.9. La media geométrica es una función de agregación promedio (idempotente)

$$f(\mathbf{x}) = \sqrt[n]{x_1 x_2 \dots x_n}.$$

Ejemplo 2.10. El producto es una función de agregación conjuntiva

$$f(\mathbf{x}) = \prod_{i=1}^n x_i = x_1 x_2 \dots x_n.$$

Definición 2.11. Una función de agregación f se denomina simétrica, si su valor no depende de la permutación de sus argumentos, es decir,

$$f(x_1, x_2, \dots, x_n) = f(x_{P(1)}, x_{P(2)}, \dots, x_{P(n)})$$

Ejemplo 2.12. Las medias aritmética y geométrica y el producto en los ejemplos 2.8-2.10 son funciones de agregación simétricas. Una media aritmética ponderada con pesos no iguales w_1, w_2, \dots, w_n , no negativos y que suman la unidad, es no simétrica

$$f(\mathbf{x}) = \sum_{i=1}^n w_i x_i = w_1 x_1 + w_2 x_2 + \dots + w_n x_n.$$

La permutación de los argumentos es muy importante en agregación, ya que ayuda a expresar la simetría, así como en la definición de otros conceptos. Una permutación de $(1, 2, \dots, 5)$ es una tupla como $(5, 3, 2, 1, 4)$. Existen $n! = 1 \times 2 \times 3 \times \dots \times n$ posibles permutaciones de $(1, 2, \dots, n)$.

Denotaremos un vector cuyos componentes hayan sido reorganizados en un orden dado por una permutación P por $\mathbf{x}_P = (x_{P(1)}, x_{P(2)}, \dots, x_{P(n)})$; en nuestro ejemplo, $\mathbf{x}_P = (x_5, x_3, x_2, x_1, x_4)$. A menudo haremos uso de las siguientes permutaciones especiales de los componentes de \mathbf{x} .

Definición 2.13. Denotamos por \mathbf{x}_\nearrow el vector obtenido reorganizando los componentes en orden creciente de \mathbf{x} , es decir, $\mathbf{x}_\nearrow = \mathbf{x}_P$ donde P es la permutación tal que $x_{P(1)} \leq x_{P(2)} \leq \dots \leq x_{P(n)}$.

Del mismo modo denotamos por \mathbf{x}_\searrow el vector obtenido reorganizando los componentes en orden decreciente de \mathbf{x} , es decir, $\mathbf{x}_\searrow = \mathbf{x}_P$ donde P es la permutación tal que $x_{P(1)} \geq x_{P(2)} \geq \dots \geq x_{P(n)}$.

Definición 2.14. Una función de agregación f es estrictamente monótona creciente si

$$x \leq y \text{ pero } \mathbf{x} \neq \mathbf{y} \text{ implica } f(\mathbf{x}) < f(\mathbf{y}) \text{ para todo } \mathbf{x}, \mathbf{y} \in [0, 1]^n.$$

La monotonía estricta es una propiedad bastante restrictiva. Notar que no existen funciones de agregación conjuntivas y disyuntivas estrictamente monótonas; esto es debido a que toda función conjuntiva coincide con $\min(\mathbf{x})$ para aquellos valores \mathbf{x} que tienen al menos un componente igual a cero, y el mínimo no es estrictamente monótono (del mismo modo las funciones de agregación disyuntivas coinciden con el $\max(\mathbf{x})$ para aquellos valores \mathbf{x} que tienen al menos un componente $x_i = 1$). Sin embargo, la monotonía estricta en conjuntos semiabiertos $]0, 1]^n$ es considerada a menudo para funciones de agregación conjuntivas ($[0, 1]^n$ respectivamente para funciones disyuntivas). Existen múltiples funciones de agregación promedio estrictamente monótonas, como por ejemplo medias aritméticas.

Definición 2.15. Una función de agregación f tiene un elemento neutro $e \in [0, 1]$, si para todo $t \in [0, 1]$ en cualquier posición se cumple

$$f(e, \dots, e, t, e, \dots, e) = t.$$

Para funciones de agregación extendidas, existe una versión más fuerte de esta propiedad, que relaciona funciones de agregación con un número distinto de argumentos.

Definición 2.16. Una función de agregación extendida F tiene un elemento neutro $e \in [0, 1]$, si para todo \mathbf{x} con $x_i = e$, para $1 \leq i \leq n$, y para todo $n \geq 2$,

$$f_n(x_1, \dots, x_{i-1}, e, x_{i+1}, \dots, x_n) = f_{n-1}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$$

Cuando $n = 2$, tenemos $f(t, e) = f(e, t) = t$; por ello, mediante iteraciones de esta propiedad obtenemos como consecuencia que todo miembro f_n de la familia tiene el elemento neutral e , es decir,

$$f_n(e, \dots, e, t, e, \dots, e) = t$$

para t en cualquier posición.

Ejemplo 2.17. La función producto $f(\mathbf{x}) = \prod x_i$ tiene como elemento neutro $e = 1$. Del mismo modo la función *mín* tiene como elemento neutro $e = 1$ y la función *máx* tiene como elemento neutro $e = 0$. La media aritmética no tiene elemento neutro.

Existe también el caso en el que un valor específico a de cualquier argumento, produce el valor de salida a . Por ejemplo, si usamos la conjunción para la agregación, si cualquier valor de entrada es 0, el valor de salida será también 0.

Definición 2.18. Una función de agregación f tiene un elemento absorbente $a \in [0, 1]$ si

$$f(x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n) = a,$$

para todo \mathbf{x} tal que $x_i = a$ con a en cualquier posición.

Ejemplo 2.19. Toda función de agregación conjuntiva tiene como elemento absorbente $a = 0$. Toda función de agregación disjuntiva tiene como elemento absorbente $a = 1$. La media geométrica tiene como elemento absorbente $a = 0$.

Definición 2.20. Un elemento $a \in [0, 1]$ es un divisor cero de una función de agregación f si para todo $i \in \{1, \dots, n\}$ existe algún $\mathbf{x} \in]0, 1]^n$ tal que su componente i -ésimo sea $x_i = a$ y mantenga $f(\mathbf{x}) = 0$, es decir, la igualdad

$$f(x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n) = 0,$$

se cumpla para algún $\mathbf{x} > 0$ con a en cualquier posición.

La interpretación de divisores cero es directa: si uno de los valores de entrada toma el valor a o un valor menor, el valor de agregación puede ser cero para algún \mathbf{x} ; por lo que es posible tener el valor de agregación cero, incluso en el caso de que todos los valores de entrada sean positivos. El valor más alto de a juega el papel de umbral, el límite más bajo en todos los valores de entrada que garantiza un valor de salida no nulo. Es decir, si b no es un divisor cero, entonces $f(\mathbf{x}) > 0$, si para todo $x_i \geq b$.

Definición 2.21. Un elemento $a \in [0, 1]$ es un divisor unidad de una función de agregación f si para todo $i = 1, \dots, n$ existe algún $\mathbf{x} \in [0, 1]^n$ tal que su componente i -ésimo sea $x_i = a$ y mantenga $f(\mathbf{x}) = 1$, es decir, la igualdad

$$f(x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n) = 1,$$

se cumpla para algún $\mathbf{x} < 1$ con a en cualquier posición.

La interpretación es similar: el valor de cualquier valor de entrada mayor que a puede convertir el valor de salida $f(\mathbf{x}) = 1$, incluso en el caso de que ninguno de los valores de entrada sea igual a 1. Por otro lado, si b no es un divisor unidad, el valor de salida no puede ser uno si todos los valores de entrada son menores que b .

Definición 2.22. Una función f de dos argumentos es asociativa si $f(f(x_1, x_2), x_3) = f(x_1, f(x_2, x_3))$ se cumple para todo x_1, x_2, x_3 en su dominio.

Consecuentemente la función de agregación n -aria puede ser de un modo único aplicando iterativamente f_2 como

$$f_n(x_1, \dots, x_n) = f_2(f_2(\dots f_2(x_1, x_2), x_3), \dots, x_n).$$

Ejemplo 2.23. Las funciones producto, mínimo y máximo son funciones de agregación asociativas. La media aritmética no es asociativa.

Definición 2.24. Una función de agregación extendida F se puede descomponer si para todo $m, n = 1, 2, \dots$ y para todo $\mathbf{x} \in [0, 1]^m$, $\mathbf{y} \in [0, 1]^n$:

$$\begin{aligned} f_{m+n}(x_1, \dots, x_m, y_1, \dots, y_n) &= \\ f_{m+n}(\underbrace{f_m(x_1, \dots, x_m), f_m(x_1, \dots, x_m)}_{m \text{ veces}}, y_1, \dots, y_n). \end{aligned}$$

Definición 2.25. Una función de agregación extendida F es bisimétrica si para todo $m, n = 1, 2, \dots$ y para todo $\mathbf{x} \in [0, 1]^{mn}$:

$$\begin{aligned} f_{mn}(\mathbf{x}) &= f_m(f_n(x_{11}, \dots, x_{1n}), \dots, f_n(x_{m1}, \dots, x_{mn})) \\ &= f_n(f_m(x_{11}, \dots, x_{m1}), \dots, f_m(x_{1n}, \dots, x_{mn})). \end{aligned}$$

Una función de agregación extendida simétrica asociativa es bisimétrica. Sin embargo existen funciones de agregación extendidas simétricas y bisimétricas no asociativas, como por ejemplo las medias aritmética y geométrica.

Definición 2.26. Una función de agregación $f : [0, 1]^n \rightarrow [0, 1]$ es invariante ante traslaciones si para todo $\lambda \in [-1, 1]$ y para todo $(x_1, \dots, x_n) \in [0, 1]^n$ se cumple

$$f(x_1 + \lambda, \dots, x_n + \lambda) = f(x_1, \dots, x_n) + \lambda$$

cuando $(x_1 + \lambda, \dots, x_n + \lambda) \in [0, 1]^n$ y $f(x_1, \dots, x_n) + \lambda \in [0, 1]$.

Definición 2.27. Una función de agregación $f : [0, 1]^n \rightarrow [0, 1]$ es homogénea si para todo $\lambda \in [0, 1]$ y para todo $(x_1, \dots, x_n) \in [0, 1]^n$ se cumple

$$f(\lambda x_1, \dots, \lambda x_n) = \lambda f(x_1, \dots, x_n).$$

Las funciones de agregación que son invariantes ante traslaciones y homogéneas se denominan *funciones de agregación lineales*.

Capítulo 3

Estudio matemático del problema

Cualquier problema en el contexto de la ciencia de datos debe considerarse desde un punto de vista matemático. En el caso en que nos ocupa, la visualización de imágenes MHE, existe relativamente poca literatura relevante.

La tarea puede, o debe, entenderse como un proceso de fusión de información. Al fin y al cabo, existe una información asociada a cada píxel cuya dimensionalidad es $N > 3$ y debe reducirse a $N = 3$. Esta información puede darse en forma de vector (multiespectral), función discretizada (hiperespectral), o cualquier forma intermedia; pero esto parece ser irrelevante desde el punto de vista de la fusión. Ahora bien, la manera en que esta fusión se aborda puede ser muy debatible. Se podría trabajar píxel a píxel, dada una función de fusión

$$f : \mathbb{R}^N \mapsto \mathbb{R}^3 ,$$

de la misma manera que se podría trabajar con operadores de área local que consideren vecindarios alrededor de cada píxel. Por ejemplo,

$$f : (\mathbb{R}^{(k \times k)})^n \mapsto \mathbb{R}^3 ,$$

donde $k \times k$ es el tamaño del área considerada alrededor de cada píxel.

En cualquier caso, existe una duda más allá del modelo de fusión elegido, y no es otra que el objetivo del mismo. ¿Qué debe cumplir una visualización de una imagen MHE? ¿Qué debe preservar como identificable/indistinguible? ¿Cuáles son los criterios de calidad que determinarán el éxito del proceso?

La literatura contiene relativamente pocos ejemplos de autores interesados en estas preguntas, a pesar de que parece un problema apremiante. De hecho, la mayor parte de los trabajos basados en tecnología MHE muestran algún tipo de visualización de sus datos. Sin embargo, estas visualizaciones no parecen seguir estándares o criterios de calidad.

Un trabajo relevante, y totalmente enfocado en la visualización de imágenes MHE es [4]. En este trabajo los autores proponen un estudio de las objetivos que debe tener un proceso de visualización de imágenes hiperespectrales. En concreto, se interesan por modelar *consistent representations of hyperspectral data that can facilitate understanding and analysis of hyperspectral scenes, and may be used in conjunction with task-specific visualizations* [4]. Sus objetivos se resumen en una serie de objetivos que, idealmente, una visualización de imágenes hiperespectrales deberían cumplir.

Las visualizaciones propuestas por Jacobson y Gupta se restringen a visualizaciones tri-estímulo, véase RGB o equivalente. Por ello, la Sección 3.1 se refiere a la representación visual del color, mientras que la Sección 3.2 queda dedicada al análisis de los objetivos propuestos en [4]. La sección 3.3 se dedica

al análisis crítico de dichos objetivos, y finalmente la Sección 3.4 analiza las alternativas existentes para cumplirlos.

3.1. La importancia del color

En nuestra vida diaria nos encontramos rodeados de imágenes a color en impresiones, televisión, monitores de ordenadores, fotografías y películas. La percepción del color es el resultado de la interacción entre un estímulo físico, receptores en el ojo humano que percibe el estímulo y el sistema neural y el cerebro, los cuales son los responsables de comunicar e interpretar las señales percibidas por el ojo [10]. Este proceso implica múltiples fenómenos físicos, neuronales y cognitivos, los cuales han de ser entendidos para poder comprender de forma completa la visión de color.

El estímulo físico para el color es la radiación electromagnética en la región del visible del espectro, la cual se denomina comunmente como luz. En el aire o el vacío, la región visible del espectro electromagnético se define típicamente como la región de la longitud de onda entre $\lambda_{min} = 360\text{nm}$ y $\lambda_{max} = 830\text{nm}$. La luz estimula los receptores de la retina en el ojo, que causan el fenómeno de la visión y la percepción de color.

El color es un fenómeno percibido y no una dimensión física como pueden ser la longitud o la temperatura, aunque la radiación electromagnética del espectro visible es mensurable como una cantidad física [11]. El observador puede percibir dos sensaciones de color diferentes como completamente iguales o como metaméricas. La identificación de colores a través de la información obtenida del espectro, no es útil para etiquetar colores que han sido medidos fisiológicamente y evaluados en la mayoría de los casos con un número pequeño de sensores. Una forma apropiada de representación debe ser obtenida mediante el almacenamiento, visualización y procesamiento de las imágenes a color. Esta representación debe adecuarse a los requerimientos matemáticos del algoritmo de procesamiento de las imágenes a color, a las condiciones técnicas de la cámara, impresora o monitor, y a la percepción humana del color. Estos múltiples objetivos no pueden ser alcanzados por igual simultáneamente; por este motivo, se utilizan diferentes representaciones en el procesamiento de imágenes a color dependiendo del objetivo que deseemos lograr.

Los espacios de color indican los sistemas de coordenadas de color en los cuales los valores píxeles de la imagen están representados. La diferencia entre dos valores imagen en un espacio de color se denomina distancia de color. Los números que describen las diferentes distancias de color en el respectivo modelo de color, son por lo general distintos a las diferencias de color percibidas por los humanos.

3.1.1. El sistema de color estándar

El modelo de *mezcla de color aditiva* se utiliza cuando luz formada por distintas longitudes de onda alcanza un lugar idéntico en la retina o sensor de la imagen. Los distintos estímulos de color se combinan en un solo color mediante solapado. Debido a la Primera Ley de mezcla aditiva de colores de Grassmann, cualquier estímulo de color puede ser expresado de forma única por un conjunto particular de tres estímulos de colores primarios, mientras cada uno de los estímulos primarios sea independiente. Usando la notación vectorial, los vectores unitarios de los colores primarios pueden ser vistos como una base (no necesariamente ortonormal) de un espacio vectorial. La Primera Ley de Grassmann de la mezcla de color puede ser expresada como

$$\mathbf{M} = R\mathbf{R} + G\mathbf{G} + B\mathbf{B}. \quad (3.1)$$

Las cantidades de colores primarios \mathbf{R} , \mathbf{G} , \mathbf{B} en el color mezclado \mathbf{M} están indicadas como R , G y B . Sin embargo no todos los colores pueden ser producidos con un único conjunto de colores primarios.

Esto no representa ningún tipo de contradicción con la Primera Ley de mezcla de color de Grassmann ya que las cantidades de colores primarios pueden considerarse negativas en la combinación de colores, por ejemplo

$$\mathbf{M} + R\mathbf{R} = G\mathbf{G} + B\mathbf{B}.$$

El principio de combinación de color mencionado también se puede extender a un conjunto \mathbf{M} de n colores, donde aplicaríamos las relaciones

$$\mathbf{M} = \{\mathbf{M}_i | i = 1, \dots, n\} \text{ con } \mathbf{M}_i = R_i\mathbf{R} + G_i\mathbf{G} + B_i\mathbf{B} \text{ y} \quad (3.2)$$

$$\mathbf{M}_{1\dots n} = \mathbf{R} \cdot \sum_{i=1}^n R_i + \mathbf{G} \cdot \sum_{i=1}^n G_i + \mathbf{B} \cdot \sum_{i=1}^n B_i$$

Analizando la longitud de onda del visible en n intervalos estrechos, cada uno de ellos se correponderá de forma aproximada con un estímulo de un color espectral. Remarcar que la frecuencia de banda en general se encuentra entre 2 y 10 nanómetros. Si \mathbf{M}_i en las ecuaciones anteriores representa valores de intervalos espectrales de color, cualquier estímulo de color puede ser expresado a partir de una mezcla aditiva de un subconjunto de \mathbf{M} que se corresponde con la composición espectral. Si aproximamos la banda de longitud de onda a cero, obtendremos un estímulo de color continuo $M(\lambda)$ de la forma

$$M(\lambda) = R(\lambda) \cdot \mathbf{R} + G(\lambda) \cdot \mathbf{G} + B(\lambda) \cdot \mathbf{B}. \quad (3.3)$$

En aras de la estandarización, la Comisión Internacional de la Iluminación (CIE) estableció los valores primarios monocromáticos usados, así como las funciones de igualación o correspondencia del color, en 1931 como la definición del observador colorimétrico estándar CIE 1931 de 2°.

Las tres funciones indican la cantidad de cada primario que son necesarias para igualar el color de un vatio de potencia radiante de la longitud de onda indicada. Han sido definidas de modo que las áreas situadas debajo de las tres curvas sean iguales entre sí para que los valores del triestímulo del blanco equienérgico sean iguales. Las funciones vendrán dadas por la expresión

$$\bar{m}(\lambda) = \bar{r}(\lambda) \cdot \mathbf{R} + \bar{g}(\lambda) \cdot \mathbf{G} + \bar{b}(\lambda) \cdot \mathbf{B}. \quad (3.4)$$

Para poder obtener una mezcla de color que cubra todo el área de las longitudes de onda correspondientes a la región del visible, el CIE definió unos valores primarios virtuales X , Y , Z , ya que presentan una mayor aplicación en colorimetría. La conversión de las curvas de valores espectrales reales a funciones de igualación virtuales $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ se obtiene a través de una transformación lineal de las funciones de igualación de la forma

$$\begin{pmatrix} \bar{x}(\lambda) \\ \bar{y}(\lambda) \\ \bar{z}(\lambda) \end{pmatrix} = \begin{pmatrix} 0.49000 & 0.31000 & 0.20000 \\ 0.17697 & 0.81240 & 0.01063 \\ 0.00000 & 0.01000 & 0.99000 \end{pmatrix} \cdot \begin{pmatrix} \bar{r}(\lambda) \\ \bar{g}(\lambda) \\ \bar{b}(\lambda) \end{pmatrix}.$$

3.1.2. Espacio de color RGB

El espacio de color más comunmente usado en tecnología computacional es el espacio de color RGB, el cual está basado en la mezcla aditiva de tres colores primarios \mathbf{R} , \mathbf{G} y \mathbf{B} . Se trata de un espacio de color orientado técnicamente. Las longitudes de onda estandarizadas internacionalmente son 700.0 nm, 546.1 nm y 435.8 nm para el rojo, verde y azul respectivamente. Remarcar que los términos rojo,

verde y azul fueron introducidos únicamente con el propósito de estandarizar la descripción de los colores primarios; los colores visibles y las longitudes de onda no son equivalentes. Para evitar posibles confusiones se puede usar la notación L , M y S para las longitudes de onda larga, media y corta respectivamente, en lugar de la notación R , G y B .

Los colores primarios son para la mayoría los colores de referencia de los sensores de escaneo. Forman una base vectorial de un espacio tridimensional ortogonal vectorial de color, donde el vector cero representa el negro; el origen también se denomina *punto negro*. Por lo tanto cualquier color puede ser visto como una combinación lineal de los vectores base en el espacio RGB ; en este espacio de color cualquier imagen a color es automáticamente tratada como una función vectorial de tres componentes. Los tres componentes del vector vienen determinados por las intensidades medidas de la luz visible en el área de onda larga, media y corta. Para una imagen digital en color \mathbf{C} (tres canales), se definen tres componentes vectoriales R , G , B para cada píxel de la imagen (x, y)

$$\mathbf{C}(x, y) = (R(x, y), G(x, y), B(x, y))^T = (R, G, B)^T. \quad (3.5)$$

Nos referimos a estos valores como valores triestímulos. Los colores que representamos mediante combinaciones explícitas de valores de los componentes vectoriales R , G , B son entidades dependientes. Todos los vectores $(R, G, B)^T$ con componentes enteros $0 \leq R, G, B \leq G_{max}$ caracterizan un único color en el espacio de color RGB . $G_{max} + 1$ indica el mayor valor permitido en cada componente del vector.

Los colores primarios rojo $(G_{max}, 0, 0)^T$, verde $(0, G_{max}, 0)^T$, azul $(0, 0, G_{max})^T$, y los colores complementarios amarillo $(G_{max}, G_{max}, 0)^T$, magenta $(G_{max}, 0, G_{max})^T$, cian $(0, G_{max}, G_{max})^T$, así como los colores acromáticos blanco $(G_{max}, G_{max}, G_{max})^T$ y negro $(0, 0, 0)^T$, representan los límites de un cubo de color, que está formado por las posibles combinaciones de valores R , G , B . Todos los vectores de color $(R, G, B)^T$ con $0 \leq R, G, B \leq G_{max}$ caracterizan un color en el espacio de color RGB . Todos los colores acromáticos (tonos grises) se encuentran en la diagonal principal $(u, u, u)^T$, con $0 \leq u \leq G_{max}$.

El espacio de color RGB es el más aplicado en la representación computacional interna de imágenes a color. Su amplia distribución se debe, entre otras cosas, a la gran estandarización de los tres colores primarios. Casi todos los colores visibles pueden ser representados por una combinación lineal de los tres vectores. Para objetos idénticos, valores de colores que difieren son generados con cámaras o escáneres distintos ya que sus colores primarios en general no coinciden. El proceso de ajustar los valores de los colores entre distintos instrumentos se denomina *gestión del color*; por ejemplo entre una cámara RGB y un monitor RGB.

En 1996, el Consorcio Internacional del Color (ICC) propuso un espacio de color estándar $sRGB$ para Internet. Los valores triestímulos son simplemente combinaciones de los valores XYZ de CIE. Pueden ser obtenidos mediante la expresión

$$\begin{pmatrix} R_{sRGB} \\ G_{sRGB} \\ B_{sRGB} \end{pmatrix} = \begin{pmatrix} 3.2410 & -1.5374 & -0.4986 \\ -0.9692 & 1.8760 & 0.0416 \\ 0.05556 & -0.2040 & 1.0570 \end{pmatrix} \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}.$$

3.1.3. Otros espacios de color

Un *espacio de color uniforme* es un espacio de color en el cual, cambios en las coordenadas de color dan lugar a cambios proporcionales en el tono y la saturación del mismo. Esto no ocurre en el caso de espacios de color físicos o técnicos, como es el caso del RGB. La CIE recomendó el uso del espacio de color LAB 1976 (CIELAB) como una aproximación a espacios de color uniformes. Es derivado del

sistema de color estándar XYZ y fue desarrollado para facilitar la medida de color de acuerdo con el sistema de color de Munsell. Este sistema se basa en tres cualidades que dispone cada color: tono o matiz, luminosidad o valor y saturación o colorido.

Los parámetros L^*a^*b representan la luminosidad de color, la posición entre el rojo y el verde, y la posición entre el amarillo y el azul respectivamente. Ya que se trata de un modelo muy extendido en imagen, las transformaciones hacia y desde CIELAB a otros espacios de color son muy usadas.

Existen espacios de colores basados intuitivamente en la percepción humana del color, que presentan un gran interés en los campos de visión por computador y computación gráfica. Un color puede ser descrito con mayor facilidad de forma intuitiva mediante los valores de tono, saturación del color e intensidad, en lugar de los componentes vectoriales en el espacio de color RGB . En el espacio de color HSI el tono, la saturación y la intensidad forman los ejes de coordenadas. Este espacio de color es muy adecuado para el procesamiento de imágenes a color y para definir características visuales locales interpretables. Definimos un color $\mathbf{q} = (R, G, B)^T$ en el espacio de color RGB . El tono H del color \mathbf{q} caracteriza el color dominante contenido en \mathbf{q} ; la saturación S del color \mathbf{q} es una medida de la pureza del color. Este parámetro depende del número de longitudes de onda que contribuyen a la percepción del color; cuanto más ancho sea el rango de las longitudes de onda, menor será la pureza del color y viceversa. La intensidad I del color \mathbf{q} se corresponde con la luminosidad relativa. Para el color $\mathbf{q} = (R, G, B)^T$ en el espacio de color RGB , la representación $(H, S, I)^T$ se encuentra en el espacio de color HSI . Una de las ventajas del espacio de color HSI es la separación de la información cromática y acromática. Una desventaja de este modelo es la existencia de singularidades.

El espacio de color HSV , también llamado HSB , es muy comunmente usado en el campo de computación gráfica. Como el espacio de color HSI se trata de un modelo basado en la percepción humana, y usamos el tono, saturación e intensidad como ejes de coordenadas. Tanto el espacio de color HSV como el HSI presentan el problema de que la transformación de una línea recta en el espacio RGB , en general no se corresponde con una línea recta en estos dos modelos. Una ventaja del espacio HSV reside en que se corresponde con el sistema de color del que haría uso un pintor al mezclar los colores, por lo que resulta muy intuitivo operar con él y su aprendizaje. En procesamiento de imágenes a color digitales, el espacio de color HSV es de una importancia menor; es usado en manipulación de los valores de color de una imagen, por ejemplo en software de edición de imágenes raster.

Además de los tres canales RGB de una imagen, existe un cuarto canal denominado *canal alpha*. Este canal de 8 bits determina la opacidad del píxel a definir, es decir, el grado de transparencia que presenta. Un valor alpha de cero indica un píxel completamente transparente, que no aporta ninguna contribución a la imagen final; un valor alpha de 255 representa un píxel completamente opaco, que oculta cualquier píxel subyacente. Los valores intermedios entre 0 y 255, representan contribuciones parciales o colores semi-transparentes. Este canal extra nos permite unir y fusionar imágenes que presenten distintas formas y tamaños. dentro de los múltiples formatos de imágenes disponibles, algunos como el JPG no soportan dicha información de canal alpha, por lo que en muchos casos no será el formato adecuado para trabajar. Por el contrario, existen algunos formatos como el TIFF, TGA o PNG que sí que pueden almacenar información en canales alpha, por lo que su uso se encuentra muy extendido.

Por último, cabe destacar que algunos autores han buscado representaciones de color más completas y complejas que el tristímulus estándar. Por un lado, están los autores que han trabajado en *gamuts* (paletas) específicas que tratan de capturar circunstancias específicas o que tratan de adaptarse a ciertas condiciones de las imágenes. Por otro lado, y de una manera más ambiciosa, se han presentado modelos de 9 canales que son capaces de discernir los tres tipos de luz con influencia en el tono percibido en un píxel o región de píxeles: (a) la luz ambiental, (b) el color del objeto en cuestión y (c) los destellos propios del material del objeto. En estos trabajos sobresalen las diferentes propuestas de Gevers *et al.*

usando el model de color de von Kries (*von Kries model*) y las posteriores extensiones de Finlayson *et al.*. Estos modelos tratan de capturar, modelar y representar cada uno de los colores para poder hacer análisis de color que sean, por ejemplo, independientes de la luz ambiental presente en el momento de la captura de una imagen.

3.2. Objetivos de diseño de Jacobson-Gupta

Consideramos a continuación los principales objetivos de diseño de visualización para la fusión de un conjunto de imágenes multi- o hiperespectrales, de forma que preservamos la mayor cantidad de información y facilitemos la interpretabilidad. No todos los objetivos de diseño presentarán la misma importancia para todas las tareas y puede no ser posible alcanzarlos todos simultáneamente; sin embargo, lograr cada uno de ellos incrementará la transmisión efectiva de la información [4, 5].

Es importante remarcar que estos objetivos pueden aplicarse de manera convencional a casi cualquier tipo de imagen de alta dimensionalidad, siendo las imágenes hiperespectrales y las imágenes multiespectrales dos ejemplos de ello.

1. **Síntesis.** La visualización debe resumir con exactitud la información original. Las funciones lineales fundamentales proporcionan un promedio ponderado de la información y en este sentido, un resumen. Para que sea más adecuado para un uso general, debería existir la posibilidad de ponderar la representación de forma arbitraria. En la mayoría de los casos todos los componentes serán representados de forma equitativa, pero en algunas aplicaciones puede ser deseable enfatizar componentes con una mayor relación señal/ruido (SNR) o ponderar en mayor medida la información reciente.
2. **Representación consistente.** La representación visual de la información debe ser consistente con la visión humana, tal que los colores mostrados tengan significados coherentes. Cualquier espectro es siempre mostrado con el mismo valor de color, de forma que se favorezca su identificación en distintas imágenes; este objetivo también facilita la comparación entre diferentes representaciones. Si los colores visualizados se corresponden con un sistema ya existente de asociación de colores, nos ofrece una nueva ventaja. Esta limitación puede ser alterada para permitir un escalado de la luminosidad de la imagen representada, lo que resultaría en un espectro representado de forma consistente con la misma tonalidad y saturación.
3. **Rapidez computacional.** El sistema computacional debe ser suficientemente eficiente para representar la visualización de forma rápida, permitiendo interactividad a tiempo real.
4. **Preservación de bordes.** Los bordes (para todas las resoluciones espaciales) en las imágenes multi- e hiperespectrales originales son representados fielmente en la visualización.
5. **Optimización del sistema.** El diseño es optimizado para adecuarse a las características de la visualización y el sistema visual humano. Representaciones triestímulas estándar pueden ser abordadas desde el espacio de color RGB e interpretadas en el contexto de la visión humana usando un espacio de color diseñado por su uniformidad perceptiva como CIELab.
6. **Punto blanco de energía equitativa.** Un píxel con el mismo valor de reflectancia para cada banda espectral se representa como una tonalidad de gris. Por ello, la falta de saturación de color se relaciona con cuánto se asemeja un objeto a un cuerpo gris reflector difuso. En los extremos, un valor cero para todos las componentes del vector de información se representa como negro y un valor máximo para todas las componentes como el punto blanco de la representación. Podemos

representar esto mediante un vector de pesos \mathbf{w} , que para cada uno de los canales RGB cumpla la relación

$$\sum_{i=1}^n w_i = 1,$$

para una imagen multi- o hiperespectral de n bandas.

7. **Características preatentivas apropiadas.** La visualización debe minimizar las características preatentivas de la imagen que distraen al espectador sin razón de ser. Por ejemplo, una pequeña región brillante de color saturada sobre un fondo de un color distinto destacará, llamando la atención del espectador.
8. **Diferencia efectiva más pequeña.** Las distinciones visuales no han de ser mayores de lo estrictamente necesario para mostrar diferencias relativas. Las distinciones mínimas reducen las agrupaciones visuales y el uso de las mínimas diferencias efectivas ayuda en el diseño de elementos secundarios y estructurales como flechas o realces. Por ello podemos definir un valor umbral ϵ que determina la distinción mínima necesaria de forma que la diferencia visual entre dos puntos Ω_1, Ω_2 cumpla

$$\Omega_2 - \Omega_1 \leq \epsilon.$$

9. **Representación de luma equitativo.** Una representación de luma equitativo es un método de especificar que a todos los componentes de la información se les da la misma importancia en la visualización. Para precisar esta idea, definimos un punto de Kronecker como un vector de N componentes

$$x[n] = \delta[n - n_0],$$

donde n nos proporciona el índice de los componentes de la información $n = 1, 2, \dots, N$, n_0 es un índice fijo $n_0 \in \{1, 2, \dots, N\}$ y $\delta[n - n_0] = 0$ excepto para $n = n_0$, en cuyo caso $\delta[n - n_0] = 1$. Si la luminosidad CIELab representada de puntos de Kronecker es la misma para todo n_0 , se denomina una representación de luma equitativo.

10. **Diferencias cromáticas equitativas.** La diferencia cromática entre dos representaciones para dos puntos de Kronecker

$$x_0 = \delta[n - n_0] \text{ y } x_1 = \delta[n - n_1],$$

debería ser la misma para todas las diferencias espaciadas equitativamente $\|n_0 - n_1\|$. Esto ayuda al espectador a distinguir de forma correcta diferencias en la representación.

11. **Paleta natural.** Las visualizaciones usan una paleta y una distribución de color consistentes con las imágenes naturales. Este objetivo se basa en que los humanos están diseñados para interpretar escenas naturales y se basa en parte en la interpretación incorrecta causada por los colores fuertes saturados.

La visualización crea una paleta natural de colores, produciendo colores preatentivos (como colores brillantes saturados) cuando es apropiado, pero sin crear distracciones preatentivas no informativas. Múltiples colores saturados y brillantes son evitados en una visualización por expertos diseñadores por ser molestos y confusos. Los colores fuertes en grandes regiones pueden inducir a efectos de contraste simultáneos que crean imprecisiones en los colores percibidos en regiones pequeñas de imágenes. Las funciones fundamentales fijadas usadas en la visualización multi- e hiperespectral pueden dar lugar a una paleta natural y pueden ser diseñadas para evitar colores brillantes saturados inapropiados que dan lugar a distracciones preatentivas. Sin embargo, el escalado de valores para la representación jugará un papel determinante en la distribución de las propiedades perceptivas del color.

12. **Invariancia ante traslaciones de longitud de onda.** Todas las longitudes de onda presentan la misma ponderación; esto permite al método de visualización trabajar de forma correcta con cualquier número de bandas espectrales en cualquier rango espectral. Esto convierte al método de visualización en fácilmente adaptable a zooming espectral, panning espectral o nueva instrumentación. Un requisito fundamental es que un pico monocromático $\delta(\lambda)$ en cualquier longitud de onda λ se representa con la misma luminosidad. Otra característica deseada es que la diferencia percibida en una propiedad del color (como el tono) entre dos espectros monocromáticos dados $\delta(\lambda_1)$, $\delta(\lambda_2)$ depende solo de la diferencia en la longitud de onda $\lambda_2 - \lambda_1$.
13. **Facilidad de selección de subconjunto de componentes.** Debería ser posible y consistente poder realizar un zoom a un subconjunto de las dimensiones de los datos y visualizar algunos $d < N$ componentes de la información; o bien panear respecto a las dimensiones de los datos y visualizar un subconjunto distinto de los d componentes de la información. Las funciones fundamentales se pueden estirar, comprimir o trasladar para adecuarse a las dimensiones de los datos de interés y permitir la interactividad para explorar las dimensiones de la información. La habilidad de panear o hacer zoom se ve favorecida si todas las dimensiones de los datos son tratadas de forma igualitaria en términos de cualidades perceptivas de proyección del color, como luminosidad (Objetivo de representación de luma equitativo) y diferencias cromáticas (Objetivo de diferencias cromáticas equitativas).

3.3. Evaluación crítica de los objetivos

La aproximación de Jacobson y Gupta es la más completa de las aproximaciones al problema. De hecho, ha tenido un cierto impacto en la literatura, si bien el trabajo no ha sido continuado por ningún autor. A pesar de que este aporte tiene un valor muy significativo, pueden encontrarse puntos de mejoría. En concreto, y de manera somera, podemos considerar los siguientes puntos débiles en la aproximación de Jacobson y Gupta.

- La aproximación completa propuesta por los autores es irrealizable. Es decir, es imposible realizar una propuesta que alcance todos los objetivos. De esta manera, no se puede decir que los autores propongan los criterios que definen un buen operador de visualización, sino una serie de objetivos individuales que podrían (o no) ser abordados por un operador. De alguna forma habría sido útil generar una o varias agrupaciones de los mismos que pudieran considerarse como definiciones legítimas de calidad genérica para un operador de visualización.
- Como los propios autores comentan, la propia satisfacción de los criterios generales depende tanto del contexto de aplicación del problema como de la información específica presente en cada imagen o lote de imágenes. Esto significa que una solución aceptable (es decir, satisfactoria) en cierto contexto se puede volver insuficiente en otro.
- Los autores no aportan, en la mayor parte de los casos, definiciones matemáticas de los objetivos. Esto abre la puerta a una cierta subjetividad en la interpretación de los mismos.
- Finalmente, y quizá como consecuencia de lo anterior, los criterios no vienen acompañados de medidas claras de satisfacción total o parcial de los objetivos. Dicho de otra manera, no existen maneras de medir cuánto se satisfacen los criterios dado un contexto genérico o específico.

Nótese que todos estas críticas tienen un fundamento claro, pero a la vez son comprensibles dada la naturaleza del problema. Quizá, si la propuesta hubiera venido acompañada de diferentes *datasets*

de referencia en los cuales probar propuestas, su recorrido en la literatura habría sido diferente. En todo caso, se trata de la propuesta más avanzada realizada hasta la fecha, y es por tanto la referencia para el problema.

3.4. Alternativas para la fusión

Como hemos expuesto en el capítulo 1 muchas aplicaciones relacionadas con la teledetección producen grandes conjuntos de imágenes, como imágenes multi- e hiperespectrales o secuencias de imágenes temporales [5]. Estos grandes conjuntos de imágenes son proyectados en un espacio de dimensiones menores para poder realizar tareas de procesamiento de imágenes. La información espacial es de gran importancia para algunas aplicaciones de clasificación y detección, pero las técnicas de reducción de la dimensionalidad más populares no suelen tener en cuenta la información espacial. Estudiaremos distintos métodos de visualización de estos conjuntos de imágenes mediante diferentes alternativas en el proceso de correspondencia de los canales R,G,B con las bandas del espectro electromagnético.

El análisis de componentes principales (PCA) es quizás el método de proyección lineal más usado [8]. Las bandas adyacentes en una imagen multi- o hiperespectral obtenida mediante teledetección suelen por lo general presentar cierta correlación. Esto implica que existe redundancia en la información y se encuentra repetida entre distintas bandas; esta repetición de la información se ve reflejada en las inter-correlaciones entre bandas.

Si dos variables x e y están perfectamente correlacionadas, sus correspondientes medidas en x e y se representarán en una línea recta cuya pendiente tienda hacia la esquina superior derecha. En el caso de que las variables no se encuentren perfectamente correlacionadas, puede existir de todos modos una dirección dominante de dispersión o variabilidad; si tomamos esta dirección dominante como el eje principal, podemos dibujar un eje menor perpendicular al anterior. Una representación usando estos nuevos ejes, en lugar de los ejes convencionales x e y , pueden mostrar en muchos casos estructuras más reveladoras presentes en los datos.

Debemos remarcar una clara distinción entre el número de variables (bandas espectrales) en el conjunto de datos de la imagen y la dimensionalidad intrínseca del conjunto de datos. En casos de correlación entre dos variables x e y , la dimensionalidad de la información es uno, a pesar de que el número de variables sea dos; en este caso el uso de un único eje mayor en la dirección dominante, en lugar de los ejes convencionales x e y , logra dos objetivos: reducir el tamaño del conjunto de datos ya que un único eje reemplaza a las dos coordenadas en el eje x e y , y la información transmitida por las coordenadas en el nuevo eje es mayor que la información transmitida mediante medidas en los ejes x o y individualmente. En este contexto la información significa variancia o dispersión respecto a la media; también puede relacionarse con el rango de estados o niveles en los datos.

Conjuntos de datos de imágenes multi- e hiperespectrales generalmente tienen una dimensionalidad menor que el número de bandas espectrales que presentan. El objetivo del análisis de componentes principales es definir el número de dimensiones presentes en los datos y fijar los coeficientes que especifican las posiciones del conjunto de ejes que marcan las direcciones de mayor variabilidad de los datos; estos ejes siempre están no correlacionados. Por lo tanto una transformada de componentes principales de una imagen multi- o hiperespectral debe llevar a cabo las siguientes tareas: definir la dimensionalidad del conjunto de datos e identificar los ejes principales de variabilidad de la información.

Estas propiedades del análisis de componentes principales (también denominado la transformada de Karhunen-Loève) pueden ser de gran utilidad si deseamos comprimir el conjunto de datos. También relaciones entre distintos grupos de píxeles que representan diferentes cubiertas de suelos pueden verse con mayor claridad, si son visualizados en el sistema de referencia del eje principal en lugar de en términos de las bandas espectrales originales, especialmente si la variancia del conjunto de datos

está concentrada en unos relativamente pocos componentes principales. La propiedad de compresión de los datos es de utilidad si disponemos de más de tres bandas espectrales. Un sistema de visualización convencional *RGB* relaciona una banda espectral con uno de los tres colores de entrada (rojo, verde y azul).

El Landsat TM provee siete bandas de información, por lo que debemos tomar la decisión de qué tres de esas siete bandas escogemos para poder realizar una composición de una imagen a color. Si la dimensionalidad básica de los datos es tres, podremos expresar la mayor parte de la información contenida en las siete bandas en términos de los tres componentes principales. Las imágenes de componentes principales pueden ser utilizadas para generar una composición RGB en falso color con el primer componente principal mostrado en rojo, el segundo mostrado en verde y el tercero en azul. Esta imagen contiene más información que cualquier otra combinación de tres bandas espectrales. El análisis de componentes principales es un método de proyección lineal y adaptable a los datos, por ello puede dar lugar a visualizaciones inconsistentes que pueden ser difíciles de interpretar.

Los autovectores asociados con los componentes principales definen unos ejes de coordenadas en el espacio de información multidimensional que contiene la información de las imágenes multi- e hiperespectrales. Esta información puede ser reexpresada en términos de un nuevo conjunto de ejes de coordenadas y las imágenes resultantes presentan ciertas propiedades que pueden ser más adecuadas para ciertas aplicaciones. La transformada de Fourier opera en imágenes de una sola banda (escala de grises), no en conjuntos de datos multi- e hiperespectrales. El objetivo de este método es descomponer la imagen en sus componentes espaciales, que se definen como ondas sinusoidales con diversas amplitudes, frecuencias y direcciones. Las coordenadas del espacio bidimensional en las cuales se definen los componentes escalares se escriben en términos de la frecuencia. Esto se denomina el dominio de frecuencia en oposición al sistema de coordenadas en el cual las imágenes se expresan comúnmente, que se denomina el dominio espacial. La transformada de Fourier se utiliza para convertir una imagen unibanda desde la representación en su dominio espacial a su representación equivalente en el dominio de frecuencias y viceversa.

La idea detrás de la transformada de Fourier es que los valores de escala de grises que forman una imagen unibanda pueden ser vistos como una superficie tridimensional de intensidad, con las filas y columnas representando dos ejes y la intensidad del nivel de gris en cada píxel representando la tercera dimensión. Una serie de ondas de frecuencia creciente y con diferentes orientaciones es ajustada a la superficie de intensidad y la información asociada a cada onda es obtenida. La transformada de Fourier nos proporciona detalles de la frecuencia de las componentes escalares ajustadas a la imagen y la proporción de información asociada a cada componente de frecuencia. Si extendemos estas nociones a una función definida sobre una rejilla bidimensional, los componentes escalares serían ondas bidimensionales y cada componente escalar estaría caracterizado por la orientación así como por su amplitud. La principal aplicación de las transformadas de Fourier en teledetección es el filtrado en el dominio de frecuencia, así como la caracterización de tipos particulares de terreno mediante sus transformadas de Fourier.

Mediante las transformadas de Fourier podemos representar la información contenida en una imagen en el dominio espacial y el dominio de frecuencia. A pesar de que la representación en el dominio de frecuencia contiene información sobre la presencia de diferentes ondas que forman la imagen en escala de grises, no nos indica el lugar de la imagen en el cual una onda específica con una frecuencia particular tiene lugar; asumimos que el mix de frecuencias es el mismo en toda la imagen. Otra desventaja del dominio de frecuencia es que asumimos que existe estacionaridad estadística, lo que conlleva que la media y la variancia de los valores píxel son constantes en todas las regiones de la imagen. La transformada discreta wavelet (DWT) representa un compromiso entre las representaciones en el dominio de frecuencia y el dominio espacial, mejorando así la transformada de Fourier. Es imposible

medir simultáneamente las frecuencias presentes en una imagen en escala de grises y la localización de esas frecuencias; esto se trata de una extensión del principio de incertidumbre de Heisenberg. Sin embargo es posible transformar una imagen en una representación que combina bandas de frecuencias y áreas espaciales específicas.

La transformada wavelet puede ser considerada como la aplicación de una secuencia de pares de filtros (de paso alto y de paso bajo) a una serie de datos. Según vamos avanzando en este proceso de filtrado, iremos creando distintos niveles de datos, cuyo número de muestras irá disminuyendo. Ya que el suavizado progresivo implica una reducción continua de los datos de alta frecuencia, la representación de la información de frecuencia alta mayor será el primer nivel y la de frecuencia alta menor será el último nivel. La transformada wavelet bidimensional se calcula del mismo modo que la unidimensional, esto es, realizando una transformada wavelet a lo largo de las filas de la imagen para obtener una matriz intermedia y a continuación a las columnas de esta matriz. Las aplicaciones de transformadas wavelet en procesamiento de imágenes incluyen desde métodos de *pan-sharpening* (combinación de imágenes multi- y pancromáticas) hasta eliminación de niebla en imágenes de muy alta resolución.

Como hemos expuesto con anterioridad el análisis de componentes principales (PCA) no tiene en cuenta la información espacial, trata el conjunto de imágenes espectrales como un conjunto no ordenado de píxeles de altas dimensiones [12]. El método Wavelet es un modo eficiente y práctico de representar bordes e información de la imagen en múltiples escalas espaciales. Incorporando información espacial en el proceso de reducción de la dimensionalidad, beneficiaríamos los algoritmos de clasificación más que en el caso de usar una reducción de la dimensionalidad invariante espacialmente. Las características de una imagen a una escala dada, como casas o carreteras, pueden ser realizadas mediante un filtrado de los coeficientes wavelet. En muchas tareas, este método puede ser una representación más útil que los píxeles. Por ello consideramos un método de reducción de la dimensionalidad PCA con coeficientes wavelet para maximizar la información de los bordes en el conjunto de imágenes cuya dimensión ha sido reducida. Remarcar que la transformada wavelet actuaría espacialmente sobre cada una de las bandas de la imagen, mientras la transformada PCA tendría lugar sobre un conjunto de imágenes, es decir, las dos transformadas operarían sobre diferentes dominios.

Otro método de fusión de imágenes multi- e hiperespectrales en el que nos centraremos con mayor profundidad en este trabajo, es mediante operadores de agregación. Este método se trata de un método lineal y parametrizable, y nos ofrece un modo de procesamiento de imágenes adaptativo a cada conjunto de datos. Presentamos en el capítulo 4 la fusión basada en operadores de agregación.

Capítulo 4

Propuesta

4.1. Operadores OWA

Como primer método de fusión para imágenes multi- e hiperespectrales, estudiaremos los operadores OWA como una subclase dentro de los operadores de agregación presentados en la sección 2.2. Gracias a ellos podremos obtener tres funciones lineales mediante la aplicación de un conjunto de pesos a los datos, que podremos relacionar con cada uno de los canales *RGB* y obtener una visualización consistente con los objetivos presentados en la sección 3.2.

Los operadores OWA (*Ordered Weighted Averaging*) forman parte de la clase de funciones de agregación de promedio [9]. Los operadores OWA son funciones de agregación simétricas que asignan pesos en función del valor de entrada, por lo que pueden enfatizar el mayor, el menor o el valor de entrada del rango medio. Tienen una gran importancia y popularidad en las técnicas de conjuntos difusos.

Reutilizaremos la notación presentada con anterioridad \mathbf{x}_{\downarrow} , que denota el vector obtenido de \mathbf{x} reordenando sus componentes en un orden decreciente $x_{(1)} \geq x_{(2)} \geq \dots \geq x_{(n)}$.

Definición 4.1. Para un vector de pesos dado \mathbf{w} , $w_i \geq 0$, $\sum w_i = 1$, la función OWA viene dada por

$$OWA_{\mathbf{w}}(\mathbf{x}) = \sum_{u=1}^n w_u x_{(u)} = \langle \mathbf{w}, \mathbf{x}_{\downarrow} \rangle .$$

Definición 4.2. Dada una función OWA de la forma $OWA_{\mathbf{w}}$, la función inversa viene dada por $OWA_{\mathbf{w}_d}$, con el vector de pesos $\mathbf{w} = (w_n, w_{n-1}, \dots, w_1)$.

Los operadores OWA tienen una serie de propiedades principales que comparten todas las funciones del conjunto.

- Al igual que todas las funciones de agregación de promedio, los operadores OWA son no decrecientes (estrictamente crecientes si todos los pesos son positivos) e idempotentes.
- El dual de una función OWA es la función OWA *inversa*, con el vector de pesos $\mathbf{w} = (w_n, w_{n-1}, \dots, w_1)$.
- los operadores OWA son continuos, simétricos, homogéneos e invariantes ante traslaciones.
- los operadores OWA no tiene elemento neutro ni absorbente, excepto por los casos especiales de mínimo y máximo.

Existen distintas medidas para caracterizar un vector de pesos y el tipo de agregación que ejecuta. La primera medida que presentaremos nos da información sobre el carácter actitudinal del operador, también denominado *orness*, y se trata de una característica importante de las funciones de agregación.

Básicamente esta magnitud se trata de una medida de cuan lejos una función de agregación dada se encuentra de la función máximo, que es considerada la función disjuntiva más débil.

Definición 4.3. Sea f una función de agregación OWA, entonces la medida de su carácter actitudinal viene dada por

$$orness(OWA)_{\mathbf{w}} = \sum_{i=1}^n w_i \frac{n-i}{n-1} = OWA_{\mathbf{w}}(1, \frac{n-2}{n-1}, \dots, \frac{1}{n-1}, 0). \quad (4.1)$$

La medida de carácter actitudinal conlleva una serie de propiedades.

- La medida orness de una función OWA y su dual se encuentran relacionadas por

$$orness(OWA_{\mathbf{w}}) = 1 - orness(OWA_{\mathbf{w}_d}).$$

- Una función OWA es su propia función dual si y solo si $orness(OWA_{\mathbf{w}}) = \frac{1}{2}$.
- En el caso especial de $orness(\text{máx}) = 1$, $orness(\text{mín}) = 0$ y $orness(M) = \frac{1}{2}$. Es más, la medida del carácter actitudinal es solo 1 si y solo si se trata de la función máximo, y 0 si y solo si se trata de la función mínimo. Sin embargo la medida orness puede ser $\frac{1}{2}$ para una función OWA distinta de la media aritmética, la cual sin embargo nunca es su propia función dual.
- Si el vector de pesos es no decreciente, es decir, $w_i \leq w_{i+1}$, $i = 1, \dots, n-1$, la medida de orness $orness(OWA_{\mathbf{w}}) \in [\frac{1}{2}, 1]$; si el vector de pesos es no creciente, la medida de orness $orness(OWA_{\mathbf{w}}) \in [0, \frac{1}{2}]$.
- Si dos funciones OWA con vectores de pesos \mathbf{w}_1 , \mathbf{w}_2 tienen sus respectivos valores orness O_1, O_2 , y si $\mathbf{w}_3 = a\mathbf{w}_1 + (1-a)\mathbf{w}_2$, $a \in [0, 1]$, entonces la función OWA con el vector de pesos \mathbf{w}_3 tiene el valor orness dado por

$$orness(OWA_{\mathbf{w}}) = aO_1 + (1-a)O_2.^1$$

Otra cantidad importante que puede ser asociada a los vectores de pesos es su dispersión, también denominada entropía.

Definición 4.4. Para un vector de pesos dado \mathbf{w} su medida de dispersión viene dada por

$$Disp(\mathbf{w}) = - \sum_{i=1}^n w_i \log w_i, \quad (4.2)$$

con la convención $0 \cdot \log 0 = 0$.

La dispersión de pesos mide el grado en el que una función de agregación de pesos f tiene en cuenta todos los valores de entrada, es decir, el grado en el cual toda la información es usada en el proceso de agregación.

- Si la medida orness no está especificada, el máximo de la dispersión $Disp$ se consigue para $w_i = \frac{1}{n}$, es decir, para la media aritmética y $Disp(\frac{1}{n}, \dots, \frac{1}{n}) = \log n$.

¹Para determinar un vector de pesos OWA con el valor orness deseado, se pueden usar múltiples combinaciones de \mathbf{w}_1 y \mathbf{w}_2 , los cuales darían lugar distintos valores de \mathbf{w}_3 pero con el mismo valor orness.

- El mínimo valor de la entropía, 0, se alcanza si y solo si $w_i = 0$, $i \neq k$, y $w_k = 1$, es decir, para estadísticas de orden.
- La entropía de una función OWA y su dual coincide, es decir,

$$Disp(\mathbf{w}) = Disp(\mathbf{w}_d).$$

Los operadores OWA presentados son un método de fusión de información N -dimensional con $N > 3$, con una amplia aplicación en el campo de procesamiento de imágenes. Este sistema de reducción de la dimensionalidad no depende de las magnitudes particulares de los valores de entrada; por ello se trata de un método general para la obtención de una primera visualización de la información disponible, independiente de la tarea particular a la que deseemos aplicarlo. A partir de esta representación inicial podremos obtener los rasgos más importantes de la escena, lo cual nos ayudará en fases más avanzadas del procesamiento, en las cuales aplicaremos métodos más adaptados a nuestro trabajo concreto.

Mediante el uso de las funciones de agregación OWA podremos alcanzar los objetivos de diseño presentados en la sección 3.2. Los operadores OWA nos proporcionan un promedio ponderado de la información, es decir, un resumen de los datos disponibles debido a su carácter de función lineal. Debido a su definición matemática y sencillez, permite una rapidez computacional no alcanzable con otros sistemas de fusión de imágenes multidimensionales más complejos. Ya que la visualización hace uso de valores de reflectancia en todo el espectro electromagnético, distintas cubiertas presentarán diferentes comportamientos espectrales, preservando de este modo los bordes originales y características representadas en las imágenes multi- e hiperespectrales originales. El vector de pesos de los operadores OWA cumple la condición del punto blanco de igual energía, ya que cumple la relación

$$\sum_{i=1}^n w_i = 1,$$

para todo vector de pesos w_i .

Al hacer uso de un vector de pesos ponderados w_i , evitamos la presencia de características preatentivas y de diferencias exageradas en las distinciones visuales que distraerían al observador en la visualización; esto es debido a que realizamos un suavizado de los datos elegidos mediante el uso del vector \mathbf{w} . El método de funciones OWA es adaptable a cualquier número de dimensiones de los datos y nos permite visualizar un subconjunto de ellos, concediendo la misma importancia a todos los componentes de la información en la representación.

Los operadores OWA nos conceden la libertad de definir un vector de pesos que mejor se adapte a la visualización que deseemos obtener. Nos permiten aplicar un mayor peso a aquellos datos que nos aporten una mayor información y mejoren nuestra representación. Estos datos se corresponderán con los valores máximos y mínimos de reflectancia en el espectro electromagnético, es decir, aquellos que nos muestran las diferencias existentes entre distintos puntos del espacio de forma más clara; esto es debido a la distinta respuesta espectral entre cubiertas del suelo. Cada tipo de cubierta interactúa de un modo distinto con la radiación procedente del Sol en su mayoría, mostrando una respuesta característica denominada *firma espectral*. Los operadores OWA utilizan estos datos para poder realizar una representación consistente.

Las funciones de agregación OWA dependen en gran medida de los vectores de pesos ponderados que elijamos para nuestro estudio. Como ya hemos expuesto esto nos proporciona una gran libertad de diseño, pero también restricciones en la obtención del método más óptimo; para ello haremos uso de las medidas que caracterizan los vectores de pesos. Otra inconveniencia asociada a estas funciones, es la no utilización de todos los componentes de entrada en caso de existir pesos de ponderación nulos,

dando así lugar a una pérdida de información en la fusión, inherente a todos los métodos de reducción de la dimensionalidad.

Como operadores de agregación OWA para nuestro estudio, hemos elegido tres funciones con pesos w_i distintos para los tres canales de visualización *RGB*. Hemos decidido decantarnos por tres funciones que tengan en cuenta los valores máximos, mínimos y mediana de reflectancia en el rango total del espectro electromagnético. La primera de nuestras funciones tendrá asociado un vector de pesos ponderado dado por

$$\mathbf{w}_{\text{máx}} = (0.6, 0.3, 0.1, 0, \dots, 0).$$

Mediante este vector de pesos aplicado a un vector ordenado decreciente \mathbf{x}_{\searrow} , tomaremos los tres valores mayores a los que aplicaremos la ponderación para cada posición píxel en todas las bandas del espectro electromagnético. Obtendríamos el resultado dado por

$$OWA_{\mathbf{w}_{\text{máx}}}(\mathbf{x}) = \langle \mathbf{w}_{\text{máx}}, \mathbf{x}_{\searrow} \rangle = (0.6 \cdot x_1, 0.3 \cdot x_2, 0.1 \cdot x_3, 0, \dots, 0).$$

La segunda de nuestras funciones vendrá dada por un vector de pesos ponderado dado por

$$\mathbf{w}_{\text{mín}} = (0, \dots, 0.1, 0.3, 0.6).$$

Al contrario que con la primera función de agregación OWA, en este caso tomaremos los tres valores menores para cada posición píxel en todas las bandas del espectro electromagnético. Obtendríamos el resultado dado por

$$OWA_{\mathbf{w}_{\text{mín}}}(\mathbf{x}) = \langle \mathbf{w}_{\text{mín}}, \mathbf{x}_{\searrow} \rangle = (0, \dots, 0.1 \cdot x_{n-2}, 0.3 \cdot x_{n-1}, 0.6 \cdot x_n)$$

donde n representa el número de bandas espectrales recogidas.

La tercera y última función de agregación OWA usada vendrá dada por dos vectores de pesos distintos, dependiendo del número de bandas espectrales n que tenga nuestro conjunto de datos. Si se trata de un número par, el vector de pesos ponderado vendrá dado por

$$\mathbf{w}_{\text{mediana}_{\text{par}}} = (0, \dots, 0.5, 0.5, \dots, 0).$$

Por el contrario si se trata de un conjunto de datos de n bandas, con n impar, el vector \mathbf{w} vendrá dado por

$$\mathbf{w}_{\text{mediana}_{\text{impar}}} = (0, \dots, 0.25, 0.5, 0.25, \dots, 0).$$

En estos dos casos tomaremos los valores que nos proporcionan la mediana de los conjuntos de datos, a la cual le aplicaremos un vector de ponderación. Obtendríamos los siguientes resultados en cada uno de los casos

$$OWA_{\mathbf{w}_{\text{med}_{\text{par}}}}(\mathbf{x}) = \langle \mathbf{w}_{\text{mediana}_{\text{par}}}, \mathbf{x}_{\searrow} \rangle = (0, \dots, 0.5 \cdot x_{n/2}, 0.5 \cdot x_{n/2+1}, \dots, 0),$$

$$OWA_{\mathbf{w}_{\text{med}_{\text{impar}}}}(\mathbf{x}) = \langle \mathbf{w}_{\text{mediana}_{\text{impar}}}, \mathbf{x}_{\searrow} \rangle = (0, \dots, 0.25 \cdot x_{n/2-0.5}, 0.5 \cdot x_{n/2+0.5}, 0.25 \cdot x_{n/2+1.5}, \dots, 0)$$

donde n representa el número de bandas espectrales recogidas.

Las funciones de agregación OWA presentadas presentan unas mejoras muy claras respecto a los métodos de fusión de imágenes que hacen uso de las funciones clásicas de máximo, mínimo y promedio para la reducción de la dimensionalidad. Estas funciones básicas también pueden ser escritas en términos de funciones de agregación OWA del siguiente modo

$$\mathbf{w}_{\text{máx clas}} = (1, 0, \dots, 0) \rightarrow OWA_{\mathbf{w}_{\text{máx clas}}}(\mathbf{x}) = \langle \mathbf{w}_{\text{máx clas}}, \mathbf{x}_{\searrow} \rangle = (x_1, \dots, 0)$$

	Orness Máximo Clásico	Orness Máximo	Dispersión Máximo Clásica	Dispersión Máxima
<i>Salinas</i>	1	0.8982	0	0.8979
<i>PaviaU</i>	1	0.8961	0	0.8979
	Orness Mínimo Clásico	Orness Mínimo	Dispersión Mínimo Clásico	Dispersión Mínimo
<i>Salinas</i>	0	0.0058	0	0.8979
<i>PaviaU</i>	0	0.0127	0	0.8979
	Orness Media Clásico	Orness Mediana	Dispersión Media Clásico	Dispersión Mediana
<i>Salinas</i>	0.5	0.5000	5.4116	0.6931
<i>PaviaU</i>	0.5	0.2485	4.6367	1.0397

Tabla 4.1: Comparación de parámetros de estudio.

$$\mathbf{w}_{\text{mín clas}} = (0, \dots, 0, 1) \rightarrow OWA_{\mathbf{w}_{\text{mín clas}}}(\mathbf{x}) = \langle \mathbf{w}_{\text{mín clas}}, \mathbf{x}_{\setminus} \rangle = (0, \dots, x_n)$$

$$\mathbf{w}_{\text{media clas}} = (1, 0, \dots, 0) \rightarrow OWA_{\mathbf{w}_{\text{media clas}}}(\mathbf{x}) = \langle \mathbf{w}_{\text{media clas}}, \mathbf{x}_{\setminus} \rangle = \frac{1}{n} \sum_{i=1}^n x_i$$

donde n representa el número de bandas espectrales recogidas.

Mostramos en el capítulo 5 las visualizaciones obtenidas por ambos métodos para distintas escenas tomadas mediante diferentes sensores con un número variable de bandas del espectro electromagnético.

Para poder comparar estos dos métodos de forma más cuantitativa, haremos uso de los parámetros presentados con anterioridad para el estudio de los operadores de agregación OWA, el carácter actitudinal y la entropía. Calcularemos estos parámetros para dos conjuntos de datos distintos presentados en el capítulo 5 a modo de comparación para nuestro estudio, que denominaremos *Salinas* y *PaviaU*. Presentamos los resultados obtenidos en la tabla 4.1, obtenidos a partir de las ecuaciones 4.3 y 4.4.

Comparamos para los valores clásicos de máximo, mínimo y promedio, con los valores correspondientes obtenidos para las funciones de agregación OWA propuestas. Para el caso de la función máximo, observamos que los valores orness obtenidos para los operadores OWA son menores que para la función clásica, pero con un valor próximo a la unidad, por lo que conserva su proximidad al máximo. El valor de la dispersión en el caso OWA es mucho mayor que en el caso clásico, por lo que el grado de información utilizada en la agregación será más elevado.

En el caso de la función mínimo observamos que los valores orness se encuentran muy próximos en ambos casos, como era esperable ya que se tratan de las funciones opuestas al máximo. Al igual que en el caso anterior, el valor de la dispersión es mucho más elevado en el caso de los operadores OWA propuestos, incluyendo del mismo modo una mayor cantidad de información que en el caso clásico.

La última función de estudio compara los resultados obtenidos con la media aritmética clásica y la mediana obtenida por operadores OWA. Entre los dos conjuntos de datos existen discrepancias remarcables debido al uso del operador mediana par para el caso de *Salinas* e impar para *Pavia*. En el caso impar el carácter actitudinal se encuentra muy alejado del valor correspondiente al máximo, aproximándose más al mínimo., mientras el valor orness del caso par presenta la misma magnitud que la media aritmética. La dispersión en los casos clásicos es mucho mayor ya que hace uso de todos los componentes del conjunto de datos.

Hacemos una representación de los vectores de pesos ponderados w_i para los dos ejemplos presentados en la tabla 4.1, a modo de visualización más intuitiva del método aplicado. Mostramos los resultados en las figuras 4.1 y 4.2.

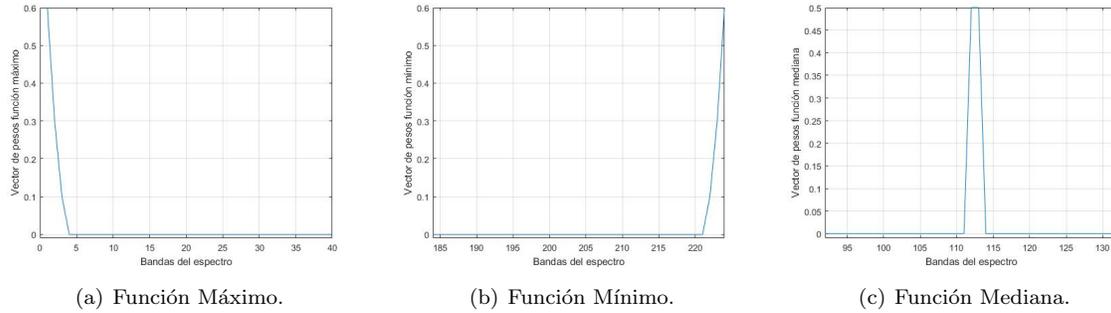


Figura 4.1: Representación funcional de diferentes vectores de pesos para operadores OWA en la imagen *Salinas*. Nótese que, en este caso, los operadores OWA se configuran de tal manera que se recuperan operadores clásicos.

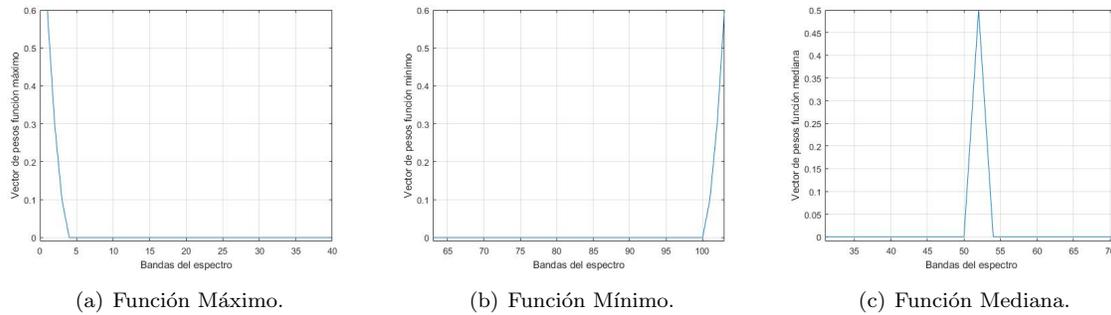


Figura 4.2: Representación funcional de diferentes vectores de pesos para operadores OWA en la imagen *PaviaU*. Nótese que, en este caso, los operadores OWA se configuran de tal manera que se recuperan operadores clásicos.

4.2. Ejemplos prácticos

En esta sección mostraremos cómo se puede convertir una imagen multi- o hiperespectral mediante el uso de operadores OWA, en un vector RGB que podamos visualizar correctamente. Para ello tomaremos diferentes espectros de las escenas de ejemplo presentadas en el capítulo 5, y les aplicaremos cuatro métodos distintos basados en funciones básicas y funciones OWA, de modo que podamos observar su diversa influencia en los datos.

Un espectro dado de una imagen multi- o hiperespectral, se trata de un vector de N elementos con $N > 3$. Por ello, mediante la aplicación de los operadores OWA presentados en la sección 4.1 anterior, a partir de un vector \mathbf{v} al que aplicaremos tres vectores de pesos w_i , podremos obtener tres componentes que conformarán el vector final *RGB*.

Mostramos en la figura 4.3 cuatro ejemplos prácticos aplicados a escenas presentadas en el capítulo 5: *PaviaU*, *Botswana*, *Salinas* e *Indian Pines*. Hemos hecho uso de cuatro funciones distintas para poder observar la influencia que la elección de estas mantiene sobre la selección del vector RGB y, por tanto, su correspondiente visualización final.

En el primero de los casos, hemos aplicado las funciones clásicas expuestas en la sección anterior, formadas por *máximo*, *mínimo* y *promedio*. Los tres casos siguientes se trata de operadores OWA, en los que hemos ido variando la cantidad de elementos a tener en cuenta en el vector de pesos w_i . En el primero de estos casos hemos tomado los pesos presentados en la sección anterior dados por

$$\begin{aligned}\mathbf{w}_{\text{máx}} &= (0.6, 0.3, 0.1, 0, \dots, 0), \\ \mathbf{w}_{\text{mín}} &= (0, \dots, 0, 0.1, 0.3, 0.6), \\ \mathbf{w}_{\text{mediana}_{\text{par}}} &= (0, \dots, 0.5, 0.5, \dots, 0), \\ \mathbf{w}_{\text{mediana}_{\text{impar}}} &= (0, \dots, 0.25, 0.5, 0.25, \dots, 0).\end{aligned}$$

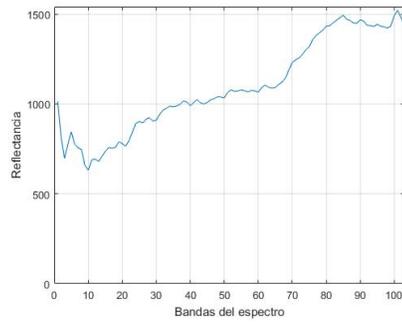
En el segundo de los casos correspondiente a los operadores OWA, hemos tomado un vector de pesos w_i , con cinco elementos no nulos, dados por las expresiones

$$\begin{aligned}\mathbf{w}_{\text{máx}} &= (0.4, 0.25, 0.20, 0.15, 0.05, 0, \dots, 0), \\ \mathbf{w}_{\text{mín}} &= (0, \dots, 0, 0.05, 0.15, 0.20, 0.25, 0.4), \\ \mathbf{w}_{\text{mediana}_{\text{par}}} &= (0, \dots, 0.5, 0.5, \dots, 0), \\ \mathbf{w}_{\text{mediana}_{\text{impar}}} &= (0, \dots, 0.25, 0.5, 0.25, \dots, 0).\end{aligned}$$

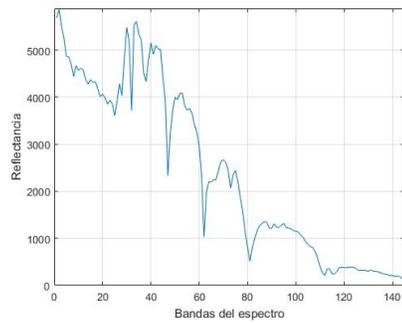
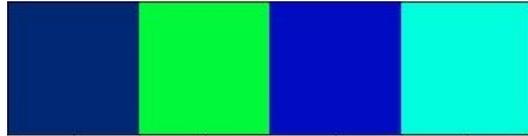
En el tercero y último de los casos correspondiente a los operadores OWA, hemos tomado un vector de pesos w_i , con diez elementos no nulos, dados por las expresiones

$$\begin{aligned}\mathbf{w}_{\text{máx}} &= (0.35, 0.25, 0.15, 0.10, 0.10, 0.05, 0.04, 0.03, 0.02, 0.01, 0, \dots, 0), \\ \mathbf{w}_{\text{mín}} &= (0, \dots, 0, 0.01, 0.02, 0.03, 0.04, 0.05, 0.10, 0.10, 0.15, 0.25, 0.35), \\ \mathbf{w}_{\text{mediana}_{\text{par}}} &= (0, \dots, 0.5, 0.5, \dots, 0), \\ \mathbf{w}_{\text{mediana}_{\text{impar}}} &= (0, \dots, 0.25, 0.5, 0.25, \dots, 0).\end{aligned}$$

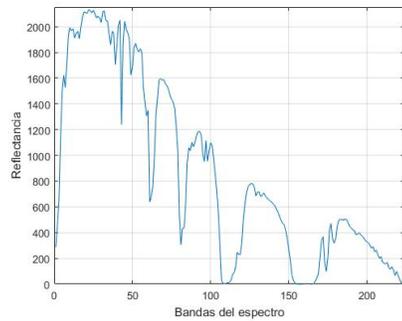
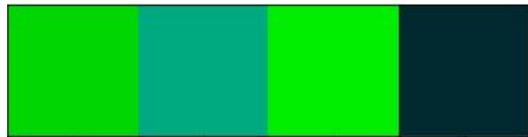
Como hemos expuesto anteriormente, presentamos en la figura 4.3 los resultados obtenidos. Podemos observar una clara diferencia respecto a las distintas funciones escogidas para la fusión de información, demostrando así la gran influencia que poseen los vectores de pesos escogidos para la visualización. Del mismo modo observamos que los valores presentan la misma tonalidad en la mayoría de los casos, es decir, el color predominante en el vector *RGB*, siendo tan solo diferentes en un mismo espectro en dos ocasiones.

(a) Espectro correspondiente a *PaviaU*

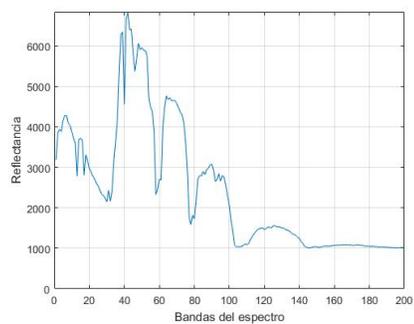
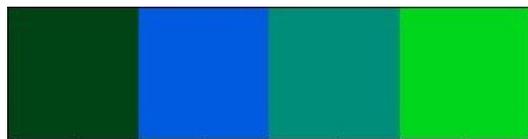
(b) Visualizaciones propuestas

(c) Espectro correspondiente a *Botswana*

(d) Visualizaciones propuestas

(e) Espectro correspondiente a *Salinas*

(f) Visualizaciones propuestas

(g) Espectro correspondiente a *Indian Pines*

(h) Visualizaciones propuestas

Figura 4.3: Representación de un espectro dado en las imágenes *PaviaU*, *Botswana*, *Salinas* e *Indian Pines*, aplicando distintos vectores de pesos. De izquierda a derecha, los operadores OWA aplicados en cada una de las escenas se corresponden con: 1. *Funciones clásicas*, 2. *Considerando tres vectores de pesos*, 3. *Considerando cinco vectores de pesos* y 4. *Considerando diez vectores de pesos*.

Capítulo 5

Experimentos

La estimación de la calidad de una imagen no es sencilla. En la mayor parte de los casos, las medidas de calidad se hacen en base a la adición de ruido (o distorsiones) en una imagen y su posterior limpieza. Sin embargo, para el presente trabajo no existe manera de cuantificar si una imagen está bien o mal visualizada [13]. Tampoco existe una imagen con la que comparar, más allá de un *ground truth* para cada imagen en la cual los colores son figurados. Esto es, en términos de cuantificación de sus objetivos, uno de los problemas que Jacobson y Gupta no pudieron acometer. Sí existen métricas de calidad de una imagen (como BRISQUE [14]) que pueden estimar la calidad de una imagen sin referencia (*no-referenceless image quality*). Sin embargo, estos métodos se basan en el estudio de los tonos y su distribución local y semi-local en la imagen. Esto nos lleva a descartarlos por dos motivos: primero, se basan en la presencia de unos priors que no se corresponden con los listados por Jacobson y Gupta. Segundo, asumen características de la distribución de información en el espectro visible, pero esto no tendría necesariamente que mantenerse en el espacio de imágenes hiperespectrales.

Así pues, los experimentos que se presentan a continuación deben ser restringidos a la inspección visual de los resultados de visualización de diferentes imágenes, sin posibilidad de generar una comparación cuantitativa.

Para ilustrar nuestra técnica de visualización propuesta, hemos usado distintas imágenes hiperespectrales del Grupo de Inteligencia Computacional de la Universidad del País Vasco (UPV/EHU).

Las figuras de la imagen 5.1 han sido tomadas mediante el sensor AVIRIS sobre el lugar de pruebas Indian Pines en el noroeste de Indiana. Se trata de una imagen de 145*145 píxeles y 224 bandas espectrales de reflectancia en el rango de longitudes de onda de 400 a 2500 nanómetros. Esta escena contiene dos terceras partes de cubiertas agrícolas y una tercera parte de bosque u otro tipo de vegetación natural perenne. Existen dos autopistas de doble carril, una línea de ferrocarril, así como un núcleo urbano de baja densidad, otras construcciones y carreteras menores. Dado que la escena ha sido tomada en junio algunos de los cultivos presentes, como maíz y soja, se encuentran en estados fenológicos poco desarrollados con menos de un 5% de cobertura. Tras aplicar una corrección a la escena y eliminar las bandas que abarcan la región de absorción del agua, obtenemos un conjunto de información de 200 bandas. En la visualización observamos una mayor nitidez en la imagen obtenida mediante las funciones propuestas, con un delimitado de parcelas mucho más definido en comparación con el obtenido mediante funciones clásicas, aproximándose más al proporcionado por el *ground truth*.

Las figuras 5.2 han sido obtenidas mediante el sensor ROSIS durante una campaña de vuelo sobre Pavia, en el norte de Italia. El número de bandas espectrales es de 102 para la zona centro de Pavia y el tamaño de la imagen es de 1096*1096 píxeles; el número de bandas para la zona de la Universidad

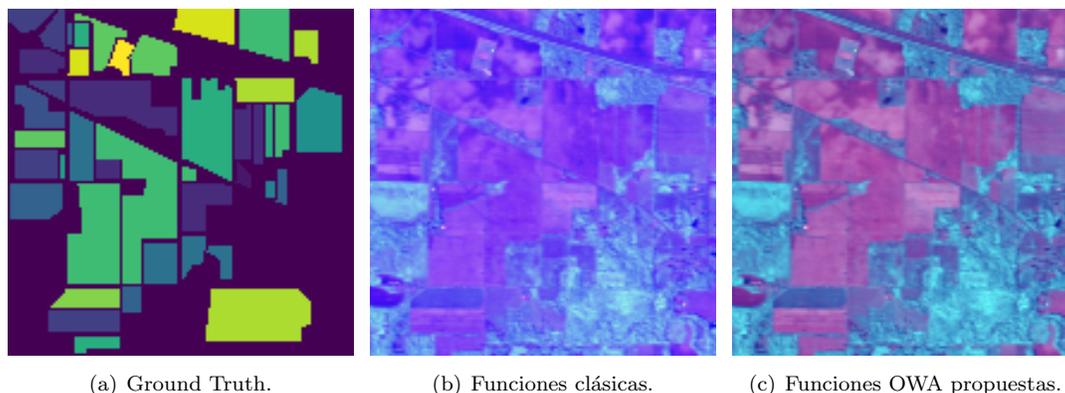


Figura 5.1: Resultados en la imagen *Indian Pines*. La figura contiene (a) El *ground truth*, tal y como se muestra en el dataset original, (b) la visualización usando las funciones clásicas (mínimo, máximo, media aritmética), y (c) la visualización usando los operadores OWA propuestos.

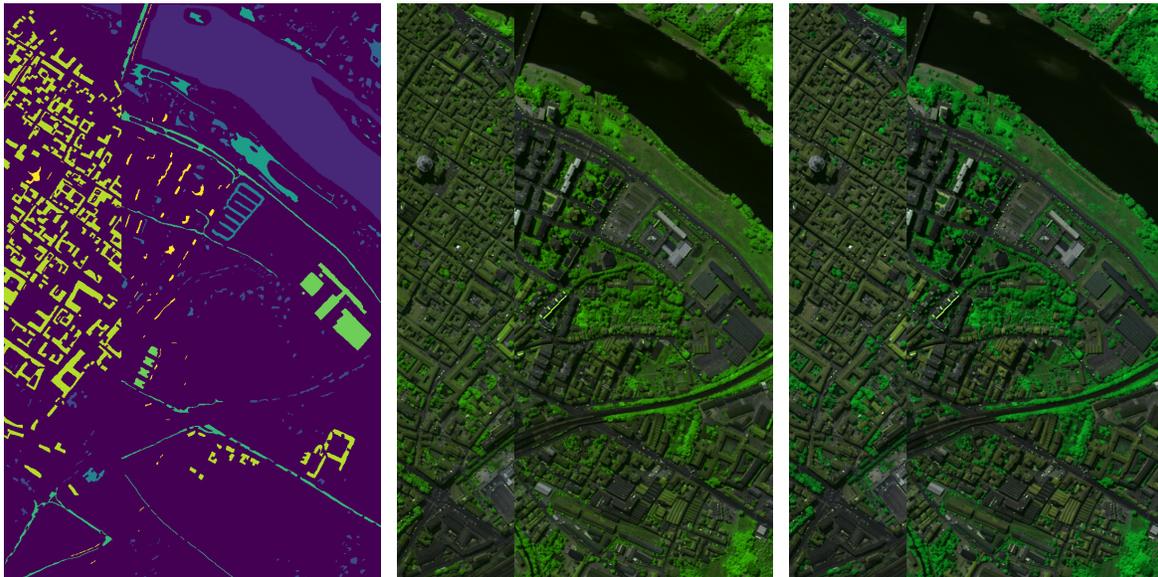
de Pavia es de 103 y presenta unas dimensiones de 610*610 píxeles. En ambas imágenes algunas de las muestras no contienen información y han sido descartadas antes del análisis. La resolución espacial es de 1.3 metros en ambos casos. En esta imagen observamos más que una clara mejora de la definición, un cambio en los colores representados en el vector *RGB* en la visualización mediante funciones propuestas frente a las funciones clásicas.

Las figuras 5.3 han sido medidas mediante el sensor AVIRIS sobre el valle de Salinas, en California. Están caracterizadas por una alta resolución espacial de 3.7 meter por píxel y tienen 224 bandas espectrales.. El área cubierta se compone de 512 líneas por 217 columnas. Como en el caso de las imágenes 5.1, se descartaron las 20 bandas de la absorción del agua. La escena incluye vegetales, suelos desnudos y campos de viñedos.

Las figuras 5.4 forman parte de una pequeña subescena de la zona del valle de Salinas presentada con anterioridad. Presenta las mismas características que la imagen completa, abarcando una zona de 86*83 píxeles. Las funciones OWA propuestas frente a las funciones clásicas, consiguen una clara mejora de la definición de los límites parcelarios, no alcanzada mediante el otro método, en ambos conjuntos de datos pertenecientes a la zona de Salinas. Podemos cotejar los resultados obtenidos con la información previa dada por el *Ground Truth*, observando un correcto ajuste de la información con las funciones propuestas.

El instrumento NASA AVIRIS adquirió los datos correspondientes a las figuras 5.5 en la zona del Centro Espacial de Kennedy (KSC), en Florida en 1996. El sensor AVIRIS adquiere datos en 224 bandas de 10 nanómetros de ancho en la región del espectro entre 400 y 2500 nanómetros. Los datos del KSC fueron adquiridos desde una altitud de aproximadamente 20 kilómetros y tiene una resolución espacial de 18 metros. Tras eliminar las bandas de absorción del agua y las bandas de baja SNR (*signal to noise ratio*), se usaron 176 bandas paara el análisis. La escena tiene un tamaño de 512*614 píxeles. Observamos que mediante el uso de las funciones OWA propuestas, podemos detectar y diferenciar una mayor superficie de cubiertas de suelo, pudiendo así discernir procedencias distintas.

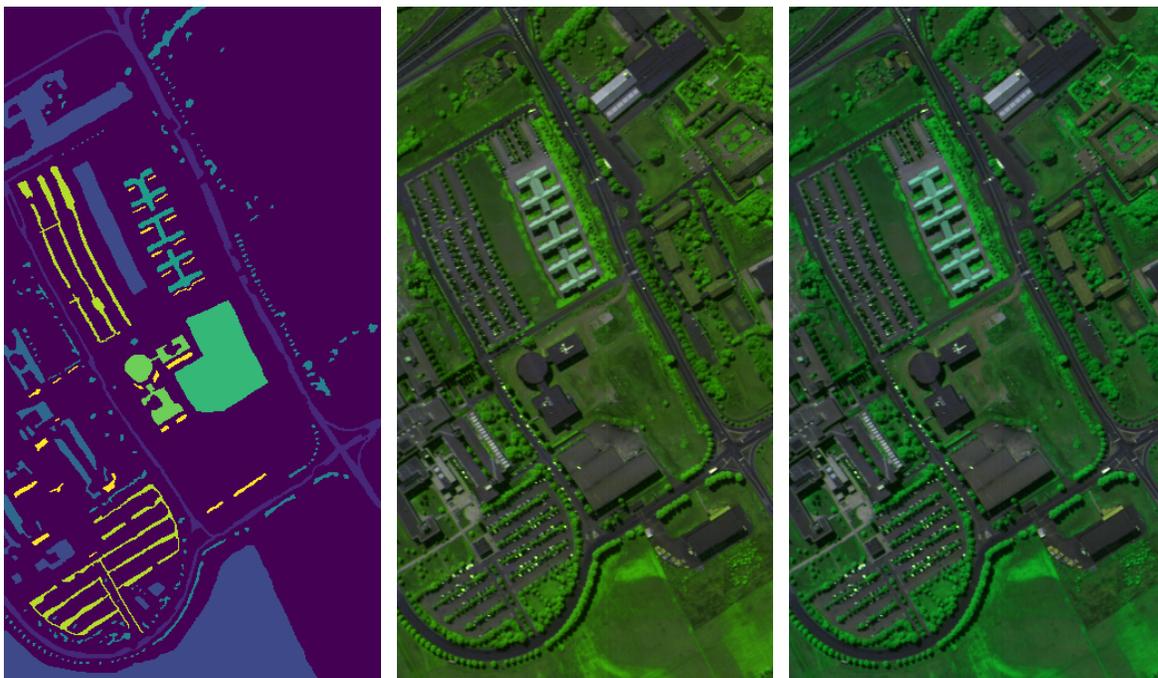
El satélite NASA EO-1 adquirió una serie de datos sobre el delta de Okavango en Botswana en 2001, representados en la figura 5.6. El sensor Hyperion en el EO-1 adquiere datos a una resolución



(a) Ground Truth.

(b) Funciones clásicas.

(c) Funciones OWA propuestas.



(d) Ground Truth.

(e) Funciones clásicas.

(f) Funciones OWA propuestas.

Figura 5.2: Resultados en la imagen *Pavia*. La figura contiene (a) El *ground truth*, tal y como se muestra en el dataset original, (b) la visualización usando las funciones clásicas (mínimo, máximo, media aritmética), y (c) la visualización usando los operadores OWA propuestos.

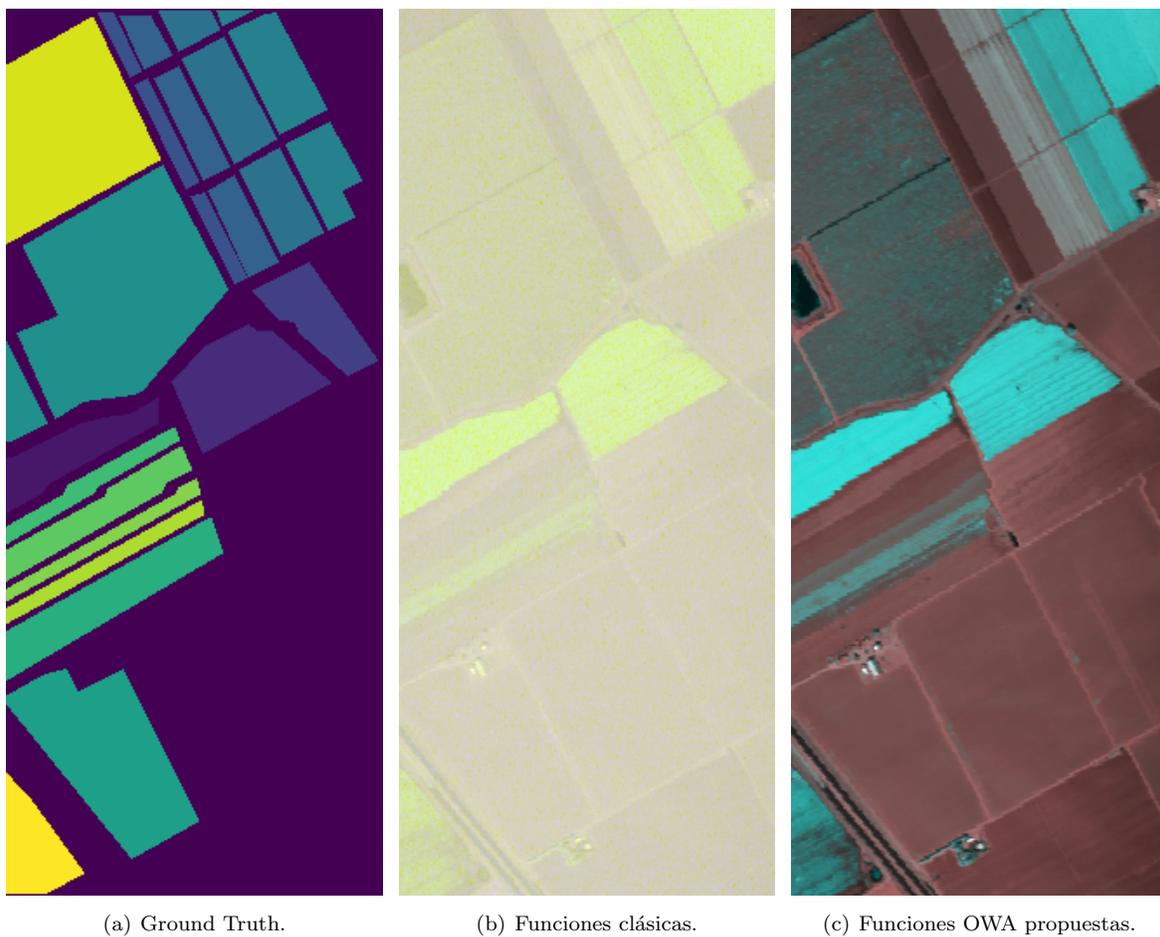


Figura 5.3: Resultados en la imagen *Salinas*. La figura contiene (a) El *ground truth*, tal y como se muestra en el dataset original, (b) la visualización usando las funciones clásicas (mínimo, máximo, media aritmética), y (c) la visualización usando los operadores OWA propuestos.

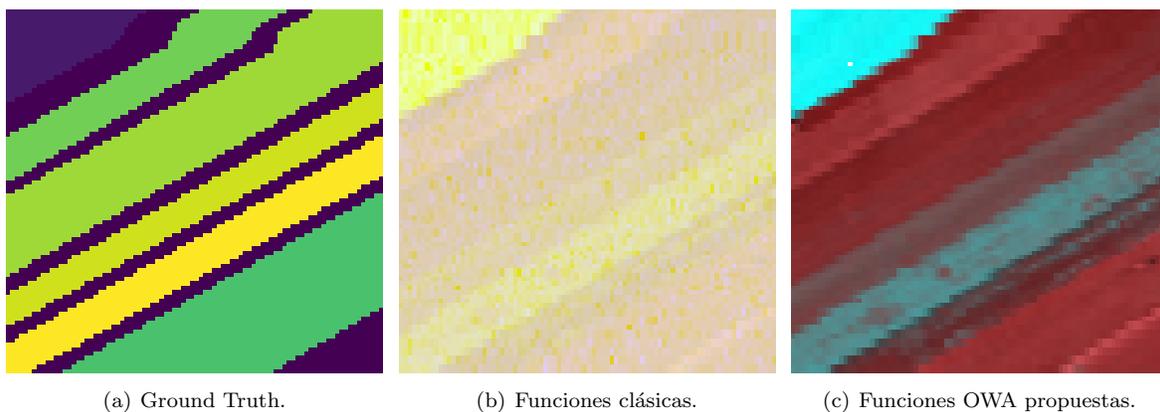


Figura 5.4: Resultados en la imagen *Salinas A*. La figura contiene (a) El *ground truth*, tal y como se muestra en el dataset original, (b) la visualización usando las funciones clásicas (mínimo, máximo, media aritmética), y (c) la visualización usando los operadores OWA propuestos.

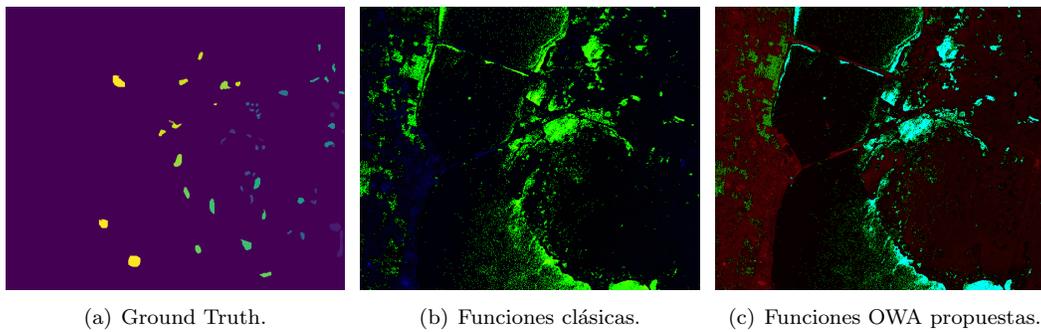


Figura 5.5: Resultados en la imagen *Kennedy Space Center*. La figura contiene (a) El *ground truth*, tal y como se muestra en el dataset original, (b) la visualización usando las funciones clásicas (mínimo, máximo, media aritmética), y (c) la visualización usando los operadores OWA propuestos.

de 30 metros por píxel sobre una línea de 7.7 kilómetros en 242 bandas cubriendo una porción del espectro de 400 a 2500 nanómetros en ventanas de 10 nanómetros. Bandas ruidosas y no calibradas que comprenden las características de absorción de agua fueron eliminadas, dejando 145 bandas para el estudio. La escena comprende pantanos estacionales, ciénagas ocasionales y bosques secos localizados en la porción distal del delta. En esta imagen podemos observar una mejora de la definición en el segundo de los casos, observando con una mayor claridad el delta, siguiendo así el curso del río.

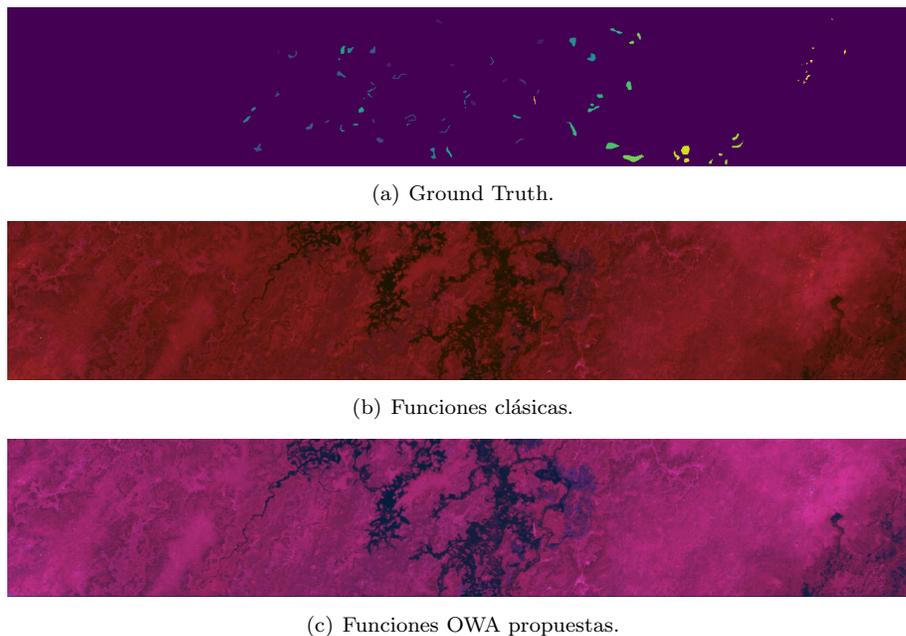


Figura 5.6: Resultados en la imagen *Botswana*. La figura contiene (a) El *ground truth*, tal y como se muestra en el dataset original, (b) la visualización usando las funciones clásicas (mínimo, máximo, media aritmética), y (c) la visualización usando los operadores OWA propuestos.

Capítulo 6

Conclusiones y líneas futuras

El presente trabajo se centra en la visualización de imágenes multi- e hiperespectrales (MHIs) en monitores tri-estímulo. Dentro de este contexto, el trabajo tiene un doble objetivo. Inicialmente, hemos realizado el estudio de todas las áreas con cierto impacto en el contexto, desde la propia definición del problema a las diferentes representaciones de color existentes en la literatura. Como segundo objetivo hemos planteado el desarrollo de un método de visualización de MHIs basado en operadores de agregación, específicamente, en operadores OWA. Esto se propone como alternativa a métodos clásicos de fusión de información. Para el diseño de la propuesta de visualización se utilizaron funciones OWA dentro de los operadores de agregación, evaluando los vectores de pesos w_i usados y realizando una aplicación práctica sobre un conjunto de ejemplos prácticos obtenidos de distintas escenas proporcionadas por la Universidad de País Vasco.

Una ventaja importante del procesamiento de imágenes MHIs mediante funciones OWA frente a otros procedimientos, es que no requiere de información previa concerniente a la escena. Este método se trata de una aplicación independiente del tipo de datos contenidos en la escena, es decir, nos brinda un sistema general, aplicable a cualquier caso práctico, proporcionándonos de este modo una primera visualización inicial de la información contenida en el conjunto de datos.

Este método, al no estar ligado a un contexto o conjunto de datos específico, presenta una mayor simplicidad, y con ello una mayor rapidez computacional, frente a métodos clásicos específicos, que muchas veces son diseñados *ad hoc* para cada tarea particular.

En los estudios numéricos y prácticos llevados a cabo se observó, como era esperable, un mejor comportamiento en la visualización frente a la obtenida mediante procedimientos clásicos. En líneas generales, podemos afirmar que el método propuesto es, como poco, tan competitivo como los anteriores. En ciertos casos, además, es claramente superior. Nótese en este sentido, que las alternativas se aplicaron a distintas escenas para poder observar las posibles diferencias existentes. Los resultados alcanzados mediante operadores OWA en las imágenes finales obtenidas, presentan en general una mejor definición.

Si se comparan los resultados entre un escenario con funciones clásicas y escenarios con un número variable de vectores de pesos w_i , observamos la alta influencia de los parámetros en los resultados. Diferentes parametrizaciones pueden tener llevar a resultados muy distintos. Achacamos esto al hecho de que un operador OWA puede usar sólo un conjunto pequeño de las longitudes de onda medidas, con lo cual la incorporación o eliminación de pesos extra puede dar lugar a diferencias significativas en los tonos RGB. De igual manera, la variación de unos pocos valores en los espectros puede también tener un gran impacto, por lo cual debería explorarse el uso de vectores de pesos con más valores (véase, menos ceros), que atenúen el impacto de ruido o distorsiones.

A modo de resumen, podemos concluir que este trabajo cumple sus dos objetivos fundamentales. Además los resultados prácticos aportados en este trabajo ponen de manifiesto las ventajas de la utilización de métodos de visualización a partir de operadores de agregación, concretamente de funciones OWA. Como complemento a esta investigación, en un futuro podrían abordarse modificaciones sobre los operadores OWA utilizados. Por ejemplo, buscando un segundo método de visualización adaptado a tareas específicas de visualización, que previsiblemente mejorarían su comportamiento cuando la información contenida en el conjunto de datos es conocida de antemano. En este sentido cabría plantearse la utilización de operadores de agregación con vectores de pesos w_i con valores no nulos en aquellas bandas que sabemos que contienen la mayor cantidad de información más relevante para el estudio concreto. Asimismo, podrían explorarse nuevos campos de estudio de la información, aplicando un análisis estadístico a los datos obtenidos en la visualización, de modo que podamos ampliar los conocimientos concernientes a la escena, mediante un examen de la información final obtenida.

Bibliografía

- [1] T. Lillesand, R. W. Kiefer, and J. Chipman, *Remote sensing and image interpretation*. John Wiley & Sons, 2014.
- [2] R. Kancheva, “Remote sensing terminology: past experience and recent needs,” in *Remote Sensing for Agriculture, Ecosystems, and Hydrology XV*, vol. 8887, p. 88871Z, International Society for Optics and Photonics, 2013.
- [3] J. R. Jensen and K. Lulla, *Introductory digital image processing: a remote sensing perspective*. Taylor & Francis, 1987.
- [4] N. P. Jacobson and M. R. Gupta, “Design goals and solutions for display of hyperspectral images,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 11, pp. 2684–2692, 2005.
- [5] N. P. Jacobson, M. R. Gupta, and J. B. Cole, “Linear fusion of image sets for display,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 10, pp. 3277–3288, 2007.
- [6] D. A. Socolinsky and L. B. Wolff, “A new visualization paradigm for multispectral imagery and data fusion,” in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, vol. 1, pp. 319–324, IEEE, 1999.
- [7] D. A. Socolinsky and L. B. Wolff, “Multispectral image visualization through first-order fusion,” *IEEE transactions on image processing*, vol. 11, no. 8, pp. 923–931, 2002.
- [8] P. M. Mather and M. Koch, *Computer processing of remotely-sensed images: an introduction*. John Wiley & Sons, 2004.
- [9] G. Beliakov, A. Pradera, and T. Calvo, *Aggregation functions: A guide for practitioners*, vol. 221. Springer, 2007.
- [10] G. Sharma and R. Bala, *Digital color imaging handbook*. CRC press, 2002.
- [11] A. Koschan and M. Abidi, *Digital color image processing*. John Wiley & Sons, 2008.
- [12] M. R. Gupta and N. P. Jacobson, “Wavelet principal component analysis and its application to hyperspectral images,” in *ICIP*, pp. 1585–1588, 2006.
- [13] A. Hore and D. Ziou, “Image quality metrics: Psnr vs. ssim,” in *Pattern recognition (icpr), 2010 20th international conference on*, pp. 2366–2369, 2010.
- [14] A. Mittal, A. K. Moorthy, and A. C. Bovik, “No-reference image quality assessment in the spatial domain,” *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695–4708, 2012.