



# UNIVERSIDAD DE GRANADA

---

## Descriptores estadísticos aplicados al análisis de escenas naturales bajo diferentes condiciones de iluminación.

---

Tesis doctoral presentada por Juan Ojeda García dentro del  
Programa de Doctorado en Física y Ciencias del Espacio.

Dirigida por D. Francisco Javier Romero Mora y D. Juan Luís  
Nieves Gómez

Curso 2022-2023

Editor: Universidad de Granada. Tesis Doctorales  
Autor: Juan Ojeda García  
ISBN: 978-84-1117-671-2  
URI: <https://hdl.handle.net/10481/79675>



*A mi padre.*





## AGRADECIMIENTOS

Cuando mis circunstancias laborales, el destino o como quieran llamarlo me volvieron a poner en contacto con el profesor Juan Luis Nieves Gómez, tras casi 20 años de nuestra licenciatura y con una trayectoria profesional ya definida, comprendí que se me ofrecía una nueva oportunidad de poder trabajar en el campo de la investigación del área de la física que más me había interesado siempre: la óptica. Este interés fue debido en gran medida a la persona responsable de impartirnos las asignaturas relacionadas con la misma, el profesor Javier Romero Mora.

Para mí ha sido un privilegio que ambos aceptaran codirigir este trabajo. Quisiera agradecer de forma particular al primero su apoyo y su guía desde el momento en que le propuse la descabellada idea de iniciar el periplo que hoy culmina, a pesar de las dificultades añadidas que conllevaba dirigir una tesis de alguien que llevaba fuera del mundo universitario y alejado de la investigación más de dos décadas y que tuvo que empezar cursando los estudios de postgrado actuales al no tener hechos los cursos equivalentes de doctorado de su época. Gracias por tu amistad y tu confianza en mí.

Además de por despertar ese amor por la óptica, quisiera agradecer al profesor Javier Romero Mora sus sugerencias y orientaciones que han permitido superar las dificultades que surgían en el desarrollo de esta investigación.

Querría extender los agradecimientos también al grupo de investigación Color Imaging Lab por sus valiosos materiales. Gracias a todos por haberme hecho participe de la comunidad investigadora.

Finalmente me gustaría agradecer a los que, fuera del departamento, me han dado su apoyo, a mi hermana Raquel por su ayuda en aspectos generales y administrativos, a Virgilio Cara Valero por hacer más legible esta investigación y, en particular, al dedicatario de este trabajo, mi padre, sin cuya inquebrantable confianza en mí no lo hubiera podido llevar a buen puerto.



---

## ÍNDICE GENERAL

---

|  |    |
|--|----|
| Capítulo 1.- INTRODUCCIÓN Y OBJETIVOS .....  | 1  |
| Capítulo 2.- PERCEPCIÓN DEL COLOR: FACTORES FÍSICOS Y BIOLÓGICOS... 7                    |    |
| 2.1 Factores físicos de la percepción del color .....                                    | 8  |
| 2.2 Factores biológicos de la percepción del color .....                                 | 9  |
| 2.2.1 Procesamiento en la retina.....  | 10 |
| 2.2.2 Transmisión al córtex visual.....  | 10 |
| 2.2.3 Procesamiento cortical.....  | 12 |
| Capítulo 3.- FUNDAMENTOS DE LA ESTADÍSTICA DE IMÁGENES<br>NATURALES .....                | 17 |
| 3.1 La estadística de imágenes naturales y los iluminantes .....                         | 22 |
| Capítulo 4.- RESULTADOS SOBRE ESTADÍSTICA DE PRIMER ORDEN EN<br>IMÁGENES NATURALES.....  | 27 |
| 4.1 Conjunto de datos empleado.....  | 27 |
| 4.2 Estadística de primer orden de las imágenes naturales.....                           | 29 |
| Capítulo 5.- RESULTADOS SOBRE ESTADÍSTICA DE SEGUNDO ORDEN EN<br>IMÁGENES NATURALES..... | 39 |
| 5.1 Estadística entre píxeles adyacentes. ....   | 40 |
| 5.2 Estadística entre píxeles no adyacentes. ....  | 46 |
| 5.3 Análisis de componentes principales (PCA). ....                                      | 55 |

|  |     |
|--|-----|
| Capítulo 6.- RESULTADOS SOBRE ESTADÍSTICA DE ORDEN SUPERIOR EN IMÁGENES NATURALES .....                        | 69  |
| 6.1 Análisis de componentes independientes (ICA).....  | 74  |
| 6.2 Índice de agrupamiento (Patchy index) .....  | 81  |
| Capítulo 7.- INFLUENCIA DEL ILUMINANTE EN LA ESTADÍSTICA DE ESCENAS NATURALES.....                             | 87  |
| 7.1 Estadística de primer orden.....   | 92  |
| 7.2 Estadística de segundo orden.....  | 100 |
| 7.2.1 Estadística entre píxeles adyacentes.....  | 100 |
| 7.2.2 Estadística entre píxeles no adyacentes.....   | 104 |
| 7.3 Estadística de orden superior. ....  | 110 |
| 7.3.1 Análisis de componentes independientes.....  | 113 |
| 7.3.2 Índice de agrupamiento. ....   | 114 |
| Capítulo 8.- DESCRIPCIÓN ESTADÍSTICA DE LOS COLORES RELEVANTES EN ESCENAS NATURALES.....                       | 117 |
| 8.1 Regiones salientes y diversidad cromática .....  | 118 |
| 8.2 Cuantificación del número de colores relevantes en escenas pictóricas.....                                 | 120 |
| 8.3 Número de colores relevantes en imágenes naturales. ....   | 123 |
| 8.3.1 Codificación eficiente de los colores relevantes que describen una escena natural.....                   | 129 |
| Capítulo 9.- CONCLUSIONES.....   | 135 |
| BIBLIOGRAFÍA.....  | 143 |
| ANEXOS.....  | 159 |
| ANEXO I: PROCEDIMIENTO PARA COMPARAR LAS MEDIAS DE MÚLTIPLES DEISTRIBUCIONES DE UN DESCRIPTOR ESTADÍSTICO..... | 161 |
| ANEXO II: LISTA DE FIGURAS .....   | 163 |
| ANEXO III: LISTA DE TABLAS .....   | 171 |

---

## Capítulo 1

# INTRODUCCIÓN Y OBJETIVOS

---

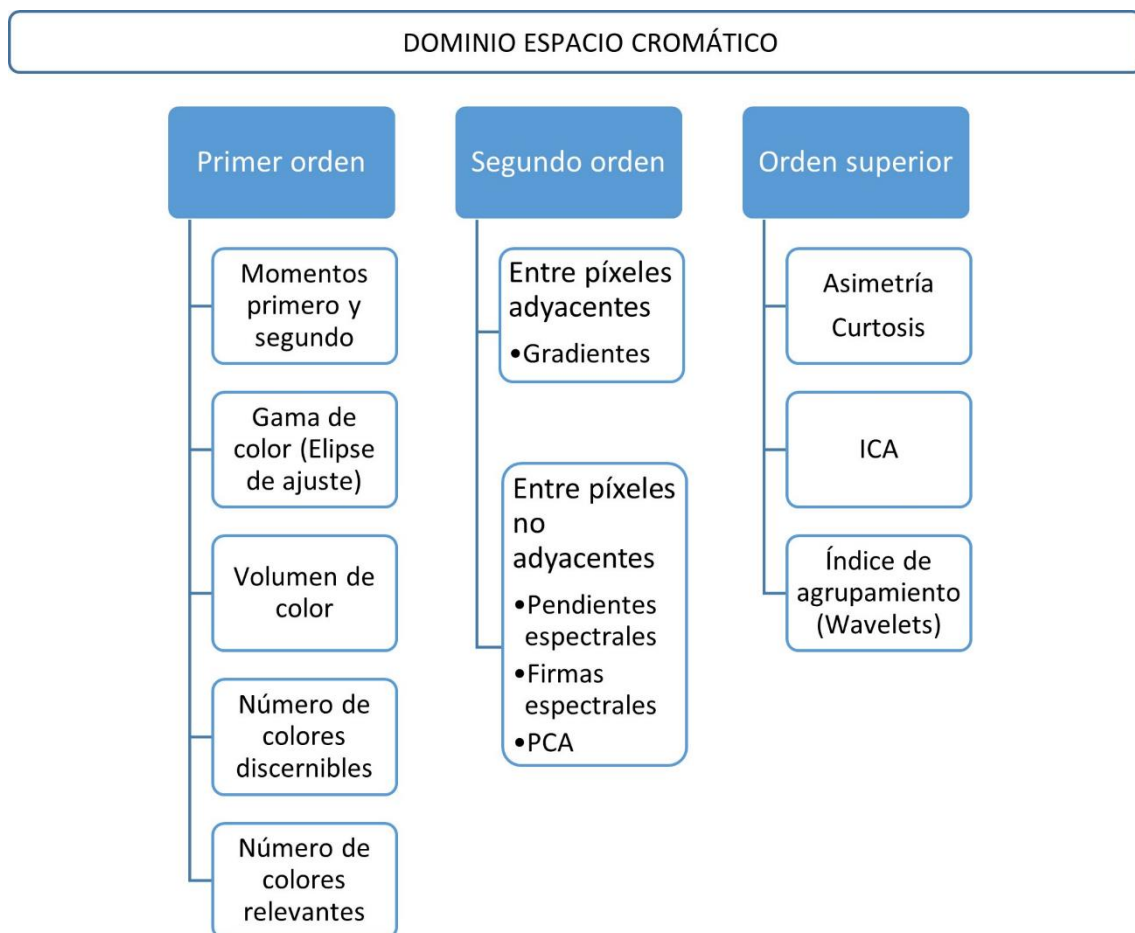
El proceso evolutivo de los sistemas sensoriales, y en particular del sistema visual humano debe entenderse como una adaptación para procesar eficientemente los estímulos recibidos del entorno. Dicha evolución viene determinada por tres factores: las limitaciones biológicas, las tareas que se desarrollan y el entorno en el que el organismo las realiza; de ahí que se considere que las imágenes de escenas naturales son los estímulos para los que el sistema visual se ha adaptado mejor.

Como ha indicado Geisler [Gei08] la caracterización más adecuada de los estímulos visuales debe ser en términos estadísticos, para tener en cuenta la relación probabilística entre estos y las tareas que se van a realizar, las fluctuaciones cuánticas inherentes a la absorción de la luz y a la naturaleza ruidosa de los procesos químicos neuronales.

Este enfoque del proceso de selección natural establece una estrecha conexión entre el diseño del sistema perceptivo de un organismo y las propiedades del entorno físico en el que este desarrolla su actividad. No obstante, no pudo establecerse una relación cuantitativa entre ambos hasta que F. Attneave [Att54] basándose en la teoría de la información de C.E. Shannon y W. Weaver [Sha49] estableció que la finalidad del sistema visual humano era la de producir una representación eficiente de la señal entrante, para lo cual proponía que en las primeras etapas del sistema visual se codificaban las señales de entrada con la finalidad de reducir la estructura redundante de las imágenes naturales. Este planteamiento ha permitido avanzar en la propuesta de marcos teóricos para comprender la percepción visual, en el diseño de paradigmas experimentales que emulen los estímulos naturales y en el desarrollo de modelos computacionales basados en la estadística del entorno [Gei08,Sim01].

El estudio de esta conexión puede hacerse usando dos enfoques diferentes [Sim01]: analizando estadísticamente las respuestas neuronales bajo condiciones de estimulación natural [Lau81, Bad98, Vin00] o modelando el procesamiento sensorial temprano a partir del estudio de algún descriptor estadístico de las entradas visuales y de las transformaciones que lo optimizan [Ati92a, Ols96, Bel97, vHa98, Sim99]. Será esta segunda metodología la que utilizaremos en este trabajo.

Dado que la redundancia de las imágenes naturales puede presentar correlaciones que pueden ser de primer orden (valores de intensidad parecidos localmente), de segundo orden (bordes orientados) o de orden superior (texturas homogéneas) [Dau89], los descriptores estadísticos pueden clasificarse también en dichos ordenes atendiendo al número de píxeles que se relacionan (píxeles individuales, pares de píxeles y grupos de más de dos píxeles).



**Figura 1.1:** Relación de los descriptores estadísticos en el dominio espacio-cromático analizados en el presente trabajo.

Thomson et al. [Tho00] y Reinhardt et al. [Rei01], consideran que el estadístico de orden enésimo más simple que se puede tener en cuenta es el momento (potencia) de dicho orden de los valores de la intensidad de la señal. Por esa razón, dichos autores consideran que la asimetría y la curtosis son los estadísticos más simples de orden tercero y cuarto (orden superior) que son sensibles a los cambios de fase en la imagen. En este trabajo, como puede verse en la figura 1.1, hemos adoptado ese criterio a pesar de que muchos libros de estadística los catalogan como momentos de primer orden.

El primer objetivo de este trabajo (capítulos 4, 5 y 6) ha sido la sistematización del cálculo de los diferentes estadísticos de primer, segundo y orden superior (que aparecen en la figura 1.1) asociados a la estructura espacio - cromática de un mismo conjunto de imágenes naturales en color y sus implicaciones en los modelos del sistema visual temprano.

En su adaptación al entorno, el sistema visual humano proporciona una percepción estable del color, independientemente de los cambios, tanto locales como globales, de la luz día que ocurren en las imágenes naturales [Fos11]. Este fenómeno es conocido como constancia de color. Dado que la señal de color que caracteriza la luz que entra en el ojo, depende tanto de la reflectancia espectral de los objetos como de la distribución de potencia espectral de la luz que los ilumina, es posible plantearse cómo la visión del color, a partir de las regularidades que presenta la estructura espacio-cromática de las imágenes naturales, puede mostrar cierta dependencia con las diferentes condiciones de iluminación bajo las que se observe una escena. El segundo objetivo de esta investigación ha sido el estudio de la influencia de la distribución espectral de potencia (SPD) de la luz día, que caracterizaremos por su temperatura correlacionada de color (CCT), en los diferentes estadísticos (capítulo 7).

La diversidad cromática de una imagen suele venir cuantificada por el número de colores discernibles (NDC) definido este como el número de cubos unidad distinguibles en un espacio de color que contienen al menos un píxel [Poi98, Mar07]. Los valores de este descriptor estadístico pueden variar desde los pocos miles hasta los dos millones (para valores teóricos), valores que un observador no podría distinguir a pesar de la capacidad del ojo humano de percibir diferencias de color con gran resolución de entornos complejos [Lin08, Mas13]. En la práctica, cuando vemos imágenes naturales de entornos humanos o incluso obras pictóricas, nos centramos en una pequeña fracción del gran número potencial de colores discernibles, aunque todavía no se ha establecido una relación entre la totalidad de ellos y los colores que atraen realmente la atención del observador.

Se consideran salientes aquellos puntos de una escena en los que un observador centrará su atención por diferir de sus vecinos en ciertas características como bordes, contrastes, luminancia o color. Son varios los modelos abajo-arriba o bottom-up, conocidos como mapas de saliencia que se han desarrollado a partir del modelo posible



biológicamente propuesto por Koch y Ullman [Koc85] para establecer las regiones salientes de una imagen.

Los profesores Nieves y Romero estudiaron heurísticamente en [Nie18] los efectos de la saliencia visual utilizando distintos modelos en la estimación del número de colores discernibles en escenas naturales, aplicable también a obras pictóricas. En su análisis determinaron que la saliencia visual podía constituir un filtro que limita el número de colores en los que un observador puede fijar su atención, así como la utilización de este número limitado de colores para la creación de imágenes segmentadas automáticamente en función de dicha diversidad. Estos resultados sugieren que, a pesar de la complejidad relativa del contenido espacio-cromático de las escenas, estos algoritmos son capaces de realizar una buena segmentación de las mismas.

No obstante, la estimación de los colores salientes que aparecen en una escena sería el resultado de un proceso conducente por una tarea arriba-abajo o top-down. En una propuesta en la que no es necesario introducir categorías de color predefinidas ni determinar qué áreas son visualmente salientes en la escena, Nieves et al. [Nie20] establecen el concepto de colores relevantes, definido como los colores discernibles categóricos que describen la diversidad cromática de una obra pictórica basándose en el contenido colorimétrico de cada imagen. Un algoritmo computacional, que simula en un proceso top - down cómo los observadores mirarían a aquellas zonas que son más relevantes tanto por ser colorimétricamente discernibles como por describir la diversidad cromática de la escena, permite estimar una paleta de colores fiable para cada una de ellas.

Los resultados obtenidos se han visto recientemente respaldados por un experimento psicofísico en el que distintos observadores determinaban los colores relevantes a partir de sus propias impresiones subjetivas [Nie21]. De esta forma, se obtuvieron gamas de colores similares a las obtenidas heurísticamente y que seguían una codificación eficiente al coincidir en gran medida con las direcciones de color preferidas, obtenidas al aplicarle un análisis de componentes independientes (ICA) a las imágenes observadas.

El último objetivo de este trabajo (capítulo 8) trata de analizar con más profundidad la relación entre la diversidad cromática de una escena y las potenciales regiones de fijación visual de un observador que la estuviera mirando. Para ello se analizará desde un punto de vista heurístico el descriptor número de colores relevantes (NRC) en distintos conjuntos de imágenes naturales, su relación con la codificación eficiente de la información por parte del sistema visual humano, sus posibles diferencias con los resultados contrastados experimentalmente y obtenidos para las imágenes de obras pictóricas ya analizadas.

De esta manera, un pequeño número de colores relevantes describe heurísticamente la diversidad cromática de una escena a partir de su contenido colorimétrico sin necesidad de introducir categorías de color predefinidas ni determinar

qué áreas de la escena son salientes. Al segmentar las imágenes usando los colores relevantes, estas mantienen el contenido cromático más reseñable de las escenas. Por estas razones, podría utilizarse dicho algoritmo para el reconocimiento de patrones cromáticos o para reducir los procesos de inferencia de aprendizaje profundo necesarios en los procesos de visión computacional integrados en la inteligencia artificial.



---

**Capítulo 2**

**PERCEPCIÓN DEL COLOR: FACTORES FÍSICOS Y BIOLÓGICOS**

---

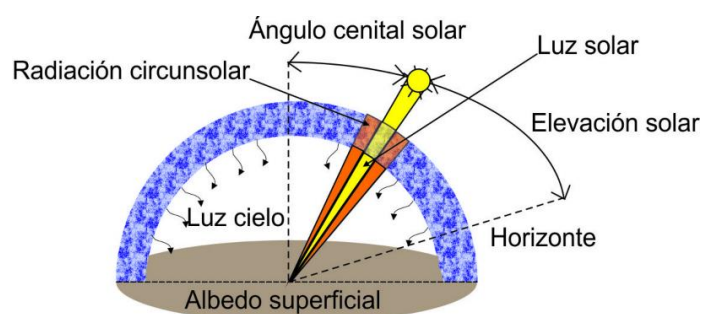
Podemos considerar la visión humana como el proceso perceptivo de formación de imágenes bidimensionales en la retina, su transformación primero en señales químicas y eléctricas después, cuyo tratamiento al llegar al córtex visual permite la percepción de colores, formas, texturas y movimientos de los objetos y las escenas de nuestro entorno a partir de la luz que estos emiten o reflejan.

El color, junto con la forma, la textura y el tamaño, es una de las características más importantes para el reconocimiento de objetos. La CIE (Commission Internationale de l'Eclairage), en la definición 17-198 de su vocabulario internacional de iluminación [CIE11] describe el color percibido como aquella “característica de la percepción visual que puede ser descrita por una serie de atributos como el tono, luminosidad (claridad) y colorido (o saturación o croma)”; asimismo, en la nota 2 a dicha definición añade “el color percibido depende de la distribución espectral del estímulo de color, del tamaño, forma, estructura y entorno de la zona del estímulo, del estado de adaptación del sistema visual del observador y de la experiencia que este último posee de las condiciones de observación en que se encuentra o en condiciones semejantes”. En esta segunda apostilla a su definición la CIE describe los factores principales de los que depende la visión del color:

- Factores físicos: la distribución espectral de la luz que entra en el órgano sensitivo y las características visuales de la zona de estímulo.
- Factores biológicos y evolutivos (o de aprendizaje): la evolución y el aprendizaje a partir de experiencias previas del sistema visual, así como su diseño y funcionamiento.

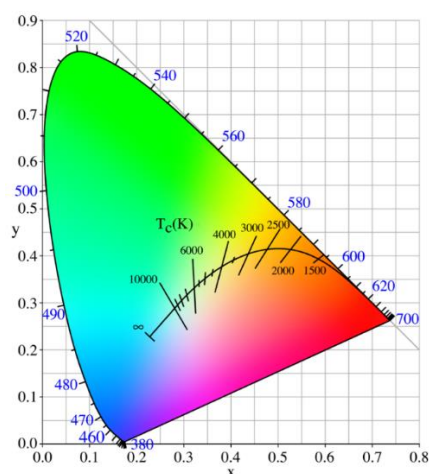
## 2.1 Factores físicos de la percepción del color

El color de los objetos aislados depende tanto de la reflectancia espectral en cada punto de los mismos como de la distribución de potencia espectral (SPD) de la iluminación bajo la que se observan, estando ambos factores confundidos en lo que se denomina *señal de color*, definida como el producto de los mismos.



**Figura 2.1:** Esquema de los elementos y componentes de la luz día. [Pey16]

Como muestra la figura 2.1, en imágenes naturales o de entornos humanos con iluminación no artificial, la distribución espectral de la radiación que alcanza la tierra y que es responsable de su iluminación, se conoce como radiación global de luz día hemisférica (a partir de ahora luz día) [Her01], y es la suma de los espectros correspondientes a la radiación directa del sol a nivel del suelo (luz solar) y a la radiación difusa dispersada a través de la atmósfera en el cielo y por el albedo de las superficies reflectantes de la región circundante llamada luz de cielo [Pey16].



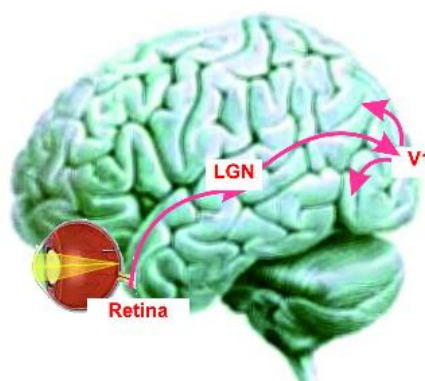
**Figura 2.2:** Diagrama de cromaticidad CIE xy 1931, lugar geométrico de las cromaticidades del cuerpo negro y varias rectas de temperatura de color correlacionada constante [Pla12].

La distribución de potencia espectral de la luz día se caracteriza por su temperatura de color correlacionada (CCT) en el rango de longitud de onda 400-700 nm y se define como la “temperatura de un radiador Planckiano (cuerpo negro) con una cromaticidad próxima a la cromaticidad asociada con la distribución espectral sobre un diagrama donde (CIE 1931 es el patrón en el que está basado el observador)  $u'$ ,  $2/3 v'$  son las coordenadas del lugar de los estímulos Planckianos y del lugar donde el estímulo de prueba está representado” [CIE11]. La figura 2.2 muestra el diagrama de cromaticidad CIE xy 1931, el lugar geométrico de las cromaticidades del cuerpo negro y varias rectas de temperaturas de color correlacionadas constante.

Por otro lado, partiendo de la idea de que los sistemas visuales biológicos han evolucionado para procesar de manera eficiente el tipo de estímulos que ellos reciben [Bar59], es lógico pensar que las imágenes naturales o las imágenes de escenas naturales sean las que presenten una estructura estadística similar a aquellas para las que el sistema visual está adaptado. Aunque en el presente el ser humano vive y desarrolla su actividad en otro tipo de entornos (urbano, industrial,...), que hemos llamado entornos humanos, en una escala de tiempo evolutiva, este periodo temporal no ha sido suficiente para generar adaptaciones significativas en nuestro sistema visual.

## 2.2 Factores biológicos de la percepción del color

La percepción del color por el sistema visual humano consiste en una serie de pasos jerárquicos que podemos agrupar en tres etapas: procesamiento retiniano, transmisión de las señales de color al córtex visual a través del núcleo geniculado lateral (que a partir de ahora denotaremos por LGN) y procesamiento cortical. En la figura 2.3 se puede ver un esquema del recorrido que sigue la información visual desde que la luz llega a la retina hasta que alcanza el córtex visual en el área V1.

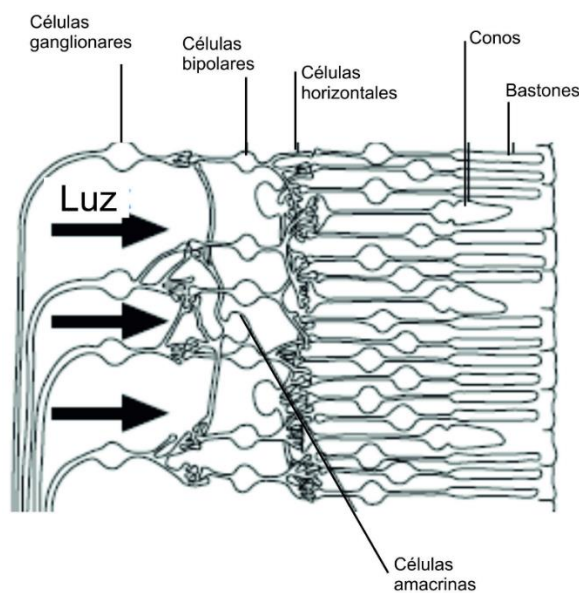


*Figura 2.3: Esquema del sistema visual*

### 2.2.1 Procesamiento en la retina

El procesamiento de la información visual en general y del color en particular comienza con la proyección en el mosaico de fotorreceptores de la imagen formada en la retina. Hay dos tipos de fotorreceptores: bastones y conos, responsables de la visión bajo condiciones de luminosidad baja y alta (visión escotópica y fotópica) respectivamente; además, estos últimos son los responsables de la visión del color. Basándose en sus sensibilidades espectrales los fotorreceptores se clasifican en conos que presentan un máximo de la misma en una longitud de onda corta (conos S), en una media (conos M) y en una larga (L) del espectro visible respondiendo preferiblemente a estímulos de color que denotaríamos como azul, verde y rojo respectivamente.

Un único cono es en sí mismo ciego al color porque su activación depende tanto de la longitud de onda como de la intensidad del estímulo. Por tanto, el requerimiento básico computacional de un sistema de visión del color es la comparación de las señales de distintas clases de fotorreceptores. [Con10]. Esta comparación se consigue mediante la reorganización de la información de los fotorreceptores a través de los tres tipos de neuronas retinianas llamadas horizontales, bipolares y amacrinas, hasta alcanzar la capa de las células ganglionares. (Figura 2.4).



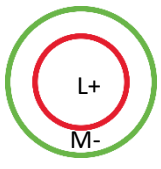
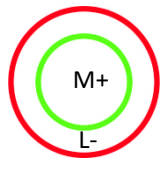
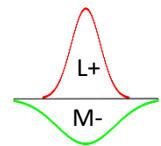
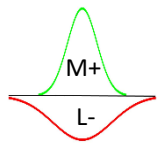
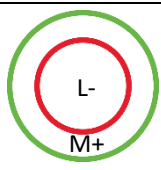
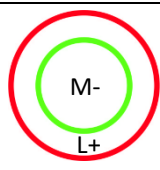
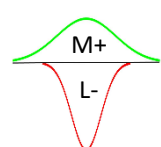
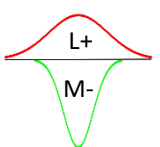
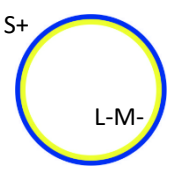
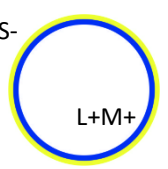
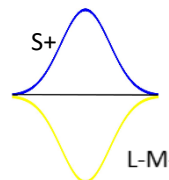
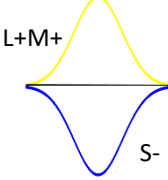
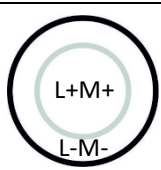
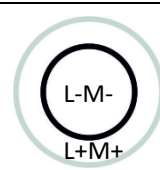
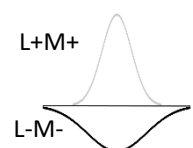
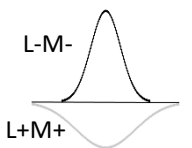
*Figura 2.4: Estructura celular de la retina humana [Fri01]*

### 2.2.2 Transmisión al córtex visual

Las células ganglionares son las encargadas de enviar la información visual como luminancia e información cromática de manera color oponente [Gao13] fuera del ojo a través de sus largos axones (que forman el nervio óptico) hasta el LGN del tálamo.

Aunque, de acuerdo con sus propiedades fisiológicas (tamaño y región del LGN a la que llegan), estas células se suelen clasificar en tres clases principales llamadas células magnocelulares (M), parvocelulares (P) y koniocelulares (K), nosotros utilizaremos la clasificación que Wiesel y Hubel [Wie66] hicieron en células de Tipo I, Tipo II y Tipo III atendiendo al tipo de sinapsis que la célula debía de tener con los conos.

Tradicionalmente, se llama campo receptivo (RF) de una neurona a la región del espacio en la que la presencia de un estímulo induce un cambio en el patrón de respuesta en la célula. Los campos receptivos de las células ganglionares, de las células del LGN, y buena parte de los de las células del córtex estriado tienen forma y simetría circular; es decir, que cualquier estímulo a una misma distancia,  $r$ , del centro del círculo, produce la misma respuesta en la célula [Cho11]. La función campo receptivo que describe la respuesta de la célula en cada punto del espacio puede modelizarse mediante una combinación lineal de gaussianas [Rod65].

|                                       |   |   |  |   |
|---------------------------------------|---|---|--|---|
| Tipo I<br>Células<br>parvocelulares   |   |   |   |   |
|                                       |  |  |  |  |
| Tipo II<br>Células<br>koniocelulares  |  |  |  |  |
| Tipo III<br>Células<br>magnocelulares |  |  |  |  |

**Figura 2.5:** Campos receptivos y sus funciones correspondientes de las células ganglionares y del LGN

Las células de Tipo I (células parvocelulares) son células oponentes en las que el centro y la periferia del campo receptivo están estimulados por conos L y M distintos (L en el centro y M en la periferia o viceversa), y además, las sinapsis con la célula tienen signos contrarios. Son células que presentan un antagonismo de tipo rojo-verde. En la



primera fila de la figura 2.5 podemos ver un esquema de sus campos receptivos y las funciones que lo describen.

Las células de Tipo II (células koniocelulares) son células oponentes simples co-extensas en las que no hay distinción entre centro y periferia utilizadas para inducir un antagonismo de tipo azul-amarillo cuya estructura original se muestra en la segunda fila de la figura 2.5.

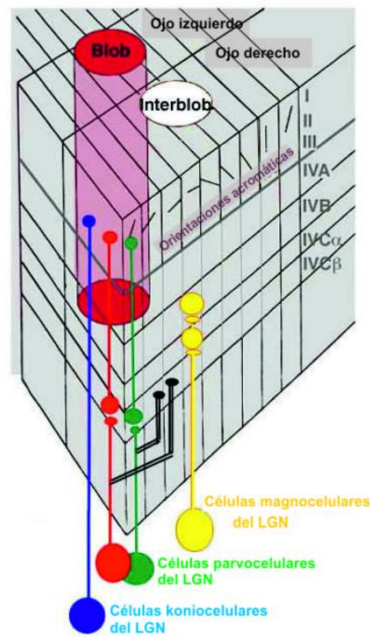
Por último, en las células de Tipo III (células magnocelulares) el centro y la periferia reciben señales de entrada de los mismos conos (L y M), pero las sinapsis del centro y de la periferia tienen signos contrarios. Muestran una oposición que codificaría la información en luminancia (ver figura 2.5 tercera fila)

Ninguno de los tres tipos de células ganglionares presenta antagonismo espacial a las variaciones de color, aunque las de Tipo I y III sí lo presentan a variaciones de luminancia (entendiendo el antagonismo espacial de un campo receptivo de una célula a un determinado estímulo como la capacidad de respuesta de signos contrarios ante dicho estímulo dependiendo de si se introduce en el centro o en la periferia del campo receptivo) [Cho11].

La información de los tres tipos de células ganglionares viaja por el nervio óptico hasta el LGN, en donde la información es redistribuida atendiendo al semicampo visual de la que procede pero que no altera la forma en que son transmitidas las señales de la retina y llegan hasta a la primera región del córtex visual.

### **2.2.3 Procesamiento cortical.**

Una vez que la información cromática y de luminancia procedente del LGN llega al córtex estriado en la región occipital, ésta se organiza para dar lugar a los mecanismos acromáticos y cromáticos que permiten al observador la percepción de las características visuales de su entorno. El área cortical que recibe la mayor cantidad de la información proyectada del LGN es el córtex visual primario, también llamado V1 o córtex estriado. Presenta una arquitectura en capas y columnas [Hub77, Bla86, Gri86, Ohk06] que podemos ver en la figura 2.6 [Gou09]. En este modelo la información de cada ojo entra por regiones distintas espacialmente alternas transmitiéndose de las capas más internas a las más externas (de la VI a la I). Además, cada camino visual entra por una capa distinta (Ver figura 2.6); las células magnocelulares (Tipo III) entran por la subcapa IVC $\alpha$  y las células parvocelulares (Tipo I) lo hacen por la capa IVC $\beta$ , donde establecen sinapsis con células corticales con su misma estructura que llevan las señales a las capas más externas. Las células koniocelulares (Tipo II) acceden directamente a los grupúsculos de células con características diferentes, que analizaremos más adelante distribuidos de manera irregular en las capas más superficiales del córtex visual primario y que se denominan “blobs” [Con10, Cho11].



**Figura 2.6:** Esquema de la estructura del córtex visual estriado. [Gou09]

Verticalmente, el córtex visual se organiza por columnas que, en los espacios interblobs, están formadas por células sintonizadas a orientaciones de luminancia similares que varían gradualmente a través de la capa. Las células que encontramos en estos espacios interblobs de las capas superficiales del córtex visual se caracterizan por sus campos receptivos alargados y orientados que, por no estar relacionadas con la visión del color, no describiremos aquí.

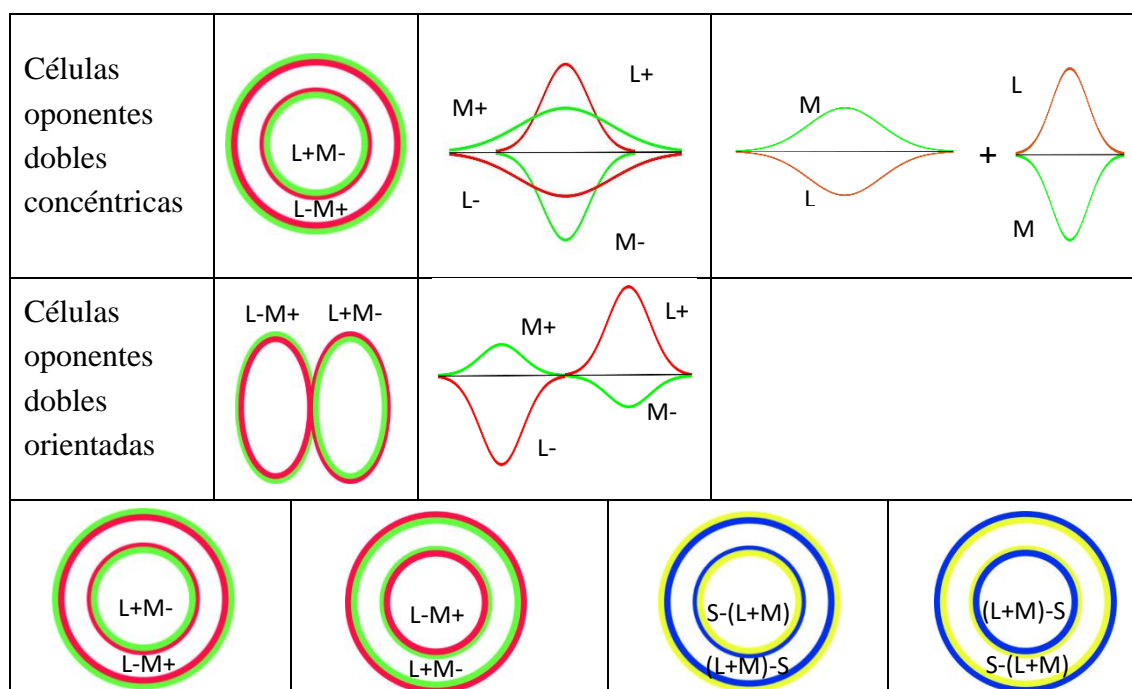
Aunque no hay consenso acerca de la organización horizontal de las señales cromáticas [Con10], es en los blobs donde tienden a coincidir las células cromáticas [Liv84] y en donde, además de encontrar células que tienen las mismas propiedades que las de Tipo II y Tipo III pero con origen parvocelular (células Tipo II-like y Tipo III-like), aparecen células oponentes dobles.

Las células oponentes dobles reciben este nombre ya que son oponentes tanto espacial como espectralmente. Según la estructura de su campo receptivo, las células oponentes dobles pueden ser de RFs concéntricos (centro-periferia) u orientados (ver las dos primeras filas de la figura 2.7 para células oponentes dobles rojo-verde). En ambos casos los campos receptivos pueden construirse usando las salidas de dos células oponentes simples de tipo II con escalas y radios diferentes (en el primer caso) o donde la contribución de cada tipo de conos no es igual (segundo caso). [Sha11]

En el caso de las células doble oponentes concéntricas, el centro del campo receptivo recibe señales de al menos dos clases de conos con las sinapsis de signos contrarios, mientras que la periferia las recibe de las mismas clases de conos pero con los signos de las sinapsis cambiados, de manera que las sensibilidades espectrales del centro

y la periferia son iguales pero opuestas (la tercera fila de la figura 2.7 y muestra los cuatro tipos de células de este tipo) [Cho11].

La principal diferencia con las células anteriores es que los aferentes de estas células dobles son las proyecciones de las células oponentes simples y sus proyecciones. De este modo, el centro del campo receptivo de una célula oponente doble con centro L+M- está formada por la superposición espacial de los centros de las células oponentes simples L+M- aferentes de las mismas [Gou91] y, del mismo modo, la periferia está formada por los centros de aferentes de centros de células oponentes simples M+L-. Para el caso de las células oponentes dobles amarillo - azul los aferentes son directamente las células de Tipo II con ese tipo de antagonismo (y su opuesto).



**Figura 2.7:** Campos receptivos y sus funciones correspondientes para una célula oponente doble rojo-verde concéntrica (primera fila) y orientada (segunda fila). La tercera fila muestra las posibles topologías de las células oponentes dobles concéntricas cromáticas[Cho11].

La percepción del color probablemente dependa tanto de las neuronas oponentes simples como oponentes dobles, y del procesado ulterior de sus señales [Vla07]. Las células oponentes simples, que no son selectivas espacialmente para el color, parecen ideales para indicar el color de una región que cubra el campo receptivo, mientras que las células oponentes dobles cuyos campos receptivos pueden responder tanto a contrastes cromáticos espaciales (es decir bordes de color) como a las regiones de color [Gao13],

son capaces de señalar contornos y bordes cromáticos y contribuir al fenómeno de la constancia de color (capacidad del sistema visual humano por la cual el color percibido de una superficie permanece constante a pesar de las variaciones en intensidad y composición espectral de la iluminación que incide en él [Fos11]) [Con10].

La existencia de neuronas oponentes dobles con campos receptivos selectivos a la orientación tanto para patrones acromáticos como para cromáticos, sin tener en cuenta la constitución de las entradas de los conos, permiten responder a los bordes y contornos de distintas formas [Tho84, Joh08, Con02, Hei05, Hor07].

Un modelo heurístico del procesamiento cromático propone que muchos de los procesos relacionados con la visión del color tienen lugar en regiones posteriores del córtex, como la región V4 (donde se han encontrado células constantes al color [Con10, Geg03]) y el córtex temporal inferior, a cuyas islas de células llamadas “globs” llegan las señales que salen del V1 por la corriente visual ventral y atraviesan la región V2 a través de sus bandas delgadas. [Des85, Kom98, Zek98, Con07, Yas10].



---

Capítulo 3

**FUNDAMENTOS DE LA ESTADÍSTICA DE IMÁGENES  
NATURALES**

---

Como ya comentamos en la sección de factores físicos de la percepción del color (sección 2.1), nuestro sistema visual, al igual que el resto de los sistemas sensoriales, se ha adaptado a través de la evolución (que nos proporciona una serie de adaptaciones fijas y programadas) y de la experiencia (que altera o ajusta las anteriores a lo largo de la vida), para procesar eficientemente los estímulos recibidos de los entornos en los que desarrollamos nuestra actividad. Con respecto al proceso evolutivo, este viene marcado por tres condicionantes: las tareas que se van a desarrollar, las limitaciones biológicas (como las restricciones metabólicas o de conexiones neuronales, la competencia con otros órganos y sistemas por el espacio...) y el entorno en que el organismo vive. Este último aspecto es el que induce a pensar que sean las imágenes de escenas naturales aquellas para las que el sistema visual se ha adaptado.

El proceso de la selección natural garantiza una fuerte conexión entre el diseño del sistema perceptivo de un organismo y las propiedades del entorno físico en el que el organismo vive. El establecimiento de una relación cuantitativa precisa entre las regularidades estadísticas del entorno y el procesamiento neuronal permite [Gei08,Sim01]:

- Proporcionar un marco para comprender la percepción visual.
- Diseñar paradigmas experimentales que reflejen las características importantes de los estímulos naturales.
- Obtener nuevos modelos computacionales basados en la estadística del entorno.

La caracterización de los estímulos naturales en términos estadísticos se establece, por un lado, por la relación probabilística que existe entre las fuentes de información (estímulos) y la tarea que se va a realizar (reconocimiento de objetos, movimientos...), y, por otro lado, a la naturaleza ruidosa de cualquier medida visual debido a la aleatoriedad inherente a la absorción de la luz (fluctuaciones cuánticas) y a los procesos químicos neuronales. [Gei08]

Como señalan Simoncelli y Olshausen [Sim01], son dos las metodologías básicas utilizadas para comprobar y refinar la relación entre las regularidades estadísticas del entorno y el procesamiento sensorial. La primera es analizar estadísticamente las respuestas neuronales bajo condiciones de estimulación natural [Lau81, Bad98, Vin00]. La segunda formulación es la de obtención de un modelo para el procesamiento sensorial temprano que describa las propiedades de la respuesta de un conjunto de neuronas a partir del estudio de algún descriptor estadístico de los estímulos visuales y de las transformaciones derivadas que lo optimizan [Ati92a, Ols96, Bel97, vHa98, Sim99]. Es esta segunda metodología la que utilizamos en este trabajo.

Se pueden obtener descriptores estadísticos en diferentes etapas (dominios) del camino visual desde el entorno físico a la respuesta conductual. Geisler [Gei08] establece cuatro dominios: el entorno natural, la imagen retiniana, la respuesta neuronal y la respuesta conductual; y dos tipos de mediciones: la estadística dentro del dominio y la estadística entre dominios. El propósito de la primera es caracterizar la distribución de probabilidad de determinada propiedad en un dominio específico, mientras que el fin de la estadística entre dominios es determinar la distribución de probabilidad conjunta de dos propiedades perteneciente cada una a un dominio distinto. Este último tipo de medición es esencial para analizar la estadística de escenas naturales dentro del marco del observador ideal Bayesiano [Gei08, Hyv09]. En el presente trabajo analizaremos descriptores estadísticos en el dominio correspondiente a la imagen retiniana.

No obstante, la relación cuantitativa entre la estadística del entorno y el procesamiento neuronal no pudo establecerse hasta que no se utilizó como enlace entre ambas el concepto de codificación eficiente de la teoría de la información, que C. E. Shannon y W. Weaver habían desarrollado en 1948 [Sha49] para cuantificar las variables que intervienen en la transmisión de un mensaje a través de un canal de comunicación y, que desde entonces ha desarrollado un papel fundamental en el modelado y comprensión de los sistemas neuronales [Sim01]. Una de las ideas clave de esta teoría, en el caso de los estímulos visuales, es que la memoria necesaria para almacenar una imagen es a menudo menor que la necesaria en un código donde cada píxel es almacenado utilizando un número fijo de bits. Esta idea surge de que parte de la memoria es consumida por la estructura redundante de la imagen. Por tanto, los contenidos de cualquier imagen pueden dividirse en información y en redundancia. Las imágenes naturales son redundantes presentando correlaciones que pueden ser de primer orden (valores de intensidades

parecidos localmente), de segundo orden (bordes orientados) o de orden superior (texturas homogéneas) [Dau89].

Fue F. Attneave quien en 1954 sugirió que la finalidad del sistema visual humano era la de producir una representación eficiente de la señal entrante, hablando explícitamente de la redundancia de las imágenes naturales, para lo cual proponía que en las primeras etapas del sistema visual se recodificaban las señales de entradas con la finalidad de reducir dicha redundancia [Att54]. Desde ese trabajo han sido muchos los autores que han obtenido importantes resultados usando variantes de la hipótesis de codificación eficiente [Lau81, Sri82, Ati92b, Bar61, Bar01, Fie94; vHa92]. Ahora bien, asumimos que para reducir la redundancia el primer paso debe ser identificarla, es decir, descubrir previamente la estructura subyacente en la imagen. Quizás el ejemplo más clásico de este enfoque de la codificación eficiente es el de la transformación de la información cromática de los conos a mecanismos oponentes de color de modo que corrigen el solapamiento de las sensibilidades espectrales de los conos [Buc83, Rud98].

Desde un punto de vista neurobiológico, sería necesario un ingente número de neuronas para la representación de cualquier patrón visual que pudiéramos encontrar. En este contexto H.B. Barlow [Bar12] propuso el procesamiento eficiente de la información como una restricción fundamental del procesamiento neuronal, de manera que un grupo de neuronas debería codificar y transmitir tanta información como sea posible dados los recursos disponibles (número de neuronas, rango dinámico de respuesta de las mismas...) así como de la estadística de las imágenes naturales que constituyen las señales de entrada al sistema. [Gei08, Sim01]. En el caso de neuronas individuales, la distribución óptima de sus respuestas dependerá de alguna restricción impuesta a los valores de las mismas ya que, si no tuviera ninguna restricción, la cantidad de información que podría codificar sería ilimitada. Si la restricción impuesta es el valor máximo de su respuesta, una neurona eficiente mostraría una distribución de las respuestas uniforme (debería usar todos los niveles de respuesta por igual); si la limitación impuesta es una varianza fija, la distribución óptima de repuestas sería una gaussiana y, si la restricción es una media fijada, la distribución que optimiza su respuesta sería una exponencial [Sim01].

El rango dinámico de las neuronas para una codificación eficiente, que asegure una respuesta en todo el rango y evite las que se dan fuera del mismo, debería ajustarse al rango dinámico de las señales que deben codificar, lo que puede hacerse ecualizando el histograma de las mismas para que con imágenes naturales responda de igual manera a todos los niveles de la señal de entrada. Encontramos confirmaciones de esta hipótesis entre la distribución de probabilidad del contraste (aunque dependiendo de como esté definido éste [Fra06]) y las funciones respuesta a la luminancia de distintos tipos de células de la retina, LGN y V1 de distintos animales [Lau81, Bra00, Cla03, Tad00].

Para el caso en el que un conjunto de neuronas trata conjuntamente la información de un estímulo, para que la codificación sea eficiente además de que la respuesta individual de cada una de ellas sea óptima, estas deben ser estadísticamente



independientes (los niveles de repuesta de todas las posibles combinaciones deben ser usados por igual) lo que no requiere otro tipo de restricciones. [Sim01]. Olshausen y Field [Ols96] mostraron que campos receptivos parecidos a los de las neuronas de V1 se obtenían de la optimización de la representación dispersa (es decir, una representación en la que la distribución de probabilidad es muy picuda en 0 y presenta colas pesadas) estadísticamente independiente de las respuestas a imágenes naturales [Gei08].

La limitación que la computación impone por el efecto Hughes [Oom08] a la dimensionalidad del descriptor estadístico, objeto de estudio y la dificultad de saber con anticipación qué estadísticas proporcionarán más información para entender el sistema visual no ha impedido el desarrollo de descriptores con dimensionalidad baja y de técnicas estadísticas que permiten el avance en el conocimiento de aspectos de la visión (incluidos aquellos con distribuciones de probabilidad con dimensión alta)[Gei08].

Un ejemplo de estas técnicas lo encontramos en el análisis de componentes principales (PCA) que bajo la hipótesis de normalidad permite describir la estructura de las distribuciones de probabilidad a través de los ejes principales (direcciones que contienen mayor varianza) de la misma. Lo importante de este análisis es que la mayoría de la varianza está contenida en unas pocas de las primeras componentes permitiendo describir características de alta dimensionalidad en un espacio dimensional bastante más pequeño. De este modo, magnitudes espectrales dependientes de la longitud de onda se han podido describir con solo unos pocos números; es el caso del espectro de potencia cromático de las fuentes de luz naturales [Jud64, Dix78] o del espectro de reflectancia de materiales reales [Mal86a, Mal86b]. Si consideramos que las sensibilidades de los conos descomponen el espectro de manera fija en tres respuestas, la aplicación del PCA al logaritmo de estas (ver sección 5.3) [Buc83, Rud98] mostró que las direcciones principales de las mismas se correspondían con mecanismos de oposición de color [Sim01].

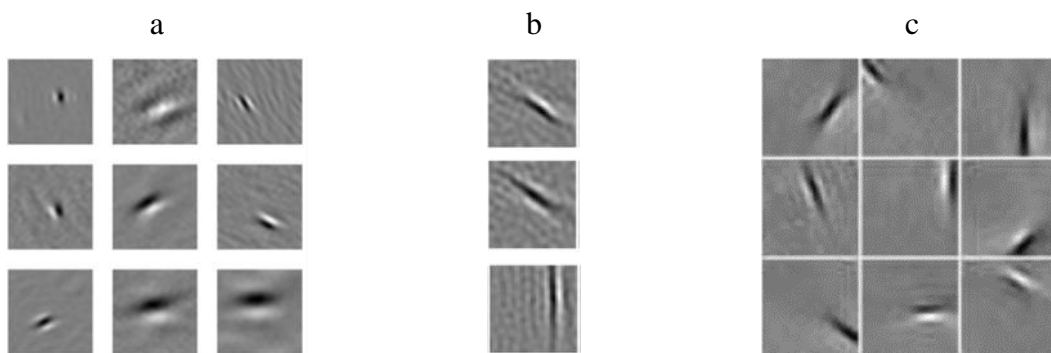
Las imágenes naturales están fuertemente correlacionadas en la intensidad de sus valores cromáticos disminuyendo ésta con la distancia entre píxeles. El estudio de la correlación en función de la separación relativa entre píxeles muestra que las imágenes naturales son invariantes a la traslación (ver sección 5.2), lo que implica que las imágenes pueden ser decorrelacionadas en el dominio de la frecuencia pudiendo expresarse el espectro de potencia como una ley de potencia (con exponente alrededor de 2) de la frecuencia espacial [Sim01, Rud94]. Distintas experiencias han mostrado mecanismos y estructuras fisiológicas de las neuronas visuales que decorrelacionan las señales en las primeras etapas del mismo [Sri82, Ati92b, vHa92, Bal00].

Otra estrategia para caracterizar la estructura espacial de las imágenes naturales es medir las propiedades generales de la distribución de probabilidad conjunta de píxeles de un vecindario espacial de un tamaño dado. Aplicando PCA a trozos de imágenes naturales [Rud98, Wac01] se muestra que se necesita un número considerable de ejes principales para obtener una fracción sustancial de la varianza en la distribución de los

trozos de imágenes y, por tanto, no conduce a un resumen compacto de la estructura espacial de las imágenes naturales de la misma manera que lo hacía para el espectro natural de irradiancia y reflectancia. [Gei08].

El problema que nos encontramos con la aplicación del PCA es que se basa en la covarianza (estadística de segundo orden) por lo que puede no dar buenos resultados si la distribución de las fuentes a la que se aplica es no gaussiana, que es lo que les ocurre a las imágenes naturales. El análisis de componentes independientes (ICA) en vez de asumir que el espacio de señales viene descrito por una única distribución gaussiana multidimensional, considera que este viene descrito por una suma de distribuciones estadísticamente independientes que representan diferentes fuentes desconocidas (en el caso de imágenes naturales contornos producidos por bordes, sombras...). El ICA estima el vector correspondiente a cada fuente a partir de una gran colección de muestras [Gei08].

Los filtros obtenidos de la aplicación del ICA a trozos de imágenes naturales se asemejan mucho espacialmente a los campos receptivos de las células oponentes simples orientadas encontradas en el córtex visual de los mamíferos, ya que están espacialmente localizados, orientados y constituyen un filtro paso banda a diferentes bandas de frecuencias espaciales [Bel97, vHa98, Lew99]. Un ejemplo de esta similitud podemos encontrarlo en la figura 3.1 [Hyv09], en donde se comparan los campos receptivos de células oponentes simples orientadas de un macaco (a), con tres vectores peso encontrados en la maximización de la dispersión en imágenes naturales (b) y con algunas de las características obtenidas al aplicar ICA a trozos de imágenes naturales monocromáticas (c).



**Figura 3.1** .- Campos receptivos de células oponentes simples orientadas de un macaco (a), ejemplos de tres vectores peso encontrados en la maximización de la dispersión en imágenes naturales (b) y algunas de las características obtenidas al aplicar ICA a trozos de imágenes naturales monocromáticas (c) [Hyv09].

A pesar de los grandes avances que ICA ha aportado al conocimiento del sistema visual hay que señalar que las fuentes obtenidas con el mismo nunca son completamente independientes, ya que se trata de un modelo que describe las imágenes como una combinación lineal de las fuentes y las imágenes realmente no están constituidas por tales dependencias lineales (por ejemplo la luz procedente de varios objetos que se tapan parcial o totalmente) [Sim01,Sim97].

Aunque en la retina y el LGN se blanquee<sup>1</sup> una imagen, su versión decorrelacionada todavía contiene estructuras obvias (ver Figura 3.2). Simoncelli y Schwartz [Sim99] mostraron que las dependencias todavía existentes en imágenes decorrelacionadas pueden eliminarse mediante una forma de procesamiento no lineal usando modelos de normalización divisiva. Son varios los trabajos [Rei79, Bon89, Gei92, Hee92, Car97] en los que estos modelos se han utilizado para tener en cuenta el comportamiento no lineal de las neuronas corticales y su estrecha relación con la no gaussianidad de las imágenes naturales.



*Figura 3.2 .- Ejemplo de una imagen natural y de su correspondiente versión blanqueada en la que todavía se aprecian estructuras claras.*

### **3.1 La estadística de imágenes naturales y los iluminantes**

La capacidad del sistema visual humano por la que el color percibido de una superficie permanece constante a pesar de las variaciones en intensidad y composición espectral de la iluminación que incide en él, es conocida como constancia de color [Fos11]. Puede que esta característica de nuestro sistema visual sea la causa por la que la influencia de la luz día en el modelado de la estadística de las imágenes naturales ha recibido poca atención. No obstante, la constancia de color no es perfecta [Fos11], y los

---

<sup>1</sup> El "whitening" o blanqueamiento de datos es una transformación de la información de partida en un conjunto de nuevas variables de modo que estén decorrelacionadas y tengan media cero y varianza unidad.

efectos de diferentes iluminaciones sobre el color y la visibilidad de los objetos puede ser grande sobre todo en aspectos de la reproducción cromática [Sch07, Ma13]. Solamente aspectos relacionados con la codificación eficiente de la distribución espectral de la señal de color y la diversidad cromática han sido estudiados bajo diferentes iluminantes naturales por algunos autores, como vamos a ver a continuación.

Chiao et al. [Chi00] estudiaron la señal de color de imágenes naturales hiperespectrales de entornos boscosos y submarinos a distintas profundidades utilizando el PCA de las mismas. Sus resultados mostraron que las tres primeras componentes obtenidas de las imágenes terrestres bajo cualquiera de los iluminantes proporcionaban cantidades de varianza similares (casi el 90% la primera y mas del 98% las tres primeras) y además eran bastante parecidas espectralmente. Estos resultados no se ven reflejados en las escenas submarinas, en donde los cambios del espectro de la iluminación con la profundidad, debidos a la absorción selectiva espectralmente del agua, limitan el ancho de banda de las componentes principales aumentando la cantidad de varianza que porta la primera componente. Por esta razón, concluían que las variaciones de iluminación natural en la señal de color de los entornos terrestres tenían poca influencia en las componentes principales que se obtenían de las mismas lo que favorecía la constancia de color en estos entornos frente a los subacuáticos.

Por lo tanto, cualquier mecanismo en el que se sustente una percepción de color estable debería tener en cuenta las variaciones de iluminación locales que se producen en las imágenes naturales y que son del mismo orden de magnitud que las globales producidas por los cambios en la luz día. Nascimento et al. [Nas16] aplicando PCA a los espectros de imágenes naturales y teniendo en cuenta sus variaciones de iluminación local obtuvieron, aunque para rangos espectrales diferentes, componentes principales con perfiles parecidos a los encontrados para variaciones globales durante el día [Her01] y que las tres primeras componentes principales proporcionaban un 99'9% de la varianza. Lo que los llevo a considerar que un mecanismo que puede contribuir a la percepción estable del color de una superficie teniendo en cuenta las bruscas variaciones locales de iluminación podría ser la razón espacial de la excitación de los conos, que se obtiene por la luz reflejada de distintas superficies en la escena y que son invariantes bajo los cambios de iluminación [Fos94, Nas02a].

Nieves et al. [Nie12] estudiaron cómo la iluminación natural podía influir a los bordes cromáticos en imágenes naturales hiperespectrales. Sus resultados mostraron ligeras diferencias (en torno al 2% del contraste máximo) en los canales luminancia y cromáticos con la CCT (2700 K -30000 K) del iluminante. Los contrastes de los tres mecanismos visuales, definidos como la razón entre el valor absoluto del gradiente y la intensidad de cada plano de la imagen en cada píxel, descendían marcadamente hasta los iluminantes caracterizados por una CCT de 9000 K manteniéndose prácticamente constantes por encima de los 10000 K y siendo más acentuado este cambio en el canal oponente amarillo-azul. Aunque obtuvieron resultados más estables al aplicar logaritmos

a las respuestas postreceptivas de los tres mecanismos visuales para tener en cuenta la adaptación local a la luz de los conos, los resultados no permitieron concluir que los mecanismos modificados compensaran mejor los cambios de luz en la detección de bordes, ni que hubiera una CCT de luz día óptima para explicar los contrastes en los tres canales.

En cuanto a la dependencia de la variedad cromática con el iluminante, Martínez-Verdú et al. [Mar07] estudiaron los sólidos de color en espacios de color uniformes bajo distintos iluminantes. Sus resultados mostraron que tanto el contenido como la forma de estos dependían claramente del iluminante utilizado. De este modo, las gamas de color correspondientes a iluminantes caracterizados por CCTs en el intervalo 5500-10000 K (salvo que tengan un ancho de banda muy estrecho) eran mayores que para las de los iluminantes fuera de ese rango. Independientemente del tamaño de la gama de color había colores discernibles bajo un iluminante que no lo eran bajo otro, lo que, dada la variedad de iluminantes posibles, el número de colores discernibles sería ilimitado. En su trabajo estos autores consideran que la constancia de color, basada en sus primeras etapas en la adaptación cromática, puede entenderse como un mecanismo que reduzca el número de colores discernibles que no son comunes con otros iluminantes.

Masaoka et al. [Mas13] proponen que el número de colores discernibles dependerá del modelo de apariencia de color (tono, luminosidad y croma), el espacio cromático y el umbral de las diferencias de color utilizadas. Sus resultados muestran que en el rango de temperaturas del iluminante 2000K - 10000K el número de colores discernibles es máximo en los espacios de color CIELAB y CIE94 alrededor de los 4000 K cuando no se ha aplicado ninguna transformación Von Kries de adaptación cromática a los datos; por el contrario, en los espacios CIECAM02 y CAM02-UCS la diversidad cromática aumenta con la temperatura del iluminante. Pero, salvo para el espacio de color CIECAM02, para los demás la variación de la dependencia del número de colores discernibles con la temperatura viene determinado principalmente por el blanco de referencia o las adaptaciones cromáticas utilizadas en cada uno de ellos (cuyo rango de aplicación en términos del cambio de la temperatura de color desconocemos para estos cálculos).

Los aspectos principales que hay que considerar en la reproducción cromática son la fidelidad o el grado de naturalidad de los colores reproducidos bajo un iluminante, el atractivo o preferencia relacionado con la calidad de la imagen y la discriminación de color. Todos ellos se caracterizan por una serie de índices de reproducción [Guo04]. basados en supuestos perceptivos y en datos psicofísicos.

Masuda y Nascimento [Ma13] analizaron el grado de naturalidad y la preferencia de 12 imágenes hiperespectrales de mostradores con alimentos en función de la CCT de iluminantes luz día y metámeros de los mismos (iluminante D65) físicamente existentes. Para ambos tipos de iluminantes encontraron una clara dependencia de estos índices con la CCT de la fuente de iluminación obteniéndose para la luz día que los

colores más naturales se obtenían a una CCT de 6040 K y que los colores más preferidos a una temperatura de 4410 K, siendo ligeramente superiores (6200 K y 4550 K respectivamente) para los metámeros del iluminante D65. Además, en todas las circunstancias las CCTs de los iluminantes bajo los que los colores parecían más naturales eran siempre superiores a la temperatura de la iluminación para los colores preferidos. Cuando se relacionaron estos descriptores con la gama cromática correspondiente a cada iluminante, se comprobó que los iluminantes elegidos para la preferencia cromática se relacionaban con los de mayor gama de colores, mientras que los seleccionados para optimizar la naturalidad presentaban unas gamas más simétricas (con una relación de aspecto más próxima a la unidad). Estos resultados se vieron confirmados en una segunda experiencia cuando permitieron a los observadores una plena libertad en la selección del espectro a la hora de elegir la iluminación, que resultaba más natural o preferida incluso cuando el espectro no es natural o es artificial. De modo que los iluminantes ideales usados para hacer que los productos parezcan naturales o atractivos no serían necesariamente luces día sino iluminantes con picos en sus espectros en las posiciones espectrales adecuadas para estos propósitos [Nas12].





---

## Capítulo 4

# RESULTADOS SOBRE ESTADÍSTICA DE PRIMER ORDEN EN IMÁGENES NATURALES

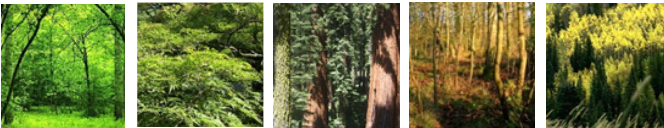




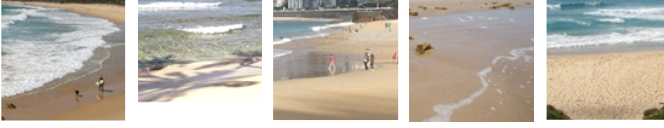


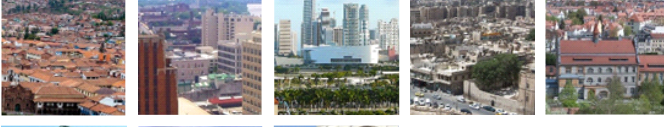
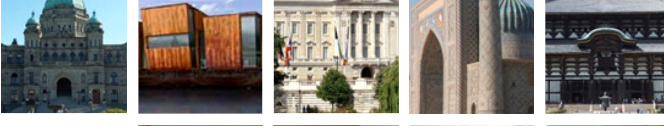
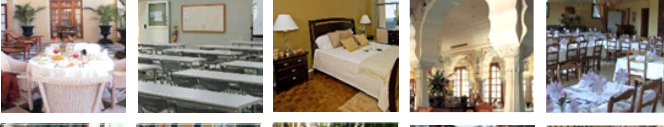

---

### 4.1 Conjunto de datos empleado.

En el presente trabajo se han usado 600 imágenes en color clasificadas en doce categorías semánticas (de 50 imágenes cada una) dependiendo del contenido de las imágenes: Bosques y parques, campo, costa, flores y frutas, montaña, playa, ríos y cascadas, autopistas, ciudades, edificios, interiores y calles. A su vez se agrupan en dos categorías generales [Ros76] llamadas imágenes rurales usando el nombre utilizado en [Nas02b] (las siete primeras clases) y entornos humanos (las cinco últimas). La figura 4.1 muestra algunos ejemplos de las escenas utilizadas agrupadas según sus categorías.

Aunque son varias las bases de datos de imágenes en color disponibles en internet como las de la University of Texas en Austin [Gei11], la de imágenes naturales de la UPenn [Tka11] o la base de imágenes calibradas de Barcelona [Par10, Par09], las imágenes seleccionadas proceden de la base de datos Scene UNDERstanding (SUN) [Xia10] del MIT. Esta base de imágenes era la única que nos permitía obtener suficientes imágenes de las distintas clases semánticas consideradas, pues contiene en la actualidad 131067 imágenes clasificadas en 908 categorías de escenas. No obstante, en las categorías semánticas en las que no se han encontrado suficientes ejemplos se han completado con fotografías tomadas por el autor con diversas cámaras digitales compactas Fujifilm con la configuración predeterminada de las mismas. Una vez seleccionadas, las imágenes fueron recortadas a un tamaño de 400x 400 píxeles y guardadas como imágenes RGB de 8 bits por canal en formato jpeg.



|                              |                   |  |
|------------------------------|-------------------|--|
| Imágenes rurales             | Bosques y parques |    |
|                              | Campo             |    |
|                              | Costa             |    |
|                              | Flores y frutas   |    |
|                              | Montaña           |    |
|                              | Playa             |   |
|                              | Ríos y cascadas   |  |
| Imágenes de entornos humanos | Autopistas        |  |
|                              | Ciudades          |  |
|                              | Edificios         |  |
|                              | Interiores        |  |
|                              | Calles            |  |

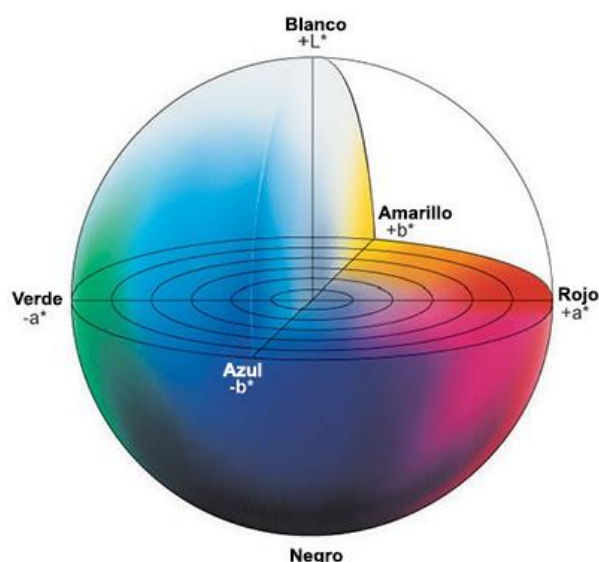
*Figura 4.1: Ejemplos de imágenes pertenecientes a las doce categorías semánticas utilizadas en este trabajo.*

Para aprovechar la propiedad de los espacios de color uniformes en que las propiedades cromáticas del entorno se corresponden con las de su representación visual [McD12], las imágenes fueron transformadas al espacio cromático CIELAB. [Ruz20]. Para ello se utilizó la función RGB2LabNew de Matlab que se basa en la recomendación BT.709 de ITU-R que usa el punto blanco de referencia correspondiente al iluminante D65 con coordenadas triestímulo XYZ (0.9505;1;1.0890). La transformación consiste en transformar la señal RGB normalizada a la unidad a señales triestímulo usando la matriz

$$\begin{pmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{pmatrix} \quad (4.1)$$

normalizándolas dividiéndolas por los valores triestímulo del punto blanco y transformándolas finalmente al espacio CIELAB.

En la figura 4.2 podemos ver una representación tridimensional de dicho espacio de color.



**Figura 4.2:** Espacio cromático CIELAB.

(<http://sensing.konicaminolta.com.mx/images/blogImages/ColorSphere.jpg>)

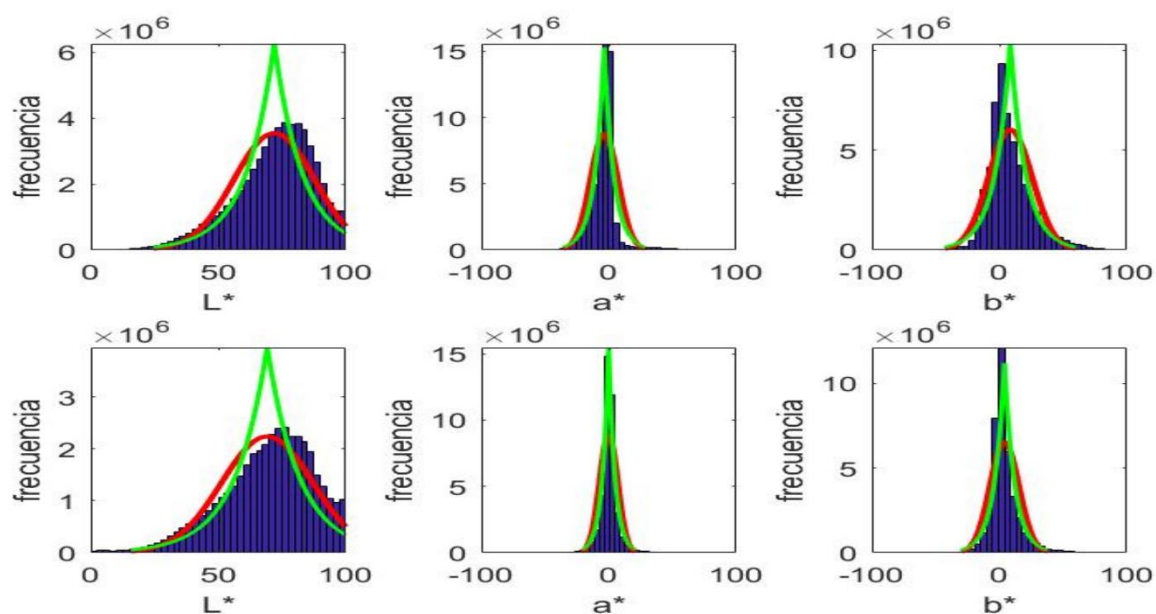
## 4.2 Estadística de primer orden de las imágenes naturales.

Para analizar las propiedades estadísticas de las imágenes naturales se aplica a un conjunto de las mismas una transformación particular que suele corresponderse con una componente, factor o propiedad relacionados con la visión humana. Según el número de píxeles afectados por las relaciones o regularidades que queremos estudiar, podemos hablar de estadística de primer orden (estudia los píxeles individualmente), estadística de

segundo orden (estudia parejas de píxeles) o estadística de orden superior (estudia las relaciones entre más de dos píxeles) [Fie87].

Las regularidades más simples que pueden estudiarse en las imágenes a través de un análisis estadístico son las propiedades individuales de los píxeles como sus coordenadas de cromaticidad. El estudio estadístico de la distribución de la intensidad de las componentes  $L^*$ ,  $a^*$  y  $b^*$  en el espacio cromático CIELAB de los píxeles viene dado a través de los histogramas de los valores correspondientes a cada una de ellas y de sus momentos.

La realización de un Test de Kolmogorov-Smirnov para determinar si nuestras muestras siguen una distribución normal o exponencial, señala que para un intervalo de confianza del 95% las muestras correspondientes al canal  $L^*$  siguen una distribución normal con valores para la  $Z$  de Kolmogorov-Smirnov de 0,774 y 0,772 y una significación de 0,588 y 0,558 respectivamente para las imágenes rurales y de entornos humanos. Del mismo modo el canal  $a^*$  de las imágenes de entornos humanos y el  $b^*$  de los dos tipos de escenas siguen una distribución exponencial ( $Z= 1,172$  y significación de 0,128 para el canal  $a^*$  de la escenas de entornos humanos y  $Z=0,888$  y  $Z=0,582$  con significación de 0,409 y 0,882 para las imágenes rurales y de entornos humanos en el canal  $b^*$ ). Este comportamiento se puede apreciar en la figura 4.3 en donde podemos observar cómo las distribuciones del canal  $L^*$  se ajustan a distribuciones normales mientras que las distribuciones correspondientes a los canales cromáticos presentan una



**Figura 4.3 :** Histogramas correspondientes a los tres canales  $L^*$ ,  $a^*$  y  $b^*$  de las imágenes rurales (fila superior) y de entornos humanos (fila inferior). Superpuesta en rojo aparece la distribución normal y en verde la distribución Laplaciana correspondientes a cada conjunto de datos.

forma más próxima al de la Laplaciana con altas frecuencias en torno al valor central. Este comportamiento disperso es semejante al de los mecanismos de oposición de color obtenidos de la decorrelación del logaritmo de la respuesta de los conos a imágenes naturales hiperspectrales [Rud98] .

Los momentos estadísticos describen de manera cuantitativa la forma de la distribución de una variable estadística  $x$  y se definen como [Pap84]:

$$m_k = E\{(x - c)^k\} = \int_{-\infty}^{\infty} (x - c)^k f(x) dx \quad (4.2)$$

donde  $k$  es el orden del momento  $m$  y  $c$  es una constante que suele ser cero (momentos respecto al origen) o la media  $\mu$  (momentos centrales).

En el caso de variables discretas los momentos pueden escribirse como una suma. Suponiendo que  $x$  toma valores  $x_i$  con probabilidad  $p_i$ .

$$m_k = \sum_{i=1}^N p_i (x_i - c)^k \quad (4.3)$$

Y en particular, para el caso de las imágenes naturales, podemos expresarlos como

$$m_k = \sum_{p=1}^N \frac{(I(p) - c)^k}{N} \quad (4.4)$$

con  $N$  el tamaño de la muestra,  $p$  un índice a través de la misma,  $I(p)$  la intensidad de cada elemento de la muestra.

Con esta definición el primer momento respecto al origen es la media:

$$m_1 = \mu \quad (4.5)$$

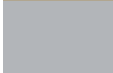
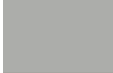
y el segundo momento central es la varianza:

$$m_2 = \sigma^2 \quad (4.6)$$

El significado de los momentos superiores es menos directo; no obstante, volveremos a ellos cuando analicemos la estadística de orden superior. La Tabla 4.1 muestra el promedio de la media y la desviación típica correspondientes a cada una de las componentes  $L^*$ ,  $a^*$  y  $b^*$  de las distintas categorías de imágenes.

La tabla 4.1 muestra los valores medios de los dos primeros momentos para cada una de las categorías semánticas de las imágenes en los tres canales  $L^*$ ,  $a^*$  y  $b^*$ . De los valores de la media vemos que en todas las categorías predominan los píxeles claros sobre los oscuros (valores medios en el canal  $L^*$  de 71,98 para las imágenes rurales y de 69,22 en las de entornos humanos), ligeramente los tonos verdes sobre los rojos (todas las categorías salvo “flores y frutas”, “interiores” y “calles” presentan valores negativos en el canal  $a^*$  entre -0,08 y -9,98) y los amarillos sobre los azules (con valores positivos en el canal  $b^*$  salvo la categoría “montaña” cuyas escenas presentan mayores regiones de cielo y de sombras azuladas). También observamos que el rango de los valores de la media en cada canal de las imágenes rurales es mayor e incluye al de las imágenes de entornos humanos. El análisis ANOVA de un factor para determinar si para un intervalo

de confianza del 95 % podrían considerarse iguales las medias de los dos primeros momentos en cada canal para las distintas clases, rechaza esta hipótesis en todas ellas.

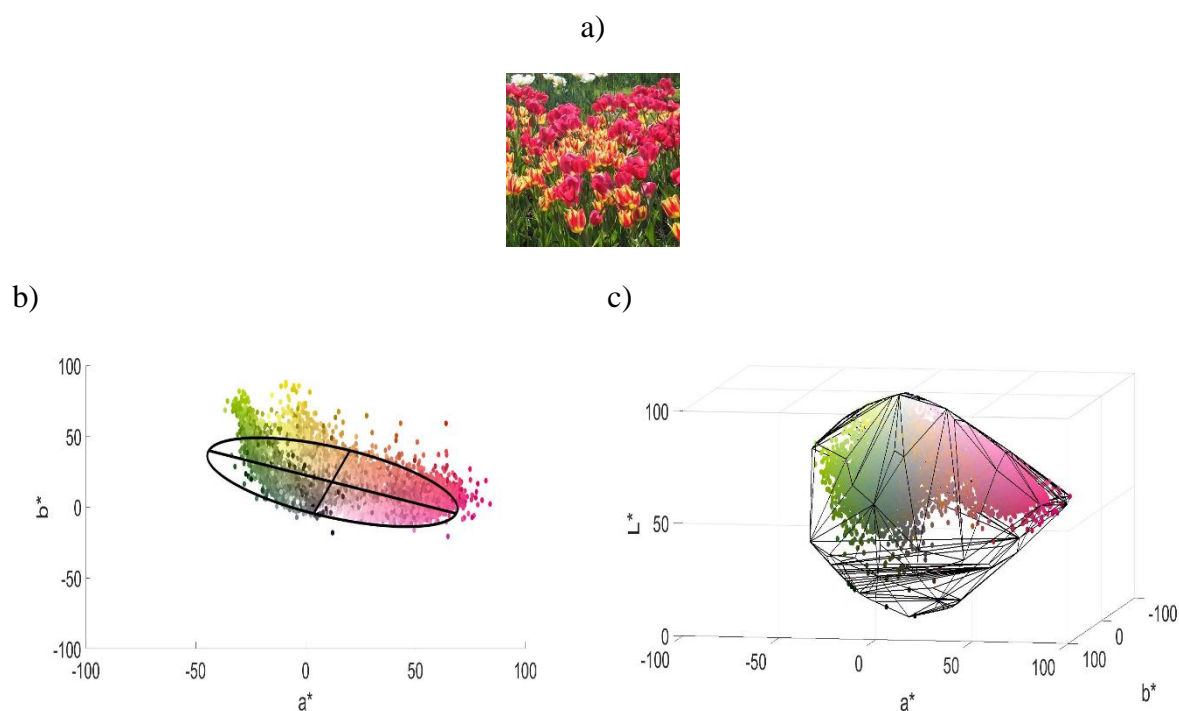
|                   | L*                 |                                   | a*                 |                                   | b*                 |                                   | Color<br>RGB  |
|-------------------|--------------------|-----------------------------------|--------------------|-----------------------------------|--------------------|-----------------------------------|---|
|                   | Media<br>( $\mu$ ) | Desviación<br>típica ( $\sigma$ ) | Media<br>( $\mu$ ) | Desviación<br>típica ( $\sigma$ ) | Media<br>( $\mu$ ) | Desviación<br>típica ( $\sigma$ ) |   |
| Bosques y parques | 65,28              | 15,10                             | -9,98              | 6,41                              | 20,03              | 11,53                             |    |
| Campo             | 73,25              | 9,75                              | -8,88              | 5,60                              | 23,43              | 9,97                              |    |
| Costa             | 72,09              | 12,82                             | -2,34              | 3,18                              | -2,20              | 7,18                              |    |
| Flores y frutas   | 68,49              | 16,67                             | 1,63               | 15,74                             | 12,74              | 15,45                             |    |
| Montaña           | 73,60              | 12,47                             | -0,78              | 2,60                              | -2,30              | 7,22                              |    |
| Playa             | 80,59              | 8,58                              | -1,04              | 2,98                              | 2,66               | 6,49                              |   |
| Ríos y cascadas   | 70,55              | 13,62                             | -3,70              | 4,07                              | 5,87               | 9,07                              |  |
| Imágenes rurales  | 71,98              | 12,72                             | -3,59              | 5,80                              | 8,61               | 9,56                              |  |
| Autopistas        | 71,88              | 11,09                             | -3,03              | 4,51                              | 3,27               | 7,12                              |  |
| Ciudades          | 70,52              | 15,53                             | -0,66              | 4,43                              | 1,15               | 7,06                              |  |
| Edificios         | 68,03              | 16,60                             | -0,08              | 4,70                              | 2,37               | 7,25                              |  |
| Interiores        | 69,25              | 17,51                             | 2,42               | 6,29                              | 10,59              | 10,17                             |  |
| Calles            | 66,43              | 18,24                             | 0,74               | 6,21                              | 2,56               | 7,91                              |  |
| Entornos humanos  | 69,22              | 15,79                             | -0,12              | 5,23                              | 3,99               | 7,90                              |  |

**Tabla 4.1:** Promedio para las distintas clases utilizadas de la media ( $m_1$ ) y la desviación típica ( $\sqrt{m_2}$ ). En la última columna se muestra el color en el espacio RGB que le correspondería a dichos valores medios.



Podemos encontrar la justificación de los colores obtenidos en la última columna de la tabla 4.1 en la suposición del “mundo gris”, que dice que el valor medio de los píxeles en una imagen está relacionado con la temperatura de color correlacionada del iluminante dominante [Pit38]. Además, a pesar de la variabilidad de las reflectancias en una escena, al promediar las intensidades medias de los tres canales de una colección de imágenes se obtiene un valor próximo al gris [Buc80]. Las diferencias respecto a este color pueden deberse a las reflectancias de los objetos predominantes en las escenas; hierba, hojas y flores en las categorías “bosques y parques”, “campo” y “flores y frutas” o a la presencia de otras fuentes de iluminación como en la clase “interiores” [Pou13].

La riqueza cromática de las imágenes puede describirse mediante los descriptores volumen de color, gama de color y número de colores discernibles (NDC). El primero se define como el volumen del sólido convexo que encierra la paleta de todos los colores de la escena en todas sus intensidades disponibles. La gama de color es la proyección del volumen de color en el plano  $a^* b^*$  del espacio de color CIELAB y el número de colores discernibles, el número de cubos unidad distinguibles en ese espacio de color que contienen al menos un píxel [Poi98, Mar07]. En la figura 4.4 se muestra una imagen correspondiente a la categoría de “flores y frutas” (a), su gama de color (b) y el sólido convexo que encierra su volumen de color (c). Por claridad se ha representado uno de cada dos píxeles en el caso de la gama de color y uno de cada cuatro para el caso del volumen de color.



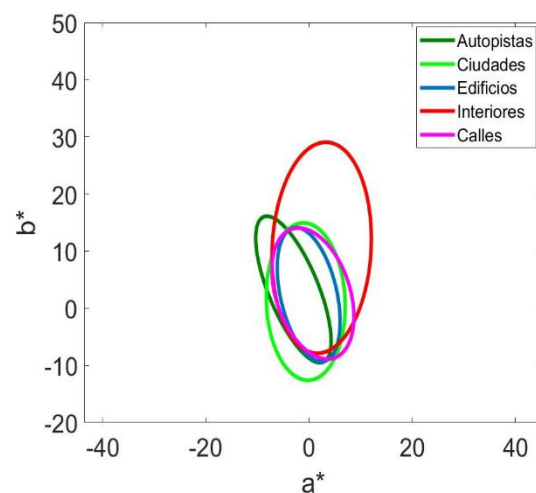
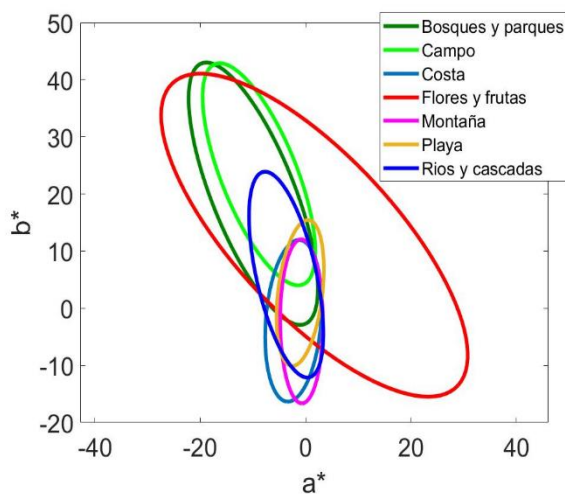
**Figura 4.4** : Una imagen original (a), su gama de color con la elipse de ajuste de mínimos cuadrados (b) y volumen de color encerrado en un sólido convexo.

Para describir los límites, la forma y la orientación de la gama de color de cada imagen se utilizaron las elipses que mejor se ajustan a los datos utilizando el criterio de mínimos cuadrados (en la figura 4.4 (b) se ha incluido la elipse de ajuste correspondiente a la gama de color representada). El área, la razón entre los ejes y el ángulo del eje mayor de la elipse con respecto al semieje horizontal del plano  $a^* b^*$  (valores positivos de  $a^*$ ) fueron empleados también para caracterizar las elipses de ajuste. La Tabla 4.2, muestra el valor medio de estos tres descriptores para las diferentes categorías semánticas utilizadas.

|                   | Área   | Ratio | Ángulo |
|-------------------|--------|-------|--------|
| Bosques y parques | 608,5  | 3,2   | 114,9  |
| Campo             | 469,3  | 3,0   | 118,5  |
| Costa             | 233,9  | 2,7   | 86,3   |
| Flores y frutas   | 1731,9 | 2,6   | 138,2  |
| Montaña           | 182,7  | 3,5   | 92,2   |
| Playa             | 170,2  | 3,1   | 87,6   |
| Ríos y cascadas   | 329,4  | 3,3   | 106,4  |

|            |       |     |       |
|------------|-------|-----|-------|
| Autopistas | 209,4 | 2,9 | 116,5 |
| Ciudades   | 330,5 | 1,8 | 92,6  |
| Edificios  | 206,2 | 2,2 | 107,8 |
| Interiores | 557,2 | 1,9 | 87,4  |
| Calles     | 268,3 | 1,7 | 116,4 |



**Tabla 4.2:** Valores medios del área (en unidades CIELAB), ratio entre los ejes y ángulo (en grados) del eje mayor con respecto al eje positivo de la componente  $a^*$  de las elipses de mejor ajuste cuadrático a la gama cromática de las distintas categorías semánticas de imágenes. Representación de dichas elipses para cada categoría.

De la observación de las elipses, así como de los valores de los distintos descriptores, todas las categorías excepto la de “flores y frutas” están elongadas en el eje amarillo-azul (eje  $b^*$ ), siendo la longitud el eje mayor entre 2,6 y 3,3 veces la del eje menor para imágenes rurales y entre 1,7 y 2,9 para las imágenes de entorno humano. Como vemos en la tabla 4.2 las escenas de entornos humanos presentan una distribución cromática más simétrica que las categorías de entornos más rurales.

De las áreas de las elipses podemos determinar qué categorías semánticas presentan mayor riqueza de colores correspondiendo claramente a la clase “flores y frutas” para el caso de imágenes rurales y a la clase “interiores” para las imágenes de entorno humano. No obstante, que dos gamas de color tengan áreas parecidas de las elipses de ajuste no significa necesariamente que sean iguales, sino que tienen límites parecidos porque los colores no están distribuidos uniformemente dentro de cada elipse [Mon16].

Un ángulo en grados del eje mayor la elipse con el semieje positivo horizontal ( $a^*$  positivo) mayor de  $90^\circ$  indica que la elipse está rotada hacia la izquierda del eje  $b$ . La mayoría de las categorías semánticas salvo “costa”, “playa” e “interiores” presentan un ángulo superior a un ángulo recto con respecto al semieje positivo horizontal.

Por último, para estimar el número de colores discernibles se segmenta el espacio de color en subvolúmenes distinguibles y se cuenta el número de estos que contienen al menos la representación cromática de un píxel de la imagen. Lógicamente, esta estimación dependerá del método de segmentación del espacio de color (cubos, elipsoides, sólidos convexos) [Mas13]. El método implementado en esta investigación es el que considera un cubo unidad como un color discernible en un espacio de color y en el que los valores  $L^*$ ,  $a^*$  y  $b^*$  de cada píxel se redondean al entero mayor o igual que su valor para que todos los valores que caigan dentro de cada cubo unidad queden agrupados en el vértice correspondiente al valor mayor de los tres canales. [Lin08].

La Tabla 4.3 muestra el rango del número de colores discernibles correspondientes a cada clase semántica. Este rango es similar en las categorías correspondientes a entornos humanos alrededor de 35000 salvo para los “interiores” que es 47500, mientras que para imágenes rurales estos oscilan de media entre 21400 y 54100; además, el número total de colores discernibles que pueden alcanzar las imágenes rurales es mayor que el de las de entornos humanos.

Si tenemos en cuenta además del número de colores discernibles el número de veces que aparece cada uno, observamos que la distribución cromática de las imágenes sigue una ley potencial [Mon16] del tipo:

$$y \sim x^{-\beta} \quad (4.7)$$

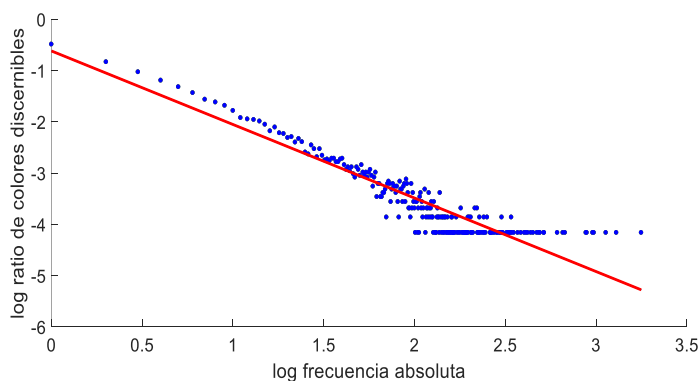
cuando representamos el logaritmo de las distintas ratios de número de colores discernibles ( $y$ ) en función del logaritmo de la frecuencia absoluta de cada una de ellas ( $x$ ), mostrando un coeficiente de correlación entre ellas en imágenes individuales de 0,69



$\rho < 0,98$  y en promedio para cada clase semántica de  $0,8 < \rho < 0,95$ . La fila superior de la Tabla 4.4 muestra un ejemplo de esta dependencia en la que al diagrama de puntos que caracteriza a la imagen que aparece se le ha añadido la recta de mejor ajuste (cuya pendiente corresponde con el exponente de la ley potencial). En la fila inferior vemos el valor medio de la pendiente de la recta de ajuste para cada una de las distintas clases semánticas.

|                     |                   |             |         |            |            |
|---------------------|-------------------|-------------|---------|------------|------------|
| Imágenes<br>Rurales | Bosques y parques | 12514-63511 | Humanos | Autopistas | 3043-39962 |
|                     | Campo             | 2842-55405  |         | Ciudades   | 2853-35985 |
|                     | Costa             | 1374-38593  |         | Edificios  | 5348-40137 |
|                     | Flores y frutas   | 20367-71956 |         | Interiores | 4178-51470 |
|                     | Montaña           | 1802-23234  |         | Calles     | 8422-44279 |
|                     | Playa             | 1131-26404  |         |            |            |
|                     | Ríos y cascadas   | 3686-57814  |         |            |            |

**Tabla 4.3:** Rango del número de colores discernibles correspondientes a las distintas categorías semánticas.

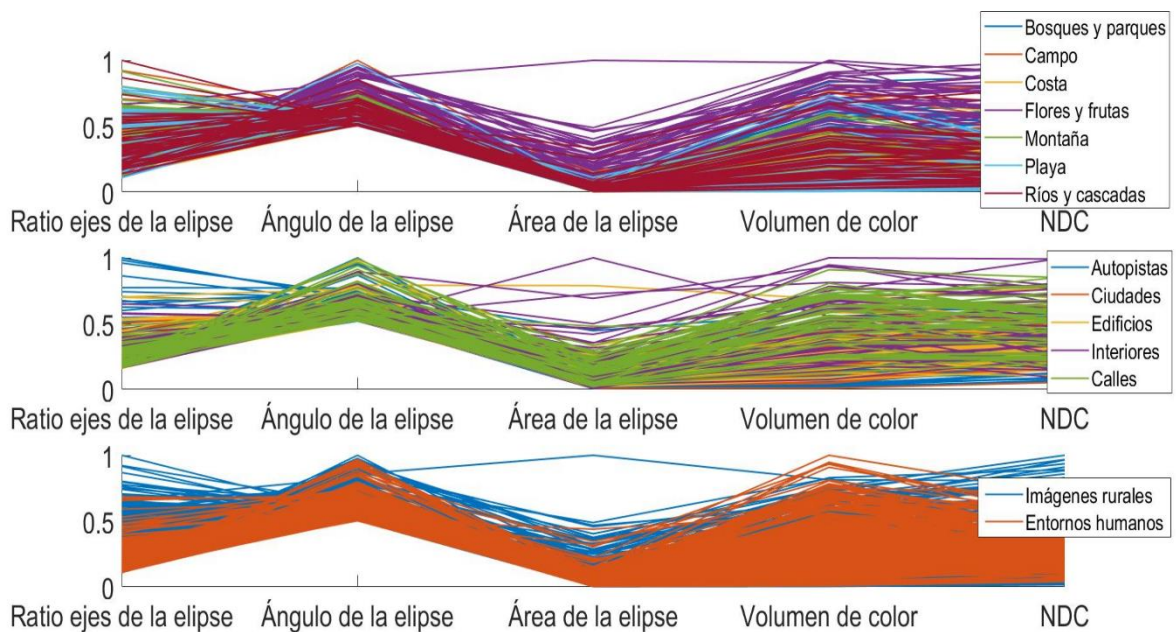


|                   |       |       |                 |         |       |                 |            |          |           |            |        |
|-------------------|-------|-------|-----------------|---------|-------|-----------------|------------|----------|-----------|------------|--------|
| Bosques y parques | Campo | Costa | Flores y frutas | Montaña | Playa | Ríos y cascadas | Autopistas | Ciudades | Edificios | Interiores | Calles |
| -2,10             | -1,61 | -1,19 | -2,04           | -1,11   | -0,89 | -1,39           | -1,05      | -1,52    | -1,39     | -1,48      | -1,39  |

**Tabla 4.4:** Fila superior: Una imagen ejemplo y el diagrama de puntos junto con la recta de ajuste correspondiente al representar el logaritmo de la ratio del número de colores discernibles frente al logaritmo de la frecuencia absoluta. Fila inferior: Pendiente media de la recta de ajuste para las distintas categorías semánticas.

Las categorías que presentan una mayor pendiente en valor absoluto serán aquellas que tienen mayor variedad cromática. De nuevo el rango de estas pendientes en las imágenes rurales es más amplio que el de las imágenes de entornos humanos. Se puede apreciar cómo las clases con pendientes mayores en valor absoluto se corresponden con las que tienen mayor número de colores discernibles.

Sin embargo, desde un punto de vista de la significación estadística, la comparación múltiple de los valores medios de los descriptores de la diversidad cromática (utilizando ANOVA de un factor para distribuciones normales o Test de Kruskal-Wallis para las que no lo son) señala que para un intervalo de confianza del 95% estas serían distintas entre las diferentes categorías de las imágenes rurales y de escenas de entornos humanos. La formación de grupos homogéneos (grupos de categorías en los que las medias de un descriptor sí se pueden considerar estadísticamente iguales) no permite discernir a qué categoría semántica pertenece una imagen o si se corresponde a una imagen natural o de entornos humanos.



**Figura 4.5:** Gráfica de coordenadas paralelas de cinco descriptores cromáticos normalizados al intervalo  $[0,1]$  (Ratio de los ejes, ángulo del eje mayor con respecto a\* positivo y el área de las elipses de ajuste de la gama de color, volumen de color y número de colores discernibles). La imagen superior representa las siete clases de imágenes rurales, la central las de entornos humanos y la inferior los dos tipos rurales y entornos humanos entre sí.

Además, la comparación de los valores medios del área, del ángulo del eje mayor con respecto al semieje positivo horizontal  $a^*$  y de la razón entre los ejes de la elipse que

mejor ajusta la gama de color, así como del volumen de color y el número de colores discernibles entre las dos supracategorías establece una igualdad entre las medias de las dos primeras (estadístico U de Mann – Whitney de 41607 y 41963 con significación de 0.306 y 0.393 respectivamente para el ángulo y el área de la elipse de mejor ajuste con un intervalo de confianza del 95%).

Esta imposibilidad de discernir si una imagen pertenece a una de las distintas clases semánticas, o incluso si esta pertenece a una imagen natural o artificial se ve reflejada en el diagrama coordenadas paralelas de la figura 4.5.

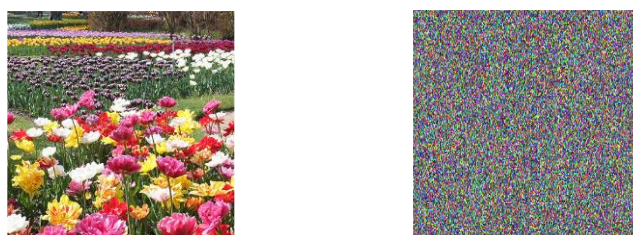
---

## Capítulo 5

# RESULTADOS SOBRE ESTADÍSTICA DE SEGUNDO ORDEN EN IMÁGENES NATURALES

---

Los píxeles de una imagen, además de los valores  $L^*$ ,  $a^*$  y  $b^*$  que nos permitirían definir su color, contienen una información extra que es su posición relativa en la escena. En la figura 5.1 vemos dos imágenes formadas por los mismos valores de los píxeles pero distribuidos de manera aleatoria en la segunda de ellas de modo que la estadística de primer orden sería igual en ambas.



*Figura 5.1 : La imagen de la derecha se obtiene distribuyendo de manera aleatoria los píxeles de la imagen natural de la izquierda de modo que las dos presentarían la misma estadística de primer orden.*

Por tanto, la estadística de primer orden, aunque aporta información sobre la diversidad cromática de las imágenes, al no capturar relaciones entre píxeles, no permite estudiar regularidades espaciales dentro de las escenas.

Una imagen está constituida enteramente por superficies planas que forman bordes donde estas se encuentran. Dado que la señal de color de los objetos depende tanto

de la distribución de la potencia espectral (SPD) de la luz que incide en ellos como de la reflectancia espectral de sus superficies, podemos pensar que dos puntos que interactúen de manera parecida con una luz dada pertenecerán a un mismo objeto o superficie. Por tanto, la búsqueda de cambios bruscos en luminancia o color entre píxeles vecinos nos permite la localización de los bordes de objetos o superficies en una imagen.

La estadística de segundo orden es la encargada de estudiar las relaciones entre pares de píxeles no necesariamente adyacentes en la imagen. Los descriptores estadísticos que caracterizan esta estadística son: los gradientes, espectros de potencia y las rúbricas espectrales.

### 5.1 Estadística entre píxeles adyacentes.

Para determinar la magnitud de la variación de una función utilizamos la derivada, que, para variables discretas, como son los valores de las intensidades de los canales  $L^*$ ,  $a^*$  y  $b^*$  en el espacio, puede aproximarse por el gradiente de dicha variable con respecto a una determinada dirección.

De los distintos procedimientos para calcular el gradiente utilizamos el del método de Sobel [Han09] que tiene en cuenta todos los píxeles vecinos de uno dado. Consiste en la convolución de los valores de cada uno de los planos de la imagen con dos “núcleos” o kernel de dimensiones  $3 \times 3$ , uno para el gradiente horizontal y otro para el gradiente vertical, dados por

$$\text{Kernel Horizontal} = k_h = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad \text{Kernel Vertical} = k_v = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (5.1)$$

Por tanto, los gradientes en cada dirección para cada píxel en cada componente de la imagen vendrán dados por:

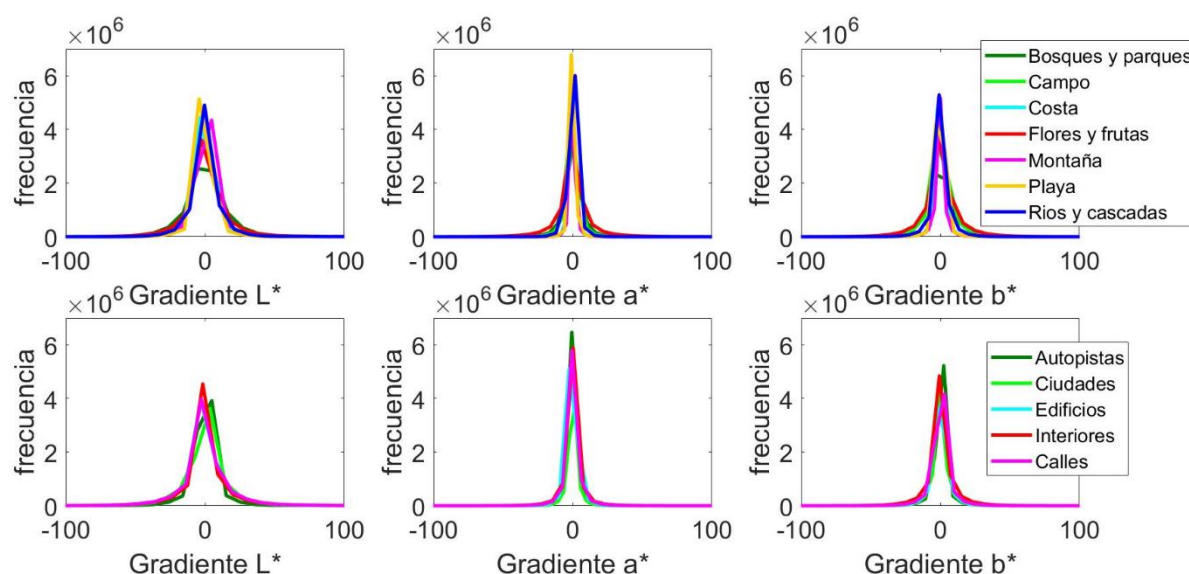
$$E_j^h = k_h * I_j \quad E_j^v = k_v * I_j \quad (5.2)$$

donde  $j$  representa los planos  $L^*$ ,  $a^*$  y  $b^*$ ,  $I_j$  la componente  $j$  de la imagen y  $*$  denota convolución. Finalmente, se hace la media de ambos para tener una estimación media en cada punto independientemente de la dirección:

$$E_j = \frac{1}{2} ( E_j^v + E_j^h ) \quad (5.3)$$

Dado que los bordes serán cambios repentinos y grandes en luminancia o color, el gradiente de las distintas componentes en esas localizaciones será por tanto grande en valor absoluto. Uno de los inconvenientes para determinar bordes a partir de los gradientes es que las superficies con texturas presentan zonas no uniformes que junto a la presencia de ruido llevarán a pequeñas fluctuaciones locales en luminancia o color y, por tanto, a fluctuaciones en el gradiente que no se corresponden con bordes reales [Pou13].

Para reducir el efecto del ruido y bordes altamente locales, las imágenes fueron previamente desenfocadas usando un filtro gaussiano con una desviación estándar de 1. Los histogramas de los gradientes correspondientes a los planos  $L^*$ ,  $a^*$  y  $b^*$  fueron calculados para cada una de las imágenes y agrupados por categorías semánticas como aparece en la figura 5.2. En todos ellos se observa un pico pronunciado en torno a cero debido a que las imágenes presentan grandes superficies que tienden a ser suaves, un rápido decaimiento por los bordes de bajo contraste y largas colas correspondientes a los bordes de alto contraste. La simetría de la distribución del gradiente alrededor del pico central puede justificarse porque los objetos similares tienden a distribuirse de la misma manera y, por tanto, la transición de un objeto a otro será parecida en cualquier dirección. El test de Kolmogorov-Smirnov con un intervalo de confianza del 95% aplicado a las distintas distribuciones de los gradientes en cada categoría nos indica que las distribuciones correspondientes a los canales cromáticos pueden modelarse por una exponencial  $e^{-\alpha x}$  ajustándose a una Laplaciana generalizada. Este ajuste al histograma de los gradientes era considerado como el mejor ajuste para estos descriptores por Huang y Mumford [Hua99].

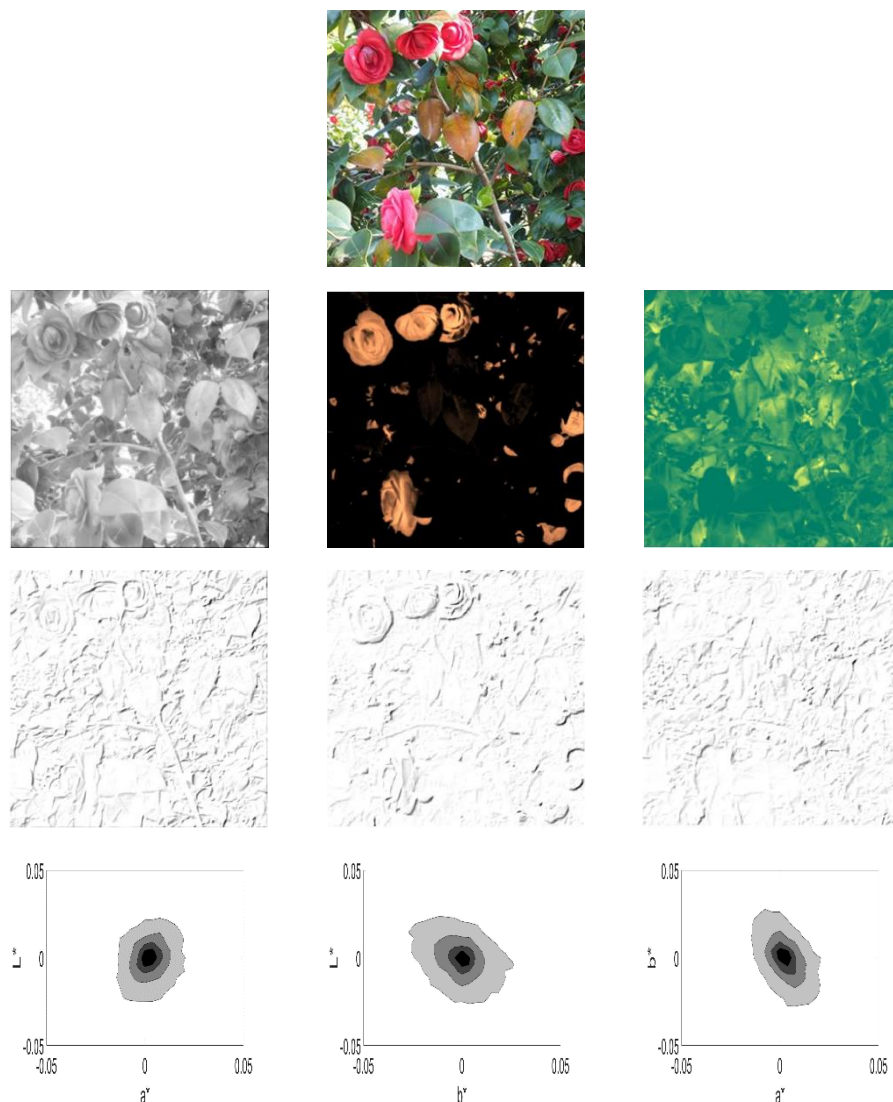


**Figura 5.2 :** Histogramas del gradiente correspondientes a las tres componentes ( $L^*$ ,  $a^*$  y  $b^*$ ) de imágenes rurales (fila superior) y de entornos humanos (fila inferior).

Para caracterizar la simultaneidad de los bordes en los tres planos  $L^*$ ,  $a^*$  y  $b^*$  y como una aproximación a la densidad de probabilidad conjunta, se calcularon los tres histogramas conjuntos correspondientes a los gradientes en los distintos planos de cada escena. El valor en la posición  $(i,j)$  de un histograma conjunto denota el número de píxeles con gradiente  $i$  en uno de los canales y  $j$  en el otro en la misma posición. En la Figura 5.3 podemos ver en la fila superior una imagen correspondiente a la categoría de “flores y



frutas”. La segunda fila representa en falso color sus componentes  $L^*$ ,  $a^*$  y  $b^*$  y la tercera fila los bordes correspondientes a los tres planos. Como puede apreciarse, muchos de los bordes de los canales cromáticos  $a^*$  y  $b^*$  también aparecen en el canal de luminancia  $L^*$ . Este comportamiento podría hacer pensar que la información cromática es redundante [Zho08] no siendo este el caso. Variaciones locales de iluminación generan sombras y penumbras que están definidas por el gradiente de luminancia y no por los cromáticos, por lo que la coincidencia de bordes en  $L^*$  simultáneamente con  $a^*$  o  $b^*$  no debe interpretarse como redundancia sino como un indicador fiable de los límites de los materiales [She08].



**Figura 5.3 :** En la fila superior, imagen original. La segunda fila representa, en falso color, sus componentes  $L^*$ ,  $a^*$  y  $b^*$ . Tercera fila, los bordes correspondientes a los tres planos. Fila inferior, histogramas conjuntos de los distintos canales.

Finalmente, la fila inferior de la figura 5.3 muestra los histogramas conjuntos de los distintos canales asociando una mayor oscuridad de gris a una mayor frecuencia en la combinación correspondiente de valores. Los histogramas se muestran en una escala logarítmica de los ejes para mejorar su representación.

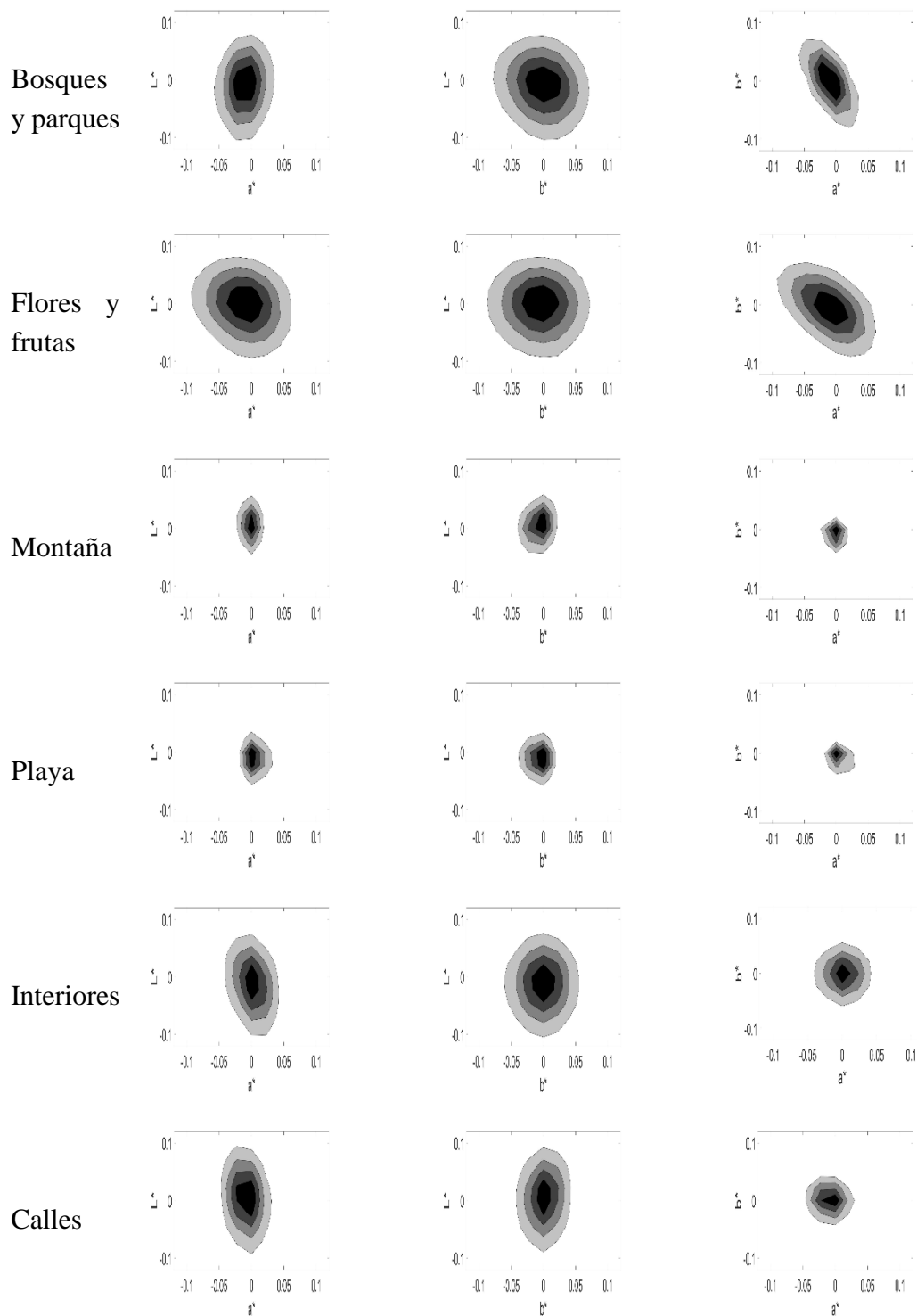
El histograma conjunto de gradientes presenta un máximo en torno al cero en ambos ejes, porque los bordes son eventos poco frecuentes, la mayoría de los píxeles en una imagen no forman parte de un borde. Si los bordes puramente cromáticos fueran menos frecuentes que los acromáticos, los histogramas deberían contraerse a lo largo de los ejes cromáticos. Dado que esto no se produce y además el ancho de las distintas regiones de las áreas grises son aproximadamente iguales en las direcciones horizontales y verticales hay tantos bordes isoiluminantes como los puramente de luminancia (o de ambos canales cromáticos). No obstante, la mayoría de los bordes combinan información de los dos canales pero en distintos grados, de modo que el valor del gradiente en una dimensión no puede predecir el valor en la otra [Han09].

Para ver el comportamiento en cada una de las categorías semánticas los gradientes se promediaron a través de las imágenes y se calcularon los histogramas conjuntos para los valores medios. De la figura 5.4 podemos decir que las formas de estos histogramas no son perfectamente simétricas y elipsoidales, siendo un patrón general la presencia de un máximo en torno al cero y un ancho de las distintas regiones de cada uno de los niveles parecidas. El alargamiento del histograma a lo largo del eje  $L^*$  frente a los canales cromáticos se debe a los diferentes métodos de escalado utilizado para estos canales. También podemos apreciar que aquellas categorías más “coloridas” como “flores y frutas” presentan valores mayores de gradientes cromáticos (mayores distancias a lo largo de los ejes  $a^*$  o  $b^*$ ) que por ejemplo las de playa o “montaña”. Semejante comportamiento se observa también en las imágenes de entornos humanos presentando asimismo muchos bordes cromáticos de alto contraste lo que permite su mejor reconocimiento [Geg00].

De nuevo observamos que la mayoría de los bordes combinan información de los dos canales. Para cuantificar el grado de independencia de los mismos se ha calculado la información mutua media entre los gradientes de los distintos canales. La información mutua media se mide en bits y nos da la cantidad de información que comparten dos variables aleatorias  $X$  e  $Y$  (con  $N$  elementos discretos cada una), valiendo cero (su valor mínimo) si ambas variables son independientes y  $\log_2 N$  (su valor máximo) si son dependientes. Se define como [Sha49]:

$$I(X, Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log_2 \left( \frac{p(x, y)}{p(x)p(y)} \right) \quad (5.4)$$





**Figura 5.4 :** Histogramas conjuntos de los gradientes entre los canales  $L^*$  y  $a^*$ (primera columna),  $L^*$  y  $b^*$ (segunda columna) y  $a^*$  y  $b^*$ (tercera columna) correspondientes a 4 categorías rurales (bosques y parques, flores y frutas, montaña y playa) y 2 de entornos humanos (interiores y calles).

donde  $p(x,y)$  es la función de distribución de probabilidad conjunta de X e Y y  $p(x)$ ,  $p(y)$  son las funciones de distribución de probabilidad marginal de X e Y respectivamente. En nuestros histogramas conjuntos X e Y estaban formadas por 64 elementos por tanto, la información mutua podrá variar entre 0 y 6 bits. La Tabla 5.1 muestra el valor de la información mutua de los gradientes entre los distintos canales para las diferentes categorías semánticas.

|                   | L*-a* | L*-b* | a*-b* |                              | L*-a* | L*-b* | a*-b* |
|-------------------|-------|-------|-------|------------------------------|-------|-------|-------|
| Bosques y parques | 0,037 | 0,048 | 0,391 |                              |       |       |       |
| Campo             | 0,045 | 0,087 | 0,237 | Autopistas                   | 0,014 | 0,015 | 0,094 |
| Costa             | 0,014 | 0,014 | 0,019 | Ciudades                     | 0,011 | 0,025 | 0,015 |
| Flores y frutas   | 0,024 | 0,016 | 0,226 | Edificios                    | 0,023 | 0,012 | 0,014 |
| Montaña           | 0,007 | 0,032 | 0,055 | Interiores                   | 0,054 | 0,016 | 0,003 |
| Playa             | 0,007 | 0,029 | 0,049 | Calles                       | 0,023 | 0,009 | 0,008 |
| Ríos y cascadas   | 0,012 | 0,011 | 0,172 |                              |       |       |       |
| Imágenes rurales  | 0,028 | 0,046 | 0,110 | Imágenes de entornos humanos | 0,015 | 0,026 | 0,003 |

**Tabla 5.1** : Información mutua de los gradientes entre los distintos canales para las diferentes categorías semánticas

A la vista de los resultados de la medida de la información mutua entre los bordes de dos canales, que se mueve entre 0,003 y 0,391 bits, podemos concluir que un gradiente en una dimensión no permite predecir el valor del gradiente más probable en la otra. Basándonos en estos datos y tal y como Gegenfurtner y Rieger señalaban en [Geg00] podemos concluir que hay un grado sustancial de independencia entre bordes cromáticos y de luminancia en escenas naturales y, por tanto, estos proporcionan fuentes independientes de información. La independencia no es perfecta, en particular, entre los bordes detectados en los dos canales cromáticos, como indica la información mutua pequeña pero mayor correspondiente a los canales  $a^*$  y  $b^*$ .

## 5.2 Estadística entre píxeles no adyacentes.

En la sección anterior vimos que las imágenes del mundo real están formadas por superficies separadas por bordes, por tanto la probabilidad de que dos píxeles adyacentes tengan valores parecidos es muy alta y esta capacidad de predicción se pierde cuando los píxeles están alejados. Para poder analizar la relación entre puntos de una imagen no necesariamente vecinos en función de su distancia haremos uso de la función de autocorrelación normalizada.

Dado un plano de una imagen  $I$  con media  $\mu_I$  y desviación típica  $\sigma_I$  y una máscara  $M$  de dimensiones (menores que las de la imagen)  $m_x \times m_y$  con media  $\mu_M$  y desviación típica  $\sigma_M$ , si consideramos que cada píxel del plano de la imagen  $(x,y)$  forma el centro de una región de tamaño el de la máscara, entonces la correlación cruzada de dicha región con la máscara nos dará una medida de la semejanza entre ambas [Pou13].

$$R(x,y) = \sum_{x_s=-m_x/2}^{\frac{m_x}{2}+1} \sum_{y_s=-m_y/2}^{\frac{m_y}{2}+1} I(x+x_s, y+y_s)M(x_s+m_x/2, y_s+m_y/2) \quad (5.5)$$

Que podemos normalizar para que  $|r(x,y)| \leq 1$  como:

$$r(x,y) = \sum_{x_s=-m_x/2}^{\frac{m_x}{2}+1} \sum_{y_s=-m_y/2}^{\frac{m_y}{2}+1} \frac{I(x+x_s, y+y_s)M(x_s+\frac{m_x}{2}, y_s+\frac{m_y}{2})}{\sigma_I \sigma_M} \quad (5.6)$$

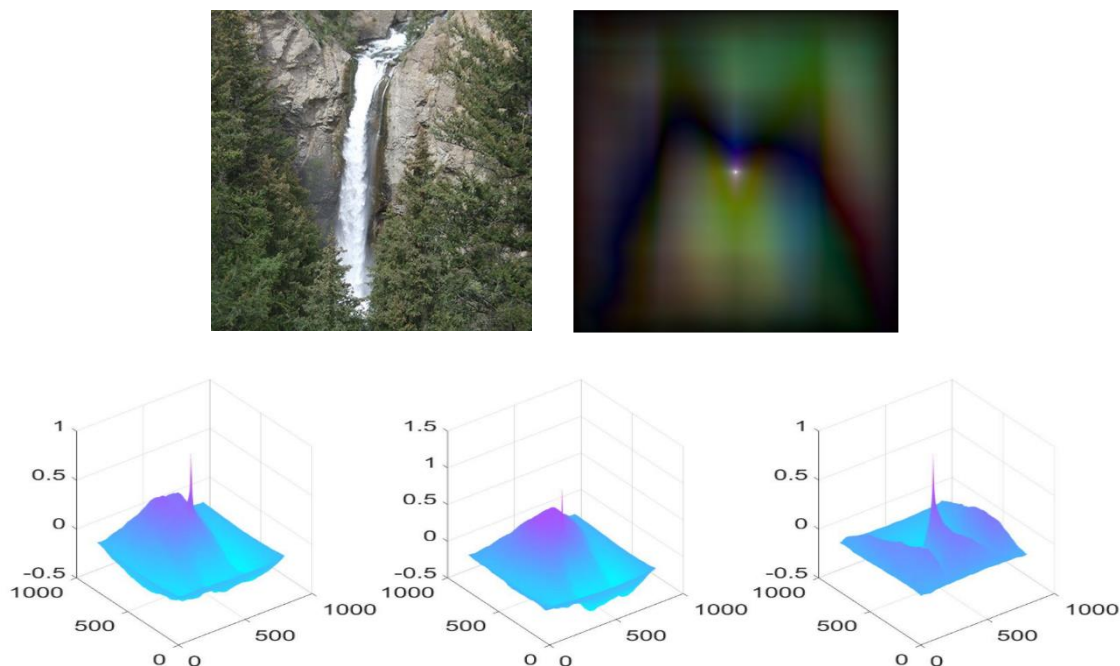
de manera que una correlación cruzada próxima a uno en valor absoluto puede ser interpretada como prácticamente una igualdad entre la región de nuestra imagen con centro  $(x,y)$  y la máscara utilizada. Esto permite la utilización de este descriptor para detectar patrones específicos en las imágenes [Gao06].

Cuando la correlación cruzada normalizada se calcula entre el plano de una imagen de dimensiones  $n_x \times n_y$  y una versión desplazada de la misma imagen, medimos su autocorrelación en función del desplazamiento (distancia entre píxeles) horizontal  $x_s$  y vertical  $y_s$ , que viene dada por:

$$r_a(x,y) = \sum_{x_s=0}^{n_x-1} \sum_{y_s=0}^{n_y-1} \frac{I(x+x_s, y+y_s)I(x,y)}{\sigma_I^2} \quad (5.7)$$

De este modo, si para un desplazamiento dado se encontrara una alta autocorrelación, podría interpretarse como una indicación de que la imagen presenta una periodicidad con una frecuencia proporcional a este desplazamiento [Pou13].

Como se aprecia en la fila inferior de la figura 5.5, las funciones de autocorrelación para cada uno de los canales presentan un máximo en su centro (desplazamiento cero) y tienden a valores menores en valor absoluto a medida que se consideran desplazamientos mayores; además, como se ve en la autocorrelación correspondiente al canal  $b^*$ , se obtiene también alguna información direccional.



**Figura 5.5 :** En la fila superior una imagen original y una representación conjunta en RGB de la autocorrelación calculada por separado para cada uno de los planos  $L^*$ ,  $a^*$  y  $b^*$ . En la fila inferior la función de autocorrelación de cada uno de los tres canales. ( $L^*$  izquierda,  $a^*$  centro y  $b^*$  derecha)

El inconveniente de trabajar con la función de autocorrelación es su alto coste computacional temporal  $O(n_x^2 n_y^2)$ . No obstante, el teorema de Wiener-Khitchine [Khi34] proporciona una herramienta matemática que permite trabajar de manera equivalente en el espacio de frecuencias que tiene un coste menor  $O(n_x n_y \log(n_x n_y))$  [Pou13]. También puede utilizarse para definir el espectro de potencia a través de la autocorrelación cuando la transformada de Fourier no converge [Lee05].

Este teorema establece que la función de autocorrelación de un plano de una imagen es igual a la transformada inversa de Fourier del espectro de potencia (cuadrado del módulo de la transformada de Fourier) de dicho canal.

$$R_a(x,y) = \mathcal{F}^{-1}(|F(u,v)|^2) \quad (5.8)$$

con  $\mathcal{F}^{-1}$  la transformada inversa de Fourier y  $F(u,v)$  la transformada de Fourier de un plano de la imagen  $I(x,y)$  que para señales discretas vienen definidas por:

$$F(u,v) = \mathcal{F}(I(x,y)) = \sum_x \sum_y I(x,y) e^{-2\pi i(ux+vy)} \quad (5.9)$$

$$I(x,y) = \mathcal{F}^{-1}(F(u,v)) = \sum_u \sum_v F(u,v) e^{2\pi i(ux+vy)} \quad (5.10)$$

No obstante debido a la falta de periodicidad de las imágenes naturales (desigualdades entre sus extremos opuestos) y a la presencia de frecuencias que no

corresponden con un armónico y que, por tanto, presentan dispersión espectral (tienen componentes en distintos armónicos) [Har78], el análisis de Fourier puede presentar resultados espectrales no deseados en la detección de frecuencias en presencia de frecuencias intensas cercanas o en la resolución de frecuencias cercanas de intensidad similar.

Para reducir estos efectos a la hora de calcular el espectro de potencia de nuestras imágenes se aplicó una función ventana adecuada para igualar los valores de los bordes y reducir la dispersión espectral. La selección de esta ventana viene dada por un compromiso entre el ancho de un lóbulo principal (que debería ser lo más estrecho posible para conseguir la mejor resolución espectral posible) y la presencia de pequeños lóbulos laterales (que deberían anularse para evitar la dispersión espectral). En nuestros cálculos, de entre los distintos tipos de funciones ventana hemos elegido la ventana de Kaiser-Bessel que, para una energía y una extensión espacial limitada, maximiza la energía en el espacio de frecuencias [Har78]. Esta ventana bidimensional se implementó multiplicando la correspondiente ventana unidimensional por su traspuesta para el tamaño de la imagen  $N$

$$w_{KB2}(x,y) = w_{KB}(x) \cdot w_{KB}'(y) \quad (5.11)$$

$$w_{KB}(x) = \frac{I_0\left(\beta \sqrt{1 - \left(\frac{x-N/2}{N/2}\right)^2}\right)}{I_0(\beta)} \quad (5.12)$$

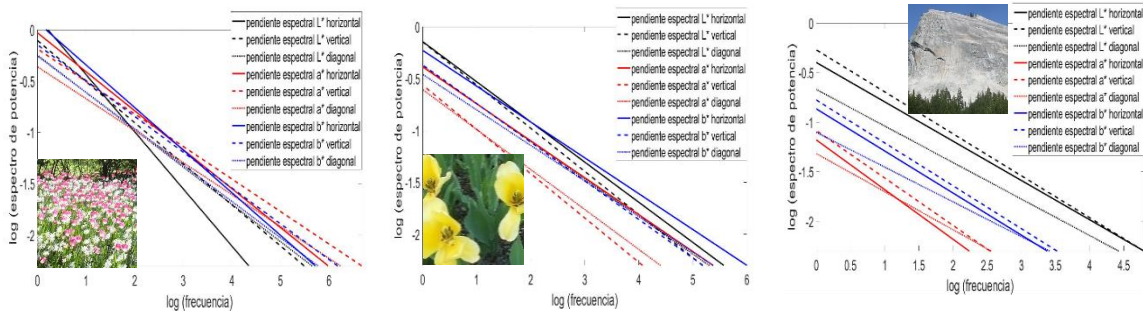
con  $0 < x < N/2$ ,  $I_0$  es la función de Bessel modificada de primera clase y orden cero y  $\beta = 2\pi$ .

Para centrarnos en las variaciones de los valores de los píxeles y descartar el valor medio de cada uno de los canales, igualamos a cero el valor medio de cada plano sustrayendo su correspondiente media a los valores originales de cada canal previamente a la aplicación de la ventana. Filtramos los tres canales de las imágenes resultantes y calculamos el cuadrado del módulo de la transformada discreta de Fourier de cada uno de ellos.

La frecuencia espacial (ciclos/grado) en una dirección dada nos indica la rapidez del cambio del estímulo en esa dirección en el espacio. El espectro de potencia codifica cuanto de cada frecuencia espacial hay disponible en una imagen y puede reducirse a una función unidimensional de la frecuencia espacial realizando una media rotacional en el plano espectral.

Cuando se trabaja con imágenes individuales, la dependencia del espectro de potencia con la frecuencia puede variar significativamente incluso entre imágenes pertenecientes a la misma categoría semántica, como se ve en las rectas de ajuste para los distintos planos de las tres imágenes de la figura 5.6 en sus distintas orientaciones ( $0^\circ$ ,  $45^\circ$  y  $90^\circ$ ); no obstante, cuando se promedian entre las imágenes de cada clase semántica se

obtienen resultados para las pendientes espectrales próximos a los descritos en los tres canales de las imágenes, como se muestra en la Tabla 5.2.



**Figura 5.6 :** Rectas de ajuste entre el logaritmo del espectro de potencia para los tres canales  $L^*$ ,  $a^*$  y  $b^*$  en tres direcciones (horizontal  $\theta=0^\circ$ , diagonal  $\theta=45^\circ$  y vertical  $\theta=90^\circ$ ) y el logaritmo de la frecuencia espacial en ciclos/imagen, para las tres imágenes perteneciendo las dos primeras a la misma categoría semántica.

Uno de los resultados fundamentales de la representación basada en frecuencias de las imágenes es que la potencia espectral ( $ps$ ) de las imágenes naturales en cualquier dirección como en promedio decae con la frecuencia,  $f$ , de acuerdo a una ley potencial,

$$ps \sim f^{-\beta} \quad (5.13)$$

con valores estimados de  $\beta$  próximos a 2. [Fie87, Bur87, Tol92, Rud94], es decir, las frecuencias más bajas contienen la mayoría de la potencia, con la potencia decreciendo en función de la frecuencia. Al mostrar la relación en una escala logarítmica, la potencia en función de la frecuencia sigue una recta de pendiente  $\beta$  (pendiente espectral).

Al estudiar la significación estadística de las diferencias encontradas en la media de las pendientes de cada canal en las tres orientaciones (horizontal, vertical y diagonal) entre cada una de las clases rurales, de las clases de entornos humanos y las dos supraclases entre sí utilizando la prueba ANOVA de un factor, los resultados obtenidos para un intervalo de confianza del 95% mostraron que solo en el plano luminancia y en dos de las categorías de entornos humanos (“autopistas” y “edificios”) podía considerarse que las pendientes eran iguales en las tres orientaciones ( $F=2,268$  y  $2,416$  y significación  $0,107$  y  $0,093$  respectivamente). Análisis similares se realizaron entre planos del espacio de color en la que solo algunas direcciones de las categorías “ríos y cascadas” y “calles” podían ser consideradas iguales (pendiente horizontal ( $F=1,366$ ,  $s=0,258$ ) y diagonal ( $F=1,707$ ,  $s=0,185$ ) de “calles” y pendientes vertical y diagonal de “ríos y cascadas” ( $F=1,505$ ,  $s=0,225$  y  $F=1,092$ ,  $s=0,338$ ) respectivamente) y entre categorías en las que solo las pendientes en la dirección diagonal en el plano de color  $L^*$ , podrían considerarse iguales dentro de las categorías de entornos humanos ( $F=2,332$  y  $s=0,056$ ). Por el contrario el análisis de las diferencias entre los valores de las pendientes espectrales

correspondientes a las dos supracategorías usando el test de Mann-Whitney para el mismo intervalo de confianza muestra que la mitad de ellas pueden considerarse iguales (es el caso de la vertical y diagonal en el plano  $L^*$  y las horizontales y verticales en el plano  $a^*$  ( $U=42829$ ,  $s=0,660$ ;  $U=40237$ ,  $s=0,093$ ,  $U=42630$ ,  $s=0,593$  y  $U=39870$ ,  $s=0,064$  respectivamente).

|                              | $L^*$      |          |          | $a^*$      |          |          | $b^*$      |          |          |
|------------------------------|------------|----------|----------|------------|----------|----------|------------|----------|----------|
|                              | Horizontal | Vertical | Diagonal | Horizontal | Vertical | Diagonal | Horizontal | Vertical | Diagonal |
| Bosques y parques            | -2,35      | -2,26    | -2,37    | -2,06      | -2,01    | -2,20    | -2,04      | -1,93    | -2,11    |
| Campo                        | -1,81      | -2,59    | -2,50    | -1,96      | -2,52    | -2,60    | -1,68      | -2,31    | -2,32    |
| Costa                        | -2,57      | -2,68    | -2,85    | -2,13      | -2,26    | -2,53    | -2,38      | -2,51    | -2,79    |
| Flores y frutas              | -2,46      | -2,63    | -2,79    | -2,67      | -2,73    | -3,02    | -2,48      | -2,55    | -2,76    |
| Montaña                      | -2,35      | -2,45    | -2,77    | -2,00      | -2,08    | -2,43    | -2,33      | -2,39    | -2,62    |
| Playa                        | -2,34      | -2,50    | -2,82    | -1,97      | -2,30    | -2,73    | -2,32      | -2,44    | -2,77    |
| Ríos y cascadas              | -2,46      | -2,25    | -2,61    | -2,31      | -2,27    | -2,56    | -2,23      | -2,48    | -2,45    |
| Imágenes rurales             | -2,33      | -2,48    | -2,67    | -2,16      | -2,31    | -2,58    | -2,21      | -2,37    | -2,55    |
| Autopistas                   | -2,54      | -2,81    | -2,98    | -2,36      | -2,40    | -2,83    | -2,36      | -2,47    | -2,93    |
| Ciudades                     | -2,12      | -2,08    | -2,77    | -2,15      | -2,23    | -2,59    | -2,17      | -2,20    | -2,67    |
| Edificios                    | -2,33      | -2,50    | -2,80    | -2,15      | -2,22    | -2,56    | -2,30      | -2,35    | -2,55    |
| Interiores                   | -2,17      | -2,49    | -2,75    | -2,02      | -2,28    | -2,51    | -2,08      | -2,37    | -2,52    |
| Calles                       | -2,23      | -2,44    | -2,77    | -2,20      | -2,28    | -2,50    | -2,04      | -2,18    | -2,46    |
| Imágenes de entornos humanos | -2,28      | -2,46    | -2,82    | -2,18      | -2,28    | -2,60    | -2,19      | -2,31    | -2,63    |

**Tabla 5.2 :** Pendientes medias de las rectas de ajuste entre el logaritmo del espectro de potencia y el logaritmo de la frecuencia espacial para los tres canales  $L^*$ ,  $a^*$  y  $b^*$  en tres direcciones (horizontal  $\theta=0^\circ$ , diagonal  $\theta=45^\circ$  y vertical  $\theta=90^\circ$ ) para las diferentes categorías semánticas.



Por lo tanto, esta disparidad de resultados no nos permite sacar ninguna conclusión del análisis de la significación estadística de estas diferencias con las orientaciones, las componentes cromáticas o las categorías semánticas comparadas.

La Tabla 5.2 también muestra que el espectro de potencia no está distribuido isotrópicamente en función de la orientación, sino que su pendiente es mayor para orientaciones diagonales que para las horizontales o verticales siendo esta dependencia angular similar para todas las clases semánticas. La presencia de mayor potencia espectral para las orientaciones correspondientes a ángulos de  $0^\circ$  y  $90^\circ$  no se debe ni a la geometría cuadrada de las imágenes ni de los píxeles sino al dominio de estructuras horizontales y verticales de las imágenes rurales [VdS98] incluso más acentuadas en las de entornos humanos.

Una de las principales características de las leyes potenciales es la invariancia a la escala que podemos caracterizar matemáticamente como:

$$f(ax) = a^n f(x) \quad (5.14)$$

y para el caso particular del espectro de potencia:

$$ps(af) = C(af)^{-\beta} = a^{-\beta} (C(f)^{-\beta}) = a^{-\beta} ps(f) \quad (5.15)$$

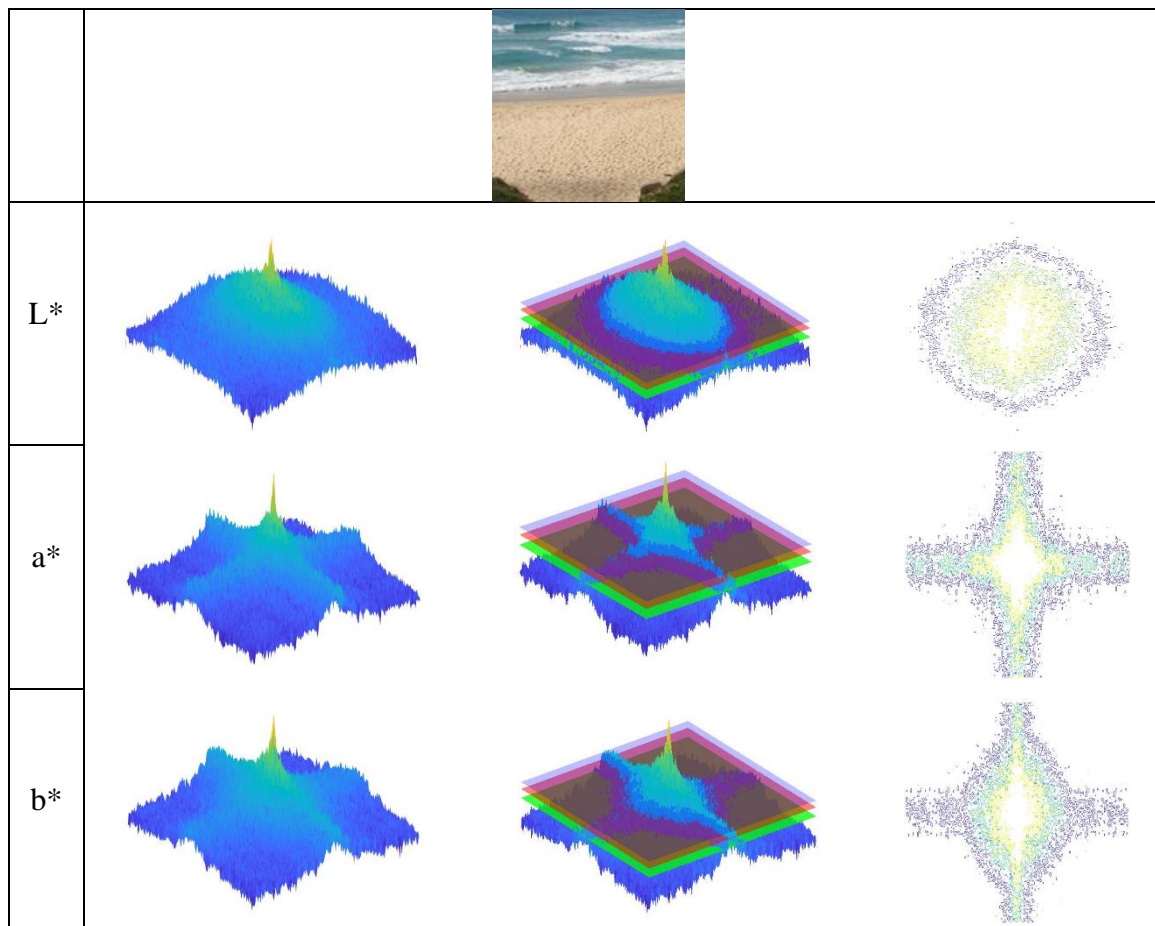
lo que implica que el espectro de potencia no varía con la escala de la frecuencia espacial (o la escala angular), es decir, con el ángulo del campo de visión, con la aparición de los objetos en las imágenes a cualquier escala angular o la distancia focal de la lente del sistema óptico de captura [Rud97,Gei08]. La invariancia a la escala es una propiedad característica de los fractales que fácilmente puede apreciarse en imágenes de follaje, bosques o paisajes. La posible desviación de este comportamiento de la potencia espectral que pueda observarse a altas frecuencias es atribuible al ruido que puedan presentar las imágenes en ese rango.

Volviendo al teorema de Wiener-Khitchine, que establece que el espectro de potencia y la función de autocorrelación forman un par transformado de Fourier, la ley potencial que muestra la potencia espectral (y por tanto la amplitud espectral con exponente la raíz cuadrada del de la potencia espectral) puede interpretarse en el dominio espacial como que la correlación entre píxeles será más fuerte para distancias entre píxeles menores.

Aunque la firma espectral de un conjunto de imágenes aparece definida como la media del espectro de amplitud para todas las imágenes del conjunto [Tor02], el término pronto se asoció al gráfico de isolíneas que se obtiene al cortar el espectro de amplitud o de potencia con distintos planos horizontales de manera que los puntos encerrados por cada isolínea contengan un determinado porcentaje de la energía total de la amplitud o el espectro [Tor03]. Si la energía total de cada plano de la imagen se obtiene sumando la respuesta espectral de todos sus píxeles, para determinar el conjunto de puntos que contiene un porcentaje dado de la energía total ordenamos los píxeles decrecientemente



según su valor espectral y sumamos el valor de cada uno con la suma de los valores de los puntos anteriores hasta que sea igual o superior al porcentaje de energía requerido. Finalmente, el valor espectral del último punto nos da la altura del plano horizontal cuya intersección con la amplitud o potencia nos proporciona cada una de las isolíneas de la firma espectral.



**Figura 5.7:** Espectros de potencia de la imagen superior para los tres canales  $L^*$ ,  $a^*$  y  $b^*$  (primera columna); intersección de planos horizontales con el espectro de potencia que nos proporcionan las regiones que encierran el 60% (plano azul), 80% (plano rojo) y 90% (plano verde) de la energía total (segunda columna) y las firmas espectrales correspondientes (60% verde claro, 80% verde oscuro y 90% azul) (tercera columna)

La figura 5.7 nos muestra en la primera columna, el logaritmo neperiano (para mejorar la representación gráfica) del espectro de potencia para los tres canales  $L^*$ ,  $a^*$  y  $b^*$  de una imagen de la categoría “playa”, en la segunda columna cada uno de los espectros es intersecado por tres planos horizontales de manera que la parte de la señal que queda por encima de ellos engloba un 60, un 80 y un 90 % de la energía total de la señal (planos azul, rojo y verde respectivamente). Finalmente, la última columna

representa la firma espectral correspondiente a los tres canales para los niveles de energía descritos (60 % verde claro, 80% verde oscuro y 90% azul).

Las estructuras geométricas que se obtienen al mantener ciertas cantidades de energía nos proporcionan información global, no localizada sobre las orientaciones dominantes y las distintas escalas espaciales que componen la imagen. Las componentes de la firma espectral en altas frecuencias indican cambios significativos y detalles de la textura de la imagen; por otro lado, las componentes en bajas frecuencias representan la estructura principal de la imagen [Fan13].

La figura 5.8 muestra la firma espectral de nuestras clases semánticas así como de los dos grandes grupos en que son clasificadas, para las tres componentes  $L^*$ ,  $a^*$  y  $b^*$ , correspondientes a niveles de energía del 60% (verde claro), del 80% (verde oscuro) y 90% (azul). Los distintos patrones que nos aparecen en la firma espectral se utilizan ampliamente en tareas de agrupamiento semántico de las imágenes [Tor02, Tor03]. De hecho, se consideran 5 familias distintas de firma espectral caracterizadas por la forma de sus orientaciones dominantes [Oli02] según:

1.- Forma horizontal. Muestra una línea dominante en el eje de frecuencia horizontal desde las bajas a las altas frecuencias (En la figura 5.8 no aparece ningún ejemplo claro de esta familia, las más parecidas serían las firmas espectrales de “calles”).

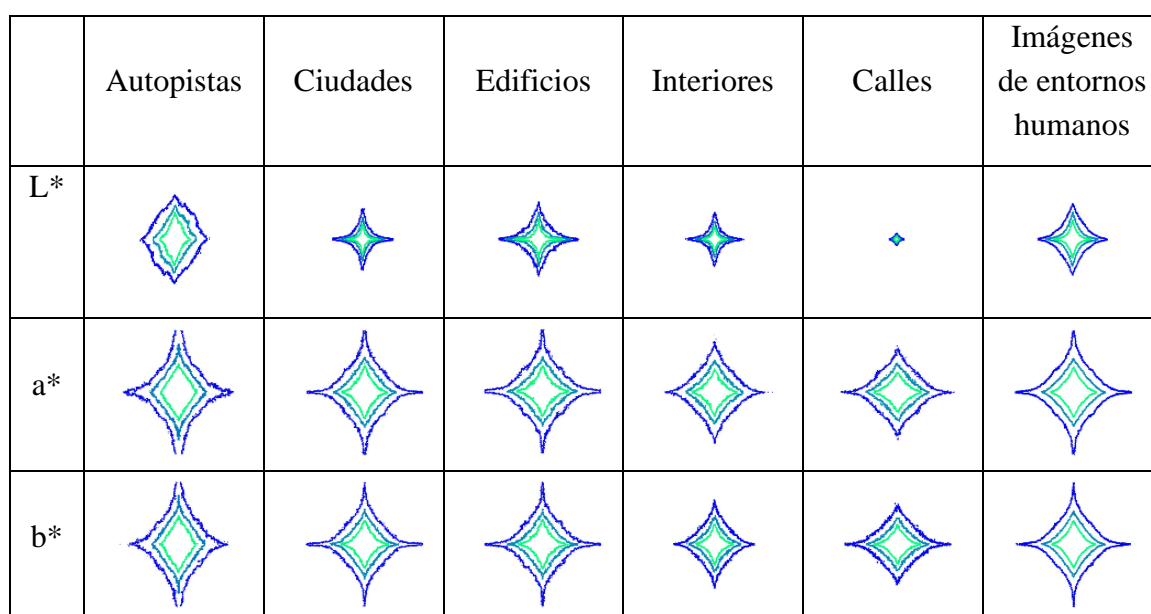
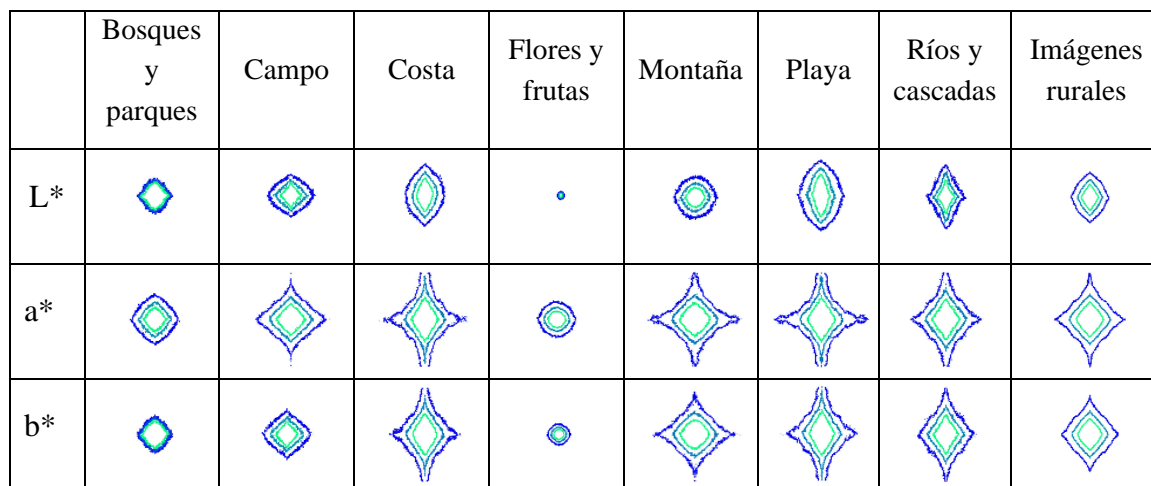
2.- Forma de cruz: Orientaciones horizontales y verticales se representan aproximadamente iguales (las de las tres componentes de “ciudades”, “edificios”, “interiores” y de las “imágenes de entornos humanos” en general).

3.-Forma vertical: muestra una línea dominante vertical revelando que la escena tiene una estructura horizontal (las de las componentes cromáticas de “costa”, “playa” y “autopistas”).

4.- Forma oblicua: Son en las que las orientaciones a 45° (más o menos 15 grados) dominan el espectro (las de las componentes luminancia de “rios y cascadas” y “autopistas”).

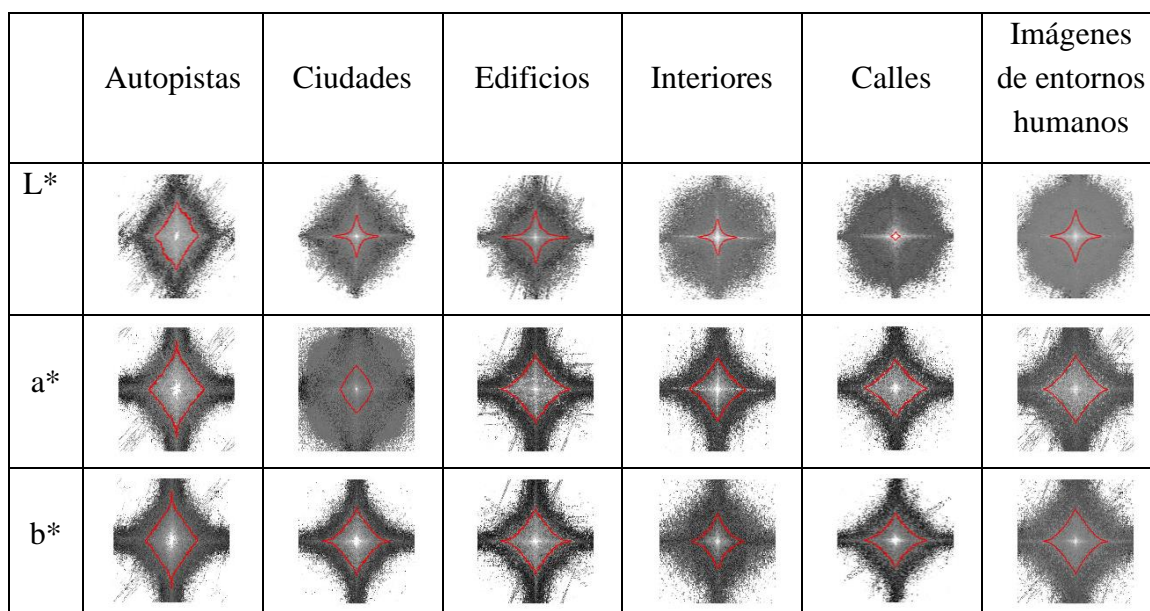
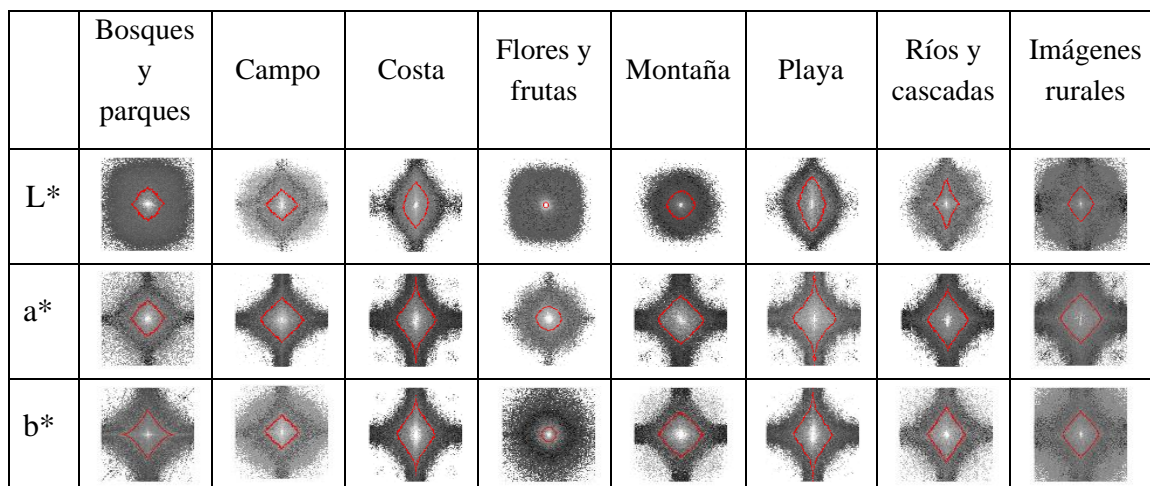
5.- Forma circular: Todas las orientaciones están representadas igualmente en la imagen dando un espectro de potencia isotrópico. (entornos altamente texturizados como “bosques” y sobre todo “flores y frutas”).

De la figura 5.8 podemos apreciar que los patrones de las firmas espectrales de las imágenes de entornos humanos son muy parecidos por lo que es difícil utilizarlos para clasificarlas semánticamente. Un comportamiento parecido encontramos en las firmas de los canales cromáticos de las imágenes rurales. Es el patrón del canal luminancia de las imágenes rurales el único que presenta diferencias significativas para utilizarlo en la clasificación de este tipo de imágenes. No obstante, el contorno de una clase de imágenes no es representativa para imágenes individuales de la clase.



**Figura 5.8:** Firma espectral de los planos L\*, a\* y b\* de las distintas clases semánticas correspondientes al 60% (verde claro), 80% (verde oscuro) y 90% (azul).

Para intentar mejorar la representación de las características globales del espectro de potencia así como para incluir la variabilidad que las imágenes individuales de cada clase introducirían en la misma, en vez de una firma espectral, algunos autores [Fan13] han propuesto un mapa de distribución que consiste en la representación de los contornos de todas las imágenes de la clase con un determinado grado de transparencia en un dominio compartido. La figura 5.9 muestra los mapas de distribución correspondientes a las firmas espectrales de la figura 5.8 junto con el contorno correspondiente al 80% de la energía (línea roja). Esta nueva representación proporciona más información sobre las orientaciones menos predominantes (distintas a la horizontal y vertical) y a altas frecuencias por lo que su utilización puede mejorar la clasificación semántica de las imágenes.



*Figura 5.9: Mapas de distribución del logaritmo neperiano del espectro de potencia de los planos  $L^*$ ,  $a^*$  y  $b^*$  de las distintas clases semánticas.*

### 5.3 Análisis de componentes principales (PCA).

No solo obtenemos importantes resultados del estudio de la estadística de segundo orden en el dominio de las frecuencias, sino que también pueden adquirirse directamente a partir de las covarianzas de los valores de los tres planos de color de los píxeles de nuestras imágenes.

El análisis de componentes principales (PCA) es una técnica estadística que consiste en una transformación lineal del sistema de coordenadas de manera que los nuevos ejes ortogonales se orienten en aquellas direcciones que capturen, decorrelacionándolos, la mayor cantidad de varianza de los datos.

Consideremos una transformación lineal de nuestros datos  $x$

$$s = w^T \cdot x \quad (5.16)$$

el problema básico del PCA puede plantearse como la obtención de aquella característica que maximiza la varianza de nuestros datos transformados

$$\begin{aligned} \max_{w: \|w\| = 1} \langle s s^T \rangle &= \max_{w: \|w\| = 1} \langle w^T x x^T w \rangle = \max_{w: \|w\| = 1} w^T \langle x x^T \rangle w \end{aligned} \quad (5.17)$$

la restricción a la norma de los pesos de la transformación lineal es necesaria para obtener resultados finitos (en caso contrario el máximo de la varianza se alcanzaría para  $w$  infinitamente grande), aunque la elección de la norma unidad es por simplicidad se obtendrían resultados iguales con cualquier otro valor [Hyv09]

Por ser la matriz de covarianza de nuestros datos originales  $\langle x x^T \rangle$  una matriz simétrica de valores reales, será también ortogonalmente diagonalizable, es decir, diagonalizable mediante una matriz  $E$  ortogonal ( $E^{-1} = E^T$ ), cuyas columnas son los vectores propios (ortonormales) de la matriz covarianza y otra matriz  $D$  diagonal formada por los valores propios  $(\lambda_1, \dots, \lambda_n)$  correspondientes [Ori17].

$$\langle x x^T \rangle = E D E^T \quad (5.19)$$

Por tanto

$$\langle s s^T \rangle = w^T \langle x x^T \rangle w = w^T E D E^T w \quad (5.20)$$

que con el cambio de variable  $v = E^T w$  podemos escribir como:

$$\langle s s^T \rangle = v^T D v = \sum_i v_i^2 \lambda_i \quad (5.21)$$

Además, por ser ortogonal la matriz  $E$ , si  $\|w\| = 1$  entonces  $\|v\| = 1$ , haciendo  $v_i^2 = m_i$  la restricción de que  $\|w\| = 1$  se reduce a que la suma de los valores de  $m_i$  (que son positivos por ser cuadrados) sea la unidad. De este modo el problema de maximizar la varianza de los datos transformados se reduce a:

$$\begin{aligned} \max_{w: \|w\| = 1} \langle s s^T \rangle &= \max_{m_i \geq 0, \sum_i m_i = 1} \sum_i m_i \lambda_i \end{aligned} \quad (5.22)$$

La solución se obtendrá cuando la  $m_i$  correspondiente al mayor  $\lambda_i$  es uno y los otros cero, es decir, la característica buscada corresponde con el vector propio (columna de  $E$ ) cuyo valor propio  $\lambda_i$  es máximo.

Después de estimar la primera componente principal, las siguientes se obtienen encontrando la característica de máxima varianza bajo la restricción de que la nueva componente debe ser ortogonal a todas las anteriores (es decir el producto escalar es 0). Dado que la matriz de vectores propios de una matriz simétrica es ortogonal, tenemos exactamente el mismo problema que para la primera componente en términos de  $m_i$ , salvo por la restricción añadida de que los  $m_i$  correspondientes a las componentes anteriores deben ser nulas; de nuevo, el óptimo en cada paso se obtiene cuando la componente



principal es el vector propio correspondiente al valor propio mayor restante. Este procedimiento puede repetirse para obtener tantas componentes como las dimensiones que hay en el espacio de datos [Hyv09].

De este modo, todas las componentes principales pueden encontrarse ordenando los vectores propios de la matriz covarianza de nuestros datos (las columnas de la matriz E) de modo que los correspondientes valores propios estén ordenados decrecientemente.

Consideremos que la matriz E está ordenada de esta manera, entonces podemos expresar  $s=W^T x$  como  $s=E^T x$ . Dada la ortogonalidad de la matriz E, la matriz covarianza de las componentes principales queda reducida a la matriz diagonal de los autovalores y por tanto estas no están correlacionadas.

$$\langle ss^T \rangle = \langle E^T x x^T E \rangle = E^T \langle x x^T \rangle E = E^T (E D E^T) E = (E^T E) D (E^T E) = D \quad (5.23)$$

Varios resultados importantes del comportamiento de la visión en color se obtienen al decorrelacionar la distribución de los valores de los píxeles correspondientes a nuestras imágenes. Para contrastar nuestros resultados con los mostrados en la literatura que hay sobre el tema [Fie94, Rud98, Wac01] no solo trabajamos en el espacio de color CIELAB, sino que también utilizaremos los espacios cromáticos RGB y LMS o de sensibilidad espectral de los conos. Para obtener los valores LMS utilizamos la matriz de transformación [Dos11]

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 17,8824 & 43,5161 & 4,1194 \\ 3,4557 & 27,1554 & 3,8671 \\ 0,0300 & 0,1843 & 1,4671 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (5.24)$$

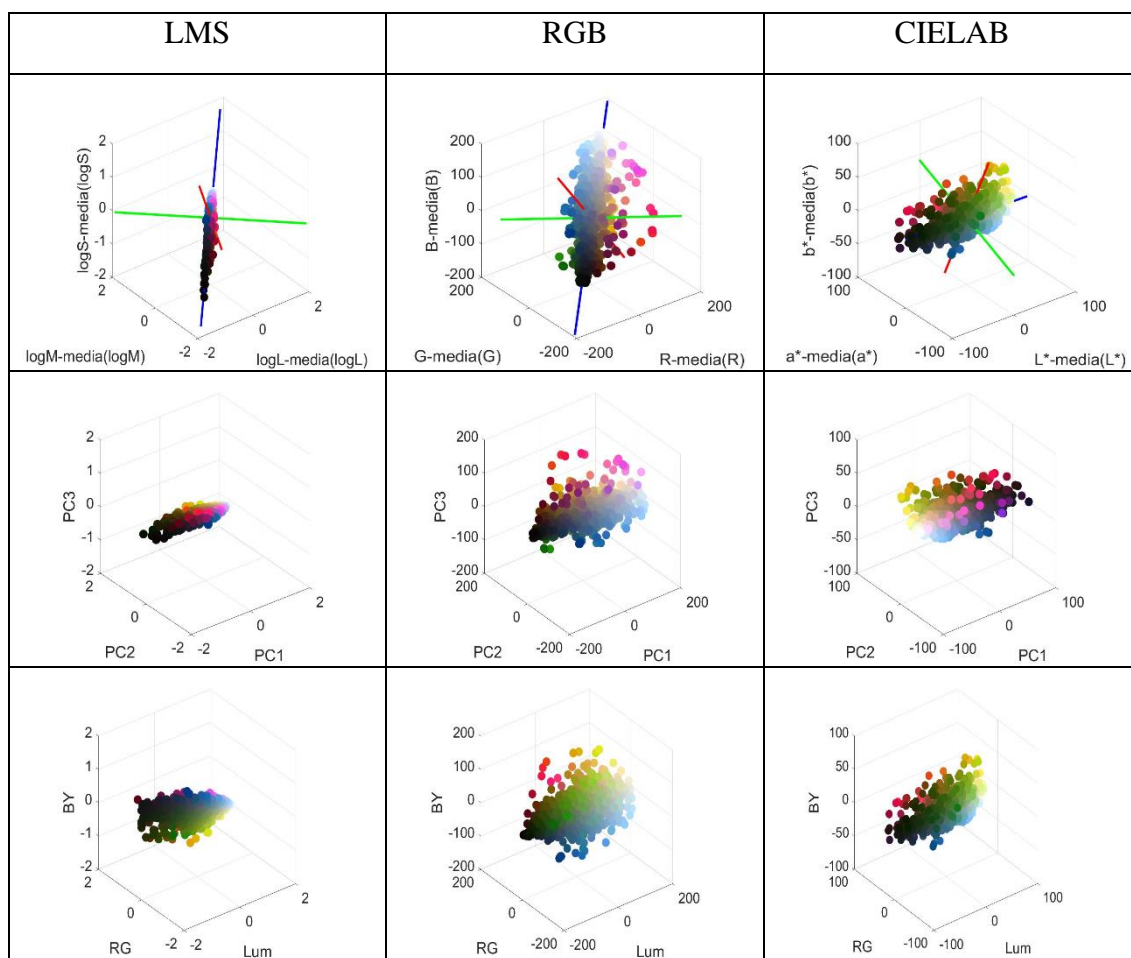
En todo el análisis que sigue en esta sección, para cada componente, al valor de cada píxel de la imagen se le restó el valor medio de todos los píxeles; de este modo, las nuevas características tienen una media cero y además podemos evaluar sus respuestas independientemente del nivel de iluminación de las mismas. Además, para trabajar en el espacio de respuesta de los conos se utilizó el logaritmo decimal de las componentes L, M y S para mejorar la asimetría de la distribución propia de esta representación [Rud98].

La primera fila de la Figura 5.10 representa las nubes de puntos de los valores de los píxeles de nuestras imágenes en los sistemas de coordenadas originales correspondientes a los tres espacios de color mencionados. En todas ellas podemos apreciar como los datos en estas distribuciones están correlacionados (la variación en una de las variables provoca variación en un mismo sentido en las otras). Las direcciones correspondientes a las tres componentes principales resultantes al aplicar PCA están representadas por las rectas azul, roja y verde respectivamente.

En la segunda fila de la Figura 5.10 al representar las mismas nubes de puntos sobre un sistema de coordenadas en las que se utilizan las componentes principales como nuevos ejes puede apreciarse ya la falta de correlación de estas. Finalmente, en la tercera fila de la Figura 5.10 se ha vuelto a representar el mismo conjunto de datos en un sistema

de coordenadas de un espacio de color oponente [Buc83] cuyos ejes se corresponden con un canal luminancia (Lum) y dos canales color oponente:Rojo-Verde (RG) y Amarillo-Azul (BY). Para definirlos en el espacio de color LMS utilizamos el espacio de color de MacLeod y Boynton [Mac79], Derrington et al. [Der84], Krauskopf et al. [Kra82] y McDermott y Webster [McD12] en el que:

$$\text{Lum} = L + M + S \quad \text{RG} = L - M \quad \text{BY} = S - (L + M) \quad (5.25)$$



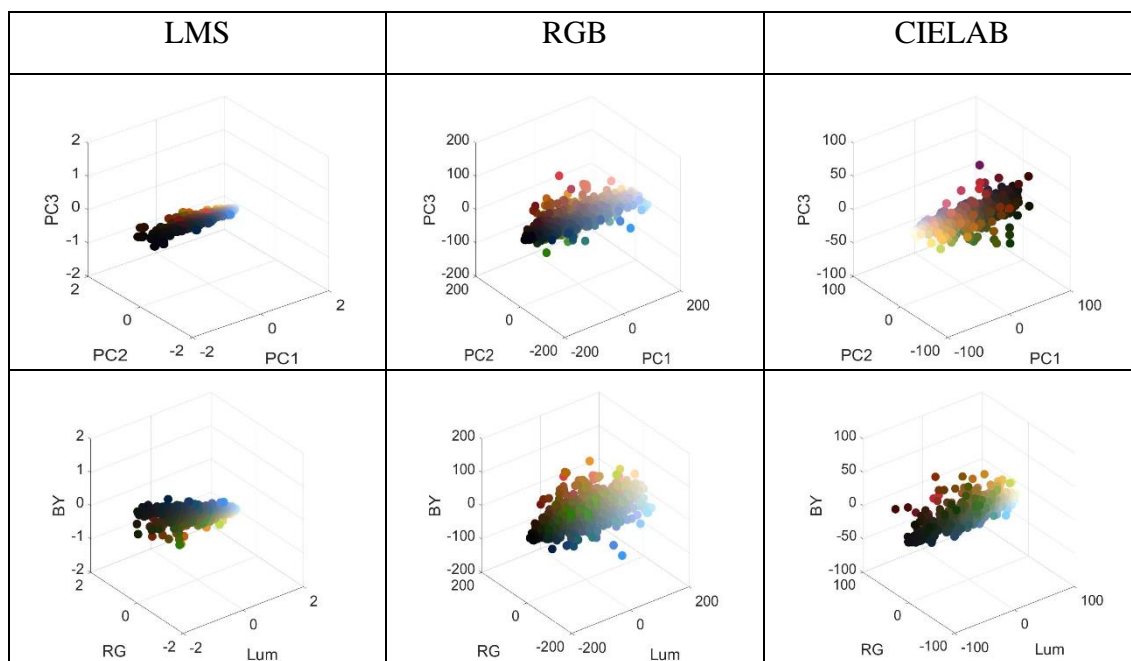
**Figura 5.10:** Representación de la nube de puntos de nuestras imágenes rurales en tres espacios de color LMS, RGB y CIELAB en tres sistemas de referencia diferentes. En la primera fila se representan en sus valores originales (en el caso del espacio LMS está representado en el espacio logarítmico). En la segunda fila en las tres componentes principales obtenidas al aplicar el PCA. En la tercera fila en los tres canales de una representación color oponente (canales luminancia (Lum), oponente rojo-verde (RG) y oponente amarillo-azul (BY)).

En el caso del espacio RGB las componentes Lum, RG y BY vienen dadas por [Gao13]:

$$\text{Lum} = R + G + B \quad \text{RG} = R - G \quad \text{BY} = R + G - 2B \quad (5.26)$$

y, por último, en el espacio Lab no hace falta ninguna transformación ya que  $L^*$  corresponde al canal Lum,  $a^*$  con el canal RG y  $b^*$  con el canal BY.

De las dos últimas filas de la figura 5.10 comprobamos que al reducir la redundancia de nuestros datos mediante la decorrelación ortogonal de los mismos a través de la aplicación del PCA obtenemos tres ejes principales próximos a las funciones cromáticas oponentes, con los canales luminancia (Lum) de mayor varianza, color oponente amarillo-azul (BY) y color oponente rojo-verde (RG) de menor varianza encontrados en el sistema visual de los primates. [Rud98, Kra82, Fla90, DeM75a, DeM75b, Rei92, Her64 y Jam55]. Los resultados obtenidos dependen solamente de la estadística de las imágenes rurales y son semejantes para los tres espacios de color. Al repetir el análisis para las imágenes de entornos humanos las componentes principales obtenidas se corresponden de nuevo con los canales Lum, RG y BY como se ve en la figura 5.11.

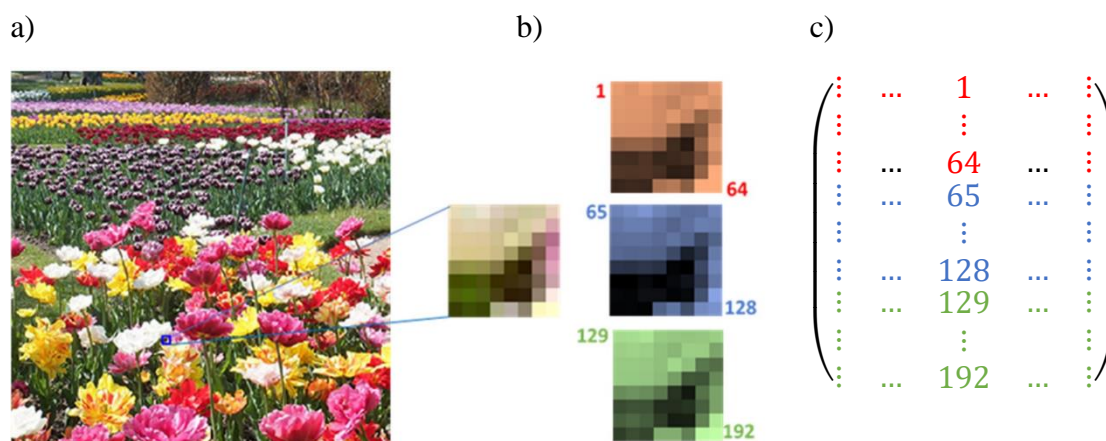


**Figura 5.11:** Representación de la nube de puntos de nuestras imágenes de entornos humanos en tres espacios de color logaritmo decimal de LMS, RGB y CIELAB. En la primera fila se representan en las tres componentes principales obtenidas al aplicar el PCA. En la segunda fila en los tres canales de una representación color oponente (canales luminancia (Lum), oponente rojo-verde (RG) y oponente amarillo-azul (BY)).

Para extender el análisis de nuestras imágenes a la estructura espacio-cromática usamos trozos cuadrados de orden 8 píxeles. Se seleccionaron aleatoriamente 50 de estos trozos en cada una de las imágenes de cada categoría semántica. De este modo, se obtuvieron 17500 trozos tanto de imágenes rurales como de imágenes de entornos humanos. Los



datos de los píxeles de cada trozo fueron ordenados como un vector columna de dimensión  $192 \times 1$  donde a los valores correspondientes a la primera dimensión del espacio de color les seguían los de la segunda y finalmente los valores de la tercera (Ver figura 5.12) [Tai00]. Es a la matriz de dimensiones  $192 \times 17500$  cuyas columnas son los datos de los 17500 trozos de las imágenes rurales (o de los entornos humanos) a la que le aplicamos el PCA [Rud98, Wac01].

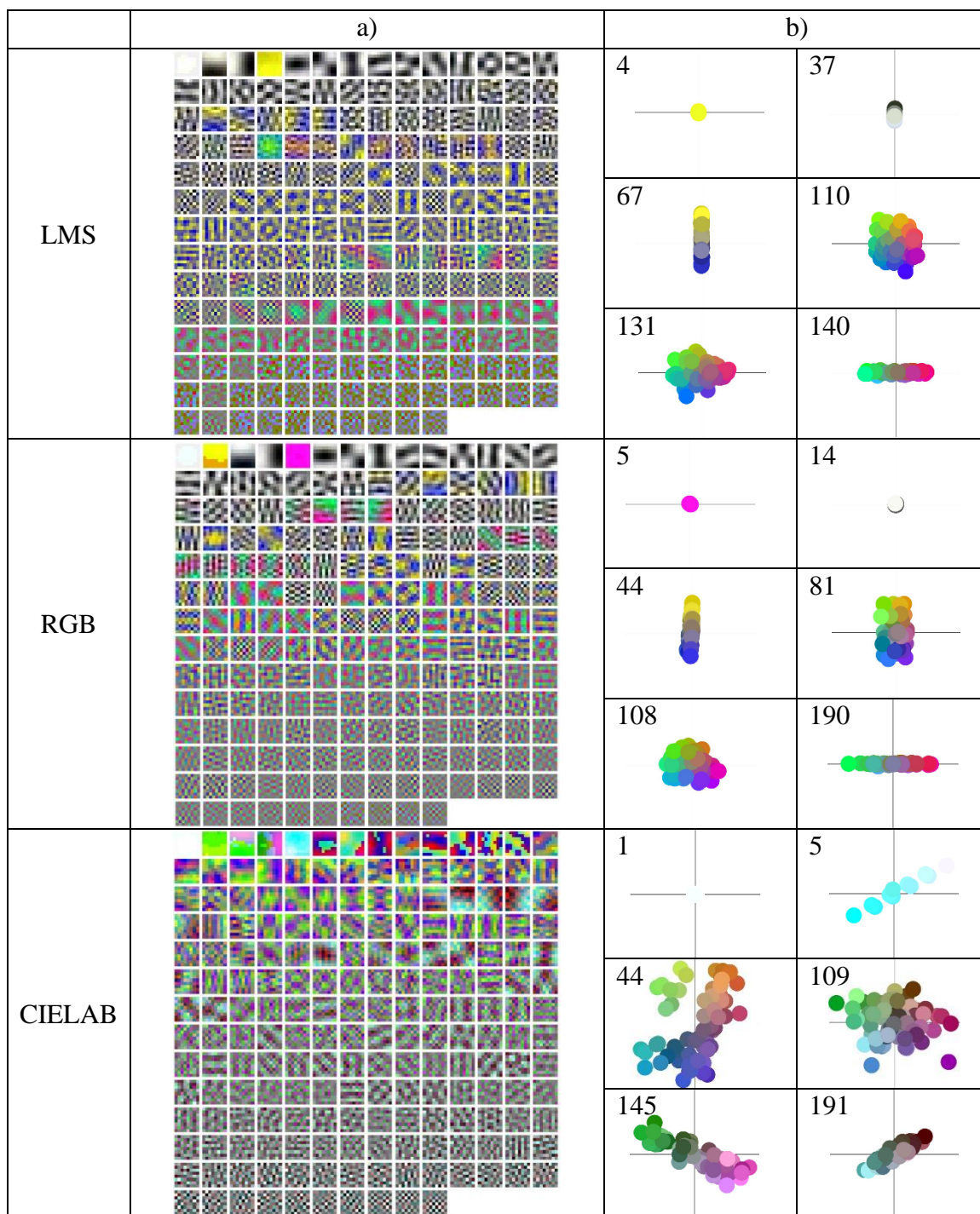


**Figura 5.12:** Ejemplo de un trozo cuadrado de una de las imágenes de trabajo de dimensiones  $8 \times 8$  (a). Representación de los valores de cada píxel del trozo seleccionado en los tres planos cromáticos de definición (b). Ordenación de los datos de cada trozo como un vector columna de dimensión  $192 \times 1$  (c) de la matriz de datos a la que le aplicamos el PCA.

Es imposible representar los datos o los ejes en este espacio de dimensión 192; no obstante, como cada detector de características es un objeto del mismo tamaño y forma que el trozo de imagen, puede representarse como un trozo de imagen en sí [Hyv09]. De este modo, cada componente principal aparece como un nuevo trozo de imagen  $8 \times 8 \times 3$  codificado cromáticamente, los valores R ,G y B de cada píxel se determinan de las coordenadas cromáticas originales normalizadas mediante la transformación lineal

$$CCN_i = 128 (CCO_i + 1) \quad (5.27)$$

siendo  $CCN_i$  la nueva coordenada cromática  $i$ -ésima en el espacio RGB y  $CCO_i$  la coordenada cromática  $i$ -ésima en el espacio de color original normalizada. Con esta codificación cromática los niveles de grises muestran las variaciones en el canal luminancia o acromático, la coloración amarilla azul las variaciones en el canal BY y finalmente la roja verde las del canal RG.



**Figura 5.13:** (a) 192 componentes principales obtenidas de nuestras imágenes rurales en los espacios de color log LMS, RGB y CIELAB dispuestas de izquierda a derecha y de arriba abajo según el orden decreciente de sus autovalores. (b) Cromaticidades de los valores de los 192 píxeles de algunas componentes en un plano isoiluminante. El eje horizontal corresponde a variaciones en el canal RG mientras que el eje vertical corresponde con cambios en el canal BY. La luminancia puede deducirse de la luminosidad de los puntos.

La Figura 5.13 (a) muestra las 192 componentes principales obtenidas de los 17500 trozos de dimensiones 8 x 8 de nuestras imágenes rurales en los espacios de color log LMS, RGB y CIELAB siguiendo la codificación cromática descrita, dispuesta de izquierda a derecha y de arriba abajo según el orden decreciente de sus valores propios (varianza). En la Figura 5.13 (b) se representan las cromaticidades de los valores de los 192 píxeles de algunas de las componentes encontradas en un plano isoiluminante correspondiendo el eje horizontal a variaciones en R - G y el eje vertical corresponde con cambios en R + G - 2B. La luminancia puede deducirse de la luminosidad de los puntos. De este modo en las componentes principales que presentan variaciones en el canal acromático (luminancia) o una cromaticidad homogénea los puntos aparecen superpuestos.

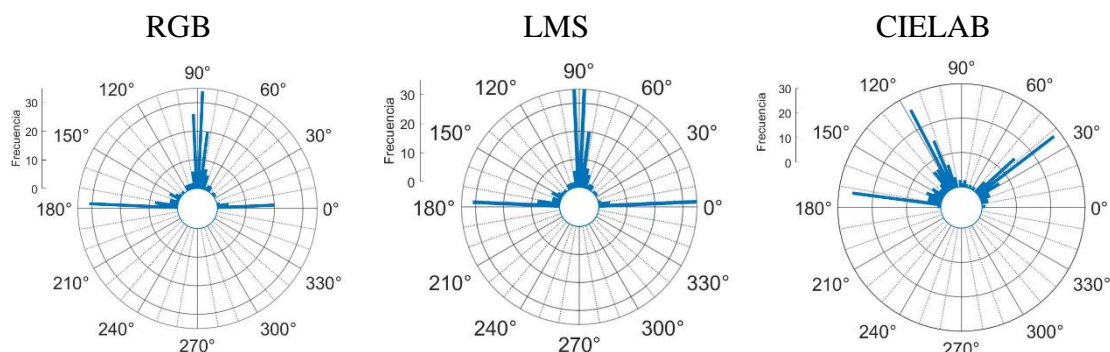
Las componentes principales obtenidas pueden clasificarse en tres categorías principales: componentes homogéneamente cromáticas, acromáticas y color oponentes [Kel13]. Las primeras tienen mayor varianza, ninguna estructura espacial y son altamente selectivas a un color (por ejemplo, la componente 4 del espacio log LMS, la 5 del espacio RGB o la 1 del espacio CIELAB). Las componentes acromáticas codifican las variaciones de luminancia (37 de log LMS y 14 de RGB) y presentan mayor varianza que las cromáticas no homogéneas que son color oponentes. Las cromaticidades de los píxeles de estas últimas al representarse en un espacio de color suelen estar agrupadas a lo largo de una recta y cruzan cuadrantes opuestos (componentes 67 y 140 de log LMS, 44 y 190 de RGB y 145 y 191 de CIELAB) aunque, en este último caso estas no coinciden con la dirección de los ejes. Dentro de estas últimas las varianzas decrecen de las características oponentes BY a las oponentes RG.

Para cuantificar las direcciones de alineación de los píxeles de las componentes principales de los trozos individuales, ejecutamos nuevamente el PCA en las coordenadas de todos los píxeles de cada componente principal y utilizamos los vectores propios para estimar las mismas. [Kel13]

$$\alpha = \text{atan} \frac{PC1_2}{PC1_1} \quad (5.28)$$

donde  $\alpha$  es el ángulo que forma la primera componente principal con el semieje positivo horizontal y  $PC1_1$  y  $PC1_2$  son las coordenadas de la primera componente principal obtenida al aplicar PCA a los valores cromáticos de los distintos trozos considerándolos en un plano isoiluminante. La figura 5.14 muestra el histograma de los ángulos obtenidos para las 192 componentes principales obtenidas de los trozos de imágenes rurales en los tres espacios de color. Vemos que la mayoría de ellos están próximos a una alineación horizontal  $0^\circ$  o vertical  $90^\circ$  ( $30^\circ$  y  $120^\circ$  en el espacio de color CIELAB), correspondiendo estas a los mecanismos de oposición de color en los distintos espacios de color utilizados. No obstante, un pequeño número de componentes (algo mayor en el espacio CIELAB) muestran ángulos diferentes así como colores de sus píxeles distintos de los principales (por ejemplo, componentes 110 y 131 del espacio log LMS, 81 y 108 del espacio RGB

de la figura 5.13 (b)) lo que indica que todavía las direcciones color oponentes están correlacionadas en algún grado en los datos [Wac01].



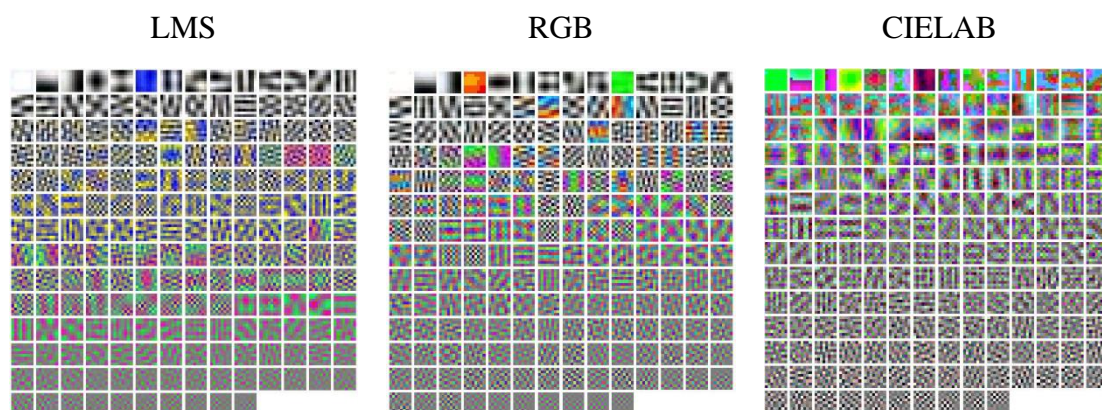
**Figura 5.14:** Histogramas circulares que muestran los ángulos de alineación (entre  $0^\circ$  y  $180^\circ$ ) de los píxeles de las componentes principales obtenidas al aplicar PCA a los trozos de imágenes rurales en cada uno de los espacios de color.

Un análisis de la significación estadística (test de Mann - Whitney) de las diferencias de las medias del ángulo de la primera componente principal de cada trozo para los distintos espacios de color señala que, para un intervalo de confianza del 95% , pueden considerarse iguales los ángulos correspondientes a los espacios RGB y CIELAB tanto para las imágenes rurales como para las de entornos humanos (U de Mann -Whitney 17743 y  $s=0,526$  para las primeras y  $U=17584$  ,  $s=0,436$  para las segundas).

Desde un punto de vista espacial, en la figura 5.13 se ve cómo las componentes principales no están localizadas ni orientadas espacialmente y su varianza disminuye cuando aumenta la frecuencia espacial. Este tipo de componentes de altas frecuencias espaciales y poca amplitud se asocian con el ruido de la señal; no obstante, en el caso del PCA no hay ninguna razón para asociar la señal con frecuencias espaciales bajas y el ruido con altas, por el contrario, muchos bordes nítidos de interés contienen muchas componentes de alta frecuencia y baja amplitud [Bel97].

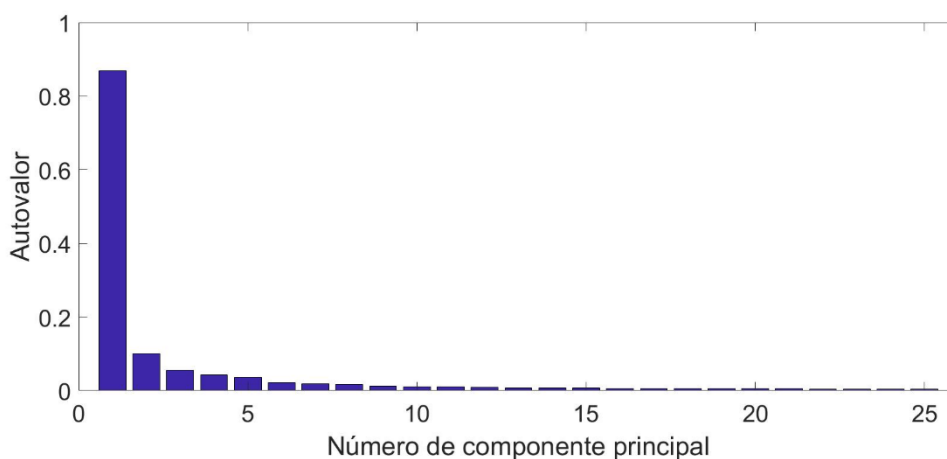
Antes de ver otros aspectos de las componentes principales, calculamos las mismas para otro conjunto del mismo número de trozos de dimensiones  $8 \times 8$  de nuestras imágenes de entornos humanos. La Figura 5.15 muestra las 192 componentes principales obtenidas en los espacios de color log LMS, RGB y CIELAB dispuestas de izquierda a derecha y de arriba abajo según el orden decreciente de sus valores propios. Podemos comprobar nuevamente que espacialmente no están localizadas ni orientadas con frecuencias espaciales crecientes al disminuir la varianza. Cromáticamente hablando volvemos a encontrar las tres categorías descritas aunque las direcciones de oposición de color se desvían de las obtenidas para imágenes rurales (componentes naranja - azul claro frente al amarillo - azul en el espacio RGB).





**Figura 5.15:** 192 componentes principales obtenidas de nuestras imágenes de entornos humanos en los espacios de color log LMS, RGB y CIELAB dispuestas de izquierda a derecha y de arriba abajo según el orden decreciente de sus autovalores.

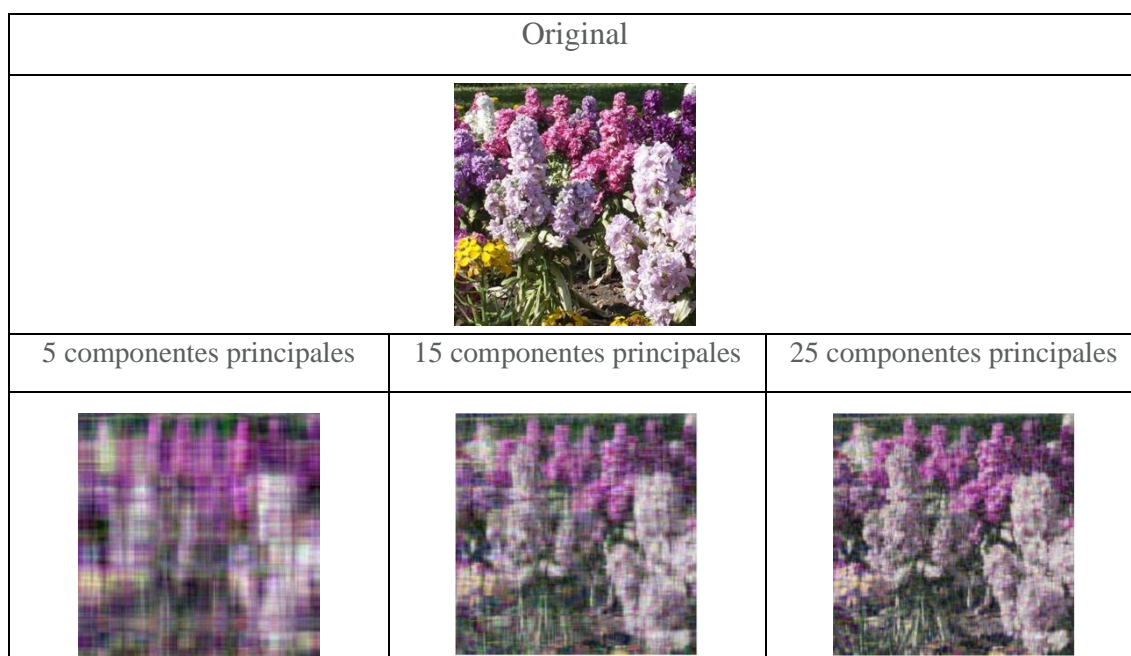
Atendiendo a los valores propios de la descomposición en vectores propios de la matriz de covarianza de los datos de los trozos de imágenes vemos cómo estos decrecen rápidamente de manera similar en los tres espacios de color (la Figura 5.16 muestra los autovalores correspondientes a las componentes principales de los trozos de imágenes rurales en el espacio CIELAB), de modo que, para nuestros datos, las primeras 16 componentes explican más del 90% y las primeras 30 componentes explican más del 95 % de la varianza de los datos.



**Figura 5.16:** Autovalores correspondientes a las componentes principales de los trozos de imágenes rurales en el espacio CIELAB

Esto permite una codificación compacta que reduce la redundancia de la estructura espacio-cromática de las imágenes naturales, consistente en la eliminación de aquellas componentes principales cuya densidad de probabilidad (varianza) es baja, es decir, reduciendo la dimensionalidad de nuestra representación y manteniendo las primeras componentes principales de modo que conservemos un porcentaje dado de la varianza total de nuestros datos con un error RMS mínimo [Gei08, Fie94].

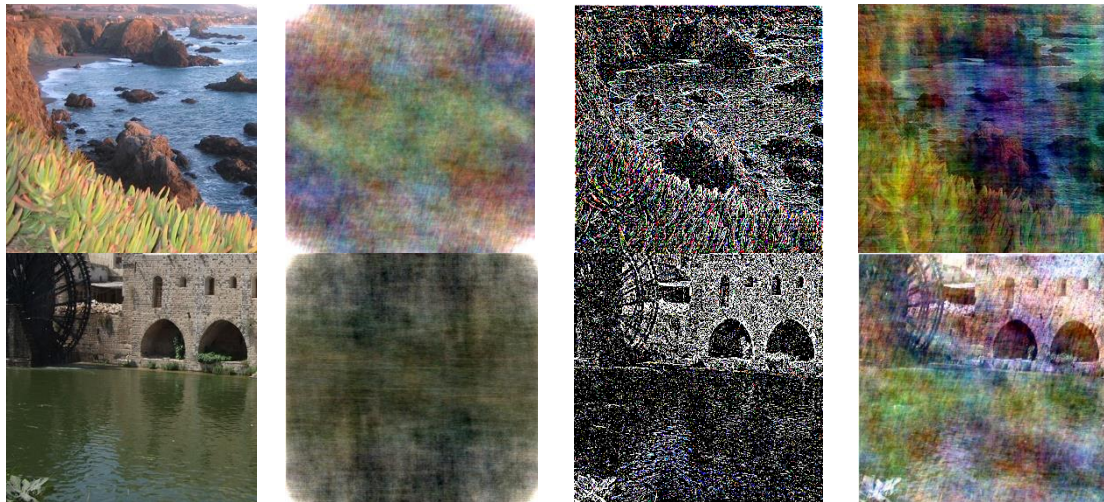
La Figura 5.17 muestra una imagen original, así como tres reconstrucciones de la misma usando distintas reducciones de la dimensionalidad: utilizando las 5, 15 y 25 primeras componentes principales obtenidas de los trozos de todas las imágenes de esa supracategoría semántica a las que les corresponden respectivamente un 78'6, un 90'8 y un 95 del porcentaje total de la varianza de los datos.



**Figura 5.17:** Una imagen natural y tres reconstrucciones de la misma usando las 5 primeras componentes principales (78,6% del total de la varianza de los datos), las 15 primeras (90,8% de la varianza total) y finalmente las 25 primeras (95% de la varianza total) de los trozos de las imágenes de estudio.

La Figura 5.18 muestra un par de ejemplos que ilustran cómo la información de la fase en el dominio de la frecuencia es más significativa para el sistema perceptivo (reconocimiento de formas, detección de bordes...) que la estructura de amplitud en dicho dominio. En la primera columna de dicha figura se representan dos imágenes originales de las que se han utilizado en este trabajo. La segunda columna es una reconstrucción de las imágenes originales que solo tiene en cuenta la información del espectro de amplitud

para lo cual se ha aplicado una transformada discreta de Fourier inversa a una función con la amplitud original y una fase de  $0^\circ$ . Por el contrario, si la transformada discreta de Fourier inversa se aplica a una función de amplitud normalizada y fase la original [Gla15], obtenemos la información que aporta exclusivamente el espectro de fase (columna 3). Finalmente, en la cuarta columna se muestran las imágenes resultantes de intercambiar los espectros de amplitud de ambas, la imagen superior corresponde a una de fase la de la primera imagen y amplitud la de la segunda imagen, mientras que la imagen inferior corresponde con una de fase la de la segunda imagen y amplitud la de la primera. Vemos que en cualquiera de las tres últimas columnas somos capaces de reconocer las imágenes originales solamente en aquellas imágenes que preservan el espectro de fase original.



**Figura 5.18:** (Primera columna) 2 ejemplos de imágenes naturales (Segunda columna) Reconstrucción de la imágenes originales usando sus respectivos espectros de amplitud y un espectro de fase de  $0^\circ$ . (Tercera columna) Reconstrucción de las imágenes originales usando un espectro de amplitud normalizado y su espectro de fase original. (Última columna) Imágenes resultantes de combinar los espectros de fase originales intercambiando los espectros de amplitud.

Podemos considerar que, en el dominio de las frecuencias, los bordes vienen descritos por una correlación local en el espectro de fase, es decir, en la posición del borde las componentes de Fourier tienen la misma fase [Fie89, Mor98].

Usando la descomposición en valores propios, se puede mostrar que las componentes principales de una matriz de covarianza típica de imágenes son sinusoides del tipo:

$$\text{sen}(ax + by + c) \quad (5.29)$$

donde las constantes  $a$  y  $b$  determinan la frecuencia y la orientación de la oscilación y la  $c$  determina la fase [Hyv09].

Este resultado se debe a que la covarianza de imágenes naturales es invariante a la traslación, es decir, depende solo de la distancia entre los píxeles, pudiendo expresarse como una función (par) de las mismas

$$\langle I I^T \rangle = \text{Cov}(I(x,y), I(x',y')) = f((x-x')^2 + (y-y')^2) \quad (5.30)$$

multiplicando la matriz covarianza por un senoide del tipo descrito tendríamos:

$$\langle I I^T \rangle \sin(ax + by + c) = f((x-x')^2 + (y-y')^2) \sin(ax + by + c) \quad (5.31)$$

haciendo el cambio de variable  $\xi = x-x'$  y  $\eta = y-y'$  tenemos

$$\begin{aligned} \langle I I^T \rangle \sin(ax + by + c) &= \sum_{\xi, \eta} f(\xi, \eta) \sin(a(\xi + x') + b(\eta + y') + c) = \\ &f(\xi, \eta) \sin(a\xi + b\eta + ax' + by' + c) \end{aligned} \quad (5.32)$$

y usando la propiedad de que  $\sin(a+b) = \sin a \cos b + \cos a \sin b$ , tenemos

$$\sum_{\xi, \eta} f(\xi, \eta) [\sin(a\xi + b\eta) \cos(ax' + by' + c) + \cos(a\xi + b\eta) \sin(ax' + by' + c)] \quad (5.33)$$

El primer término de la suma es nulo por ser la suma del producto de una función par por el seno que es una función impar, por lo que finalmente nos queda:

$$\langle I I^T \rangle \sin(ax + by + c) = [\sum_{\xi, \eta} f(\xi, \eta) \cos(a\xi + b\eta)] \sin(ax' + by' + c) \quad (5.34)$$

que nos muestra que el senoide es un vector propio de la matriz de covarianza con autovalor  $[\sum_{\xi, \eta} f(\xi, \eta) \cos(a\xi + b\eta)]$ . Este resultado es independiente del valor del parámetro  $c$  (fase) del senoide. Puesto que los pesos de las componentes principales son aproximadamente sinusoidales, la aplicación de los mismos a un trozo de imagen estará relacionado con la transformada de Fourier del trozo de imagen, es decir, con su espectro de amplitud o de potencia. [Hyv09]. Dado que las PCs no son dependientes del espectro de fase (ya que solo depende del espectro de amplitud), estas no son válidas para reflejar la estructura local de los datos [Fie94].

A pesar de las limitaciones que muestran las componentes principales:

- Puede haber componentes con varianzas iguales lo que implica multiplicidad de soluciones para un mismo autovalor [Hyv09].
- Están influidas por fluctuaciones aleatorias, pudiendo confundirse los filtros de frecuencias altas (generalmente asociadas a características opuestas RG) con ruido en la señal.
- Tienen un carácter no localizado espacialmente independiente del espectro de fase presentando solamente orientaciones a baja frecuencia por lo que no sirven como detectores de bordes.

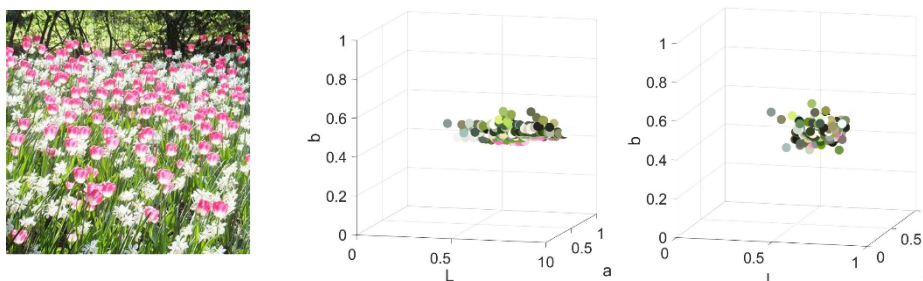


Estas limitaciones les impiden reflejar la estructura espacio-cromática de las imágenes naturales y explicar los campos receptivos de las células corticales, sí proporcionan una codificación compacta color oponente que decorrelaciona los datos y una manera óptima de transmitir la información a través del nervio óptico [Mor98] asemejando las propiedades cromáticas de las neuronas del LGN [Der84, Wac01].

## Capítulo 6

**RESULTADOS SOBRE ESTADÍSTICA DE ORDEN SUPERIOR EN IMÁGENES NATURALES**

El “whitening” o blanqueamiento es la operación que elimina las regularidades estadísticas de segundo orden [Doi05] y se utiliza como paso previo al cálculo de la estadística de orden superior para estudiar las propiedades que no dependen de las covarianzas. El whitening de datos es una transformación de la información de partida en un conjunto de nuevas variables de modo que estén decorrelacionadas y tengan media cero y varianza unidad. Para obtener datos blanqueados a partir del PCA lo único que hay que hacer es normalizar las varianzas de las componentes principales dividiéndolas por su desviación típica [Hyv09]. En la figura 6.1 se representa la nube de puntos original y la nube de puntos de los mismos datos blanqueados usando el procedimiento descrito y que es el paso previo para los análisis que se llevan a cabo en este capítulo.



**Figura 6.1:** Nube de puntos de los valores originales (izquierda) de los píxeles de la imagen en el espacio de color CIELAB y la nube de puntos de los mismos datos blanqueados (derecha).

Ya en la sección estadística de primer orden se definieron los momentos de orden  $k$  como

$$m_k = E\{(x - c)^k\} \quad (6.1)$$

por lo que el primer momento con respecto al origen era la media de la distribución y el segundo momento central se correspondía con la varianza:

$$m_1 = \mu \quad m_2 = \sigma^2 \quad (6.2)$$

Los momentos de tercer y cuarto orden se interpretan a través de los descriptores llamados asimetría y curtosis definidos respectivamente como:

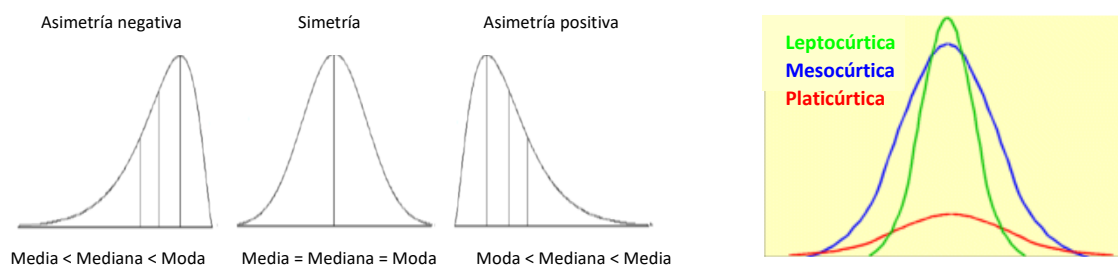
$$S = \frac{E\{(x-c)^3\}}{E\{(x-c)^2\}^{3/2}} \quad K = \frac{E\{(x-c)^4\}}{E\{(x-c)^2\}^2} \quad (6.3)$$

Estas medidas adimensionales valen por definición 0 y 3 respectivamente para distribuciones gaussianas de los datos.

Dado que la estadística de segundo orden sólo utiliza la información en amplitud en el dominio de la frecuencia, la asimetría y la curtosis constituyen los estadísticos globales más simples que pueden capturar la estructura de fase en el dominio espectral de la distribución de nuestros datos. [Rei01, Tho00]. No obstante, los momentos de orden superior, al depender de potencias de mayor exponente de los datos, son descriptores estadísticos poco robustos, es decir, muy sensibles a los valores particulares [Hua99] de nuestra distribución, por lo que se necesitan muestras de tamaño suficientemente grande para producir una estimación robusta [Tho99a]. Sin embargo, la dimensionalidad de nuestras imágenes ( $N=16000$  para cada imagen individual) permite considerar robustas las medidas obtenidas.

Thomson [Tho99b] señaló que para que la curtosis sea significativa en el análisis de imágenes naturales, los datos deberían estar decorrelacionados o blanqueados antes de cualquier cálculo de estadística de orden superior [Rei01]. Por lo que, previamente a nuestro análisis de estos descriptores, hemos blanqueado nuestros datos usando el PCA tal y como se describió en la sección anterior, lo que no solo elimina las dependencias de segundo orden, sino que nos proporciona una distribución de media cero y varianza unidad.

Tal y como su nombre indica, el descriptor asimetría caracteriza el grado de asimetría de una distribución alrededor de su media. También puede interpretarse como un indicador de la diferencia entre la media y la mediana de un conjunto de datos. En las tres distribuciones de datos de la izquierda de la figura 6.2 [Doa11] se muestran ejemplos con las tres situaciones posibles: distribuciones con asimetría negativa (cuyas colas se extienden hacia los valores negativos del conjunto de datos), simétricas y con asimetría positiva (cuyas colas se extienden hacia los valores positivos del conjunto de datos).



**Figura 6.2.-** A la izquierda se representan tres distribuciones con asimetrías negativa, cero (función simétrica) y positiva respectivamente. Las funciones asimétricas presentan una cola más larga en el sentido de la asimetría [Doa11]. A la derecha se presentan tres nuevas distribuciones simétricas con curtosis menor que 3 (distribución platicúrtica), 3 (distribución mesocúrtica) y mayor que 3 (distribución leptocúrtica) respectivamente [Čis10].

Los resultados que obtuvimos al calcular la asimetría en los tres planos  $L^*$ ,  $a^*$  y  $b^*$  de nuestras imágenes de estudio pueden verse en las columnas segunda, tercera y cuarta de la Tabla 6.1. En el canal acromático la asimetría es negativa en todas las categorías (predominio de los píxeles claros sobre los oscuros) tanto de imágenes rurales con  $|S| < 0,69$  como para imágenes de entornos humanos  $|S| < 0,84$ . Este

|                   | $L^*$ | $a^*$ | $b^*$ |                              | $L^*$ | $a^*$ | $b^*$ |
|-------------------|-------|-------|-------|------------------------------|-------|-------|-------|
| Bosques y parques | -0,42 | -0,26 | 0,84  | Autopistas                   | -0,44 | -0,01 | 0,42  |
| Campo             | -0,39 | -0,13 | 0,69  | Ciudades                     | -0,60 | 0,25  | 0,28  |
| Costa             | -0,37 | 0,15  | 0,25  | Edificios                    | -0,64 | 0,42  | 0,40  |
| Flores y frutas   | -0,69 | 0,25  | 0,79  | Interiores                   | -0,84 | 0,69  | 0,87  |
| Montaña           | -0,38 | 0,14  | 0,20  | Calles                       | -0,76 | 0,65  | 0,54  |
| Playa             | -0,45 | 0,21  | 0,38  |                              |       |       |       |
| Ríos y cascadas   | -0,41 | -0,08 | 0,68  |                              |       |       |       |
| Imágenes rurales  | -0,44 | 0,04  | 0,55  | Imágenes de entornos humanos | -0,66 | 0,40  | 0,50  |

**Tabla 6.1 :** Valores de la asimetría ( $S$ ) en los tres canales del espacio cromático CIELAB para las distintas categorías semánticas de nuestro estudio.

resultado, que parece contradecir el de trabajos previos, ya que afirman que la asimetría de la luminancia es positiva en imágenes naturales monocromas [Lau81, Bra00, Att07, Dro01, y Gra16] no lo hace si consideramos el de la versión en escala de grises de nuestras imágenes, en el que  $0,03 < S < 0,27$  (solamente la categoría “playa” muestra una asimetría negativa de -0,09).

En los canales cromáticos observamos que para el canal  $a^*$  las distintas categorías muestran menor asimetría que el canal acromático pero más alternancia en el sentido de las mismas aunque en la mayoría de las categorías es positiva. Finalmente, en el canal  $b^*$  la asimetría positiva en todos los casos (predominio del azul sobre el amarillo) vuelve a ser mayor  $S < 0,87$ . Los canales que presentan mayores valores de asimetría pueden deberse en parte a un mayor rango dinámico de las escenas rurales en los mismos [Gra16]

La curtosis, basada en el cuarto momento normalizado de las distribuciones de datos, es una medida de la falta de “normalidad” de las mismas, centrándose en la extensión (tamaño) de las colas de la distribución en relación con una gaussiana, para la que el descriptor es 3. En la parte derecha de la figura 6.2, se muestran tres ejemplos de distribuciones simétricas [Čis10]: una distribución platícúrtica que presenta una curtosis menor que 3 y una silueta más “aplanada” que una distribución normal; una distribución mesocúrtica (distribución normal) con curtosis 3 y una distribución leptocúrtica con curtosis mayor que 3 y una silueta más “acentuada” (“picuda”) que la distribución gaussiana.

Como ya se comentó en el capítulo 5, las imágenes naturales presentan muchas características locales aperiódicas situadas en posiciones dentro de la imagen que en el dominio de las frecuencias vendrían descritos por correlaciones locales en el espectro de fase [Gla15, Fie89]. Como las componentes de frecuencias espaciales más bajas de una imagen de tamaño finito muestran relativamente pocas oscilaciones por imagen, parece poco probable que contribuyan simultáneamente a la congruencia en fase local de las características locales de la imagen; por tanto, deben ser las frecuencias espaciales más altas las que interactúan para producir estas estructuras [Tho99a] y consiguientemente las imágenes naturales tendrán una distribución leptocúrtica (curtosis mayor que tres). Los resultados que obtuvimos al calcular la curtosis en los tres planos  $L^*$ ,  $a^*$  y  $b^*$  de nuestras imágenes de estudio blanqueadas pueden verse en las tres últimas columnas de la Tabla 6.2. En todas las categorías y todos los canales la curtosis es mayor que tres [Rei01, Tho00, Tho99a, Tho99b] siendo mayor en los canales cromáticos que en el canal luminancia.

En [Fie94] Field describe dos tipos de codificación eficiente a entradas constituidas por imágenes naturales: la codificación compacta y la codificación dispersa. Frente al primero comentado al analizar el PCA, en un código disperso hay una gran población de unidades codificadoras disponibles, de las cuales solo un pequeño subconjunto estará activo como respuesta a un determinado estímulo individual, es decir,

el código tiene una alta dispersión de población. Además, ante estímulos diferentes serán distintos los subconjuntos pequeños de la población de unidades codificadoras que se activarán, por lo que la codificación también es dispersa a través de la población. [Wil01].

| Curtosis (K)      |      |      |      |      |       |       |
|-------------------|------|------|------|------|-------|-------|
|                   | R    | G    | B    | L*   | a*    | b*    |
| Bosques y parques | 3,99 | 3,94 | 4,16 | 4,78 | 7,47  | 6,83  |
| Campo             | 3,66 | 3,74 | 3,72 | 4,68 | 7,27  | 6,48  |
| Costa             | 4,49 | 4,47 | 4,51 | 5,51 | 9,13  | 9,02  |
| Flores y frutas   | 5,34 | 5,36 | 5,36 | 7,47 | 9,18  | 9,20  |
| Montaña           | 4,75 | 4,75 | 4,83 | 5,61 | 9,53  | 8,57  |
| Playa             | 4,58 | 4,57 | 4,57 | 6,06 | 9,84  | 9,64  |
| Ríos y cascadas   | 4,30 | 4,29 | 4,33 | 5,36 | 9,34  | 9,27  |
| Imágenes rurales  | 4,45 | 4,45 | 4,50 | 5,64 | 8,82  | 8,43  |
| Autopistas        | 4,90 | 4,89 | 4,95 | 6,74 | 10,27 | 10,97 |
| Ciudades          | 4,61 | 4,64 | 4,65 | 6,53 | 10,94 | 11,07 |
| Edificios         | 4,81 | 4,83 | 4,81 | 6,93 | 11,94 | 11,83 |
| Interiores        | 6,06 | 6,09 | 6,12 | 9,18 | 15,65 | 13,96 |
| Calles            | 5,31 | 5,33 | 5,31 | 8,02 | 14,35 | 14,31 |
| Entornos humanos  | 5,14 | 5,15 | 5,17 | 7,48 | 12,63 | 12,43 |

**Tabla 6.2:** Valores de la curtosis (K) en los tres canales de los espacios cromáticos RGB (entradas a los fotorreceptores) y CIELAB (salidas de las células ganglionares retinianas) para las distintas categorías semánticas de nuestro estudio.

Se considera que la curtosis es un descriptor para caracterizar una codificación dispersa [Mal86a]. Además de los problemas encontrados en los descriptores basados en momentos superiores ya descritos, en el caso de la curtosis un valor no siempre se corresponde con una distribución dispersa sino que pueden corresponderse con una función de distribución de probabilidad significativamente multimodal o con una amplia distribución de varianzas locales [Tho99a, Bad96]. No obstante, las funciones de distribución de probabilidad de la mayoría de nuestras imágenes eran unimodales lo que nos permite utilizar el descriptor curtosis como una medida de la dispersión.

Un sistema visual cuya finalidad sea la optimización de la dispersión debería maximizar la curtosis a través de las diferentes etapas de su procesamiento [Bad96]. En la Tabla 6.2 se muestra cómo las curtosis de las imágenes de entrada en el espacio RGB es menor que para los mecanismos de luminancia y color oponentes correspondientes a las salidas de las células ganglionares de la retina que podrían describirse por medio de las coordenadas de color CIELAB; además, estas diferencias son más acentuadas en los canales cromáticos.

## 6.1 Análisis de componentes independientes (ICA).

El análisis de componentes independientes (ICA) es un modelo generativo [Jut91, Com94], en el que las imágenes naturales (datos observados) son generadas como transformaciones de algunas variables originales latentes (ya que no pueden ser observadas directamente). Se puede considerar el método más utilizado para resolver el problema de separación ciega de fuentes proporcionando la solución independiente (semi-local) al mismo [Bel97].

Si representamos por  $I(x,y)$  los valores de los píxeles de una imagen, o en la práctica, un pequeño trozo de imagen, ICA lo generará como una superposición lineal de algunas características  $A_i$

$$I(x, y) = \sum_{i=1}^m A_i(x, y) s_i \quad (6.4)$$

para todos los píxeles. Los  $s_i$  son coeficientes que son diferentes de trozo a trozo por lo que pueden considerarse como variables aleatorias (al igual que los trozos de imagen) mientras que las características  $A_i$  son las mismas para todos los trozos.

Usando notación matricial  $\mathbf{x}$  es el vector aleatorio cuyos elementos son los datos observados ( $I(x,y)$  para cada píxel), y  $\mathbf{s}$  el vector aleatorio con elementos  $s_1, \dots, s_n$  (ambos vectores son vectores columna) y sea  $\mathbf{A}$  la matriz con elementos  $a_{ij}$ , entonces el modelo se escribe como

$$\mathbf{x} = \mathbf{A}\mathbf{s}. \quad (6.5)$$

En donde debemos estimar tanto  $\mathbf{A}$  como  $\mathbf{s}$  a partir de los datos observados bajo las condiciones más generales que sean posibles.

Con ICA solo podemos recuperar las componentes  $s_i$  salvo una constante multiplicativa. Normalmente, simplificamos la situación definiendo las componentes para que tengan varianza unidad. Esto solo deja indeterminado el signo de las componentes. Por tanto, para cualquier componente  $s_i$  podríamos considerar también la componente  $-s_i$ . [Hyv09]

En ICA se asumen las siguientes condiciones:

1.- La suposición fundamental es que las  $s_i$  son independientes estadísticamente cuando son consideradas como variables aleatorias. Dos magnitudes son independientes estadísticamente hablando si la información de los valores de una de ellas no da información de los valores de la otra; es el caso de las variables independientes  $s_i$  pero no el de los datos originales. La independencia puede definirse a través de las densidades de probabilidad, si  $p(y_1, y_2)$  es la función de densidad de probabilidad conjunta (pdf) de  $y_1$  y  $y_2$ , y  $p_1(y_1)$  y  $p_2(y_2)$  las pdfs marginales de  $y_1$  e  $y_2$ :

$$p_1(y_1) = \int p(y_1, y_2) dy_2 \quad p_2(y_2) = \int p(y_1, y_2) dy_1 \quad (6.6)$$

entonces diremos que  $y_1$  e  $y_2$  son independientes si y solo si la pdf conjunta se puede factorizar como:

$$p(y_1, y_2) = p_1(y_1)p_2(y_2). \quad (6.7)$$

definición que puede extenderse a  $n$  variables aleatorias como el producto de  $n$  términos.

Una propiedad importante de las variables independientes es que dadas dos funciones  $h_1(y_1)$  y  $h_2(y_2)$  se cumple

$$E\{h_1(y_1) h_2(y_2)\} = E\{h_1(y_1)\} E\{h_2(y_2)\} \quad (6.8)$$

lo que permite demostrar que las variables independientes están decorrelacionadas, aunque la falta de correlación no implica dependencia. [Hyv00]

2.- Por simplicidad también asumiremos que las matrices  $\mathbf{A}$  son cuadradas lo que nos permitirá calcular su inversa y obtener las componentes independientes simplemente por:

$$\mathbf{s} = \mathbf{W}\mathbf{x}. \quad (6.9)$$

donde  $\mathbf{W} = \mathbf{A}^{-1}$

3.- Finalmente, para poder estimar el modelo, asumimos que las distribuciones de los  $s_i$  son no gaussianas. La restricción fundamental en ICA es que las componentes independientes no deben ser gaussianas. Dada la esfericidad de la distribución de probabilidad de los datos blanqueados y sabiendo que la distribución de cualquier transformación ortogonal de una gaussiana  $n$ -variable tiene exactamente la misma distribución que la original, siendo las  $n$  variables independientes, entonces en el caso de variables gaussianas el modelo del ICA solo podría llegar a estimar hasta una



transformación ortogonal de las mismas (no se podría identificar la matriz  $\mathbf{A}$  para componentes independientes gaussianas) [Hyv09,Hyv00].

La justificación de que la optimización de la no gaussianidad nos da las componentes independientes la encontramos en el teorema central del límite. Este resultado clásico en teoría de probabilidad muestra que la distribución de una suma de variables aleatorias independientes tiende, bajo ciertas condiciones, hacia una distribución gaussiana. Por tanto, la suma de dos variables aleatorias independientes normalmente tendrá una distribución más próxima a la gaussiana que cualquiera de las dos variables originales.

Consideremos una combinación lineal  $y$  de los  $x_i$  datos, entonces

$$y = \mathbf{w}^T \mathbf{x} = \sum_i w_i x_i \quad (6.10)$$

donde  $\mathbf{w}$  es un vector a determinar. Si  $\mathbf{w}$  fuera una de las filas de  $\mathbf{A}^{-1}$  entonces  $\mathbf{w}$  sería una de las componentes independientes. Haciendo el cambio de variable  $\mathbf{z} = \mathbf{A}^T \mathbf{w}$  tenemos

$$y = \mathbf{w}^T \mathbf{x} = \mathbf{w}^T \mathbf{A} \mathbf{s} = \mathbf{z}^T \mathbf{s}. \quad (6.11)$$

por tanto  $y$  es una combinación lineal de  $s_i$ , con pesos dados por  $z_i$ .

Por el teorema central del límite la suma de incluso dos variables aleatorias sería más gaussiana que las variables originales y será lo menos gaussiana posible en el caso de que sea igual a solo una de las  $s_i$ ; en este caso, solo uno de los elementos  $z_i$  de  $\mathbf{z}$  es no nulo. Tomando  $\mathbf{w}$  como un vector que maximiza la no gaussianidad de  $\mathbf{w}^T \mathbf{x}$  necesariamente se corresponde con una  $\mathbf{z}$  con una sola componente no nula o, lo que es lo mismo, a una componente independiente. De hecho, la optimización de la no gaussianidad en el espacio  $n$ -dimensional de vectores  $\mathbf{w}$  tiene  $2n$  máximos locales, dos para cada componente independiente, correspondiendo a  $s_i$  y  $-s_i$  (las componentes independientes solo pueden estimarse hasta un signo multiplicativo). Para encontrar estas componentes independientes necesitamos hallar todos estos máximos locales; esto no es difícil porque las diferentes componentes independientes están decorrelacionadas pudiendo restringir la búsqueda al espacio que da estimaciones decorrelacionadas con las ya obtenidas, proceso equivalente a la ortogonalización en un espacio transformado adecuado (blanqueado). [Hyv00]

Para usar la no gaussianidad en las estimaciones del ICA debemos tener una medida cuantitativa de la misma para una variable aleatoria  $y$  (que asumimos tiene media cero y varianza unidad). Una de las medidas que pueden usarse con este fin es la curtosis que, como ya vimos, presenta el problema de la falta de robustez y depende de la muestra medida.

La otra medida que se utiliza es la neguentropía. La entropía de una variable aleatoria (en teoría de la información) puede interpretarse como el grado de información que nos da la observación de la variable. Cuanto más aleatoria, es decir, impredecible y

desestructurada sea la variable, mayor es su entropía. La entropía para una variable discreta aleatoria  $Y$  se define como:

$$H(Y) = -\sum_i P(Y = a_i) \log P(Y = a_i) \quad (6.12)$$

Un resultado fundamental de la teoría de la información es que una variable gaussiana tiene la mayor entropía entre todas las variables aleatorias de varianza igual. [Pap84,Cov91]. Lo que significa que la entropía puede utilizarse para medir la no gaussianidad. De hecho, esto muestra que la distribución gaussiana es la más aleatoria o la menos estructurada de las distribuciones. Para obtener una medida de no gaussianidad, que es cero, para una variable gaussiana y siempre positiva, se usa una versión ligeramente modificada de la definición anterior llamada neguentropía,  $J$ , que se define como sigue:

$$J(\mathbf{y}) = H(\mathbf{y}_{gauss}) - H(\mathbf{y}) \quad (6.13)$$

donde  $\mathbf{y}_{gauss}$  es una variable aleatoria gaussiana de la misma matriz covarianza que  $\mathbf{y}$ . Debido a las propiedades antes mencionadas, la neguentropía es siempre no negativa y cero si y solo si  $\mathbf{y}$  tiene una distribución gaussiana y además es invariante a transformaciones lineales invertibles [Com94, Hyv99].

La dificultad de utilizar la neguentropía radica en su cálculo por lo que se propone una aproximación basada en el principio de entropía máxima [Hyv98], que en el caso de una sola función no cuadrática  $G$  tiene la forma:

$$J(y) \propto [E\{G(y)\} - E\{G(v)\}]^2 \quad (6.14)$$

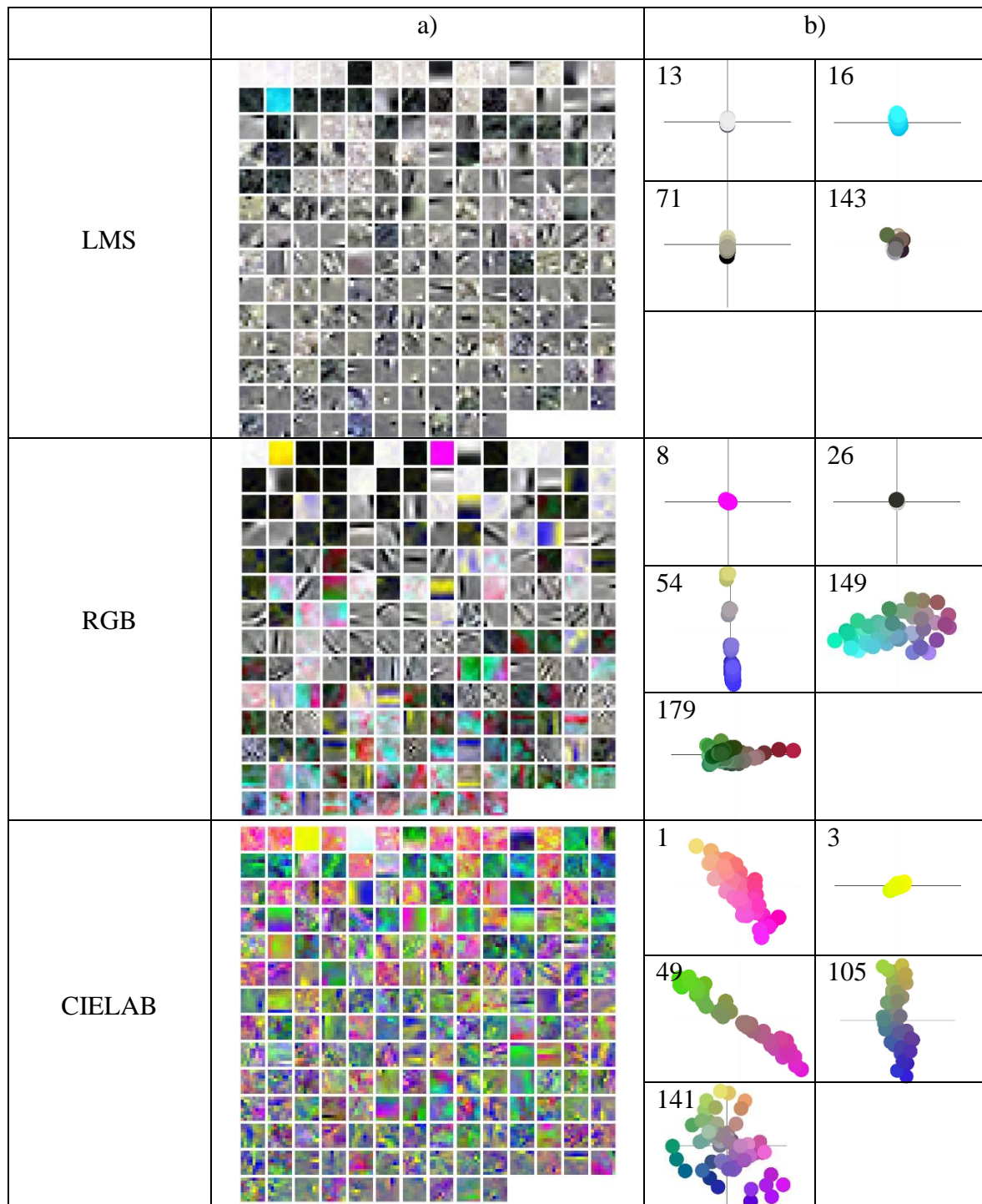
donde la variable  $y$ , y la variable gaussiana  $v$  tienen media cero y varianza unidad. La elección de una función  $G$  que no crezca demasiado rápidamente proporciona estimadores robustos que han mostrado ser muy útiles, en particular.

$$G_1(u) = \frac{1}{a_1} \log \cosh a_1 u \quad G_2(u) = -\exp(-u^2/2) \quad (6.15)$$

para  $1 \leq a_1 \leq 2$  [Hyv00].

Podemos encontrar justificaciones al modelo teórico de aproximación al ICA descrito desde el punto de vista del principio de optimización infomax a través de la estimación de máxima verosimilitud [Pha92] o a través de la maximización de la entropía de salida (o flujo de información) de una red neuronal con salidas no lineales [Bel95, Nad94] o desde el punto de vista de la teoría de la información, a través de la minimización de la información mutua [Hyv00].

Hemos usado una herramienta de Matlab libre llamada FastICA [Hyv04] para el análisis del ICA, que usa una aproximación de la entropía negativa normalizada como criterio para la independencia utilizando la función objetivo  $G_2(u) = -\exp(-u^2/2)$  y que es equivalente a una aproximación infomax de punto fijo o de máxima verosimilitud [Hyv00].



**Figura 6.3:** (a) 192 características (filas de la matriz  $A$ ) obtenidas de aplicar ICA a nuestras imágenes rurales en los espacios de color log LMS, RGB y CIELAB dispuestas de izquierda a derecha y de arriba abajo según el orden decreciente de su norma  $L_2$ . (b) Cromaticidades de los valores de los 192 píxeles de las características indicadas en un plano isoiluminante. El eje horizontal corresponde a variaciones en el canal RG mientras que el eje vertical corresponde con cambios en el canal BY. La luminancia puede deducirse de la luminosidad de los puntos.

Aplicamos el algoritmo FastICA [Hyv04] (que blanquea los datos antes de maximizar la entropía) a la misma matriz de dimensiones  $192 \times 17500$  cuyas columnas son los datos de los 17500 trozos de las imágenes rurales (o de los entornos humanos) que ya utilizamos en el PCA. Dado que no se pueden calcular las varianzas (energía) de las componentes independientes, no podemos ordenarlas como lo hacíamos con las componentes principales [Hyv00].

En la Figura 6.3 (a) se muestran las 192 características (filas de la matriz **A**) obtenidas de los 17500 trozos de dimensiones  $8 \times 8$  de nuestras imágenes rurales en los espacios de color log LMS, RGB y CIELAB siguiendo la misma codificación cromática que en la sección 5.3, dispuesta de izquierda a derecha y de arriba abajo según el orden decreciente de su norma L2 (o norma euclídea). Las funciones correspondientes a los espacios log LMS y CIELAB muestran un aspecto más ruidoso debido a que cualquier no linealidad inherente a la conversión entre el espacio de color RGB, en el que originalmente estaban nuestras imágenes, y los otros dos espacios de color de trabajo (log LMS y CIELAB) afecta al resultado del ICA [Hyv09]. La Figura 6.3 (b) representa las cromaticidades de los valores de los 192 píxeles de algunas de las características encontradas en un plano isoiluminante donde las variaciones  $R - G$  se corresponden con el eje horizontal (rojo - verde) y las variaciones en  $R + G - 2B$  con el vertical (amarillo-azul).

En los tres espacios de color podemos apreciar características localizadas y orientadas espacialmente cuya frecuencia aumenta al disminuir la norma L2. Como con las componentes principales, las características de las componentes independientes pueden clasificarse en tres categorías principales: características homogéneamente cromáticas, acromáticas orientadas (bordes de luminancia) y color oponentes (bordes de colores) [Wac01, Kel13, Lee02].

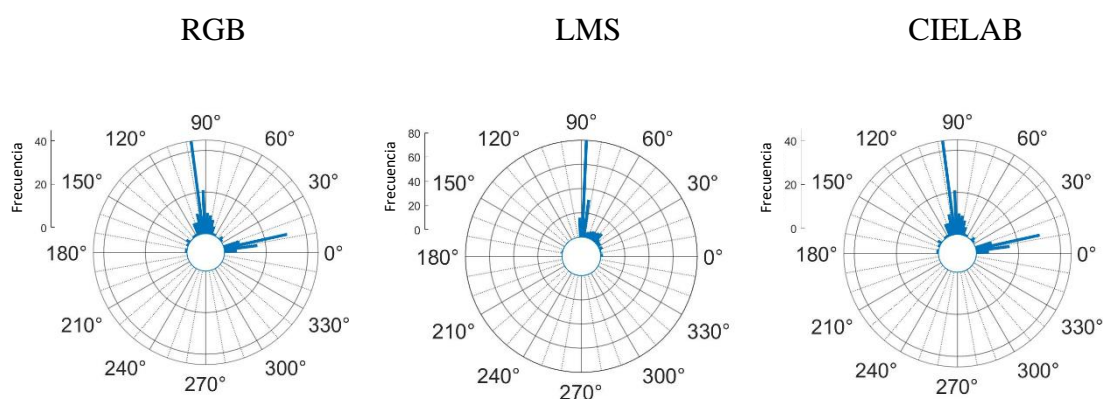
Las características cromáticas homogéneas (por ejemplo, funciones base 16 en log LMS, 8 en RGB y 3 CIELAB) están entre las que tienen una norma L2 mayor, ninguna estructura espacial definida y son altamente selectivas a un color presentando todos sus puntos prácticamente el mismo color (como se puede ver en el color casi uniforme de estas características en la parte izquierda de la figura 6.3 o por la superposición de los puntos que representan las cromaticidades de los píxeles de las mismas en un plano isoiluminante en la parte derecha de la figura 6.3).

Las características acromáticas (por ejemplo, funciones base 13 en log LMS y 26 en RGB) están localizadas y orientadas de manera parecida a las encontradas en el análisis de imágenes rurales de escala de grises [Ols96, Bel97, vHa98, Lew99] representando bordes de luminancia. [Lee02]. Como en [Kel13], se encuentran en el rango medio bajo de norma L2.

La mayoría de las características cromáticas no homogéneas (por ejemplo, características 71 en log LMS, 54 y 179 en RGB; y 1, 49 y 105 en CIELAB) son oponentes

en color, es decir, las cromaticidades de los píxeles cuando se representan en el plano isoiluminante, están alineados y cruzando frecuentemente dos cuadrantes opuestos. Su estructura espacial está localizada y orientada, es decir, codifican bordes cromáticos. No obstante, hay también un número de características cromáticas no homogéneas (por ejemplo, funciones base 143 en log LMS, 149 en RGB y 141 en CIELAB), que muestran una oposición más débil, con sus píxeles más dispersos en el espacio de color.

La figura 6.4 muestra el histograma de los ángulos obtenidos para las 192 características obtenidas al aplicar ICA a los trozos de imágenes rurales en los tres espacios de color. La falta de la restricción de normalidad de las funciones base del ICA hacen que ahora la dirección predominante de estas esté más repartida en todas las direcciones; concentrándose en las direcciones que atraviesan el primer y tercer cuadrante en los espacios de color log LMS y CIELAB (predominando los  $90^\circ$  en el espacio log LMS) y en las direcciones correspondientes a  $10^\circ$  y  $100^\circ$  en el espacio de color RGB. Un análisis de la significación estadística (test de Mann - Whitney) de las diferencias de las medias del ángulo de la primera componente principal de cada trozo para los distintos espacios de color, confirma, para un intervalo de confianza del 95%, que la media de estas direcciones se puede considerar distinta atendiendo al espacio de color tanto para las imágenes rurales como para las de entornos humanos.

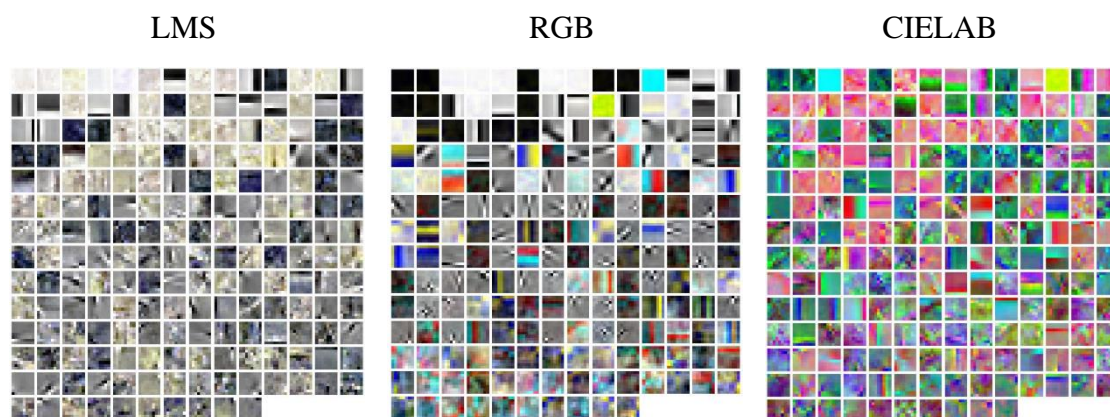


**Figura 6.4:** Histogramas circulares que muestran los ángulos de alineación (entre  $0^\circ$  y  $180^\circ$ ) de los píxeles de las características obtenidas al aplicar ICA a los trozos de imágenes rurales en cada uno de los espacios de color

Como se muestra en la Figura 6.5, obtenemos características muy parecidas al aplicar ICA a un mismo número de trozos de imágenes de dimensión  $8 \times 8$  de entornos humanos que las que obteníamos al aplicarlo a imágenes rurales. Por tanto, a partir de las propiedades estadísticas de orden superior del espectro de nuestras imágenes, ICA encuentra que las características correspondientes a señales máximamente independientes separan la información en luminancia (o detectores de bordes de luminancia) y en



información cromática color oponente (detectores de bordes de color no ortogonales) que tienen distribuciones dispersas, lo que constituye una codificación altamente eficiente. [Tai00, Lee02].



**Figura 6.5:** 192 características (filas de la matriz  $A$ ) obtenidas de aplicar ICA nuestras imágenes de entornos humanos en los espacios de color log LMS, RGB y CIELAB dispuestas de izquierda a derecha y de arriba abajo según el orden decreciente de su norma  $L_2$ .

La estructura localizada y orientada de las características de las funciones base se suele asociar con los campos receptivos de las células simples y complejas de V1 y áreas corticales superiores [Tai00]; además, las distribuciones dispersas de las mismas están relacionadas con el aumento de neuronas y, por tanto, de la dimensionalidad en estas regiones cerebrales [Kel13]. No obstante, ya en [Kel13] se enumera una serie de características de las funciones base que contradice a los resultados obtenidos para los campos receptivos de este tipo de células: no muestran simetría, la oposición rojo – verde no se corresponde con el eje L-M del espacio de color log LMS, sino que incluye contribuciones de la componente S. En el caso de imágenes de color teseladas en las que cada píxel tiene solo una medida cromática frente a las tres habituales (como es en el caso de la retina o en imágenes con mosaico de filtros de color (CFA)), Alleyson y Susstrunk [All04] muestran que los análisis de componentes principales e independientes (PCA e ICA) no son métodos adecuados para estimar la información de luminancia de las escenas naturales si tenemos en cuenta el muestreo que del espacio se realizan en estos sensores y en la retina.

## 6.2 Índice de agrupamiento (Patchy index)

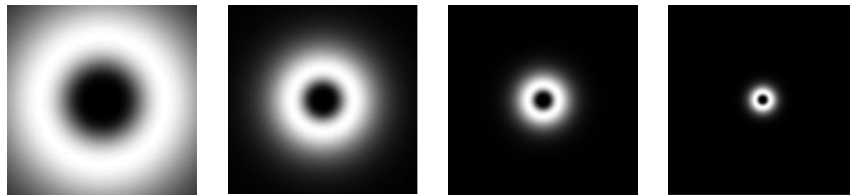
Este término fue introducido en Yoonessi et al [Yoo08] para nombrar un estadístico que describiera la dispersión estructural de la imagen y que, en vez de centrarse

en la proporción de píxeles encerrada por bordes [Han07], lo hiciera en la proporción de la imagen que contiene áreas uniformes. De este modo, una imagen con un área uniforme muy grande o con muchas regiones uniformes presentará un índice de agrupamiento mayor que una imagen con pequeñas y pocas de estas regiones. El índice de agrupamiento se define como la porción de píxeles cuyos valores (para cada uno de los canales de luminancia y crominancia) se encuentran dentro de un umbral, promediados a filtros con diferentes escalas.

En el algoritmo de tres etapas utilizado, primeramente, se filtraron en el dominio de las frecuencias los tres canales  $L^*$ ,  $a^*$  y  $b^*$  de nuestras imágenes en el espacio de color CIELAB con un banco de 4 filtros isotrópicos log Gabor de frecuencias 3, 6, 12 y 24 ciclos por grado y cuya expresión en el dominio espectral viene dada por [Fie87]:

$$\text{LGF}(f, \theta) = \exp \left\{ \left( -\log \left( \frac{R(f_i, \theta_j)}{F} \right)^2 \right) / \left( 2 \log(\sigma/F)^2 \right) \right\} \quad (6.16)$$

Donde  $f$  y  $\theta$  son las coordenadas polares de cualquier píxel,  $R$  es el radio vector,  $F$  es la frecuencia espacial central y  $\sigma$  es el ancho de banda de la frecuencia espacial del logaritmo de la gaussiana (Ver figura 6.6).






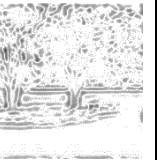

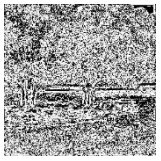
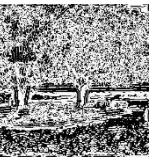

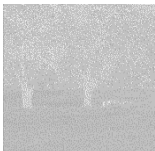

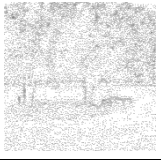
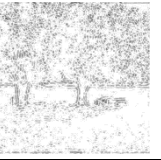
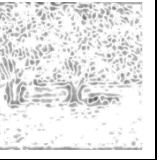
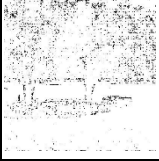
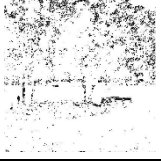
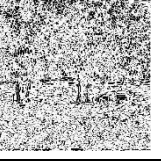
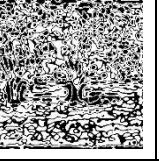



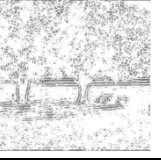
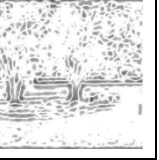

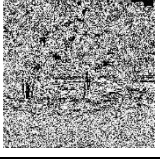
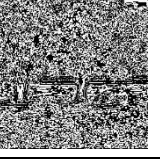



**Figura 6.6:** Los 4 filtros log Gabor de frecuencias 3, 6, 12 y 24 ciclos por grado utilizados en el estudio del índice de agrupamiento.

La elección de este tipo de filtro no vino dada solo porque fijan a cero la media de cada plano de la imagen [Laj09] ni porque reproducen bastante bien la respuesta de las células corticales [Fie87, Fie93], sino porque también son independientes del rango, es decir, su respuesta viene determinada por las relaciones entre los píxeles en vez de por el rango de la señal de entrada y, en ese sentido, el cálculo del índice de agrupamiento para los canales  $L^*$ ,  $a^*$  y  $b^*$  de una imagen no depende de los diferentes niveles de la señal en ellos.

El segundo paso consistió en transformar nuevamente al dominio espacial el espectro de amplitud de la imagen filtrada. Finalmente las imágenes filtradas se binarizaron usando un umbral fijo (más robusto al ruido) y se contabilizó la porción de píxeles que estaba entre ambos límites. Se mantuvo el valor arbitrario para el umbral de  $\pm 0,07$  que Yoonessi et al. usaron en su trabajo [Yoo08] para poder comparar nuestros resultados con los de

los mismos. La figura 6.7 muestra un ejemplo para cada uno de los canales de una imagen del resultado de su filtrado por los cuatro filtros log- Gabor, así como el resultado de su binarización usando los umbrales fijados y la porción de píxeles dentro de los mismos.

| Imagen original   | Planos $L^*, a^*, b^*$ de la imagen original   | Filtro 1  | Filtro 2  | Filtro 3   | Filtro 4  | Índice |
|---|--|---|---|--|---|--------|
|  | <br>$L^*$   |    |    |    |    | 0,63   |
|   |  |    |    |    |    |        |
|   | <br>$a^*$ |   |   |   |   | 0,79   |
|   |  |  |  |  |  |        |
|   | <br>$b^*$ |  |  |  |  | 0,58   |
|   |  |  |  |  |  |        |

**Figura 6.7:** Muestra el método utilizado para obtener el índice de agrupamiento (última columna). La primera columna muestra la imagen original utilizada en el ejemplo. La segunda columna sus canales  $L^*$  (luminancia),  $a^*$  (rojo-verde) y  $b^*$  (amarillo-azul) en el espacio CIELAB. En las cuatro columnas siguientes (correspondientes a los 4 filtros log Gabor de frecuencias 3, 6, 12 y 24 ciclos por grado utilizados). Bajo la imagen filtrada aparece la correspondiente imagen binarizada aplicando los umbrales fijos.



El índice de agrupamiento es la media de la porción de píxeles de cada canal dentro de los umbrales, promediada a todas las imágenes de cada clase semántica para cada escala y a través de las escalas. Un primer resultado común es que el índice de agrupamiento para cada imagen o categoría disminuye al aumentar la frecuencia del filtro. El resultado global para todas las imágenes fue de un índice de agrupamiento de  $0,67 \pm 0,12$ ,  $0,80 \pm 0,11$  y  $0,76 \pm 0,13$  para los canales  $L^*$ ,  $a^*$  y  $b^*$  respectivamente del total de nuestras imágenes. Se comprobó la significación de las diferencias entre las medias del descriptor en los tres canales con un intervalo de confianza del 95% a través de la prueba de Kruskal – Wallis. Los resultados mostraron que las diferencias eran significativas, con valores de Chi cuadrado de 295,102 (para el total de las imágenes), 55,096 (para las imágenes rurales) y 310,426 (para las imágenes de entornos humanos) y un valor de la significación  $s = 0$  en todos los casos. Resultados similares se obtienen si se analizan las diferencias de las medias del índice de agrupamiento entre pares de planos usando el test de Mann -Whitney

La Tabla 6.3 muestra el índice de agrupamiento de los tres canales  $L^*$ ,  $a^*$ ,  $b^*$  de las distintas categorías de nuestras imágenes ordenados en orden decreciente del índice en el canal luminancia  $L^*$ . En todas ellas (salvo en la categoría “campo” donde presenta un agrupamiento ligeramente mayor en el canal  $L^*$  que en el  $b^*$ ) el índice de agrupamiento de los canales cromáticos es mayor que en el canal luminancia; además, salvo en las categorías “costa”, “montaña” y “flores y frutas” en el resto las imágenes están más agrupadas en el canal  $a^*$  que en el  $b^*$ . Cuando las categorías semánticas se ordenan según el orden creciente o decreciente del índice de agrupamiento en uno de los tres planos de la imagen no se encuentran ordenadas en los otros dos.

Al aplicar el algoritmo a los canales luminancia, rojo-verde y amarillo-azul de las imágenes a partir del espacio log LMS se obtuvieron los siguientes resultados para el índice de agrupamiento:  $0,78 \pm 0,12$  para el canal luminancia,  $0,87 \pm 0,10$  para el canal rojo – verde y  $0,80 \pm 0,12$  para el canal amarillo-azul, siendo nuevamente significativas las diferencias encontradas entre ellos, con valores de Chi cuadrado de 203,316 (para el total de las imágenes), 76,047 (para las imágenes rurales) y 160,145 (para las imágenes de entornos humanos) y un valor de la significación  $s = 0$  en todos los casos.

A partir de los resultados en ambos espacios de color podemos ver que los planos cromáticos de las imágenes contienen más áreas uniformes que el plano de luminancia, resultados que se podían esperar dado que las diferencias cromáticas se deben a las propiedades de los objetos (color, forma...) mientras que las diferencias en luminancia no provienen solamente de las diferencias de los materiales y sus reflectancias sino de la iluminación no homogénea (con sombras, penumbras...) característica de las imágenes naturales [Rub82, Tap02, Olm04 ]. Entre los dos canales cromáticos la menor cantidad de áreas uniformes en el canal amarillo – azul puede deberse a variaciones hacia el amarillo en las sombras frecuentes de escenas de follaje abundante [Joh05] así como al

desplazamiento hacia el azul en las sombras de la iluminación de días fuertemente soleados [Chu94].

|                                     | L*                 | a*                 | b*                 |
|-------------------------------------|--------------------|--------------------|--------------------|
| Playa                               | 0,81 ± 0,15        | 0,88 ± 0,10        | 0,86 ± 0,09        |
| Autopistas                          | 0,78 ± 0,12        | 0,86 ± 0,11        | 0,84 ± 0,13        |
| Costa                               | 0,73 ± 0,11        | 0,79 ± 0,09        | 0,81 ± 0,09        |
| Rios y Cascadas                     | 0,71 ± 0,13        | 0,81 ± 0,12        | 0,77 ± 0,13        |
| Campo                               | 0,69 ± 0,15        | 0,71 ± 0,13        | 0,66 ± 0,19        |
| <b>Imágenes rurales</b>             | <b>0,68 ± 0,16</b> | <b>0,76 ± 0,14</b> | <b>0,73 ± 0,17</b> |
| Montaña                             | 0,66 ± 0,14        | 0,75 ± 0,12        | 0,75 ± 0,10        |
| <b>Imágenes de entornos humanos</b> | <b>0,65 ± 0,14</b> | <b>0,86 ± 0,11</b> | <b>0,80 ± 0,13</b> |
| Interiores                          | 0,64 ± 0,14        | 0,86 ± 0,12        | 0,76 ± 0,16        |
| Calles                              | 0,62 ± 0,09        | 0,88 ± 0,07        | 0,83 ± 0,10        |
| Edificios                           | 0,61 ± 0,11        | 0,86 ± 0,11        | 0,79 ± 0,13        |
| Flores y frutas                     | 0,61 ± 0,12        | 0,69 ± 0,13        | 0,69 ± 0,14        |
| Ciudades                            | 0,59 ± 0,10        | 0,83 ± 0,10        | 0,78 ± 0,12        |
| Bosques y parques                   | 0,55 ± 0,11        | 0,67 ± 0,13        | 0,58 ± 0,15        |

*Tabla 6.3: muestra el índice de agrupamiento de los tres canales L\*, a\*, b\* de las distintas categorías de nuestras imágenes ordenados en orden decreciente del índice en el canal luminancia L\*.*

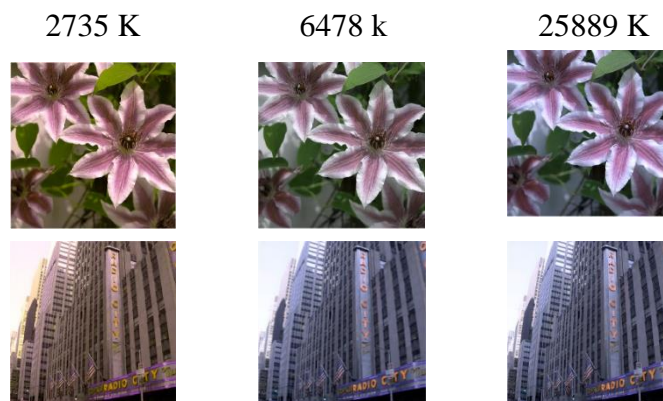
Este último párrafo muestra un ejemplo de cómo las condiciones de iluminación de las escenas naturales pueden afectar al comportamiento de los descriptores que las caracterizan, lo que enlaza perfectamente con el siguiente capítulo en el que se analizará con más profundidad cómo se comportan los distintos estadísticos estudiados en estos tres últimos capítulos bajo distintas condiciones de iluminación caracterizadas por la temperatura correlacionada de color (CCT) de sus distribuciones de potencia espectral.



## Capítulo 7

**INFLUENCIA DEL ILUMINANTE EN LA  
ESTADÍSTICA DE ESCENAS NATURALES**

El color de los objetos viene caracterizado por la llamada señal de color, que es el producto de la reflectancia espectral en cada punto de los mismos por la distribución de potencia espectral (SPD) de la iluminación bajo la que se observan. Para estudiar la influencia del iluminante en la estadística de las imágenes naturales, cada imagen RGB fue normalizada al rango (0,1) y sus valores triestímulo CIE calculados y simulados bajo un conjunto de 108 distribuciones de potencia espectral (SPD) caracterizadas por sus respectivas temperaturas de color correlacionadas (CCTs) en el rango de longitud de onda 400–700 nm. En la figura 7.1 podemos ver ejemplos de varias imágenes reproducidas bajo tres iluminantes diferentes (correspondientes a CCTs de 2735, 6478 y 25,889 K).



**Figura 7.1:** Ejemplo de una imagen natural y una de entonos humanos simuladas bajo tres iluminantes diferentes caracterizados por CCTs de 2735, 6478 y 25889 K.

La simulación fue hecha usando el algoritmo de adaptación cromática de Bradford [Süs00]. Esta transformación, basada en la adaptación cromática, modela el cambio de iluminación proporcionando un mapa de valores triestímulo XYZ para un iluminante dado bajo una fuente de luz de referencia:

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = M_b^{-1} \begin{pmatrix} R'_w/R_w & 0 & 0 \\ 0 & G'_w/G_w & 0 \\ 0 & 0 & B'_w/B_w \end{pmatrix} M_b \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (7.1)$$

donde las magnitudes  $R_w, G_w, B_w$  y  $R'_w, G'_w, B'_w$  se calculan de los valores triestímulos del iluminante de referencia y aquel para el que queremos hacer la simulación usando las funciones de ajuste de color de Judd [Jud51], y siendo  $M_b$  la matriz de Bradford [Süs00] definida como

$$M_b = \begin{pmatrix} 0,8951 & 0,2664 & -0,1614 \\ -0,7502 & 1,7135 & 0,0367 \\ 0,0389 & -0,0685 & 1,0296 \end{pmatrix} \quad (7.2)$$

Para la distribución de potencia espectral (SPD) de la fuente de iluminación se utilizaron los datos que Hernández-Andrés et al. [Her01] obtuvieron para la luz día hemisférica, es decir, basada en irradiancias espectrales globales sobre una superficie horizontal de la luz solar directa, si está presente, y del cielo entero [Pey16], bajo diferentes condiciones atmosféricas, a lo largo de dos años en Granada, España (una zona urbana no industrial, de coordenadas geográficas 37°11'N 3°37'W y 680m de altitud), y cubriendo un amplio rango de CCTs desde 3766 K hasta los 25889 K [Her01]. Las SPD de las luces día que cubren el rango de CCTs inferior a 3766 K fueron simuladas con SBDART [Pey16,Ric98] (ya que la luz día hemisférica con CCT inferior a 3600K puede observarse en raras condiciones), una herramienta informática que permite calcular la transferencia de energía radiativa dentro de la atmósfera a través de planos paralelos en la superficie terrestre bajo diferentes condiciones climáticas. De este modo, finalmente, las imágenes simuladas bajo los diferentes iluminantes se transformaron al espacio de color CIELAB. El uso de datos reales para CCTs superiores a 3766 K y simuladas para el resto provoca en muchas de las representaciones que siguen una discontinuidad en las mismas a esa temperatura de color correlacionada.

Para comprobar la validez de la transformación cromática de Bradford para reproducir imágenes bajo diferentes iluminantes hemos simulado un conjunto de imágenes hiperespectrales bajo tres iluminantes de nuestro conjunto de luz día (CCTs 2735 K, 6478 K, y 25889 K). Estas escenas serán las escenas de referencia con las que compararemos las simulaciones obtenidas de las mismas usando la transformación cromática de Bradford con una fuente de iluminación de luz día de 6478K como iluminante de referencia.

Las 13 escenas hiperespectrales, que podemos clasificar en 9 de imágenes rurales y 4 de imágenes de entornos humanos (Ver figura 7.2), se obtuvieron de las colecciones

de Foster, Nascimento y Amano (2004) y Nascimento, Ferreira y Foster (2002) [Fos06,Nas02b] disponibles on-line:

[https://personalpages.manchester.ac.uk/staff/david.foster/Hyperspectral\\_images\\_of\\_natural\\_scenes\\_04.html](https://personalpages.manchester.ac.uk/staff/david.foster/Hyperspectral_images_of_natural_scenes_04.html)

[https://personalpages.manchester.ac.uk/staff/david.foster/Hyperspectral\\_images\\_of\\_natural\\_scenes\\_02.html](https://personalpages.manchester.ac.uk/staff/david.foster/Hyperspectral_images_of_natural_scenes_02.html)



**Figura 7.2:** *Imágenes hiperespectrales utilizadas a la hora de evaluar la transformación cromática de Bradford para simular escenas bajo distintos iluminantes. Las imágenes de la primera fila pertenecen a la colección de Foster, Nascimento y Amano (2004) y el resto a la colección de Nascimento, Ferreira y Foster (2002) [Fos06, Nas02b].*

Se calcularon las diferencias de color s-CIELAB promedio para cada una de las escenas de referencia en relación con las simuladas así como para las dos categorías. Para lo cual los canales de luminancia y cromaticidad de las imágenes fueron filtrados espacialmente, transformados a espacio CIELAB y finalmente calculada su diferencia de color [Zha97], que entre dos píxeles con componentes  $(L_1^*, a_1^*, b_1^*)$  y  $(L_2^*, a_2^*, b_2^*)$  viene dada por la fórmula CIELAB 1976 como:

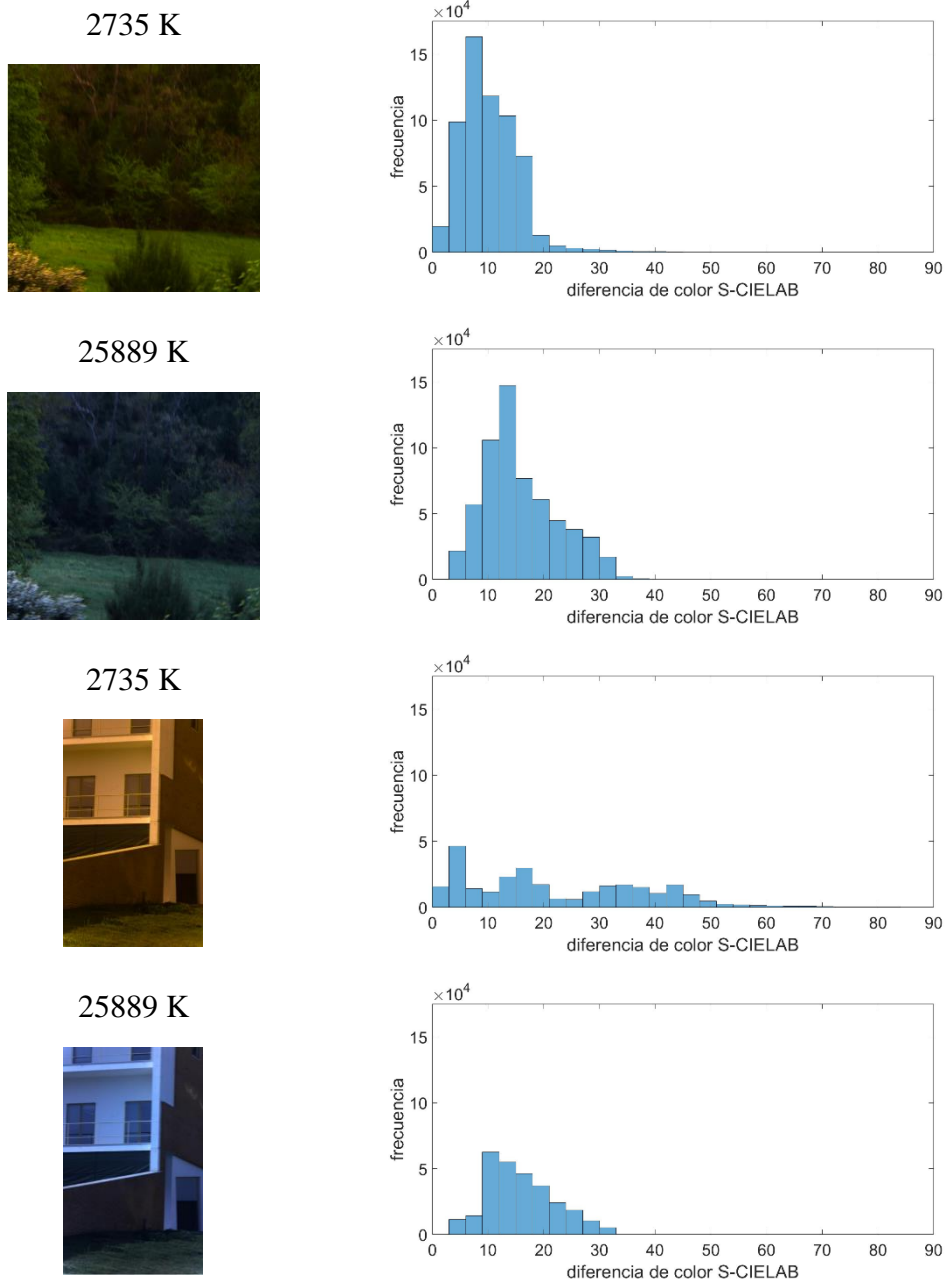
$$\Delta E = \sqrt{(L_2^* - L_1^*)^2 + (a_2^* - a_1^*)^2 + (b_2^* - b_1^*)^2} \quad (7.3)$$

La Tabla 7.1 muestra los resultados obtenidos, siendo la diferencia de color s-CIELAB de las imágenes de entornos humanos mayor que la de las imágenes rurales para los dos iluminantes de nuestro rango de trabajo; también se aprecia que el promedio de la diferencia de color en las imágenes rurales pertenecientes a la colección de 2004 (imágenes de la 1 a la 5) es menor ( $5,1 \pm 5,1 \Delta E$  para 2735 K y  $9,6 \pm 6,3 \Delta E$  para 25889 K) que la de la colección de 2002 ( $11,0 \pm 7,3 \Delta E$  para 2735 K y  $14,9 \pm 7,9 \Delta E$  para 25889 K). En función de la CCT del iluminante vemos que, finalmente, para temperaturas bajas, las imágenes rurales presentan una diferencia de color promedio más pequeña, teniendo el comportamiento opuesto las imágenes de entornos humanos.

|                              | 2735 K          | 25889 K         |
|------------------------------|-----------------|-----------------|
| Imagen 1                     | $3,8 \pm 4,9$   | $7,8 \pm 5,6$   |
| Imagen 2                     | $6,8 \pm 5,5$   | $19,5 \pm 8,8$  |
| Imagen 3                     | $6,9 \pm 6,5$   | $8,8 \pm 8,5$   |
| Imagen 4                     | $1,6 \pm 2,9$   | $1,6 \pm 3,0$   |
| Imagen 5                     | $6,3 \pm 5,9$   | $10,1 \pm 5,8$  |
| Imagen 6                     | $8,5 \pm 13,8$  | $8,4 \pm 8,7$   |
| Imagen 7                     | $8,8 \pm 4,6$   | $17,1 \pm 10,4$ |
| Imagen 8                     | $10,3 \pm 5,3$  | $15,9 \pm 6,7$  |
| Imagen 9                     | $16,5 \pm 5,4$  | $18,1 \pm 5,9$  |
| Imágenes rurales             | $7,7 \pm 6,1$   | $11,9 \pm 7,0$  |
| Imagen 10                    | $9,2 \pm 6,6$   | $19,3 \pm 11,1$ |
| Imagen 11                    | $8,2 \pm 10,1$  | $5,4 \pm 3,5$   |
| Imagen 12                    | $22,9 \pm 16,6$ | $15,9 \pm 6,1$  |
| Imagen 13                    | $20,0 \pm 13,2$ | $13,6 \pm 5,4$  |
| Imágenes de entornos humanos | $15,1 \pm 11,6$ | $13,5 \pm 6,6$  |

**Tabla 7.1:** Diferencia de color CIELAB ( $\Delta E$ ) media para cada una de las imágenes hiperspectrales y cada una de las categorías semánticas (imágenes rurales y de entornos humanos), entre la imagen original y la simulada por la adaptación cromática de Bradford para los iluminantes extremos de nuestro rango de trabajo de CCTs 2735 K y 25889 K.





**Figura 7.3:** En la columna de la izquierda aparecen dos imágenes simuladas con la transformación cromática de Bradford bajo los iluminantes de CCTs de 2735 K y 25889 K. En la columna de la derecha se muestra el histograma de las diferencias de color s-CIELAB entre la imagen original y la simulada bajo la transformación cromática de Bradford utilizada.

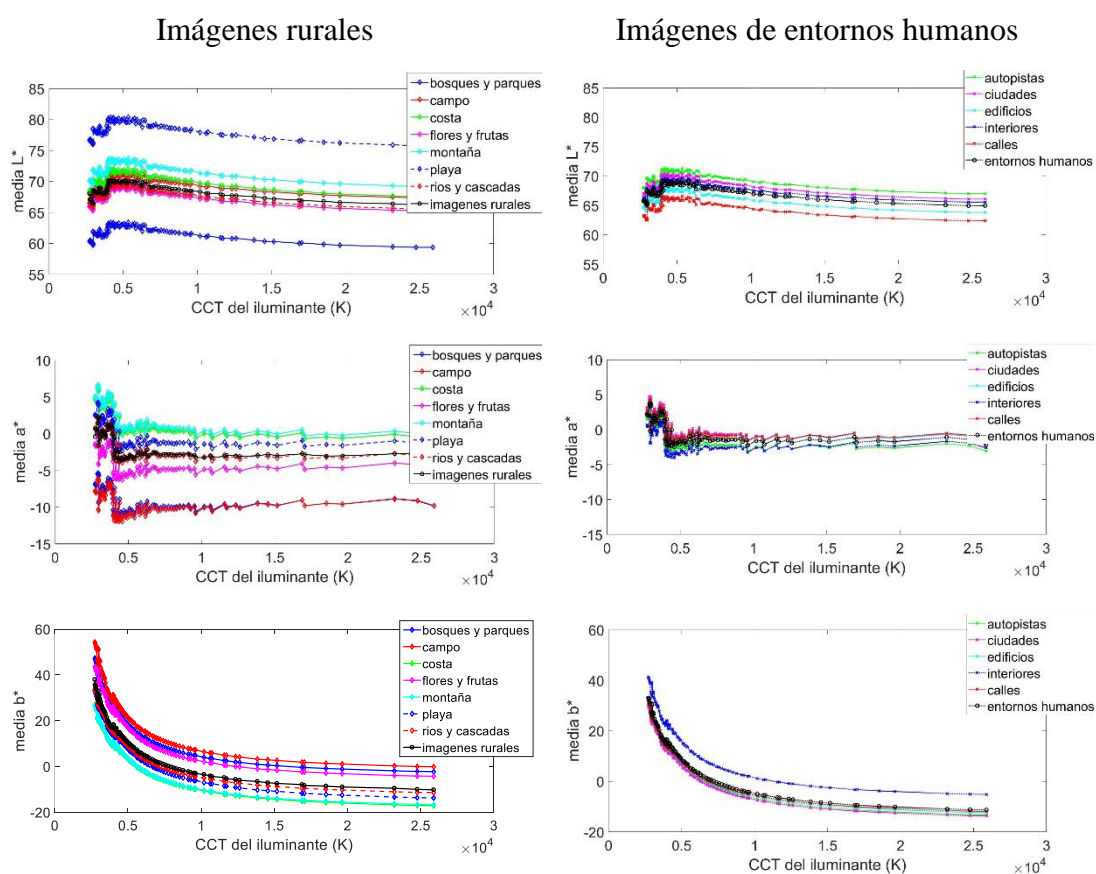
Aunque el umbral máximo de la diferencia de color  $\Delta E$  para considerar dos colores como uno solo varía, según las condiciones, los materiales o el proceso industrial al que se le aplica ( estrictamente  $\Delta E = 1$  para distinguir dos colores aislados,  $\Delta E = 5$  según la norma ISO 12647-2 encargada de los estándares de impresión), en el intento de

cuantificar diferencias de color con imágenes usando la métrica s-CIELAB, Zhang y Wandell [Zha97] tampoco establecen un valor estándar de umbral para considerarlas iguales.

La figura 7.3 muestra ejemplos de las imágenes resultantes (columna de la izquierda) al someter la imagen de referencia (bajo iluminante de 6478 K) a la adaptación cromática de Bradford utilizando los iluminantes extremos nuestro rango de trabajo, 2735 K y 25889 K. En la columna de la derecha se muestra el histograma de las correspondientes diferencias de color s-CIELAB entre las dos condiciones.

## 7.1 Estadística de primer orden.

En la figura 7.4 se muestra el valor de la media (primer momento) de la intensidad de las componentes  $L^*$ ,  $a^*$  y  $b^*$  para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante. En primer lugar observamos que las variaciones en el canal  $b^*$  en función del iluminante es mucho



**Figura 7.4:** Valor medio del primer momento (media) de la intensidad de las componentes  $L^*$ ,  $a^*$  y  $b^*$  para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante.

mayor (un porcentaje de variación con respecto al valor máximo comprendido entre 27,9% y 29,8% para las imágenes rurales y ligeramente inferior para las imágenes de entornos humanos entre 27,1 % y 28%) que en los otros dos canales (entre 5,7 % y 6,3 % para el canal  $L^*$  y entre 4,6 % y 5,6 % para el canal  $a^*$  en el caso de imágenes rurales y nuevamente algo más bajo para las imágenes de entornos humanos entre 6% y 6,3% y entre 4,6 % y 4,8% para los canales luminancia y color oponente rojo-verde) para todas las categorías de imágenes; además, la intensidad media del canal oponente amarillo-azul decrece con mayor rapidez para iluminantes con temperaturas de color más bajas que con los de temperaturas más altas.

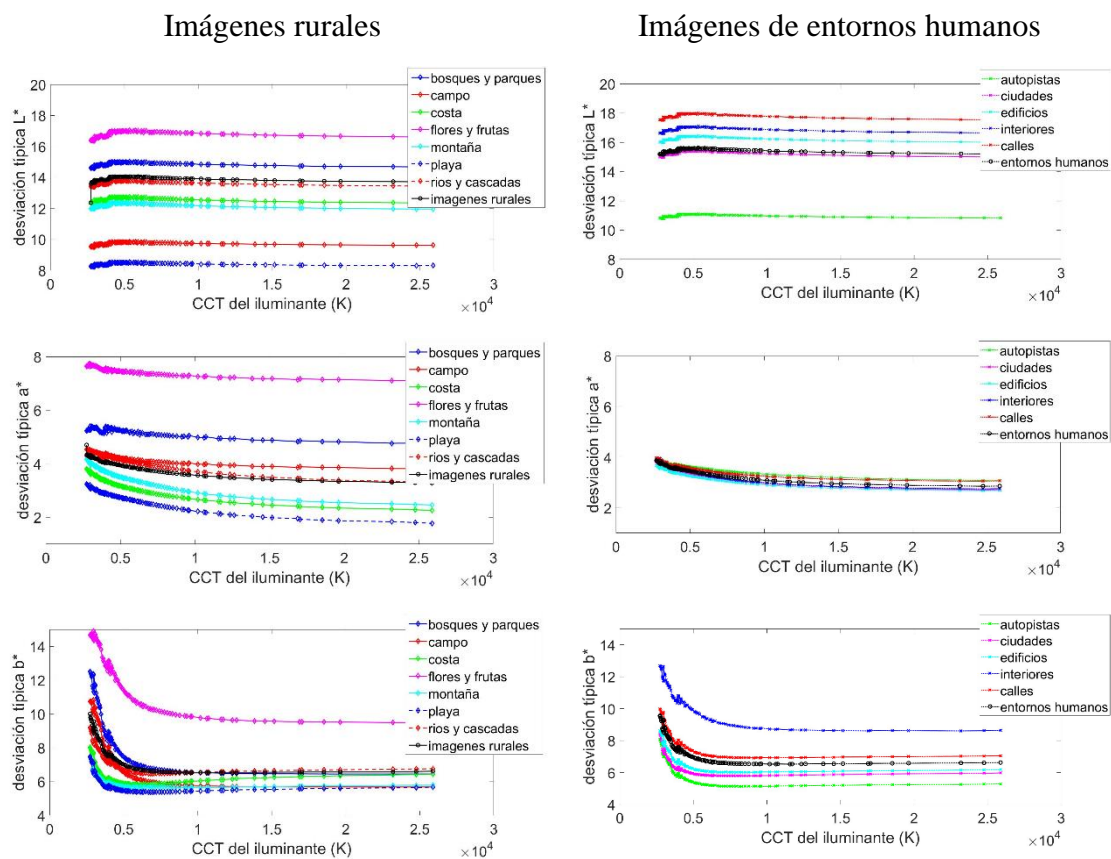
Como hemos señalado, es el otro canal cromático  $a^*$  el que muestra la menor dependencia con la CCT del iluminante pero frente a un decrecimiento continuado como el que mostraba el canal oponente  $b^*$  (salvo donde cambia el método de obtener la SPD de los distintos iluminantes), el canal oponente rojo - verde decrece rápidamente para temperaturas del iluminante bajas, presenta un mínimo local en el intervalo 4422K – 4602 K para crecer lentamente a partir del mismo en todas las categorías (tanto de las imágenes rurales como de entornos humanos). Finalmente, el canal luminancia acromático presenta un comportamiento similar para todas las clases de imágenes utilizadas y crece rápidamente para CCTs de la fuente de iluminación pequeñas, presentando un máximo local en el intervalo 5858 K y 7571 K y decreciendo más lentamente a medida que las temperaturas superan a la de dicho máximo relativo.

El análisis ANOVA de un factor, para determinar si para un intervalo de confianza del 95 % podrían considerarse iguales las medias de los primeros momentos de cada una de las categorías para los distintos iluminantes, rechaza esta hipótesis en todas ellas.

La figura 7.5 muestra los resultados obtenidos del análisis de la dependencia de la desviación típica (segundo momento) de la intensidad de las tres componentes  $L^*$ ,  $a^*$  y  $b^*$  en función de la CCT del iluminante. Estos indican una dependencia mayor de los canales cromáticos  $b^*$  y  $a^*$  (con rangos entre 24,2 % y 48,5% de su valor máximo para el primero y entre el 8,1% y 45,3% para el segundo en las imágenes rurales) con la temperatura del iluminante. Este amplio rango de variación en porcentaje es debido a que aquellas categorías que tiene mayor dispersión (una desviación típica más alta) presentan una variación absoluta de la misma con la temperatura del iluminante menor.

El canal luminancia (cuyo porcentaje de variación respecto a su valor máximo varía en el rango 2,8 % - 12%) presenta para todas las categorías de imágenes rurales un máximo relativo en el intervalo de temperaturas del iluminante 5346 K – 5931 K. Los canales cromáticos presentan una respuesta decreciente en función de la temperatura a lo largo de todo el rango de las mismas (canal  $a^*$ ) o hasta el intervalo de CCT 6305 K – 9520 K donde alcanzan un mínimo local para crecer muy lentamente a altas temperaturas (canal  $b^*$  salvo las categorías “bosques y parques”, “campo” y “flores y frutas” de las

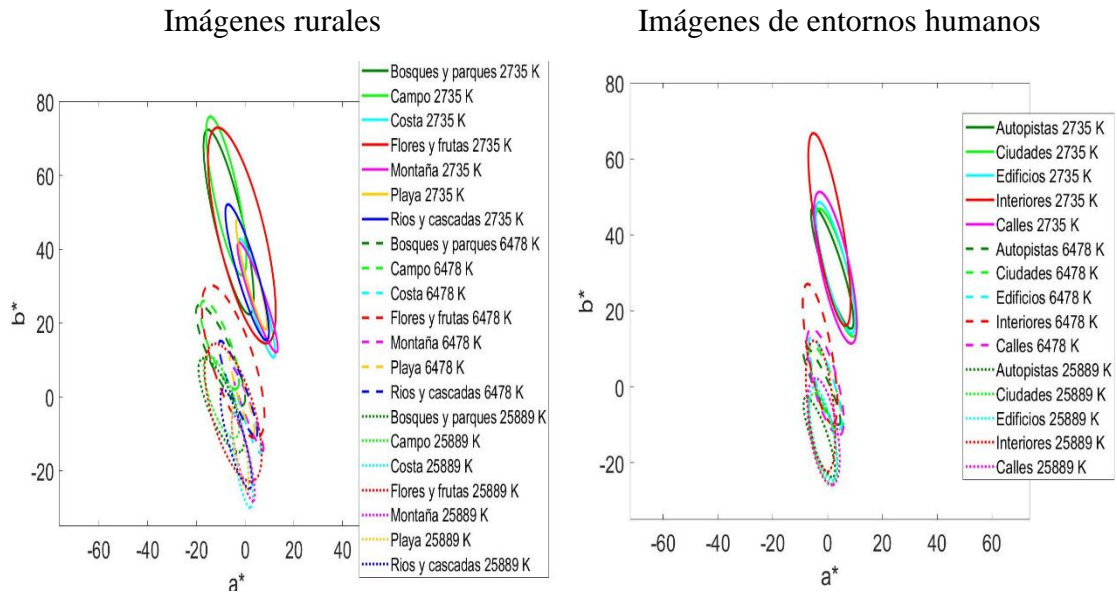
imágenes rurales que decrecen para todo el dominio de CCTs), como podemos ver en la primera columna de la Figura 7.5.



**Figura 7.5:** Valor medio del segundo momento (desviación típica) de la intensidad de las componentes  $L^*$ ,  $a^*$  y  $b^*$  para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante.

Las imágenes de entornos humanos presentan un comportamiento parecido al de las imágenes rurales en los tres planos (segunda columna de la Figura 7.5), mostrando variaciones más parecidas entre las distintas categorías lo que se refleja en rangos de variación más estrechos para cada uno de los planos (entre 2,7% y 3% , entre 22% y 30% y entre el 30 y 36% del valor máximo respectivamente para  $L^*$ ,  $a^*$ ,  $b^*$ ) y ligeramente inferiores a los correspondientes a las imágenes rurales.

Para los segundos momentos, el análisis ANOVA de un factor determina que para un intervalo de confianza del 95 % si pueden considerarse significativamente iguales las medias de las desviaciones típicas de cada una de las categorías para los distintos iluminantes, en el canal  $L^*$  no siéndolo en los canales cromáticos.

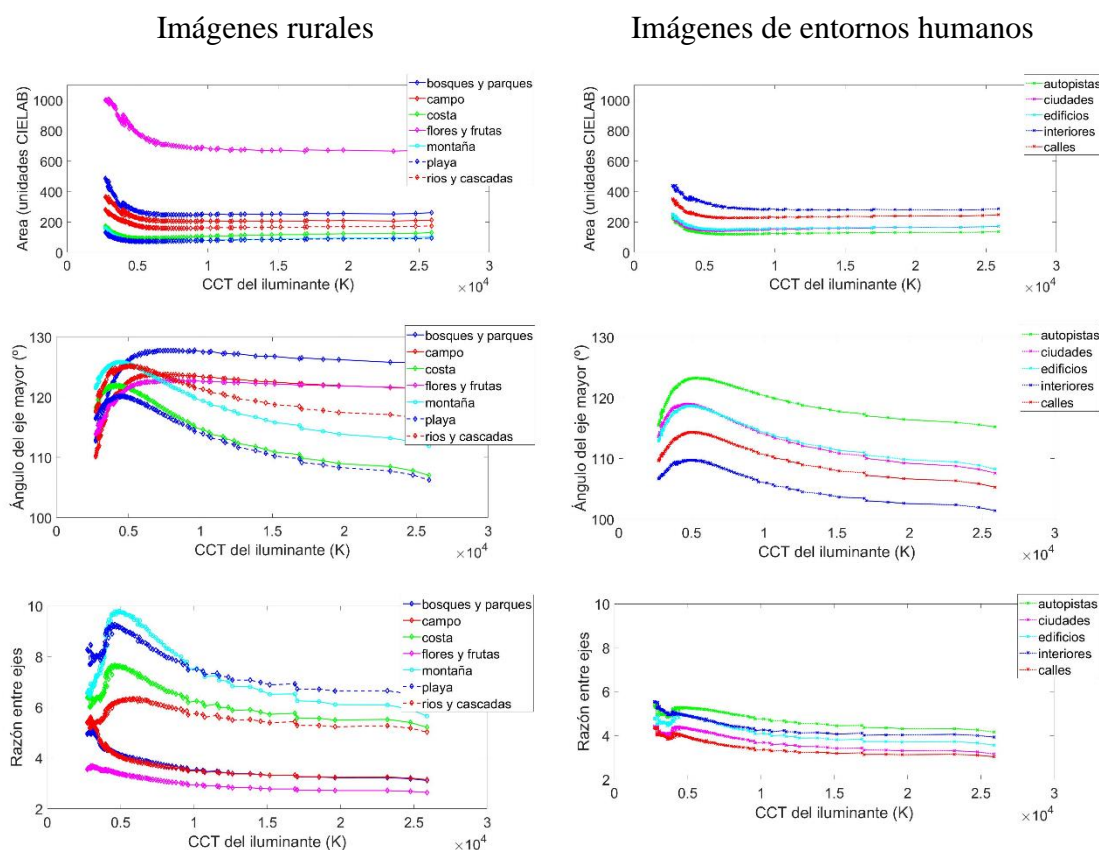


**Figura 7.6:** Elipses de ajuste de mínimos cuadrados de las gamas de color media de cada una de las categorías semánticas para las temperaturas de color de los iluminantes 2735 K (líneas continuas), 6478 K (líneas discontinuas) y 25889 K (líneas de puntos).

Con respecto a los descriptores estadísticos de la diversidad cromática, la figura 7.6 muestra las elipses de ajuste de mínimos cuadrados de las gamas de color media de cada una de las categorías semánticas (para imágenes rurales en la columna de la izquierda y para imágenes de entornos humanos en la de la derecha) para tres temperaturas de color de los iluminantes, las dos temperaturas extremas del intervalo de trabajo 2735 K (líneas continuas) y 25889 K (líneas de puntos), y la correspondiente a 6478 K (líneas discontinuas). De la imagen podemos apreciar que las áreas correspondientes a 2735 K son las mayores y las correspondientes a 6478 K las menores salvo para la categoría “flores y frutas” que es decreciente con la temperatura para los tres casos mostrados. En el caso de los ángulos de los ejes mayores de las elipses con respecto al semieje positivo de  $a^*$  estos son mayores de  $90^\circ$  en todos los casos y mayores para todas las categorías bajo el iluminante correspondiente a 6478 K y siendo los menores los correspondientes a la mayor CCT, salvo para las categorías “bosques y parques”, “campo” y “flores y frutas”. Finalmente, todas las elipses se encuentran elongadas en la dirección del canal oponente amarillo-azul presentando una razón entre sus semiejes decreciente en función de la temperatura, es decir, su elongación disminuye en función de la temperatura para todas las categorías excepto para “costa”, “montaña” y “ríos y cascadas” cuya razón entre ejes es mayor a los 6478 K. Los centros medios de las gamas de color se encuentran prácticamente centrados respecto al eje  $a^*$ , experimentando un ligero desplazamiento hacia valores negativos cuando las temperaturas del iluminante aumentan (de -0,39 a -4,42 para imágenes rurales y de 2,27 a -2,08 para imágenes de



entornos humanos), mientras que respecto al eje  $b^*$  los centros de las elipses se desplazan significativamente de valores positivos medios a valores negativos cuando las CCTs crecen (de 37,97 a -9,48 para imágenes rurales y de 33,07 a -11,27 para las de entornos humanos).

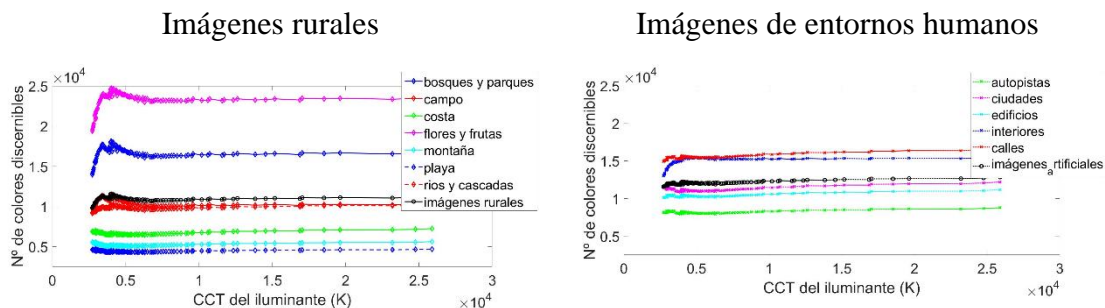


**Figura 7.7:** Variación con la CCT del iluminante de los descriptores de la elipse que mejor ajusta la gama de color: área en unidades CIELAB (primera fila), ángulo del eje mayor con el semieje positivo del eje  $a^*$  (segunda fila) y razón entre los semiejes de la elipse (tercera fila); para imágenes rurales (primera columna) y de entornos humanos (segunda columna).

Al extender el análisis de los tres descriptores a todo el rango de temperaturas observamos (ver figura 7.7) que el área de la elipse (en unidades CIELAB de área) que mejor ajusta la gama de color presenta una variación con respecto al valor máximo para cada categoría entre el 33,8% y el 55,6% en imágenes rurales y entre 36,6% y 47,7% en imágenes de entornos humanos, con un mínimo relativo en el intervalo de temperaturas 5858 K – 9119 K para las primeras y en torno a los 6304 K para las segundas. Por el contrario, las variaciones del ángulo del eje mayor de la elipse con el semieje positivo de  $a^*$  presenta un máximo local en el rango 4184 K – 9614 K en imágenes rurales y entre

4705 K - 5425 K en entornos humanos con rangos de variación para las distintas categorías entre 6,1% - 12,3% y 6,6% - 9,6% respectivamente. Finalmente la dependencia de la razón entre los ejes de la elipse con la temperatura del iluminante presenta un máximo relativo en el intervalo de temperaturas 6304 K - 8300 K para las imágenes rurales (salvo en las categorías con distribuciones cromáticas más simétricas: “bosques y parques”, “campo” y “flores y frutas” que decrecen casi de manera continuada con la CCT de la fuente de alimentación) y alrededor de 4422 K para las imágenes de entornos humanos. El rango de variación de la razón entre los ejes de la elipse de esta dependencia con la CCT de la fuente de iluminación para cada categoría es de 20,8% - 43,8% del valor máximo para las imágenes rurales y de entre 23,1% - 30,6% para las de entornos humanos.

Para los descriptores cromáticos que caracterizan la elipse que mejor ajusta la gama de color de una imagen, las pruebas estadísticas de comparación múltiple de las medias de las distribuciones para un intervalo de confianza del 95 %, señalaron que, salvo para la razón entre los ejes de la elipse de mejor ajuste de las categorías “ríos y cascadas” y “autopistas” que sí pueden considerarse significativamente iguales bajo los distintos iluminantes, la hipótesis de igualdad de las mismas vendría rechazada en el resto.



**Figura 7.8:** Valor medio del número de colores discernibles de cada categoría semántica para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante.

El último descriptor de primer orden relacionado con la información cromática de nuestras imágenes es el número de colores discernibles. La mayoría de las categorías semánticas presentan una variación en torno al 10 % del valor máximo del número de colores discernibles con la temperatura del iluminante. Como se puede ver en la figura 7.8 las categorías “bosques y parques”, “campo”, “flores y frutas” y “ríos y cascadas” de imágenes rurales e “interiores” son las que presentan mayores variaciones (entre el 15% y el 23 % para las rurales y del 16% para la categoría de “interiores”) y son las que se corresponden con las de mayor número de colores discernibles (a excepción de la categoría “calles” de imágenes de entornos humanos), presentando un máximo local en

torno a los 4016 K frente a las categorías de imágenes rurales con menor variedad cromática que presentan un mínimo relativo en el intervalo 3903 K – 6305 K. Respecto a las imágenes de entornos humanos todas las categorías presentan un máximo local en el rango 4016 K - 4949 K. Cabe resaltar el crecimiento lento (casi constante) de todas las categorías tanto de imágenes rurales como de entornos humanos para temperaturas de color del iluminante superiores a 8288 K para las primeras y de 7406 K para las segundas.

El test de Kruskal - Wallis aplicado a las distribuciones del número de colores discernibles para cada categoría y los distintos iluminantes determina que para un intervalo de confianza del 95 %, este descriptor puede considerarse significativamente igual bajo las distintas condiciones de iluminación para cada categoría semántica.

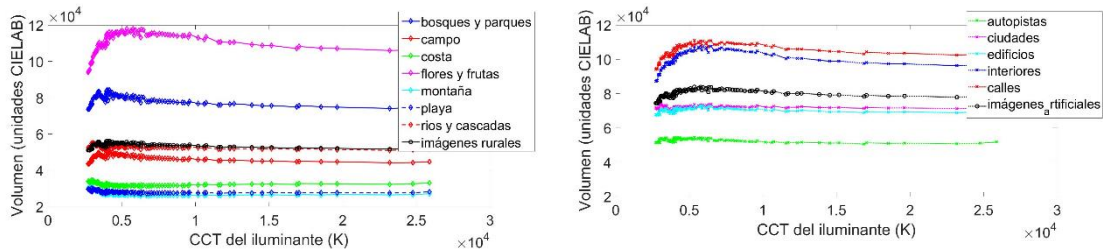
Todos los descriptores estadísticos de primer orden analizados que sólo dependen de los dos canales cromáticos presentan variaciones significativas con las CCTs de los iluminantes luz día utilizados, siempre ligeramente menores en los entornos humanos que en las imágenes rurales. En lo referente a los dos primeros momentos de la distribución de intensidad de los tres canales luminancia, y cromáticos oponentes rojo-verde y amarillo-azul, es este último el que presenta una mayor dependencia con la temperatura presentando una variación promedio del 28% del valor máximo para la media y de un 33% para la desviación típica (similar entre imágenes rurales y de entornos humanos). El canal oponente rojo-verde es el que presenta menos influencia con la temperatura del iluminante (variaciones por debajo del 5% de media de su valor máximo) pero la dispersión de los datos en torno a la media (caracterizada por la desviación típica) depende más significativamente de la CCT del iluminante que la del canal acromático, (cuya desviación típica puede considerarse estadísticamente igual para todos los iluminantes para un intervalo de confianza del 95%).

En lo referente a los descriptores de la gama de color, el ángulo del eje mayor de la elipse de mejor ajuste de dicha gama con el semieje positivo del canal  $a^*$  (que siempre tiene mayor amplitud que un ángulo recto) es el que menos varía con la temperatura, alrededor del 9% del valor máximo de cada categoría de promedio. Los otros dos descriptores el área y la razón de los semiejes de la elipse varían significativamente con el iluminante luz día de promedio por categoría un 43 % y un 31 % del valor máximo; asimismo, el centro de la elipse se desplaza muy acentuadamente a lo largo del eje  $b^*$  por lo que podemos considerar que las gamas de color sí muestran una dependencia importante con la CCT que caracteriza al iluminante.

Aunque aparentemente nuestros datos (áreas de las elipses de mejor ajuste presentan un mínimo en el rango 5500K – 10000K) parecen contradecir los resultados obtenidos por Martínez-Verdú et al. [Mar07], que determinaban que las gamas de color correspondientes a iluminantes caracterizados por CCTs en ese intervalo (salvo que tengan un ancho de banda muy estrecho) eran mayores que para las de los iluminantes fuera de ese rango, esta contradicción desaparece si, en vez de considerar el área de las elipses que mejor ajustan la gama de color, tomamos el volumen de color (teniendo en



cuenta el valor de la luminancia ( $L^*$ ) que sí presenta un máximo en dicho intervalo de temperaturas [Oje17] (Ver figura 7.9).



**Figura 7.9:** Volumen de color en unidades CIELAB de cada categoría semántica para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante.

En el mismo trabajo [Mar07] los sólidos de color en espacios de color uniformes bajo distintos iluminantes dependían claramente de estos tanto en su forma como en su contenido. Por el contrario, nosotros obtenemos que esta dependencia con la temperatura, tanto del volumen de color como del número de colores discernibles, es sensiblemente menor a la obtenida para los otros descriptores pudiendo considerarse desde un punto de vista de la significación estadística para un intervalo de confianza del 95 % como independientes de la CCT del iluminante lo que podría deberse a las primeras etapas de la constancia de color (adaptación cromática).

Nuestros resultados acerca del número de colores discernibles confirman, para la mayoría de las categorías, los resultados de Masaoka et al. [Mas13] que mostraban que en el rango de temperaturas del iluminante 2000K - 10000K el número de colores discernibles es máximo en los espacios de color CIELAB y CIE94 alrededor de lo 4000 K cuando no se ha aplicado ninguna transformación Von Kries de adaptación cromática a los datos; en nuestro caso, estos máximos se alcanzaban alrededor de 4016 K para las imágenes rurales y en el rango de temperaturas 4016 K – 4949 K.

En el trabajo de Masuda y Nascimento [Ma13] en el que analizaron el grado de naturalidad y la preferencia de imágenes hiperespectrales en función de la CCT de iluminantes luz día, encontraron una clara dependencia de estos índices con la CCT de la fuente de iluminación. Ellos concluyeron que para la luz día los colores más naturales se obtenían a una CCT de 6040 K y que los colores más preferidos a una temperatura de 4410 K lo que estaba relacionado con unas gamas de color más simétricas o mayores respectivamente. Nuestros datos muestran que muchas categorías muestran un mínimo local en la razón entre los semiejes de las elipses (mayor simetría en la gama de color) alrededor de los 4000 K aunque no es la configuración más simétrica de la gama cromática que se alcanzaría a las temperaturas más altas del rango. No obstante, también

hemos comentado ya que el rango en el que se alcanza el máximo de colores discernibles en nuestra experiencia coincidiría con su resultado de preferencia de colores (gammas más ricas) a 4410 K.

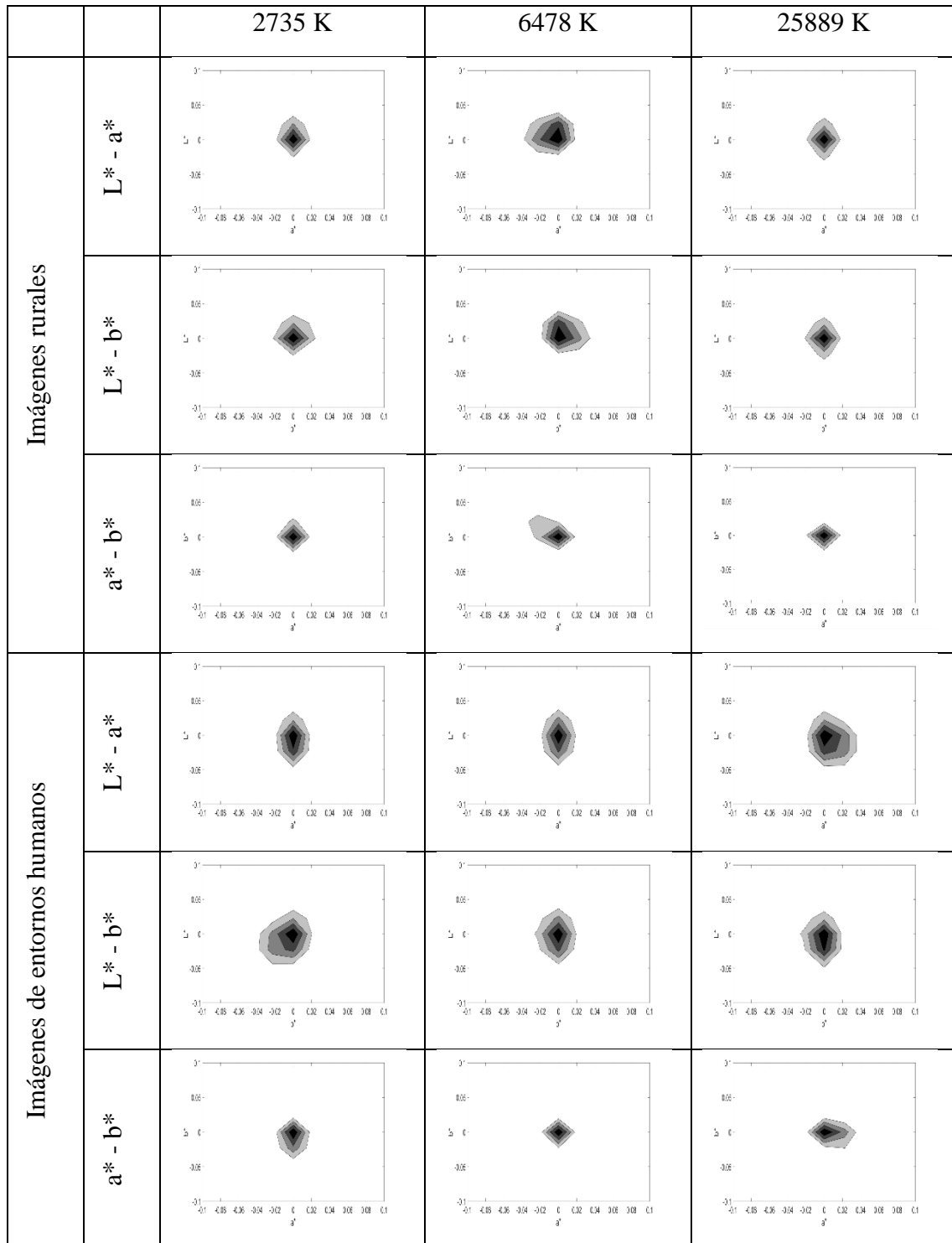
Foster [Fos21] teniendo solo en cuenta sólo las propiedades espectrales de la luz reflejada por cada superficie en las muestras de imágenes hiperespectrales de su trabajo y basándose en los principios de la teoría de la información determina computacionalmente la mínima cota superior del número de superficies que pueden considerarse cromáticamente iguales a partir de los cambios reales en la luz reflejada por las distintas superficies de la escena que se producen en un determinado periodo temporal. Mientras el número medio de superficies distinguibles era de 10000 en el espacio CIECAM02, el número de superficies que podían considerarse iguales tras un intervalo de 2 minutos era de 600, tras 10 minutos de 200 y tras una hora de 70. Esta importante reducción en el reconocimiento de las superficies por su color, observada incluso en cortos intervalos de tiempo, es mucho menor cuando se consideraban solamente cambios globales puros del espectro, para lo cual Foster [Fos21] estima nuevamente el número de superficies cromáticamente iguales bajo un iluminante de 4000 K y 6500 K obteniendo un número entre 4400 y 5500 dependiendo del modelo de ruido interno usado. Esta menor dependencia con el iluminante se asemeja más a nuestros resultados sobre volumen de color, número de colores discernibles e índice de agrupamiento que desde un punto de vista estadístico podían considerarse significativamente iguales para los distintos iluminantes; además, el intervalo que utiliza Foster en su trabajo [Fos20] coincide con el rango en el que estos descriptores presentan mayor variación, como muestran las figuras 7.8 y 7.9

## **7.2 Estadística de segundo orden.**

### **7.2.1 Estadística entre píxeles adyacentes.**

En el capítulo 5 vimos que la estadística de segundo orden era la primera aproximación al estudio de las regularidades espaciales de las imágenes, para lo que utilizamos descriptores con el fin de estudiar las relaciones entre pares de píxeles adyacentes (gradientes) o no adyacentes (espectros de potencia y rúbricas espectrales) así como el análisis de componentes principales.

Para estudiar la influencia del iluminante en la estadística de segundo orden de píxeles adyacentes se analizaron los histogramas conjuntos de los gradientes medios en los distintos planos de cada escena para las categorías imágenes rurales y de entornos humanos bajo tres iluminantes caracterizados por temperaturas de color de 2735 K, 6478 K y 25889 K.



**Figura 7.10:** Histogramas conjuntos en escala logarítmica para los tres canales  $L^*$ ,  $a^*$  y  $b^*$  de las imágenes rurales y las imágenes de entornos humanos bajo tres iluminantes caracterizados por temperaturas de 2735 K, 6478 K y 25889 K.

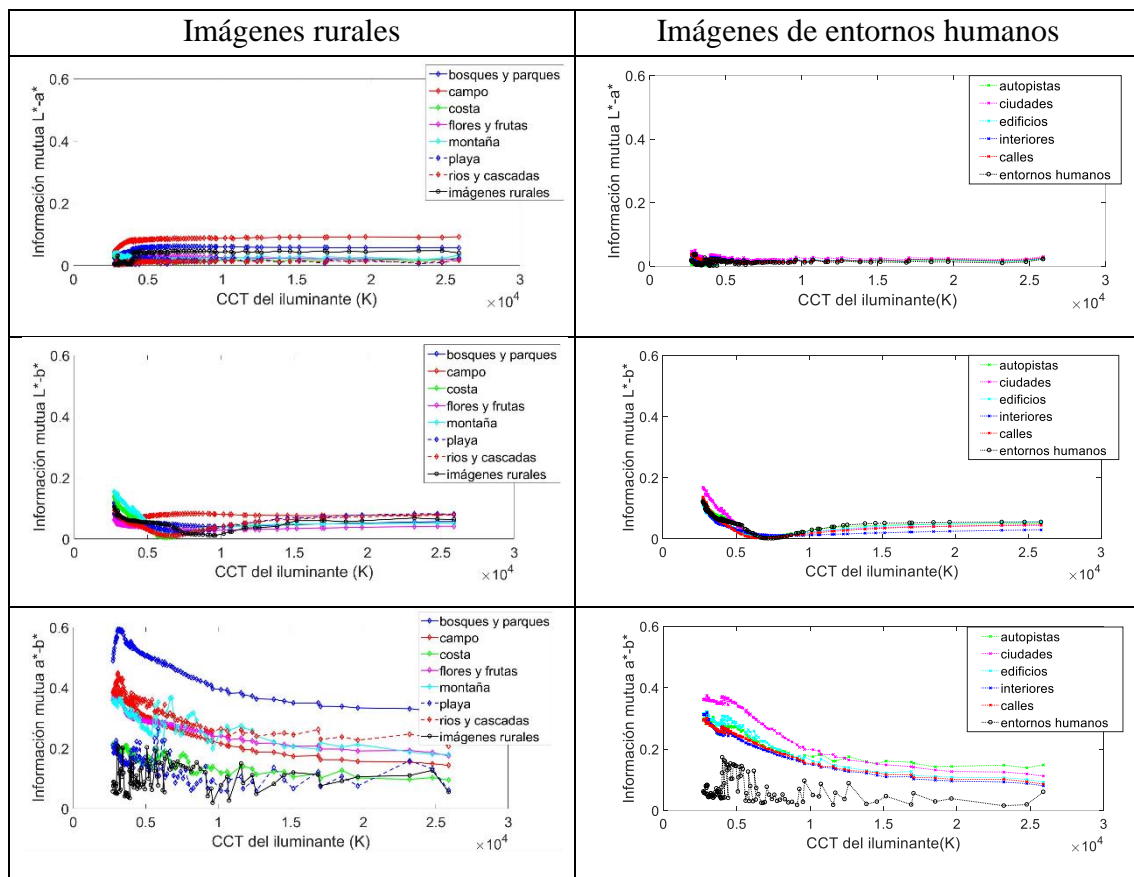
En la figura 7.10 se han representado dichos histogramas conjuntos en escala logarítmica. Como pasaba con el caso de las imágenes originales todos ellos presentan un máximo en

torno a cero (pues la mayoría de los píxeles de las imágenes tanto rurales como de entornos humanos no corresponden a bordes de las mismas independientemente de la luz día bajo la que estén sometidas), una distribución aproximadamente elipsoidal y cierto grado de simetría con un ancho parecido de las distintas regiones. No obstante, en el caso de las imágenes rurales vemos que en los histogramas conjuntos  $L^*-a^*$  y  $L^*-b^*$  el canal luminancia ( $L^*$ ) presenta una altura que permanece prácticamente constante para las tres temperaturas mientras que, para iluminantes alrededor de 6478 K, los canales cromáticos ( $a^*$  y  $b^*$ ) presentan mayores frecuencias para valores más altos en valor absoluto de sus gradientes (el histograma se ensancha en esa dirección). En el histograma conjunto  $a^*-b^*$  es el canal oponente amarillo-azul el que presenta prácticamente la misma altura siendo el canal rojo-verde el que presenta un histograma con frecuencias mayores para gradientes mayores alrededor de CCTs del iluminante de 6478 K. Este incremento de bordes no lo es en la dirección de  $b^*$  sino en la dirección de la bisectriz de los cuadrantes segundo y cuarto del histograma conjunto y por tanto con mayor presencia de bordes verdes claros. Las imágenes de entornos humanos presentan histogramas conjuntos que abarcan superficies mayores lo que indica una mayor presencia de bordes en los tres canales que en las imágenes rurales; con respecto a la influencia de estos con la temperatura del iluminante, podemos apreciar que la mayor frecuencia de gradientes más altos en el plano rojo-verde se alcanza a los 25889 K tanto en los histogramas  $L^*-a^*$  y  $a^*-b^*$ , mientras que el mayor número de bordes en el plano amarillo-azul se produce a bajas temperaturas del iluminante (2735 K) tanto en el histograma  $L^*-b^*$  como en el histograma  $a^*-b^*$ .

También se observa que la mayoría de los bordes combinan información de los dos canales. El grado de independencia de los mismos se describía con la información mutua media entre los gradientes de los distintos canales medida en bits (en nuestro caso dado que se utilizaron 64 elementos debería estar en el rango 0 - 6 bits). Al analizar la información mutua entre los gradientes de los distintos canales en función de la temperatura del iluminante esta varía entre 0,001 y 0,594 bits lo que nos permite concluir que un gradiente en una dimensión no permite predecir el valor del gradiente más probable en la otra.

En la figura 7.11 podemos apreciar que la información mutua entre los dos canales cromáticos es claramente superior a la existente entre uno de ellos y el canal luminancia, además es en la información mutua entre los canales rojo-verde y amarillo-azul donde se produce un rango de variación absoluta con la temperatura del iluminante mayor (entre 0,123 y 0,306 bits para las imágenes rurales y entre 0,158 – 0,260 bits para las imágenes de entornos humanos) decreciendo a partir de los 5858 K en todas las categorías semánticas. Esta pequeña información mutua nos indica que los bordes correspondientes a los canales amarillo – azul y rojo – verde no son totalmente independientes mostrando mayor dependencia entre ellos (principalmente a temperaturas del iluminante menores), que entre el canal luminancia y uno de los canales cromáticos. Con respecto a la información mutua entre los gradientes de los canales luminancia y los canales

cromáticos, se ve que a partir de los 13840 K estas varían muy lentamente; del mismo modo, bajo temperaturas inferiores a los 4602 K la variación es acentuada presentando un extremo local entre los 6305 K y los 9615 K (que es un mínimo) para todas las categorías semánticas en el caso de los canales acromático y amarillo-azul. Aunque el valor de la información mutua es menor en imágenes de entornos humanos para los distintos iluminantes así como las variaciones entre las distintas categorías con las CCTs de los mismos, la variación relativa de la misma con respecto al máximo de cada categoría señala valores inferiores y más similares entre los tres tipos ( $L^*-a^*$ ,  $L^*-b^*$  y  $a^*-b^*$ ) con un rango de variación del 74,35%, 75,14% y un 62,63% respectivamente, que para las imágenes de entornos humanos con una variación del 82,84%, 95,81% y 72,30%. A pesar de la importante variación relativa, dados los pequeños valores de la información mutua, no parece que esta sea significativa para el proceso de la visión. Desde el punto de vista de la significación estadística para un intervalo del 95% el test de Kruskal – Wallis establece que pueden considerarse iguales para cada categoría (excepto “playa” e “interiores”) los valores medios de la información mutua entre los canales  $L^*$  y  $a^*$  bajo los distintos iluminantes.



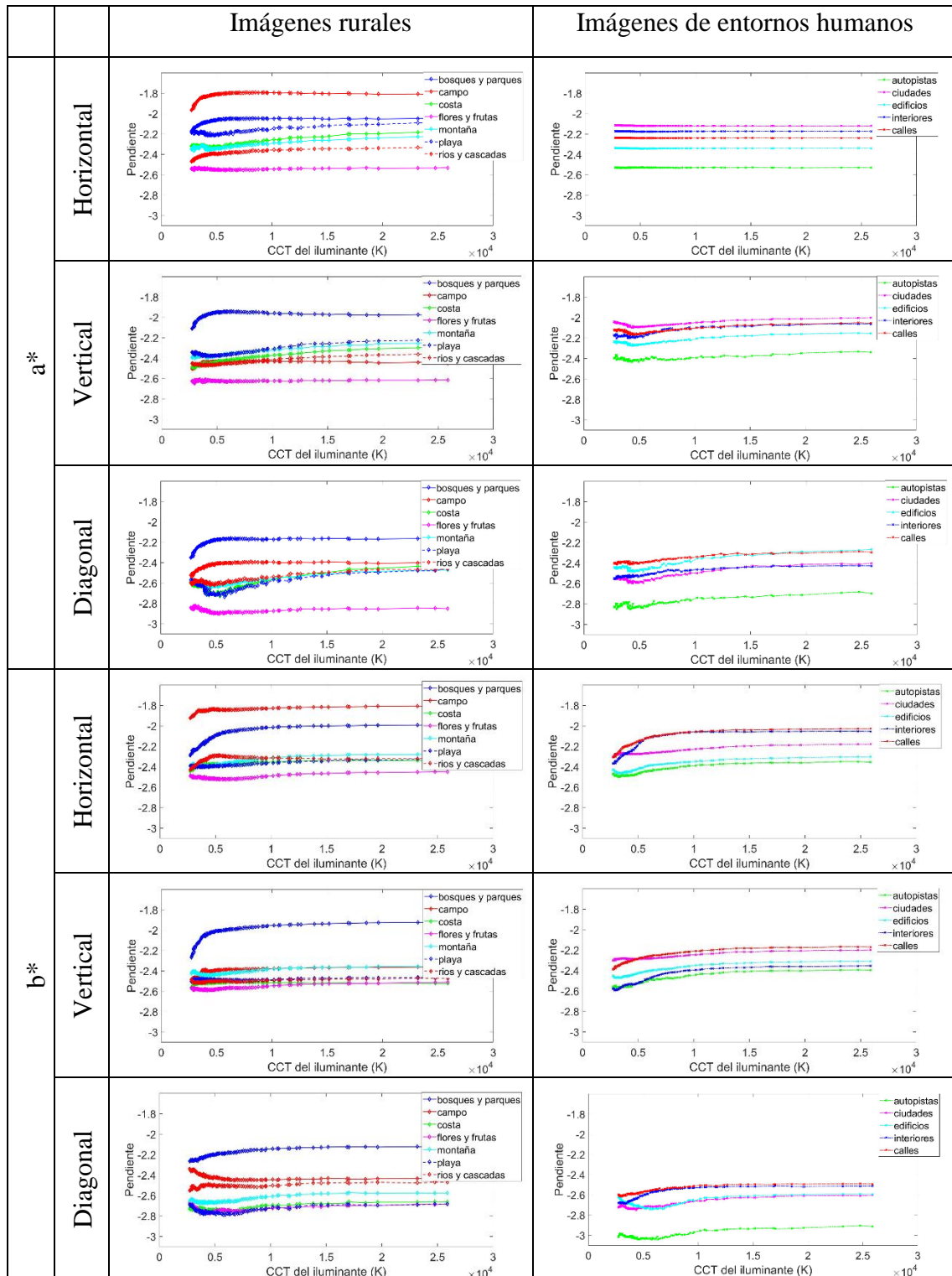
**Figura 7.11:** Información mutua entre los gradientes de las componentes  $L^*$ ,  $a^*$  y  $b^*$  para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante.

### 7.2.2 Estadística entre píxeles no adyacentes.

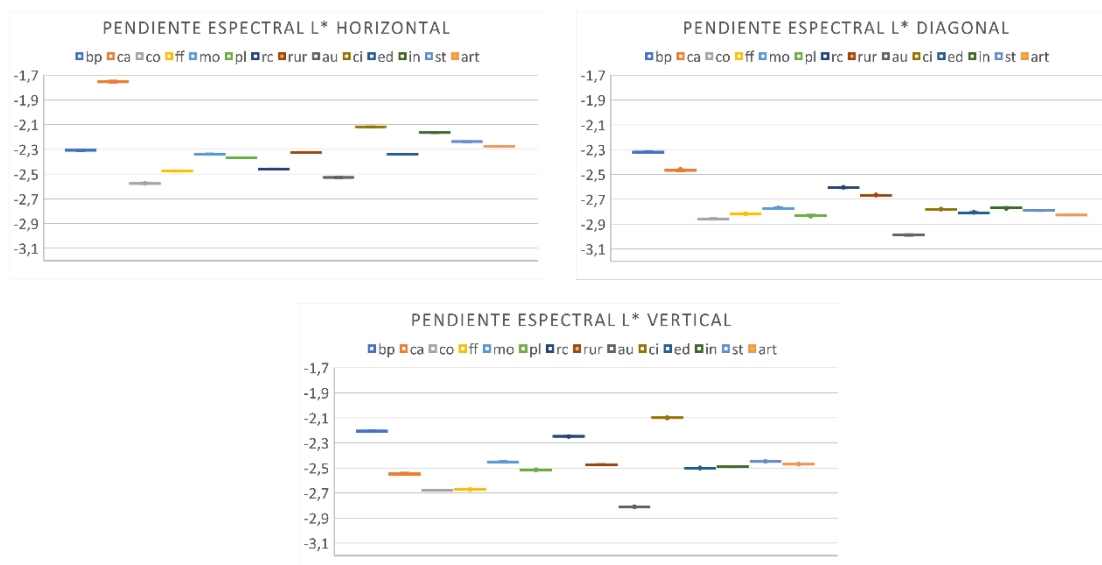
Cuando mostrábamos en una escala logarítmica la potencia espectral en función de la frecuencia espacial se obtenía una recta cuya pendiente (pendiente espectral) variaba según la orientación, el canal y la categoría semántica entre -3,02 y -1,68 en el conjunto de nuestras imágenes originales. Al estudiar el comportamiento de este descriptor bajo los distintos iluminantes caracterizados por su temperatura de color para tres orientaciones distintas ( $0^\circ$  (horizontal),  $45^\circ$  (diagonal) y  $90^\circ$  (vertical)) y las distintas categorías semánticas, observamos que los canales cromáticos parecen mostrar cierta dependencia con el iluminante variando sus rangos entre 0,027 y 0,25 para el canal  $a^*$  y entre 0,087 y 0,664 para el canal  $b^*$ . Al considerar las variaciones relativas de las pendientes con respecto a su valor absoluto máximo, éstas son casi independientes de la orientación (alrededor de un 4,5%) para el canal  $a^*$  siendo inferior la variación en diagonal (8,85%) que en las direcciones horizontales y verticales (13,96 % y 11,62% respectivamente) para el canal  $b^*$ , comportándose con menor dependencia de la CCT del iluminante pero de manera similar las imágenes de entornos humanos. Todas ellas muestran una rápida variación para temperaturas inferiores a 3934 K, ralentizándose a partir de los 9614 K y pudiendo mostrar, según las categorías, extremos relativos en dicho intervalo, tal y como se muestra en la figura 7.12. Para los canales cromáticos el valor absoluto de la pendiente correspondiente a la dirección  $45^\circ$  es mayor que el correspondiente a las direcciones horizontal y vertical. A pesar de estas variaciones, el valor de la pendiente se encuentra en el mismo rango que en las imágenes originales por lo que sigue siendo válido el análisis realizado del descriptor en el capítulo anterior de las imágenes originales para el comportamiento del mismo bajo los distintos iluminantes luz día.

El análisis de la comparación estadística de la media de las distribuciones de las pendientes en las distintas orientaciones en el plano luminancia, mediante el test de Kruskal – Wallis, mostraba que estas podían considerarse significativamente iguales para los distintos iluminantes caracterizados por la CCT para un intervalo de confianza del 95%. Esta influencia prácticamente nula de la CCT del iluminante para el canal  $L^*$  es la que se muestra en los diagramas de cajas y bigotes de la figura 7.13 en los que la representación para cada categoría queda reducida al valor de la correspondiente mediana mostrando la anisotropía de los valores de las pendientes siendo mayores en valor absoluto nuevamente las correspondientes a la orientación de  $45^\circ$  (diagonal) que las de las otras dos orientaciones más parecidas entre sí.





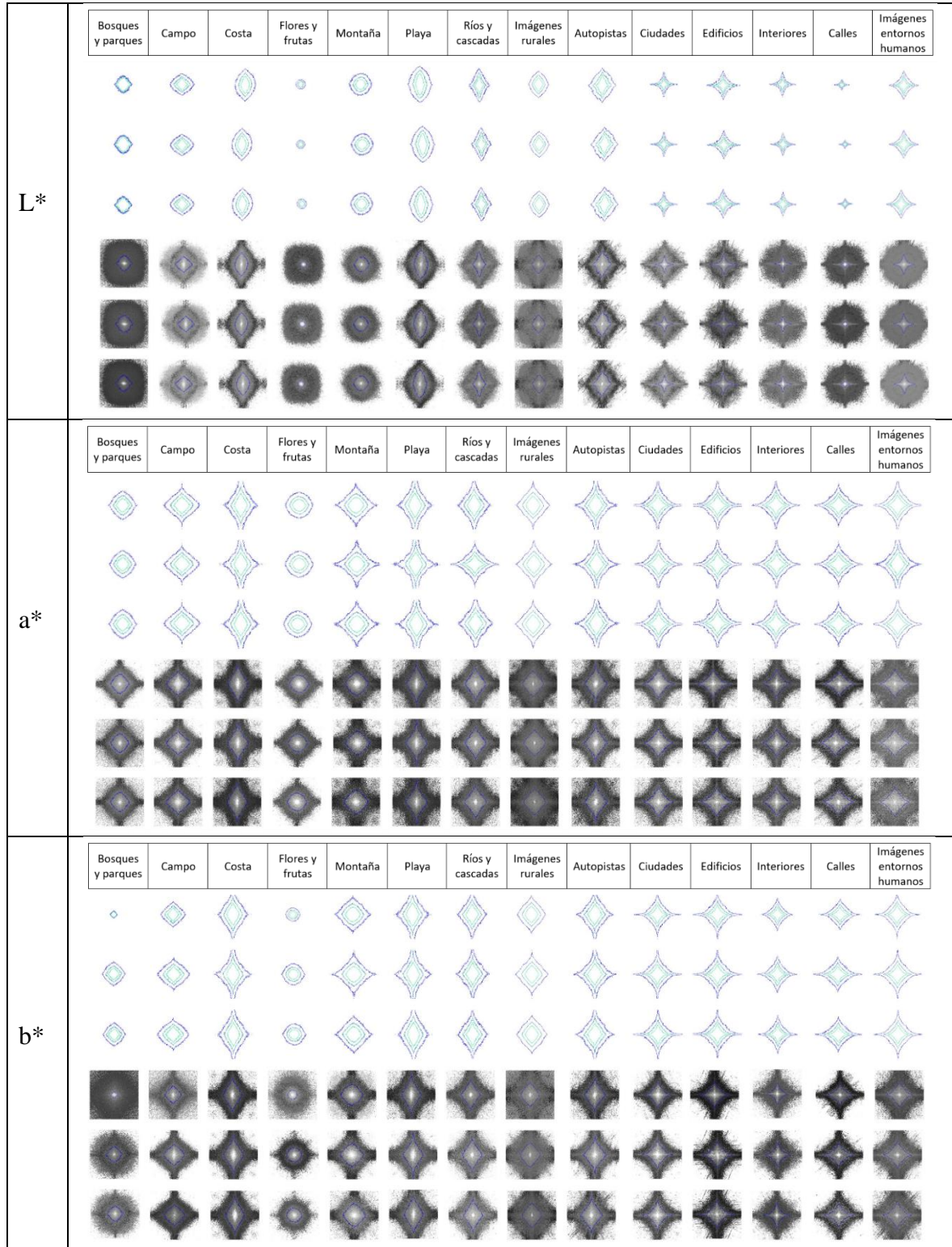
**Figura 7.12:** Pendiente espectral de las componentes cromáticas ( $a^*$  y  $b^*$ ) para tres orientaciones horizontal ( $0^\circ$ ), diagonal ( $45^\circ$ ) y vertical ( $90^\circ$ ) para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante.



**Figura 7.13:** Diagramas de cajas y bigotes de las pendientes espectrales del canal luminancia ( $L^*$ ) en función de la temperatura de color del iluminante para las distintas categorías semánticas (bp: bosques y parques, ca: campo, co: costa, ff: flores y frutas, mo: montaña, pl: playa, rc: ríos y cascadas, rur: imágenes rurales, au: autopistas, ci: ciudades, ed: edificios, in: interiores, st: calles y art: imágenes de entornos humanos) y tres orientaciones  $0^\circ$  (horizontal),  $45^\circ$  (diagonal) y  $90^\circ$  (vertical).

Finalmente, representamos las firmas espectrales (gráfico de isolíneas que al cortar el espectro de potencia con distintos planos horizontales contienen el 60% (verde claro), el 80% (verde oscuro) y 90% (azul) de la energía total del espectro) así como sus mapas de distribución (representación de los contornos de todas las imágenes de la clase con un determinado grado de transparencia en un dominio compartido) de todas las categorías semánticas para tres iluminantes correspondientes a la temperatura de color de 2735 K, 6498 K y 25889 K. Las estructuras geométricas obtenidas en la figura 7.14, al mantener ciertas cantidades de energía, muestran que estas son prácticamente independientes del iluminante para el canal luminancia, mientras que para los dos canales cromáticos se producen cambios en los detalles y la textura de la imagen (variaciones en las frecuencias altas correspondientes a direcciones distintas de  $0^\circ$  y  $90^\circ$  claramente observadas en los mapas de distribución), con una mayor dispersión de las mismas para los iluminantes correspondientes a las temperaturas más extremas que para 6478 K, más concentradas en la orientación de  $45^\circ$ , mientras que la estructura principal de la imagen (asociada a las bajas frecuencias) permanece por lo general estable, así como las frecuencias más altas en las direcciones horizontales y verticales.



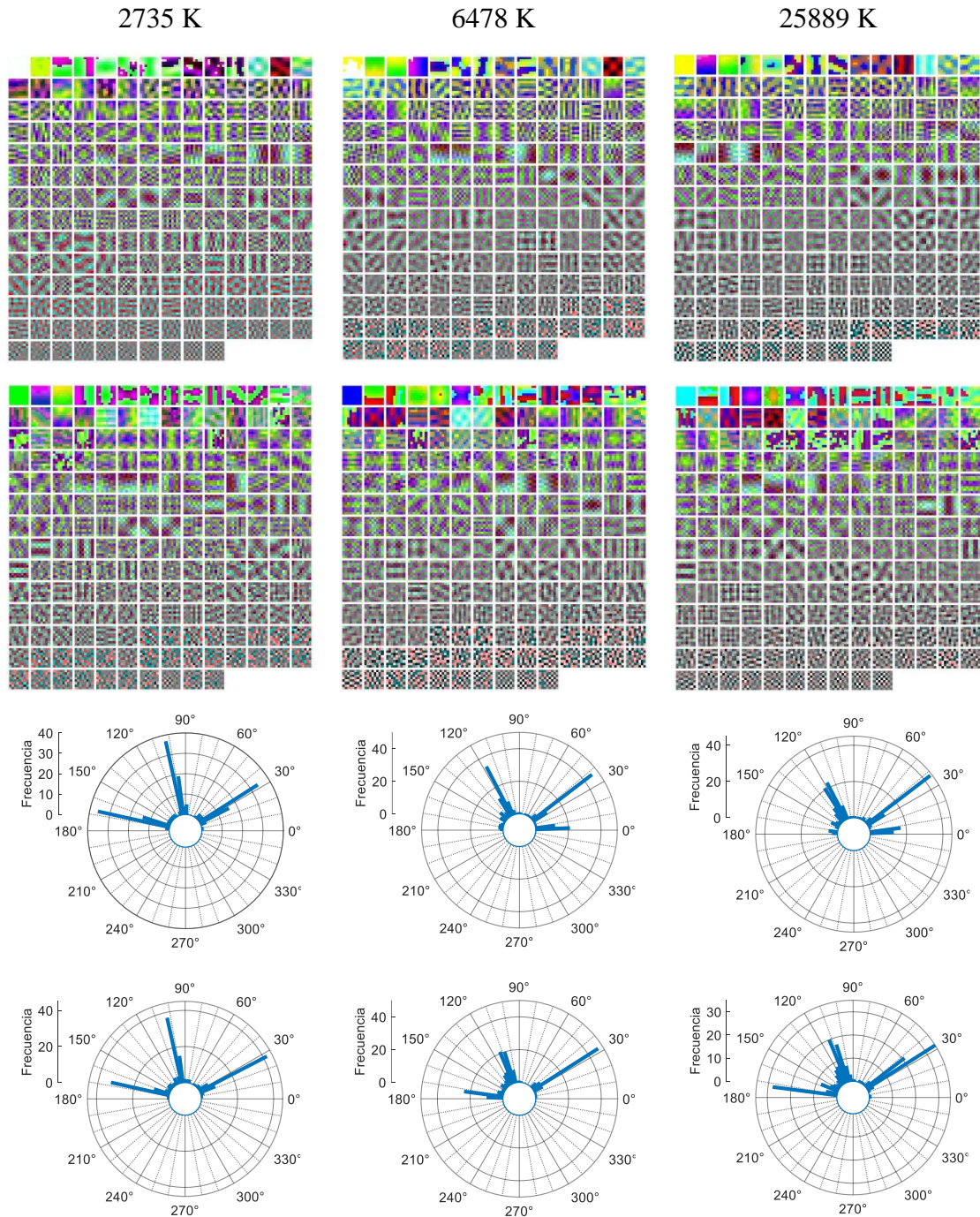


**Figura 7.14:** Firma espectral (las tres primeras filas de cada figura) y mapas de distribución (las tres últimas filas de cada figura) para los tres canales  $L^*$ ,  $a^*$  y  $b^*$  y las diferentes categorías semánticas para tres iluminantes caracterizados por las temperaturas de color 2735 K (la primera fila y la cuarta fila), 6478 K (la segunda fila y la quinta fila) y 25889 K (la tercera fila y la sexta fila).

Una herramienta que permitía el análisis de la estructura espacio-cromática de las imágenes era el análisis de componentes principales (PCA) de trozos cuadrados seleccionados aleatoriamente de las imágenes de todas las categorías semánticas. Para cuantificar las direcciones de alineación de los píxeles de dichas componentes principales de los trozos individuales con respecto al semieje positivo horizontal, ejecutamos nuevamente el PCA en las coordenadas de todos los píxeles de cada componente principal considerándolas en un plano isoiluminante calculando el arco tangente de las coordenadas de la primera componente principal.

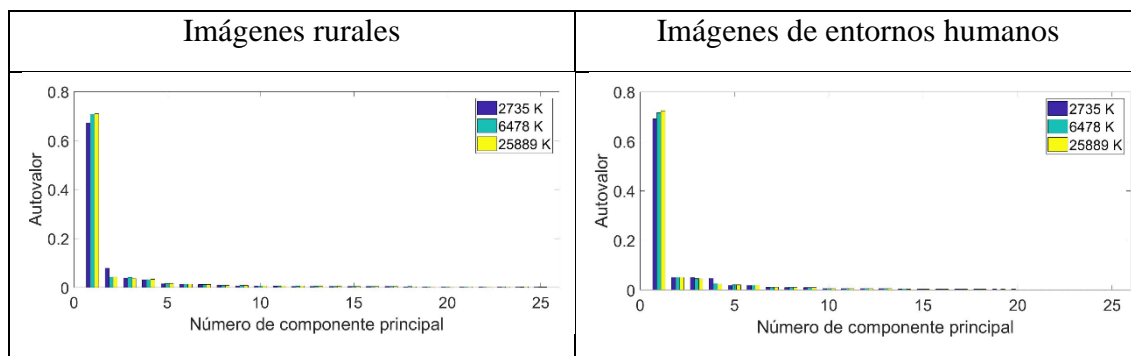
La figura 7.15 muestra en la parte superior las 192 componentes principales obtenidas de la selección de los trozos de imágenes rurales (primera fila) y de imágenes de entornos humanos (segunda fila) en el espacio de color CIELAB bajo tres iluminantes (CCTs correspondientes a 2735 K, 6478 K y 25889 K) y en la parte inferior los histogramas circulares de los ángulos (entre 0° y 180°) de las direcciones obtenidos para las mismas. Vemos que las orientaciones, tanto para las imágenes rurales como de entornos humanos, presentan tres direcciones predominantes: una primera dirección que se corresponde con valores próximos a una alineación horizontal (10° para las imágenes rurales bajo un iluminante con CCTs medias o altas o 170° en los demás casos) y dos direcciones que se aproximan a la perpendicularidad entre sí a medida que la CCT del iluminante aumenta y que varían entre los 30° - 40° y los 100° - 120° correspondiendo estas a los mecanismos de oposición de color; estas direcciones coinciden con las que presentaban las componentes principales de las imágenes originales. También para CCTs medias o altas del iluminante aparecen más componentes principales con ángulos diferentes a los más frecuentes señalados, lo que indica una mayor correlación entre las direcciones color oponentes para los iluminantes con temperaturas mayores.

Dado que el autovalor asociado a cada autovector de la descomposición en vectores propios de la matriz de covarianza de los datos explica el porcentaje de varianza asociado a esa componente principal, en la figura 7.16 se muestra el histograma correspondiente a los valores propios de las 25 primeras componentes principales de los trozos de imágenes bajo los tres iluminantes ya comentados; como era de esperar, para cada iluminante este muestra un decrecimiento rápido en las primeras componentes explicándose más del 90 % de la varianza de los datos con las 13 primeras componentes para imágenes rurales y 9 para las de entornos humanos; además, la varianza explicada por la primera componente es mayor a medida que la temperatura del iluminante aumenta (67,3%, 70,9 % y 71,0 % para las imágenes rurales y 69%, 71,5% y 72,4% para las de entornos humanos respectivamente) y las diferencias a partir de la quinta componente son inferiores al medio punto porcentual.



**Figura 7.15:** Las 192 componentes principales correspondientes a 17500 trozos de imágenes rurales (primera fila) y de entornos humanos (segunda fila) simuladas en el espacio de color CIELAB bajo tres iluminantes: 2735 K, 6478 K y 25889 K, histogramas circulares correspondientes al ángulo (entre  $0^{\circ}$  y  $180^{\circ}$ ) que forman dichas componentes con el semieje horizontal positivo para las imágenes rurales (tercera fila) y de entornos humanos (cuarta fila)





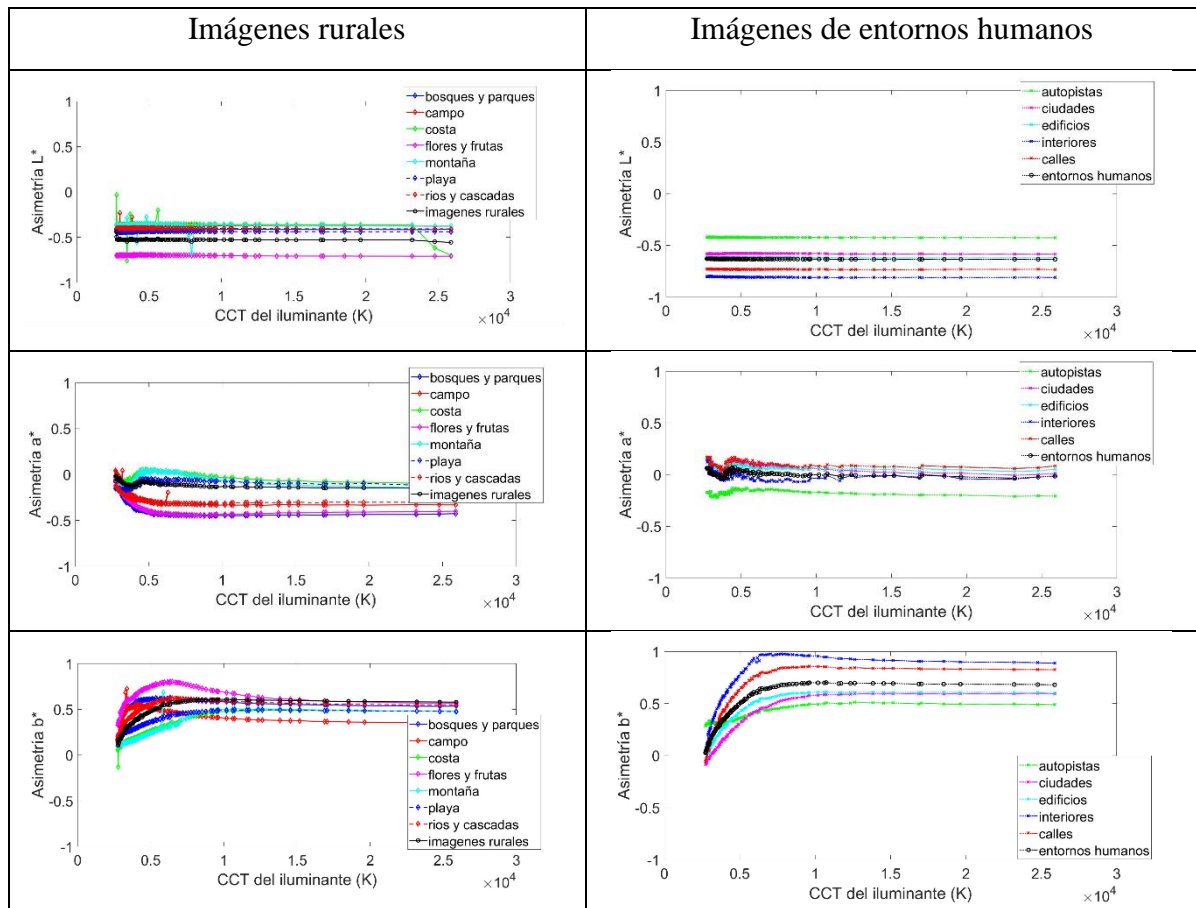
**Figura 7.16:** Autovalores normalizados de las 25 primeras componentes principales correspondientes al PCA de trozos de imágenes rurales y de entornos humanos bajo tres iluminantes: 2735 K, 6478 K y 25889 K .

Las ligeras variaciones obtenidas, tanto en las direcciones de las componentes principales como en el valor de sus autovalores al aplicar el PCA a trozos de imágenes para estudiar su estructura espacio-cromática, apuntan en la misma dirección que los resultados señalados por Chiao et al. [Chi00] y Nascimento et al. [Nas16] al aplicar este análisis a los espectros de la señal de color de imágenes naturales hiperespectrales terrestres bajo distintos iluminantes o condiciones de iluminación local diferentes. No obstante, estas pequeñas diferencias, desde el punto de vista de la significación estadística (test de Kruskal – Wallis) para un intervalo de confianza del 95 %, no son suficientes para considerar iguales las orientaciones de las componentes principales de los trozos de imagen bajo distintos iluminantes aunque sí lo son para considerar iguales los autovalores.

### 7.3 Estadística de orden superior.

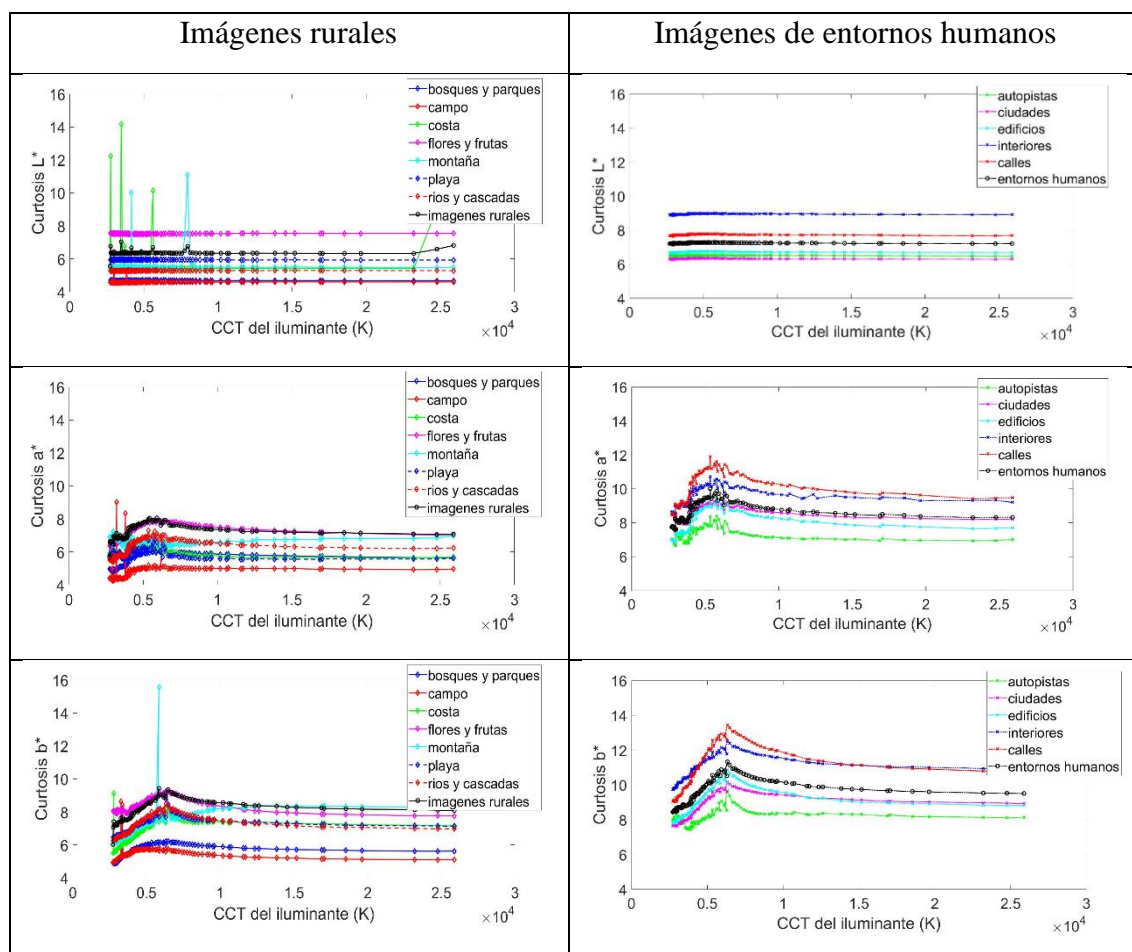
Los momentos tercero y cuarto de una distribución de datos adquieren su significación a través de los descriptores asimetría y curtosis que son los descriptores globales más simples de orden superior [Tho00, Rei01]. Con el fin de analizar la influencia del iluminante en ellos se calcularon para los datos blanqueados correspondientes a las imágenes simuladas bajo nuestros 108 iluminantes. Las dos gráficas superiores de la figura 7.17 muestran como el descriptor asimetría para el canal luminancia toma siempre valores negativos (por lo que las imágenes presentan más píxeles claros que oscuros) y apenas muestra dependencia con la temperatura del iluminante (a excepción de algunos valores atípicos para las imágenes rurales) en todas las categorías semánticas. La dependencia de los canales cromáticos muestra en todos los casos una tendencia asintótica de casi constancia con la temperatura del iluminante para temperaturas superiores a los 15000 K. El canal rojo-verde (el más simétrico de los canales) muestra valores máximos locales en el rango 4600K - 6173K para aquellas categorías de imágenes rurales con distribuciones menos asimétricas (tanto positivas

como negativas) y mínimos locales a temperaturas superiores en el resto de clases de esta categoría. En el caso de la asimetría del canal  $a^*$  para las imágenes de entornos humanos, el aparente comportamiento más errático debido a la escala menor utilizada, muestra un máximo local alrededor de una temperatura del iluminante de 4600 K, Dependiendo de la categoría ese extremo relativo se traduce (según su rango de valores) en una distribución más o menos asimétrica. Finalmente, la asimetría correspondiente al canal amarillo- azul, que toma siempre valores positivos, es la que muestra una mayor dependencia con el iluminante y un crecimiento rápido para bajas temperaturas hasta el rango 4600K - 6339K donde todas las categorías alcanzan un máximo local (más acentuado para aquellas categorías más asimétricas), decreciendo más lentamente a partir de ese extremo hasta una tendencia asintótica horizontal que muestra la escasa dependencia del descriptor con la CCT del iluminante para temperaturas altas (> 15000K).



**Figura 7.17:** Valor medio de la asimetría de la intensidad de las componentes  $L^*$ ,  $a^*$  y  $b^*$  para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante.

Al aplicar un análisis similar al descriptor curtosis obtenemos resultados parecidos; de tal manera que, todas las categorías semánticas en todos los canales muestran una curtosis media positiva y mayor en las categorías de entornos humanos. Una práctica independencia de su valor para el canal luminancia (en todo el rango de temperaturas) y para los canales cromáticos para temperaturas altas ( $>15000\text{K}$ ) con respecto a la CCT del iluminante y un marcado máximo relativo en ambos canales cromáticos en el rango  $5345\text{K} - 6340\text{K}$  como se puede ver en la figura 7.18. La variación relativa de la curtosis con respecto al valor máximo de la misma fue en promedio del 25,1% y del 33,75% para los canales oponentes rojo-verde y amarillo-azul respectivamente en las imágenes rurales y ligeramente inferior de un 24,1% y un 26,1% en las imágenes de entornos humanos.



**Figura 7.18:** Valor medio de la curtosis de la intensidad de las componentes  $L^*$ ,  $a^*$  y  $b^*$  para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante.

La aplicación de un test estadístico de comparación múltiple de medias señala que para un nivel de confianza del 95% hay suficiente evidencia para considerar iguales los valores de ambos descriptores (asimetría y curtosis) para cualquiera de las categorías en el canal luminancia bajo las distintas condiciones de iluminación.

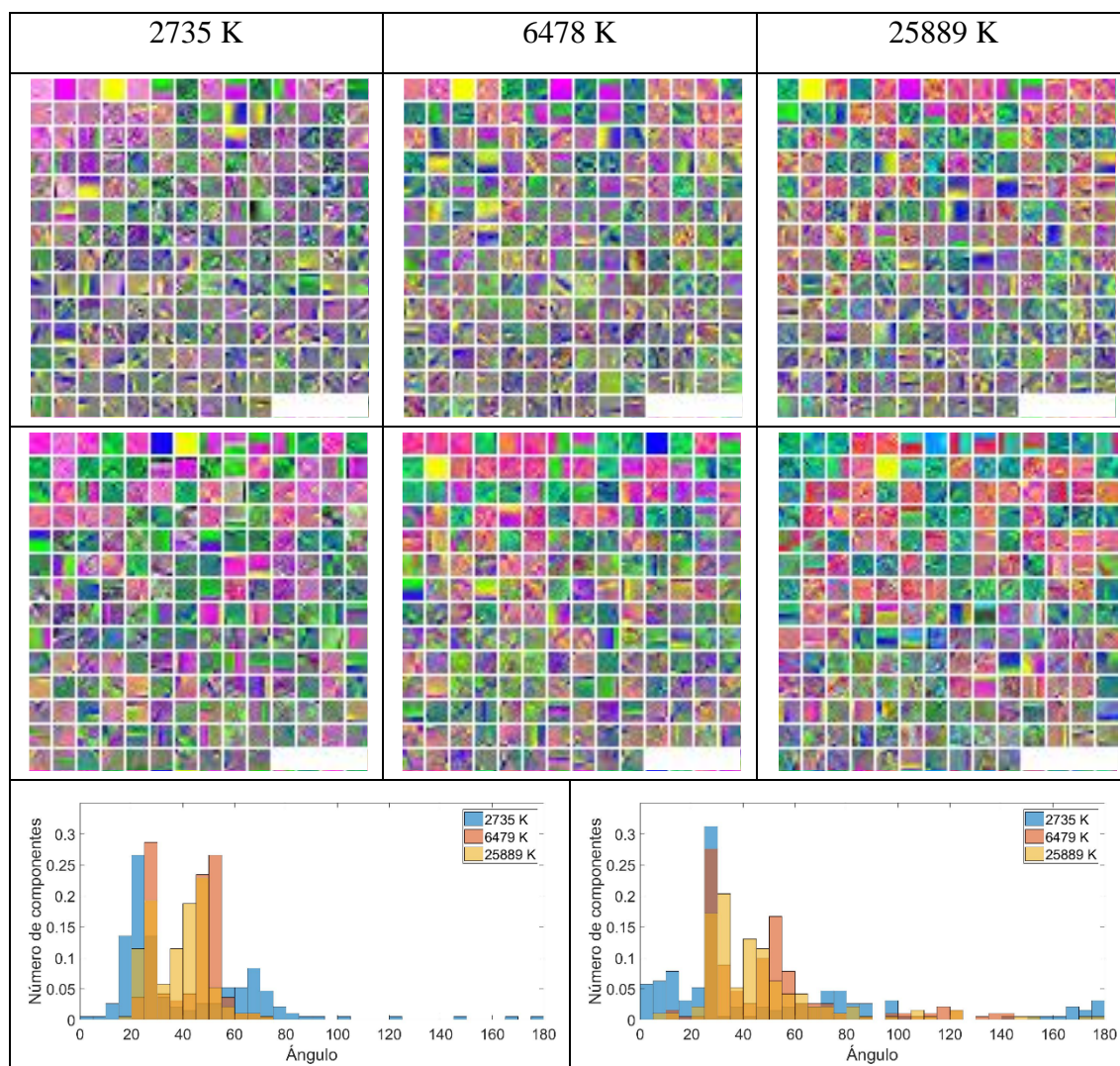
### **7.3.1 Análisis de componentes independientes.**

Un estudio similar al realizado con las componentes principales obtenidas al aplicar PCA de trozos cuadrados seleccionados aleatoriamente de las imágenes de todas las categorías semánticas, se hizo para las orientaciones de las características de las componentes independientes resultantes de aplicar ICA a los mismos trozos de imágenes. La figura 7.19 muestra nuevamente, en las dos primeras filas, las 192 características de las componentes independientes, ordenadas de izquierda a derecha y de arriba abajo en orden decreciente de su norma euclídea, obtenidas de la selección de los trozos de imágenes rurales (primera fila) y de imágenes de entornos humanos (segunda fila) en el espacio de color CIELAB bajo tres iluminantes (CCTs correspondientes a 2735 K, 6478 K y 25889 K). En la última fila de la figura 7.19 (a la izquierda para imágenes rurales y a la derecha para las de entornos humanos) se representa los histogramas comparativos de los ángulos de la primera dirección principal obtenida para cada una de las características y las distintas temperaturas de color de los iluminantes empleados.

En relación a estos histogramas vemos que, en todos los casos, las características de las componentes independientes del ICA presentan dos orientaciones más frecuentes alrededor de los  $25^\circ$  y los  $50^\circ$  ( $70^\circ$  para las temperaturas bajas del iluminante) con respecto al semieje positivo del canal oponente rojo-verde ( $a^*$ ) a la que se podría añadir los  $0^\circ/180^\circ$  en el caso de imágenes artificiales bajo el iluminante caracterizado por una CCT de 2735 K. No obstante, el ángulo de la primera componente principal de las características obtenidas al aplicar ICA muestra para las imágenes rurales y temperaturas medias y altas que puede tomar cualquier dirección en el primer cuadrante (entre  $20^\circ$  y  $75^\circ$ ) y que este rango se extiende a todos los valores posibles ( $0^\circ-180^\circ$ ) para imágenes rurales bajo iluminantes con CCT bajas y para las imágenes de entornos humanos bajo cualquier temperatura del iluminante.

Ampliando los resultados de las orientaciones de las características de las componentes independientes a todo el rango de temperaturas del iluminante y aplicándole el test de Kruskal – Wallis no se obtiene suficiente evidencia estadística para considerarlos iguales para un nivel de confianza del 95%.



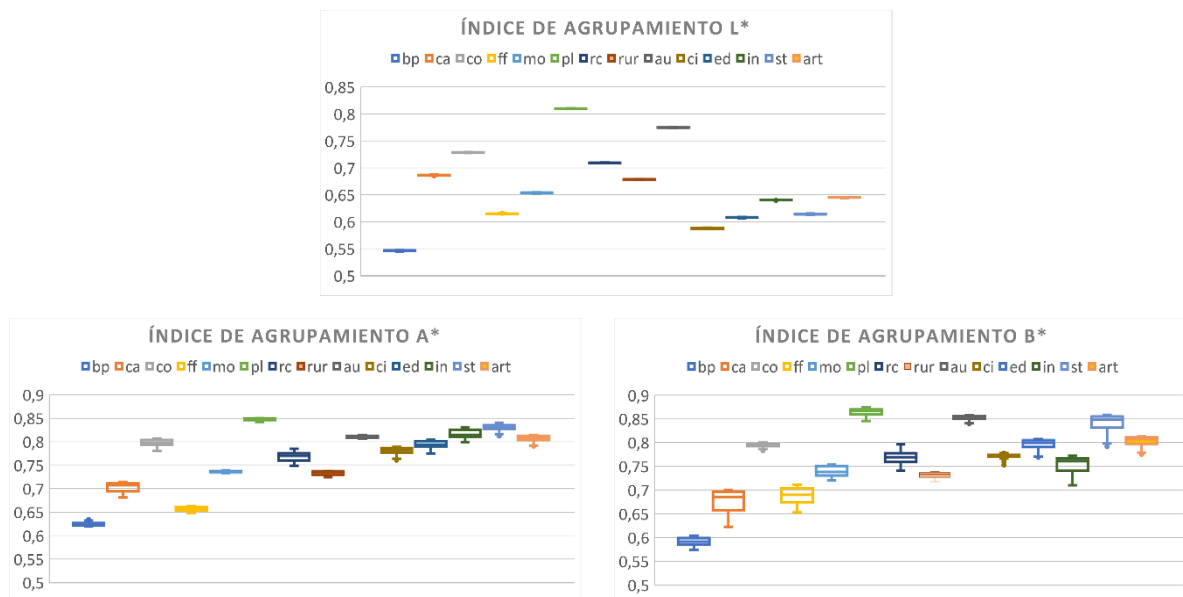


**Figura 7.19:** Las 192 características de las componentes independientes correspondientes a 17500 trozos de imágenes rurales (primera fila) y de entornos humanos (segunda fila) simuladas en el espacio de color CIELAB bajo tres iluminantes: 2735 K, 6478 K y 25889 K (parte superior) e histogramas normalizados comparativos para los tres iluminantes de los ángulos que forman dichas componentes con el semieje horizontal positivo (imágenes rurales a la izquierda, imágenes de entornos humanos a la derecha).

### 7.3.2. Índice de agrupamiento.

Finalmente, el comportamiento del descriptor índice de agrupamiento frente a la temperatura de color del iluminante aparece reflejado en la figura 7.20. En el diagrama de cajas y bigotes de la misma vemos que el índice de agrupamiento para el canal de luminancia permanece prácticamente independiente de la temperatura de color del iluminante (rango de variación entre 0,001y 0,003), siendo el canal color oponente

amarillo-azul b\* el que experimenta una mayor variación (rango entre 0,017 y 0,077, en porcentaje con respecto al valor máximo de cada categoría supone una variación media de 5,5% para imágenes rurales y un 5,09% para las de entornos humanos) frente al canal a\* color oponente rojo-verde (rango entre 0,006 y 0,036, con una variación media en tanto por ciento del valor máximo del 2,59% y el 3,03% de las dos categorías de imágenes respectivamente); además, estas variaciones de ambos canales se producen a temperaturas del iluminante pequeñas, siendo prácticamente constantes en ambos canales para todas las categorías a partir de 14 553 K y presentando muchas de ellas un extremo relativos en el rango de temperaturas entre 3020,6 K y 7205,5 K.



**Figura 7.20.-** Diagramas de cajas y bigotes del valor medio del índice de agrupamiento en función de la CCT del iluminante para cada uno de los canales luminancia L\* y cromáticos a\* y b\* para las distintas categorías semánticas (bp: bosques y parques, ca: campo, co: costa, ff: flores y frutas, mo: montaña, pl: playa, rc: ríos y cascadas, rur: imágenes rurales, au: autopistas, ci: ciudades, ed: edificios, in: interiores, st: calles y art: imágenes de entornos humanos).

El test de Kruskal – Wallis, aplicado a las distribuciones de índices de agrupamiento correspondientes a cada categoría y para cada iluminante en los distintos canales L\*, a\* y b\*, muestra que hay suficiente evidencia estadística para considerar iguales dichos índices salvo los correspondientes a las categorías “campo” y “calles” en el canal b\* para un nivel de confianza del 95%.



---

## Capítulo 8

# DESCRIPCIÓN ESTADÍSTICA DE LOS COLORES RELEVANTES EN ESCENAS NATURALES

---

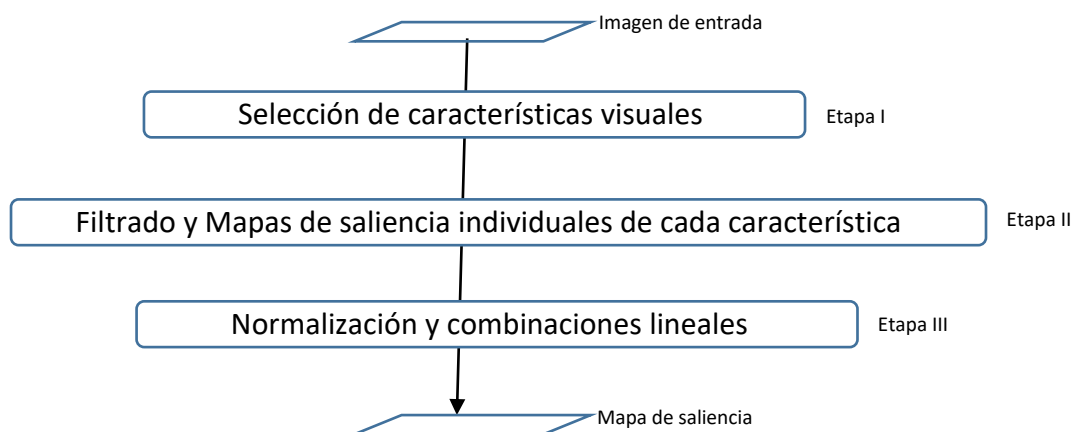
En el estudio de los descriptores estadísticos de primer orden del capítulo 4 se introdujo el número de colores discernibles como el estadístico que normalmente se utiliza para cuantificar la diversidad cromática de las imágenes. Aunque la estimación del número de colores discernibles es un problema clásico y se considera que hay alrededor de 2 millones bajo el iluminante D65 [Mas13], al estimar este número para el conjunto de nuestras imágenes utilizando el método de Linhares et al. [Lin08], obtuvimos un rango entre 1100 y 72000 colores discernibles (ver Tabla 4.3). En el método implementado segmentamos el espacio de color CIELAB en subvolúmenes distinguibles (en este caso cubos unidad de lado 1 unidad CIELAB), los valores  $L^*$ ,  $a^*$  y  $b^*$  de cada píxel se redondearon al entero mayor o igual que su valor para que todos los valores que cayeran dentro de cada cubo unidad quedaran agrupados en el vértice correspondiente al valor mayor de los tres canales y, finalmente, contamos el número de estos cubos (colores discernibles) que contenían al menos la representación cromática de un píxel de la imagen.

No obstante y a pesar de la reducción entre 2 y 3 órdenes de magnitud del número de colores discernibles estimados por este procedimiento y a la capacidad del sistema visual de percibir diferencias de color con una gran resolución en escenarios complejos, no es posible que un observador sea capaz de detectar todos los colores que producen una diferencia de 1 unidad  $\Delta E^*_{ab}$  por lo que generalmente se centrará sólo en una pequeña fracción del gran número potencial de colores discernibles cuando mira imágenes naturales o de entornos humanos u obras pictóricas. Lo que no está claro todavía es la relación entre el número de colores discernibles y el de los colores que atraen realmente la atención del observador [Nie18].

## 8.1 Regiones salientes y diversidad cromática

Está claro que, cuando un observador examina una determinada imagen, las características visuales de bajo nivel de ciertas áreas de la misma atraen su atención [Itt00]. Estas áreas que fijan la atención del observador son las llamadas regiones salientes. Basándose en el principio de que puntos de la escena serán salientes si difieren en ciertas características (bordes, contrastes, luminancia, color) de sus vecinos, Koch y Ullman [Koc85] propusieron un modelo bottom-up plausible biológicamente conocido como mapa de saliencia.

Todos los modelos de saliencia bottom-up clásicos pueden resumirse en tres etapas: selección de diferentes características visuales pertinentes (tales como color, bordes, contraste, etc), cálculo de un mapa de saliencia para cada característica (empleando un proceso de filtrado que emula al sistema visual humano) y un paso de integración final de todos los mapas de saliencia por separado en una única salida que señala esas regiones que atraerían la fijación y atención del observador (ver figura 8.1) [Nie18].



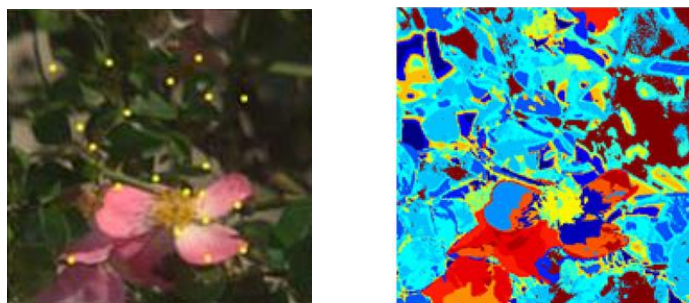
**Figura 8.1:** Esquema general del modelo de saliencia clásico bottom-up

Nieves y Romero estudiaron heurísticamente en [Nie18] los efectos de la saliencia visual (utilizando distintos modelos: Itti et al [Itt98], Harel et al [Har06], Tian et al [Tia10] y Canosa [Can05]) en la estimación del número de colores discernibles en escenas naturales (aplicable también a obras pictóricas o a escenas que contengan objetos humanos) para lo cual definieron los descriptores: colores discernibles salientes y colores salientes destacables.

Partiendo de la hipótesis de que las regiones salientes atraerán la atención del observador, este será capaz de determinar sólo el número de colores discernibles de los píxeles de estas áreas. Es este número que indica la diversidad cromática saliente al que Nieves y Romero [Nie18] llamaron colores discernibles salientes. Para todos los modelos

obtuvieron en su trabajo una reducción superior al 40% al considerar solo los colores discernibles salientes frente al número de colores discernibles obtenidos de la imagen al completo si no hubiéramos tenido en cuenta las regiones salientes. En cuanto a la distribución cromática de los colores o gama de color, no todos los algoritmos estudiados muestreaban uniformemente la gama de color de la imagen original [Nie18].

Esta aproximación heurística de la diversidad cromática saliente puede usarse para analizar qué colores atraerían naturalmente la atención de los observadores y crear imágenes segmentadas automáticamente en función de dicha diversidad. Para lo cual el mapa de saliencia es binarizado y segmentado usando un simple clustering “k-means” de los colores discernibles salientes. El algoritmo de clustering devuelve para cada color saliente una imagen etiquetada por un índice de agrupamiento y en la que los centroides de cada región etiquetada son los llamados colores salientes destacables (colores conectados en cada área saliente de la imagen). El número de colores salientes destacables se estimó contando el número de centroides en la imagen etiquetada. En la figura 8.2 vemos un ejemplo de Nieves y Romero [Nie18] en el que la imagen de la izquierda muestra la imagen original con las posiciones superpuestas (puntos amarillos) de los colores salientes destacables determinados por el algoritmo de Harel et al [Har06] y, en la imagen de la derecha, la segmentación de la imagen original asociada a esos colores salientes destacables.



**Figura 8.2:** Ejemplo procedente de Nieves y Romero [Nie18] de localización sobre la imagen original de los colores salientes destacables (puntos amarillos) obtenidos por el algoritmo de Harel et al. [Har06] (izquierda) y la imagen correspondiente segmentada según esos colores salientes destacables (derecha).

Para algunos de los métodos de saliencia utilizados como el de Itti et al. [Itt98] o el de Harel et al. [Har06] se obtuvieron un número reducido de colores salientes destacables (entre 17 y 25) con una desviación típica relativamente pequeña (entre 3 y 7). Massaro et al. [Mas12] usando un sistema de seguimiento de mirada calcularon el número de clusters involucrados en juicios estéticos de las obras pictóricas y encontraron alrededor de 13 clusters cuando se observaron cuadros que representan la naturaleza, lo



que no está muy alejado al número de colores salientes destacables para ciertos algoritmos de saliencia y que puede deberse al diferente objeto de estudio: representaciones pictóricas de la naturaleza e imágenes naturales.

Estos resultados sugieren cómo, a pesar de la complejidad relativa del contenido espacio-cromático de las escenas, estos algoritmos son capaces de realizar una buena segmentación de la imagen basada en su contenido, es decir, basada en sus colores salientes a pesar de los posibles errores cualitativos producidos por luces intensas o sombras en la escena original.

## **8.2 Cuantificación del número de colores relevantes en escenas pictóricas**

Son varios los algoritmos que han tratado de extraer los colores representativos de una imagen. Entre ellos encontramos algoritmos de visión del color basados en técnicas de clustering (agrupamiento) que necesitan, para ser implementados, una estimación previa de dichos colores o grupos [Bez81] o basados en los picos máximos de los histogramas de frecuencia de distintas características cromáticas para determinar qué píxeles tienen una relevancia mayor y asociarlos con las regiones cromáticas significativas de la imagen [Del05,Mor07].

También podemos encontrar los llamados “algoritmos de nombrado de colores” (color naming algorithms) que intentan establecer una caracterización discreta del número de colores que aparecen en una imagen, basándose en una terminología del color básica, concepto introducido por Berlin y Kay [Ber69]. Con la mayoría de estos algoritmos se obtiene un rango de colores parecido que dependiendo del léxico de colores oscila entre 11 y 15 nombres para definir todas las categorías lingüísticas de colores [Lin06, Lin09, Lin14], las 27 regiones categóricas distintas ajustadas dentro del espacio de color RGB por Griffin y Mylonas [Gri19], el uso de solo alrededor 30 nombres de color en inglés hablado [Der95] o los 50 territorios categóricos distintos encontrados en el espacio de color con anterioridad por Chapanis [Cha65].

Como ya comentamos en el apartado anterior, Nieves y Romero [Nie18] estudiaron, de manera heurística, los colores que pueden atraer la atención visual y cómo la saliencia visual puede constituir un filtro que limite el número de colores en el que un observador pueda fijar su atención. La aplicación de un modelo de saliencia basado en un procesamiento visual bottom-up permitía el cálculo de los descriptores colores salientes discernibles y colores salientes destacables. Por el contrario, la estimación de los colores relevantes que aparecen en una escena sería el resultado de un proceso conducente por una tarea top-down.

En una propuesta en la que no es necesario introducir categorías de color predefinidas ni determinar qué áreas son visualmente salientes en la escena, Nieves et al.

[Nie20] establecen, basándose en el contenido colorimétrico de cada imagen, el concepto novedoso de colores relevantes, definido como los colores discernibles categóricos que describen la diversidad cromática de una obra pictórica. Un algoritmo computacional es aplicado a las miles de pinturas de diferentes pintores y estilos de la base de datos de Khan et al. [Kha14] lo que les permite estimar una paleta de colores fiable para cada una de ellas.

La categorización de los colores discernibles se realizó mediante cuatro parámetros:

- El tamaño del cubo en el cual se divide el espacio CIELAB, que establecieron en 20 unidades de arista (frente a la arista unidad del método clásico)[Lin08]
- Un porcentaje (o umbral general) mínimo del total de píxeles en un cubo para considerarlo relevante que se estableció en el 3% del número de píxeles totales
- Para tener en cuenta que Milojevic et al. [Mil18] encontraron que los colores más saturados pueden actuar como predictores de cómo un observador categorizaría la distribución de color de objetos naturales, se consideraron como colores relevantes también los que correspondían a aquellos cubos que teniendo menos del 3% de los píxeles totales mostraban valores de  $L^*$  superiores a 80, o valores de  $C^*$  superiores al percentil 50 de la imagen.

Los colores considerados como relevantes se determinan como los valores medios de los colores de los píxeles en cada cubo seleccionado. El número de colores relevantes (NRC) medio para todos los cuadros que analizaron fue de 18 (con una SD de 6). Una vez que se han determinado los colores relevantes que aparecen en el cuadro, el algoritmo asigna a cada píxel de la imagen un color relevante (dependiendo de la distancia euclídea entre este color relevante y el color que debería asignarse originalmente al píxel) permitiendo una segmentación colorimétrica de la misma. En la figura 8.3 vemos un ejemplo de Nieves et al [Nie20] en el que la imagen de la izquierda de la fila superior muestra la pintura original, en la fila inferior la paleta de colores de la misma (los colores relevantes) y en la imagen de la derecha de la fila superior la imagen segmentada a partir de dichos colores.

En este proceso (top-down) de determinación de los colores relevantes de una escena se simulaba cómo los observadores mirarían a aquellas zonas que son más relevantes tanto por ser coloriméricamente discernibles como por describir la diversidad cromática de la escena.

Para proporcionar un respaldo psicofísico a este modelo computacional, Nieves et al. [Nie21] desarrollaron un experimento que analizaba cómo los observadores determinan los colores relevantes a partir de sus propias impresiones subjetivas de las imágenes observadas. En el experimento participaron seis observadores ajenos al mundo

del arte (para evitar potenciales sesgos en los resultados) con visión normal del color a los que se les presentaba un conjunto de 20 imágenes de pinturas pertenecientes al Museo del Prado [Mdp21] y otras 20 de la colección de la base de datos de Khan et al. [Kha14], usando PsychToolbox3 [Psy20] para la programación de la experiencia. En cada realización del experimento el observador debía seleccionar con el ratón un área dentro de la imagen que contuviera un color relevante, es decir, píxeles de un área cromáticamente relevante, pudiendo hacerlo tantas veces como quisiera. Una descripción completa del experimento puede verse en [Nie21].



**Figura 8.3.-** Ejemplo en una de las imágenes de la base de datos de Khan et al. [Kha14] de los resultados obtenidos al aplicar el algoritmo de Nieves et al. [Nie20]. En la fila superior izquierda la imagen original, en la fila superior derecha la imagen segmentada usando los colores relevantes obtenidos que aparecen en la fila inferior (paleta de color de la pintura).

El número de colores relevantes medio, determinado por los distintos observadores en el experimento, fue de  $21 \pm 5$  para las imágenes de obras pertenecientes al museo del Prado y de  $22 \pm 6$  para las de la base de datos de Khan, siendo este número reducido de colores representativo de los colores destacables que describían las distintas obras. Al aplicar el algoritmo computacional de Nieves et al [Nie20] a estas pinturas se obtiene un número de colores relevantes de  $19 \pm 6$  que está en concordancia con el número de colores relevantes subjetivos de los observadores del experimento. Además, cuando se compararon los conjuntos de dichos colores, aunque no se solapaban exactamente, mostraban gamas de color similares, donde la mayoría de los colores obtenidos heurísticamente tenían su color subjetivo correspondiente, a excepción de algunos colores experimentales caracterizados por valores de  $a^*$  y que no eran determinados por el algoritmo que puede resultar demasiado restrictivo según que pintura analicemos.

Como se indicaba en el capítulo 6, la aplicación del análisis de componentes independientes (ICA) a las imágenes naturales mostraba que una codificación oponente de la información cromática de estas era eficiente, con direcciones de oposición preferentes diferentes además de las direcciones cardinales clásicas rojo – verde y amarillo - azul [Kel13, Lee02]. Para comprobar si las impresiones subjetivas de los observadores del experimento venían derivadas de un procesamiento óptimo y eficiente de las instancias más representativas de las imágenes de las pinturas, a estas se les aplicó un ICA usando el paquete FastICA de Matlab [Hyv04]. Como se esperaba por estudios anteriores [Nie18, Nod08, VGe20] y por resultados previos (ver capítulos 6 y 7), la mayoría de las bases obtenidas con ICA mostraban una distribución espacial orientada que podía dividirse en características homogéneas, color oponentes y acromáticas. Las orientaciones de las mismas, cuando se representan en los diagramas de cromaticidad, muestran, como en el caso de imágenes naturales [Kel13], la preferencia de algunas direcciones oponentes no restringidas a los ejes cromáticos rojo – verde y amarillo – azul sugiriendo que los cuadros proporcionan algunas distribuciones de colores preferidos.

Para analizar si los colores relevantes obtenidos en el experimento psicofísico podían ser instancias de los colores de las direcciones de color preferidas en el ICA, Nieves et al. [Nie21] determinaron para cada color relevante subjetivo su dirección en el diagrama de cromaticidad como el arco tangente del cociente de sus componentes BY y RG de acuerdo a las ecuaciones

$$RG = \frac{R-G}{\sqrt{2}} \quad BY = \frac{R+G-2B}{\sqrt{6}} \quad (8.1)$$

que no son más que una normalización de las ecuaciones (5.26) para las componentes cromáticas oponentes.

La relativamente buena concordancia encontrada entre los ángulos obtenidos y las direcciones de color preferidas obtenidas con el ICA, apoyaba la hipótesis de los autores de que los observadores seleccionaron el contenido de color más relevante de las pinturas siguiendo una codificación eficiente [Nie21].

### 8.3 Número de colores relevantes en imágenes naturales.

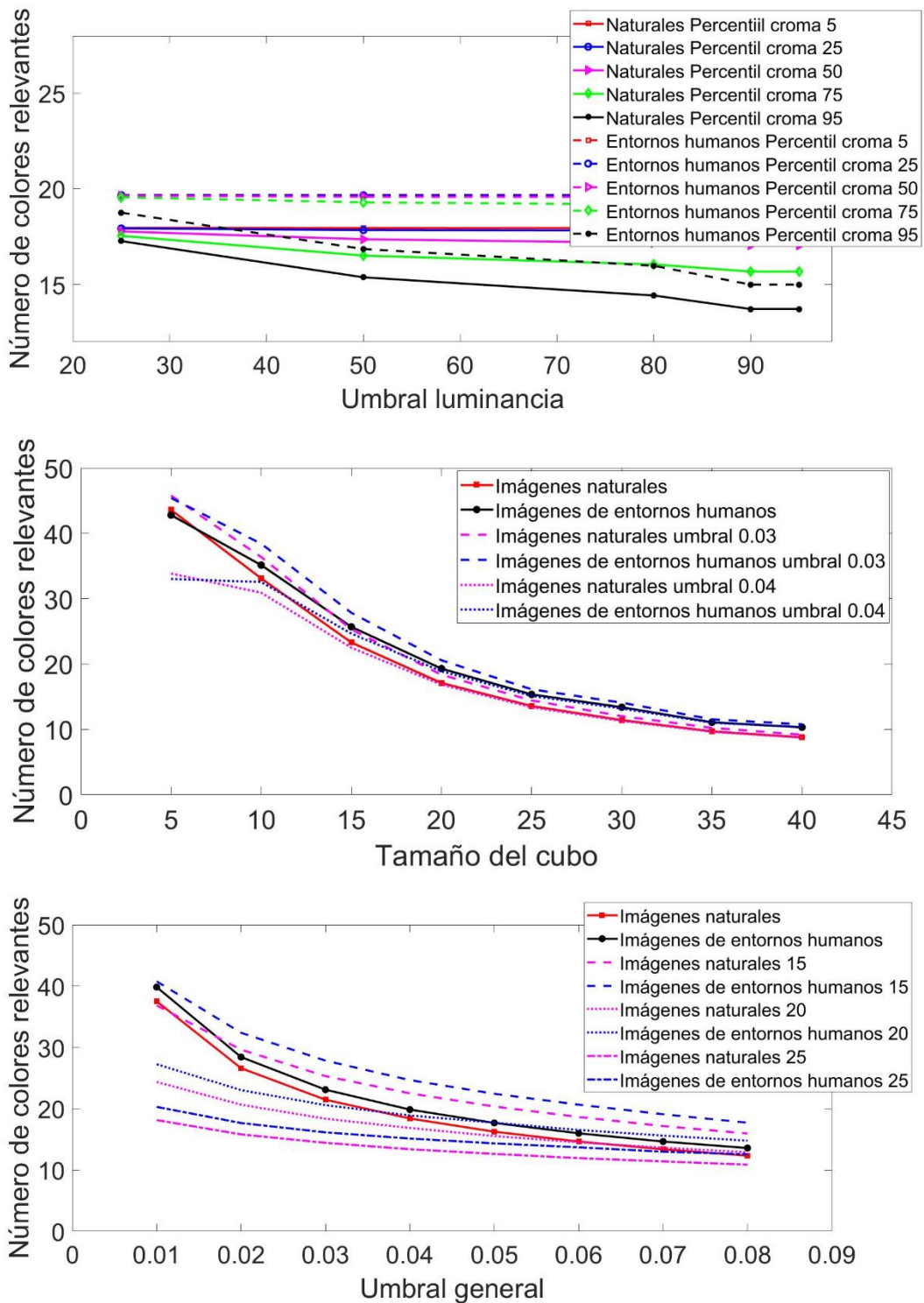
En esta sección vamos a ampliar a escenas naturales el estudio del algoritmo computacional de cálculo de colores relevantes que Nieves et al. [Nie20] definieron para imágenes pictóricas. Comenzaremos analizando los valores de los parámetros utilizados en el algoritmo para los que el número de colores relevantes muestra menor dependencia de los mismos. Seguidamente, calcularemos el número de colores relevantes a dos conjuntos de imágenes naturales diferentes (el que hemos utilizado en los capítulos anteriores del presente trabajo y el de la base de datos de McGill [Olm04]) comparando los resultados entre ellas así como con los de imágenes pictóricas del trabajo originario

de Nieves et al. [Nie20]. Finalmente, analizaremos si los colores relevantes seleccionados por el algoritmo de las escenas naturales siguen una codificación eficiente de las mismas al corresponderse o no sus direcciones en el plano cromático (de ejes RG y BY) al de las características de las componentes independientes al aplicar ICA a trozos de nuestro conjunto de imágenes.

Como indicamos en la sección anterior, el descriptor colores relevantes venía definido como los colores discernibles categóricos que describen la diversidad cromática de un cuadro. En su algoritmo para el cálculo de dichos colores relevantes estos investigadores utilizaban cuatro parámetros: el tamaño de la arista del cubo en el cual se dividía el espacio de color CIELAB, un valor umbral general consistente en un porcentaje mínimo del total de píxeles que superado en un cubo lo fija como color relevante, y un valor de luminancia y un percentil del croma del color del cubo suficiente para considerarlo saturado y tenerlo en cuenta como color relevante aunque el número de píxeles en él sea inferior al del umbral. Para establecer el valor de cada uno de ellos se ha calculado el número de colores relevantes utilizando diferentes tamaños de arista del cubo del espacio de color (5, 10, 15, 20, 25, 30, 35 y 40 unidades del espacio CIELAB), distintos umbrales generales (1%, 2%, 3%, 4%, 5%, 6%, 7% y 8% del total de los píxeles), valores de luminancia del color de cada cubo de 25, 50, 80, 90 y 95 y percentiles 5, 25, 50, 75 y 95 del croma de dicho color. La fila superior de la figura 8.4 representa la variación del número de colores relevantes medio de nuestro conjunto de imágenes rurales y de entornos humanos promediado para todos los valores de la luminancia y los percentiles del croma en función del tamaño de la arista del cubo del espacio de color. Las líneas continuas negra y roja con valores intercalados corresponden a valores promediados para todos los valores del umbral general (las llamadas en la gráfica imágenes rurales y de entornos humanos) frente a los colores relevantes promediados para un solo valor del umbral 3% o 4% del total de los píxeles en un cubo (líneas azules y magentas discontinuas y punteadas respectivamente). En la imagen vemos cómo a medida que el tamaño de la arista del cubo aumenta, el número de colores relevantes disminuye siendo menor esta disminución a medida que aumenta el tamaño de la arista del cubo. Además, para tamaños iguales o mayores a 20 el número de colores relevantes, obtenido para el promedio de los umbrales, es prácticamente el mismo que el calculado para un umbral del 4%.

La segunda fila de la figura 8.4 representa la variación del número de colores relevantes medio nuevamente promediado para todos los valores de la luminancia y los percentiles del croma en función del umbral general. Las líneas continuas negra y roja con valores intercalados corresponden a valores promediados para todos los valores del tamaño de la arista del cubo del espacio de color (las llamadas en la gráfica imágenes rurales y de entornos humanos) frente a los colores relevantes promediados para un solo valor de dicha arista de 15, 20 y 25 unidades CIELAB (líneas azules y magentas





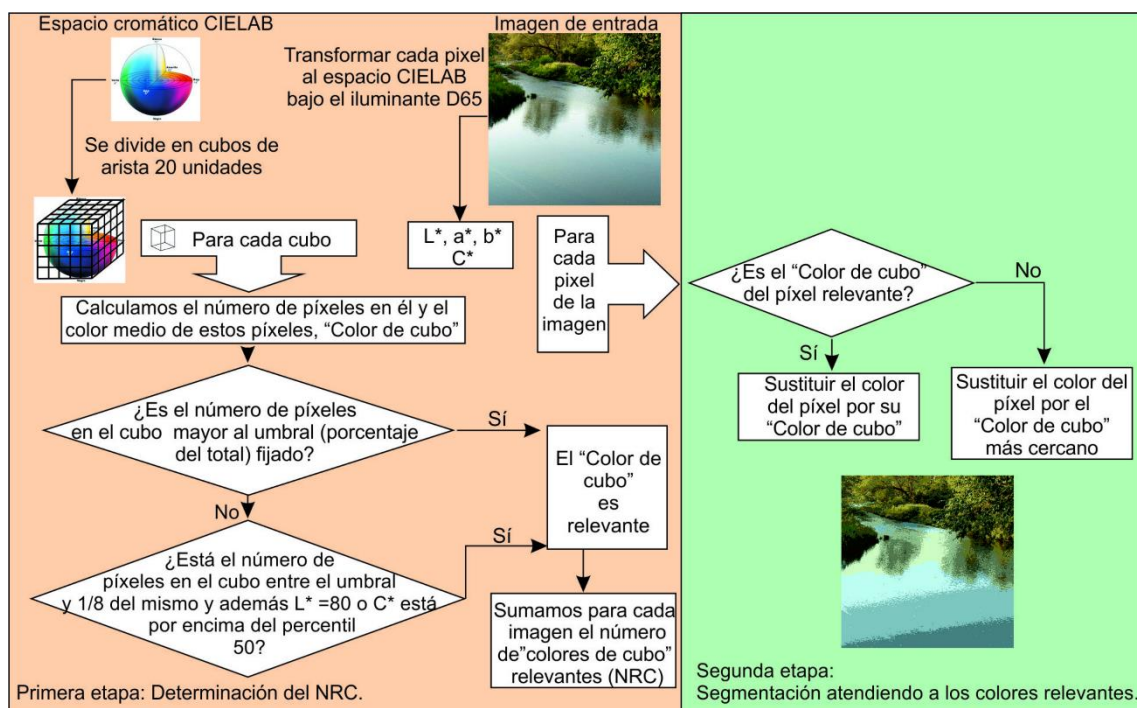
**Figura 8.4:** Análisis de la dependencia de los parámetros del algoritmo de colores relevantes. Arriba: NRC en función del tamaño del cubo del espacio de color, central: NRC en función del umbral general (porcentaje del total de los píxeles en un cubo) y abajo: NRC para un tamaño del cubo de 20 unidades y un umbral del 4% en función del valor luminancia y percentil del croma.



discontinuas, continuas y punteadas respectivamente). En la imagen vemos una dependencia con el umbral general parecida a la que mostraba el número de colores relevantes con el tamaño del cubo. Además, para un umbral mayor o igual del 4% el número de colores relevantes obtenido para un tamaño de la arista del cubo de 20 es siempre el más parecido al promediado para todos los tamaños. Estos comportamientos nos han hecho fijar los valores de estos parámetros en 20 unidades CIELAB para el tamaño de la arista del cubo y del 4% para el umbral.

Una vez dividido el espacio de color CIELAB en cubos de arista del tamaño de 20 unidades CIELAB ( $\Delta E^* = 20$ ) y transformadas las componentes RGB de cada píxel de una imagen a dicho espacio de color (componentes  $L^*$ ,  $a^*$  y  $b^*$ ), se calcula el número de píxeles que habría dentro de cada uno de los cubos en los que se ha dividido el espacio cromático y el valor medio de las componentes  $L^*$ ,  $a^*$  y  $b^*$  que asociamos al color del cubo del que determinamos también el cromatismo definido como

$$C^* = \sqrt{a^{*2} + b^{*2}} \quad (8.2)$$



**Figura 8.5:** Esquema del algoritmo del cálculo del número de colores relevantes de Nieves et al. [Nie20] que hemos aplicado para determinar el NRC de cada una de nuestras imágenes naturales.

Si el número de píxeles del cubo es mayor que un porcentaje del número total de píxeles (que hemos llamado umbral general y que hemos fijado en el 4%), entonces el color del cubo será relevante. Si el número de píxeles del cubo está comprendido entre el

valor del umbral y un octavo del mismo, también lo consideraremos relevante en el caso de que podamos considerar el color del cubo como un color saturado, es decir, cuando se cumpla que  $L^* \geq 80$  y el percentil ( $C^*$ )  $\geq 50$ . El número de cubos que para cada imagen cumplen estas condiciones es el número de colores relevantes y el color de cada uno de los cubos, el color de la “paleta” de colores cromáticamente relevantes en la correspondiente imagen (Figura 8.5 izquierda).

El algoritmo permite también segmentar la imagen original de una manera sencilla con solo comprobar para cada píxel de esta si se corresponde con un color relevante. Si es ese el caso, el color del píxel se sustituye por su correspondiente color relevante; en caso contrario, el color del píxel se sustituirá por el del cubo más cercano (desde un vista euclídeo) en el espacio cromático CIELAB (Figura 8.5 derecha).

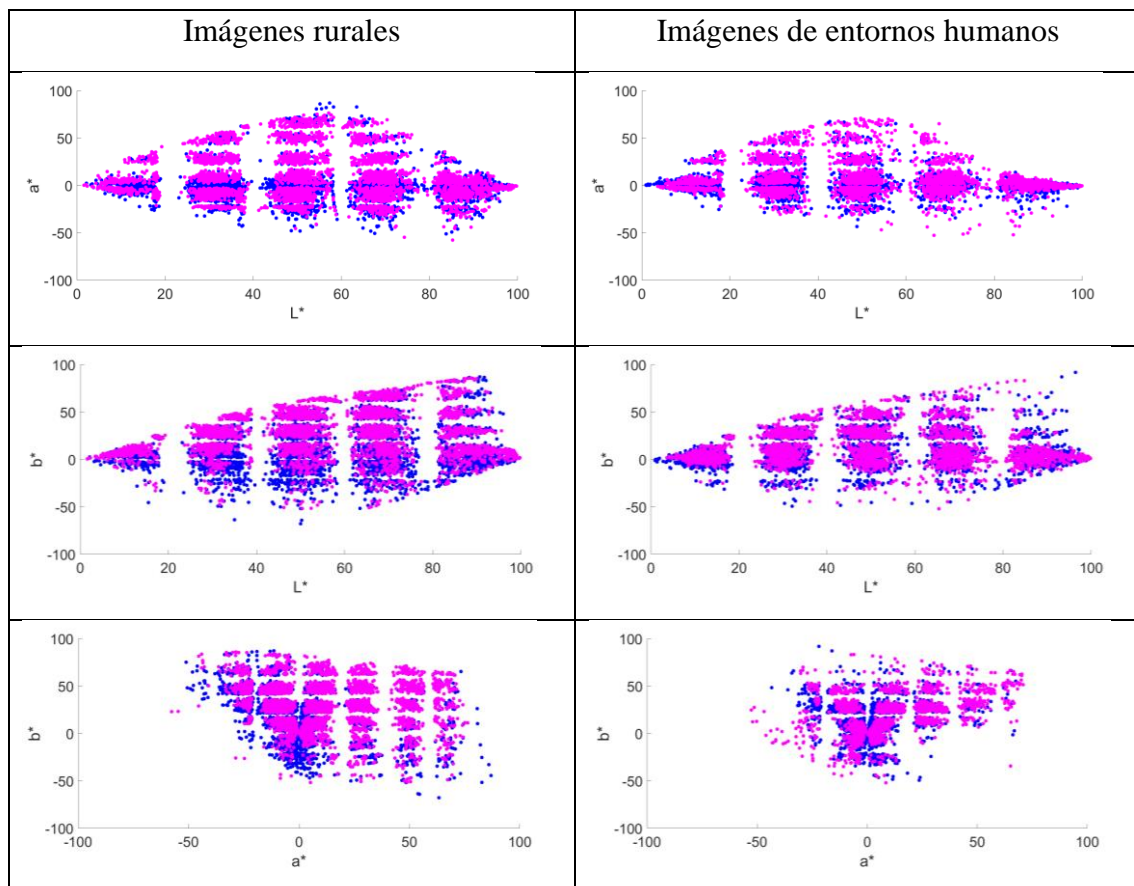
| Imágenes de este trabajo     |        | Imágenes de la base de datos McGill |        |
|------------------------------|--------|-------------------------------------|--------|
|                              | NRC    |                                     | NRC    |
| Bosques y parques            | 15 ± 4 |                                     |        |
| Campo                        | 10 ± 4 | Orillas                             | 16 ± 7 |
| Costa                        | 15 ± 6 | Flores                              | 19 ± 6 |
| Flores y frutas              | 22 ± 6 | Follaje                             | 20 ± 5 |
| Montaña                      | 15 ± 5 | Frutas                              | 16 ± 7 |
| Playa                        | 10 ± 4 |                                     |        |
| Ríos y cascadas              | 15 ± 5 |                                     |        |
| Imágenes rurales             | 15 ± 6 |                                     | 19 ± 6 |
| Autopistas                   | 14 ± 5 |                                     |        |
| Ciudades                     | 20 ± 4 |                                     |        |
| Edificios                    | 18 ± 5 |                                     |        |
| Interiores                   | 19 ± 5 |                                     |        |
| Calles                       | 22 ± 5 |                                     |        |
| Imágenes de entornos humanos | 19 ± 6 |                                     | 19 ± 5 |

**Tabla 8.1:** NRC obtenidos para los distintos conjuntos de imágenes naturales.

La Tabla 8.1 muestra el número de colores relevantes obtenidos para las distintas categorías de los dos conjuntos de imágenes ya descritos. Para nuestro conjunto original de imágenes vemos que el NRC de las imágenes rurales (15) es inferior al del conjunto de imágenes de la base de datos de McGill, en la que este número coincide para las dos

supracategorías semánticas de este conjunto de imágenes. Esta coincidencia también se produce entre las escenas de entornos humanos de ambos conjuntos (19). El número medio obtenido es prácticamente el mismo que obtuvieron Nieves et al [Nie20] para las obras pictóricas de  $18 \pm 6$ .

La figura 8.6 representa el total de colores relevantes en los planos  $L^* - a^*$ ,  $L^* - b^*$  y  $a^* - b^*$  del espacio de color CIELAB obtenidos con nuestro algoritmo para los dos conjuntos de imágenes (puntos azules para las imágenes de este trabajo y puntos magenta para las imágenes de la base de datos McGill) rurales y de entornos humanos. Las gamas de colores relevantes para las dos supracategorías de ambos conjuntos son semejantes destacando que la correspondiente a las imágenes de entornos humanos es más reducida que la de las escenas rurales que presentan tonalidades magentas, moradas y amarillos verdosos que no son consideradas relevantes por el algoritmo en el caso de la primera categoría.



**Figura 8.6:** Total de colores relevantes en el espacio de color CIELAB de los dos conjuntos de imágenes (puntos azules para las imágenes de este trabajo y puntos magenta para las imágenes de la base de datos McGill) rurales y de entornos humanos.

La segunda parte del algoritmo permitía la segmentación de la imagen original basándose en los colores relevantes obtenidos en el algoritmo, ejemplos de las cuales podemos ver en las figuras 8.3 y 8.5. Para ver qué grado de la imagen original podemos recuperar con dicha segmentación calculamos el coeficiente de correlación entre cada plano de color del espacio cromático RGB de la imagen original y de la imagen segmentada. Los valores medios obtenidos para cada categoría se muestran en la Tabla 8.2. Observamos mayor correlación con menor dispersión en las imágenes de entornos humanos frente a las imágenes rurales. Estos resultados de entornos humanos son parecidos a los obtenidos por Nieves et al. [Nie21] para obras pictóricas (de  $0.94 \pm 0,006$  para las imágenes del museo del Prado y  $0,96 \pm 0,002$  para las imágenes de la base de datos de Khan) basándose en la segmentación de las imágenes proporcionada por los colores relevantes subjetivos señalado por los observadores de su experimento. Nuestros resultados muestran que los coeficientes de correlación entre los distintos planos de color RGB de las imágenes originales y de las segmentadas obtenidos por el algoritmo presentan mucha más dispersión que los obtenidos por Nieves et al. [Nie21] utilizando los colores relevantes subjetivos sobre imágenes de cuadros. Aunque los resultados de la tabla 8.2 no sugieren una recuperación perfecta de la imagen original, sí podemos considerar que las imágenes segmentadas mantienen el contenido cromático más reseñable de las escenas rurales o de entornos humanos.

|                              | Imágenes de este trabajo |                 |                 | Imágenes de la base de datos McGill |                 |                 |
|------------------------------|--------------------------|-----------------|-----------------|-------------------------------------|-----------------|-----------------|
|                              | R                        | G               | B               | R                                   | G               | B               |
| Imágenes rurales             | $0,88 \pm 0,11$          | $0,91 \pm 0,09$ | $0,88 \pm 0,11$ | $0,88 \pm 0,20$                     | $0,88 \pm 0,20$ | $0,84 \pm 0,21$ |
| Imágenes de entornos humanos | $0,91 \pm 0,07$          | $0,94 \pm 0,05$ | $0,92 \pm 0,05$ | $0,93 \pm 0,07$                     | $0,95 \pm 0,06$ | $0,94 \pm 0,07$ |

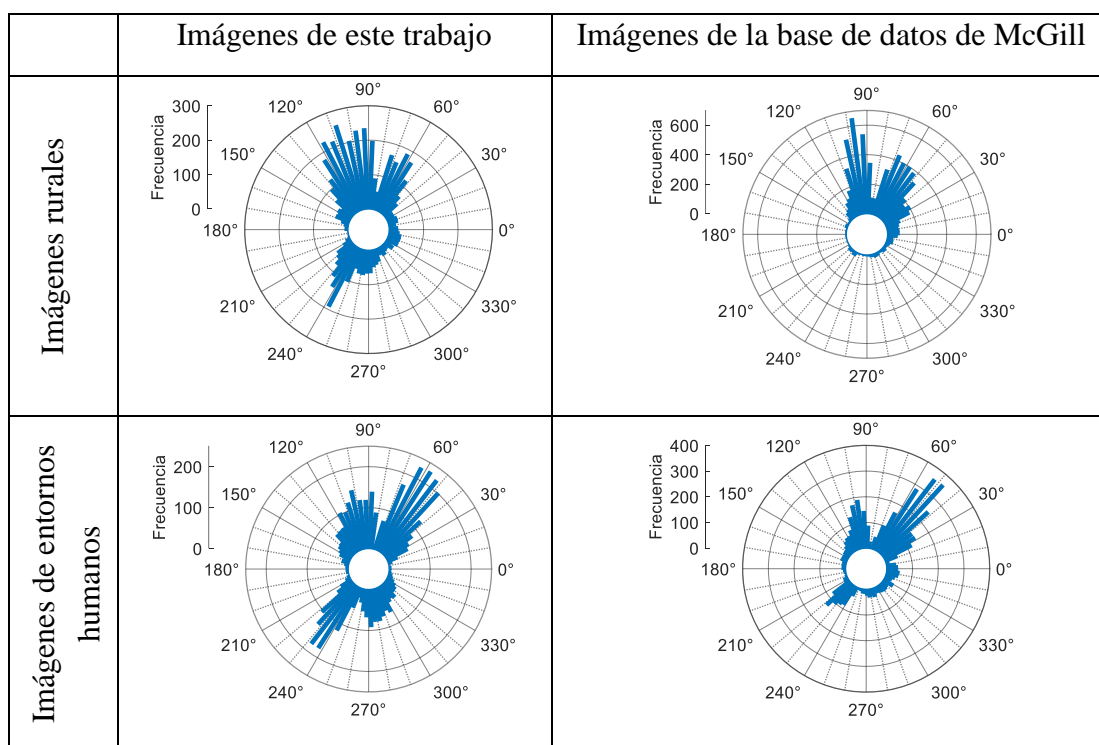
**Tabla 8.2 :** Coeficientes de correlación para cada plano del espacio RGB entre las imágenes originales y las imágenes segmentadas de los dos conjuntos de imágenes.

### 8.3.1. Codificación eficiente de los colores relevantes que describen una escena natural.

Como ya hemos visto en las secciones 6.1 y 7.3.1 el ICA ha sido utilizado para encontrar una transformación lineal de la señal de entrada que permita una representación óptima de la misma que sea tan estadísticamente independiente como sea posible. En dichas

secciones vimos cómo la codificación oponente de la información visual era eficiente para codificar escenas naturales en color encontrando diversas direcciones de oposición cromática además de las rojo-verde y las amarillo-azul [Wac01, Kel13].

Para comprobar si las direcciones de codificación eficiente tienen alguna relación con las de los colores relevantes obtenidos con nuestro algoritmo comenzamos calculando, tal y como se describía en la sección 8.2, el ángulo de la dirección de los colores relevantes obtenidos para los dos conjuntos de imágenes como el arco tangente del cociente de sus componentes BY y RG, de acuerdo a las ecuaciones (8.1). La figura 8.7 muestra el histograma circular de las direcciones obtenidas para los colores relevantes calculados para las imágenes rurales y de entornos humanos de los dos conjuntos de imágenes utilizados. En ellos podemos ver algunas direcciones que destacan ampliamente y en las que se agruparían los colores relevantes; estas son las descritas por ángulos entre  $50^\circ$  y  $70^\circ$ , entre  $90^\circ$  y  $120^\circ$  y en nuestras imágenes y en las de entornos humanos de la colección de McGill entre  $220^\circ$  y  $240^\circ$ .

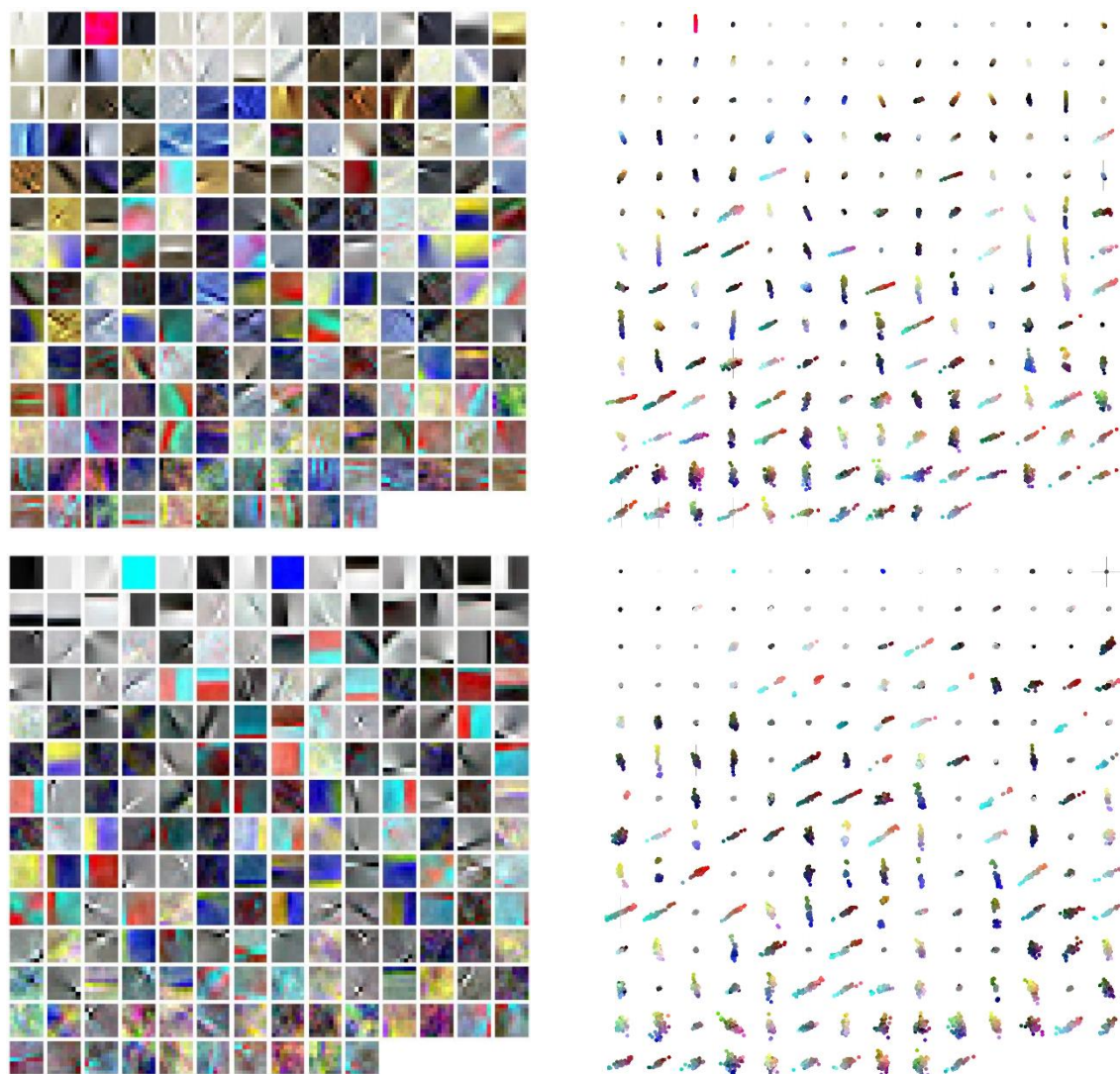


**Figura 8.7 :** Histogramas circulares de los ángulos de las direcciones de los colores relevantes en el plano RG y BY para los dos conjuntos de imágenes y las categorías rurales y de entornos humanos.

En la sección 6.1 aplicamos el ICA en nuestro conjunto de imágenes naturales cuyos resultados con respecto a las características obtenidas para el espacio de color RGB se mostraban en la fila central de la figura 6.2 y las direcciones de las cromaticidades de



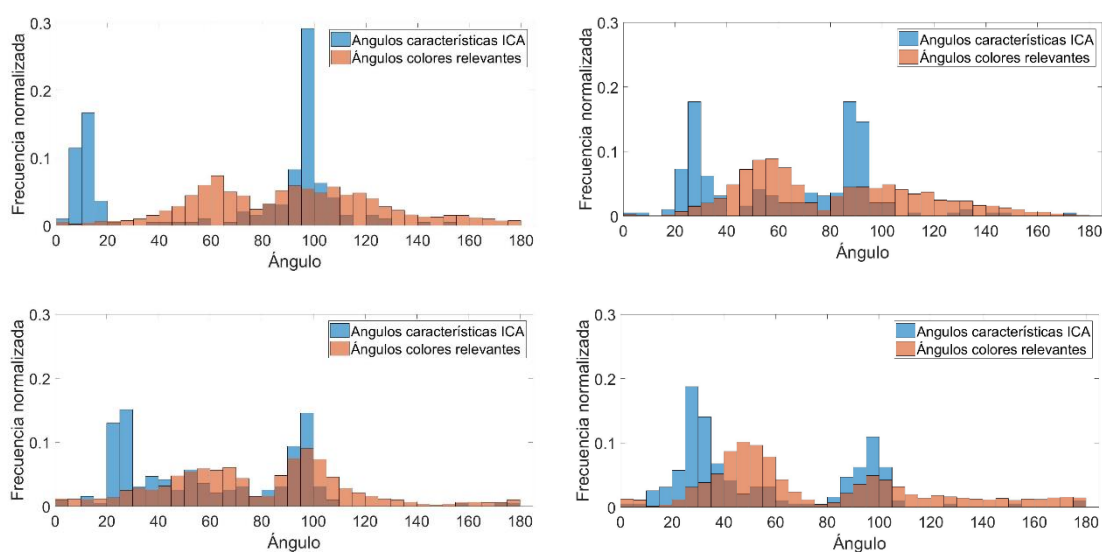
sus píxeles por los ángulos mostrados en el histograma circular de la izquierda de la figura 6.3. Hemos aplicado nuevamente ICA a las imágenes de la base de datos de McGill obteniendo las características mostradas (en orden decreciente de la norma L2 de izquierda a derecha y de arriba abajo) en la parte izquierda de la figura 8.8 cuyas cromaticidades para los píxeles de cada una de ellas se muestra en la parte derecha de la misma figura.



**Figura 8.8 :** (Izquierda) Las 192 características (filas de la matriz  $A$ ) obtenidas de aplicar ICA a las imágenes de la base de datos McGill en el espacio de color RGB (imágenes rurales en la fila superior y de entornos humanos en la fila inferior). (Derecha) Cromaticidades de los valores de los 192 píxeles de cada una de las características indicadas en un plano isoiluminante.



Al igual que para nuestro conjunto de imágenes originario, las características obtenidas con las imágenes de esta nueva colección muestran distribuciones espaciales orientadas y localizadas pudiendo ser cromáticamente homogéneas, acromáticas orientadas y color oponentes localizadas y orientadas espacialmente cuya frecuencia aumenta al disminuir la norma L2. Como con las componentes principales, las características de las componentes independientes pueden clasificarse en tres categorías principales: características homogéneamente cromáticas, acromáticas orientadas (bordes de luminancia) y color oponentes (bordes de colores) que, en muchas de ellas, no coinciden con las direcciones de oposición cromática rojo-verde y amarillo-azul.



**Figura 8.9:** Histogramas conjuntos normalizados de los ángulos (entre  $0^\circ$  y  $180^\circ$ ) de las direcciones de los colores relevantes y de la primera componente principal de cada una de las características obtenidas con el ICA para nuestras imágenes (fila superior) y de la base de datos de McGill (fila inferior) para imágenes rurales (izquierda) y de entornos humanos (derecha).

Con el fin de mostrar las direcciones de las cromaticidades de las características obtenidas del ICA para ambos conjuntos de imágenes y de comprobar si existe alguna relación con los colores relevantes obtenidos de nuestro algoritmo computacional, la figura 8.9 muestra los histogramas conjuntos normalizados de los ángulos de las direcciones de los colores relevantes tal y como se describió en la sección anterior (ahora reducidos a ángulos entre  $0^\circ$  y  $180^\circ$ ) y los de la primera componente principal (con respecto al semieje positivo del canal RG) de la distribución de los 192 píxeles de cada una de las características obtenidas con el ICA. La coincidencia en muchas de las direcciones obtenidas (principalmente en la base de datos de McGill y en las imágenes de entornos humanos de ambos conjuntos de escenas) sugiere que la elección de los colores

más relevantes de las escenas realizada por el algoritmo podría corresponderse con una codificación eficiente de las mismas, lo que está de acuerdo con los resultados obtenidos de la experiencia psicofísica realizada por Nieves et al. [Nie21] que justificaba los resultados heurísticos para imágenes de obras pictóricas del algoritmo aquí aplicado.



---

## Capítulo 9

# CONCLUSIONES

---

La consideración de que el SVH ha evolucionado para reducir la estructura redundante de los estímulos que lo alcanzan y, por tanto, producir una representación eficiente de los mismos establece una conexión fuerte con las propiedades de dichos estímulos. En este trabajo, a partir del estudio de diferentes descriptores estadísticos de las entradas visuales, de las transformaciones que pueden optimizarlos, de la influencia del iluminante y de la atención del observador en los mismos, se ha intentado caracterizar los descriptores y establecer relaciones con los modelos del procesamiento temprano del sistema visual humano [Ati92a, Ols96, Bel97, vHa98, Sim99].

El análisis de los estadísticos de primer orden, estudiados en el capítulo 4, muestra cómo el promedio de los valores de la intensidad de los píxeles en los tres canales  $L^*$ ,  $a^*$  y  $b^*$  para cada categoría semántica se corresponden con un valor próximo al gris (tabla 4.1) con ligeras diferencias debidas a las reflectancias de los objetos predominantes en las escenas o a la presencia de otras fuentes de iluminación [Pou13], lo que permite justificar la suposición del “mundo gris” que dice que, a pesar de la variabilidad de las reflectancias de una escena, el valor medio de la intensidad de sus píxeles estaría relacionado con la CCT del iluminante predominante dando un valor próximo al gris [Pit 38].

La descripción de la riqueza cromática de las imágenes mediante el número de colores discernibles nos proporciona una cifra comprendida entre 21400 y 54100 para las imágenes rurales y alrededor de 35000 para las imágenes de entornos humanos (excepto para la categoría “interiores” que es de 47500), valores que tienen el mismo orden de magnitud que el número distinto de colores de superficies determinado por Marín-Franch

y Foster [Mar10] al utilizar métodos de teoría de la información basados en el descriptor información mutua para tener en cuenta tanto las frecuencias de los distintos colores de superficies como la incertidumbre en la estimación del observador de los mismos que fue de 7300 en el espacio de color CIECAM02.

Al tener en cuenta tanto el número de colores discernibles como el número de veces que aparece cada uno, la distribución cromática sigue una ley potencial entre la ratio del número de colores discernibles y su frecuencia absoluta [Mon16], de manera que aquellas imágenes o categorías con mayor exponente en valor absoluto presentan mayor variedad cromática (son las que también presentan mayor número de colores discernibles).

El análisis estadístico de la diferencia de medias entre las distintas categorías para los diferentes descriptores cromáticos para un intervalo de confianza del 95% señala que estos son diferentes para el total de las categorías; no obstante, la formación de distintos grupos homogéneos impide utilizarlos para clasificar una imagen dentro de una determinada categoría, tal como se ve reflejado en el diagrama coordenadas paralelas de la figura 4.5.

En el capítulo 5, el estudio de los bordes (cambios repentinos y grandes en luminancia o color) a través del descriptor gradiente de las intensidades de los canales  $L^*$ ,  $a^*$  y  $b^*$  de nuestras imágenes desenfocadas, para evitar las fluctuaciones locales que las zonas no uniformes de las texturas o la presencia de ruido puede introducir en el mismo, muestra distribuciones exponenciales centradas en cero. La representación de los tres histogramas conjuntos de los gradientes en los distintos planos de cada escena, como aproximación a la densidad de probabilidad conjunta, no muestra que los bordes puramente cromáticos sean menos frecuentes que los de luminancia (figuras 5.3 y 5.4); no obstante, la mayoría de los bordes combinan información de cada dos canales en distintos grados, de modo que el valor en una dimensión no puede predecir el valor en la otra [Han09], datos que cuantitativamente se confirman al calcular la información mutua entre los bordes de cada pareja de canales (comprendida entre 0,003 y 0,391 bits). A la vista de estos resultados, hay un grado sustancial de independencia entre bordes cromáticos y de luminancia en escenas naturales y, por tanto, estos proporcionan fuentes independientes de información [Geg00]. Los valores de la información mutua de los histogramas conjuntos de los gradientes, aunque pequeños no son nulos, lo que indica que la independencia entre los bordes detectados entre cada pareja de canales no es perfecta, en particular, entre los bordes de los dos canales cromáticos de las imágenes rurales que presentaban un valor de la información mutua mayor.

Uno de los resultados fundamentales de la estadística de segundo orden entre píxeles no adyacentes basado en la frecuencia espacial de las imágenes es que la potencia espectral de las imágenes naturales en promedio y en cualquier dirección decae con la frecuencia espectral según una ley potencial de exponente un valor próximo a -2 [Fie87, Bur87, Tol92, Rud94]. Para las imágenes de estudio dicho valor varió entre -2,16 y -2,67

según tres orientaciones ( $0^\circ$ ,  $45^\circ$  y  $90^\circ$ ), los tres canales  $L^*$ ,  $a^*$  y  $b^*$  para las imágenes rurales y de entornos humanos (Tabla 5.2). Al estudiar la significación estadística de las diferencias encontradas en la media de la pendiente con los tres parámetros anteriores (categoría semántica, orientación y canal) para un intervalo de confianza del 95%, se obtuvieron resultados tan dispares que no permiten sacar conclusiones del mismo. La Tabla 5.2 también muestra cómo, para todas las clases semánticas, el espectro de potencia no está distribuido isotrópicamente en función de la orientación, sino que su pendiente es mayor para orientaciones diagonales que para las horizontales o verticales siendo esta dependencia angular similar. Este comportamiento se debe al dominio de estructuras horizontales y verticales de las imágenes rurales [VdS98], incluso más acentuadas en las de entornos humanos.

Al representar las firmas espectrales (gráfico de isolíneas que se obtiene al cortar el espectro de potencia por distintos planos horizontales de manera que cada una de ellas encierre los puntos correspondientes al 60, 80 y 90% de la energía total de la señal), solamente el patrón del canal luminancia de las imágenes rurales (figura 5.9) presenta diferencias significativas para utilizarlo en la clasificación de este tipo de imágenes en las distintas categorías semánticas establecidas; no obstante el contorno de una clase de imágenes no es representativa para imágenes individuales de la misma.

Si consideramos que las sensibilidades de los conos descomponen el espectro de manera fija en tres respuestas, la aplicación del PCA a sus logaritmos [Buc83, Rud98] muestra que las direcciones principales de las mismas se corresponden con mecanismos de oposición de color [Sim01]. La aplicación a nuestro conjunto de datos (imágenes rurales y de entornos humanos) del PCA para reducir la redundancia de estos mediante su decorrelación ortogonal reproduce el resultado anterior en tres espacios de color diferentes (logaritmo de las sensibilidades de los conos: LMS, RGB y CIELAB) proporcionando en todos ellos tres ejes principales próximos a las funciones luminancia y cromáticas oponentes, con la varianza decreciendo del canal luminancia (Lum), al canal oponente rojo-verde (RG) pasando por el canal oponente amarillo-azul (BY).

Al aplicar PCA a trozos de imágenes en los tres espacios de color, para obtener información de la estructura espacio-cromática, se obtenían una serie de componentes principales que pueden clasificarse en tres categorías principales (Figura 5.14): componentes homogéneamente cromáticas (de mayor varianza, sin estructura espacial y altamente selectivas a un color), acromáticas (que codifican las variaciones de luminancia) y color oponentes (que presentan la menor varianza y decrecen desde las componentes BY a las RG) [Kel13]. Las cromaticidades de estas últimas presentan alineaciones que en su mayoría corresponden con los mecanismos de oposición de color ( $0^\circ$  y  $90^\circ$ ) aunque esas son de  $30^\circ$  y  $120^\circ$  en el espacio CIELAB (Figura 5.15). Un pequeño número de componentes muestran diferentes ángulos, lo que indica que las direcciones color oponentes están correlacionadas en algún grado en los datos [Wac01].



De los resultados de la asimetría de la tabla 6.1 el canal  $a^*$  es el más simétrico aunque el sentido (signo de esta) varía predominando el signo positivo, el canal  $b^*$  presenta siempre una asimetría considerable. (predominio del azul sobre el amarillo) y positiva mientras que el canal  $L^*$  tiene asimetría negativa en todas las categorías (predominio de los píxeles claros sobre los oscuros). Este resultado, que parece contradecir el de trabajos previos ya que afirman que la asimetría de la luminancia es positiva en imágenes naturales monocromas [Lau81, Bra00, Att07, Dro01, y Gra16], no lo hace si consideramos el de la versión en escala de grises de nuestras imágenes, en el que  $0,03 < S < 0,27$  (siendo solamente la categoría “playa” la que sigue mostrando una asimetría pequeña negativa de -0,09).

Las tres columnas de la derecha de la figura 6.2 muestra los resultados obtenidos al calcular la curtosis en los tres planos  $L^*$ ,  $a^*$  y  $b^*$  de nuestras imágenes de estudio blanqueadas que es mayor que 3 en todas las categorías y en todos los canales [Rei01, Tho00, Tho99a, Tho99b] siendo mayor en los canales cromáticos que en el canal luminancia. Si el sistema visual humano debe optimizar la dispersión, debería maximizar la curtosis a través de las diferentes etapas de su procesamiento. En la tabla 6.2 se puede comprobar que la curtosis en los diferentes planos del espacio de color RGB es menor que para los mecanismos  $L^*$ ,  $a^*$  y  $b^*$  correspondientes a las salidas de las células ganglionares de la retina, siendo más acentuadas en los canales cromáticos.

Las características obtenidas al aplicar ICA a un mismo número de trozos de imágenes de dimensiones  $8 \times 8$  de las escenas rurales y de entornos humanos que se muestran en las figuras (6.3 y 6.5) a partir de las propiedades estadísticas de orden superior del espectro de dichas imágenes se corresponden con señales máximamente independientes que separan la información en luminancia (o detectores de bordes de luminancia) y en información cromática color oponente (detectores de bordes de color no ortogonales) presentando distribuciones dispersas lo que constituye una codificación altamente eficiente. [Tai00, Lee02]. Como mostraron Olshausen y Field [Ols96] la optimización de la representación dispersa, estadísticamente independiente de las respuestas a imágenes naturales (como la que realiza el ICA), proporcionaba filtros que se asemejan mucho espacialmente a los campos receptivos de las células oponentes simples orientadas, encontradas en el córtex visual de los mamíferos ya que están espacialmente localizados, orientados y constituyen un filtro paso banda a diferentes bandas de frecuencias espaciales [Bel97, vHa98, Lew99].

El último estadístico de orden superior analizado, el índice de agrupamiento, muestra que los planos cromáticos de las imágenes contienen más áreas uniformes que el plano de luminancia, como cabía esperar dado que las diferencias cromáticas se deben a las propiedades de los objetos mientras que las diferencias en luminancia no provienen solamente de las diferencias de los materiales y sus reflectancias sino de la iluminación no homogénea característica de las imágenes naturales [Rub82, Tap02, Olm04]. Entre los dos canales cromáticos la menor cantidad de áreas uniformes en el canal amarillo –

azul puede deberse a variaciones hacia el amarillo en las sombras frecuentes de escenas de follaje abundante [Joh05], así como al desplazamiento hacia el azul en las sombras de la iluminación de días fuertemente soleados [Chu94].

A continuación, describimos los aspectos comunes o mayoritarios que los descriptores estadísticos de los distintos órdenes analizados muestran con la variación de la distribución de potencia espectral de la luz día (capítulo 7) caracterizada por su temperatura correlacionada de color. Un primer aspecto destacable es el hecho de que todos los descriptores presentan un comportamiento asintótico horizontal al establecer la relación de los mismos con la temperatura correlacionada de color del iluminante; para CCTs superiores a 15000K la dependencia de los estadísticos con la misma es pequeña o nula.

Por el contrario, para los iluminantes del espectro caracterizados por las menores temperaturas de color es para los que se producen las mayores variaciones de los descriptores con la CCT (siempre que muestren una dependencia con ella). Esta dependencia no se debe a la utilización de dos conjuntos de datos para caracterizar la distribución de potencia espectral de la fuente de iluminación, los de Hernández-Andrés et al. [Her01] para el rango de CCTs desde 3766 K hasta los 25889 K y las simuladas con SBDART para temperaturas inferiores a 3766 K, dada la dificultad de ser observadas [Pey16, Ric98]. El efecto de este cambio de datos se refleja en la mayoría de las figuras del capítulo 7 como una discontinuidad en la gráfica para los valores de las temperaturas en que estos cambian.

Dado el comportamiento que presentan los descriptores en función de la temperatura de color del iluminante en los dos extremos del rango de temperaturas utilizadas, en caso de no ser estrictamente monótonas, salvo la discontinuidad ya comentada, presentan un extremo relativo en el intervalo de CCTs 3900 K-9600K. Estos máximos o mínimos locales pueden estar relacionados con otros índices de reproducción [Guo04] basados en supuestos perceptivos y en datos psicofísicos. Como comentamos en la sección dedicada a los estadísticos de primer orden, el valor de la temperatura correlacionada de color del iluminante alrededor de la que se alcanzaba el mayor número de colores discernibles en nuestro experimento coincidía con la CCT para la cual Masuda y Nascimento [Ma13] obtenían el mayor grado de preferencia (gammas más ricas de color). La localización de estos extremos relativos en un rango centrado alrededor de los 6500K (CCT asociado a la luz día) sugiere que otros aspectos relacionados con la visión del color pueden estar optimizados para estas iluminaciones [Oje17].

Si analizamos el grado de dependencia de los distintos descriptores estadísticos con la temperatura del iluminante para los distintos mecanismos de visión en el espacio CIELAB luminancia (canal  $L^*$ ), canal cromático oponente rojo-verde ( $a^*$ ) y canal cromático oponente amarillo-azul ( $b^*$ ), el primero es el que muestra menor dependencia con la CCT, habiendo evidencias estadísticas suficientes para que en un intervalo de confianza del 95% se consideren los descriptores desviación típica, asimetría, curtosis

(momentos segundo tercero y cuarto de la distribución de la señal de color) y pendientes espectrales independientes de la misma. De los dos canales cromáticos  $a^*$  y  $b^*$ , es siempre este último el que muestra una mayor sensibilidad a las variaciones del iluminante. Lovell et al. [Lov05] justifican este comportamiento de los canales oponentes rojo-verde y amarillo-azul a la proximidad de los picos de absorción espectral de los conos L y M, lo que hace que los cambios graduales sobre el ancho del espectro tengan un efecto diferencial pequeño sobre los conos. Desde un punto de vista evolutivo esta menor sensibilidad a los cambios con el iluminante del canal  $a^*$  permite a los sistemas visuales tricromáticos una detección óptima de frutas u hojas comestibles (maduras) de las que no lo son durante todo el día (ignorando sombras y cambios de iluminación) salvo al amanecer y la puesta de sol donde funciona mejor el canal  $b^*$  [Lov05]).

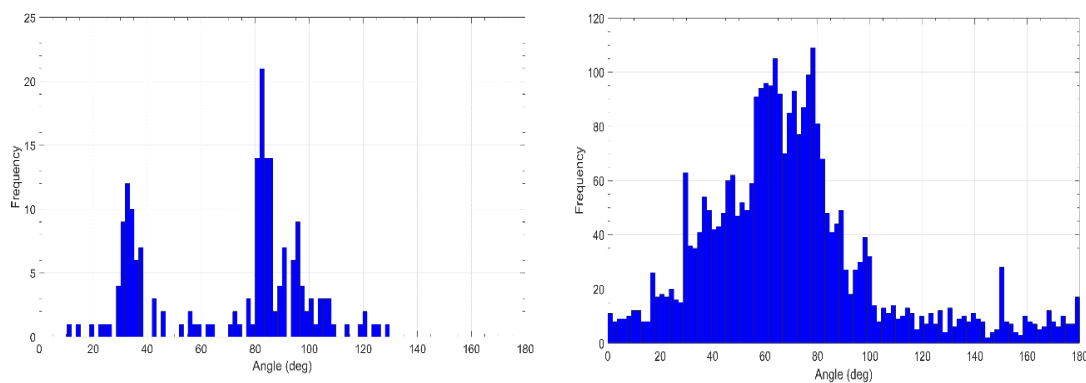
Por último, debo señalar que los valores medios de descriptores volumen de color, número de colores discernibles e índice de agrupamiento para los distintos iluminantes pueden considerarse significativamente iguales desde un punto de vista estadístico lo que viene a destacar que la paleta de todos los colores de la escena en todas sus intensidades disponibles, así como el número de cubos unidad en que podemos segmentar el espacio de color que contiene al menos un píxel y la proporción de la imagen que contiene áreas uniformes en los tres planos del espacio CIELAB para un intervalo de confianza del 95% pueden considerarse invariantes lo que podría deberse al proceso de adaptación cromática de las primeras etapas involucradas en el fenómeno de la constancia de color.

En la sección 8.3 hemos intentado ver el comportamiento del algoritmo propuesto por Nieves et al. [Nie20] para calcular el número de colores relevantes (NRC) cuando se aplica a imágenes naturales para lo que se utilizaron dos conjuntos de imágenes: aquella que hemos usado en todo el trabajo presente y 494 imágenes de la base de datos McGill [Olm04] divididas en cinco categorías semánticas: orillas, flores, follaje, frutas (imágenes rurales) y entornos humanos. El NRC obtenido para nuestras imágenes rurales es menor  $15 \pm 6$  que el obtenido para el de la base de datos McGill  $19 \pm 6$ , pero coincide en ambos conjuntos de escenas en el caso de imágenes de entornos humanos donde es 19 (Tabla 8.1). Valores prácticamente iguales a los obtenidos en trabajos previos al aplicar el algoritmo a obras pictóricas [Nie20, Nie21].

Cuando se representa el total del número de colores relevantes obtenidos al aplicar el algoritmo a nuestros dos conjuntos de escenas, las gamas de colores relevantes obtenidas tanto para categorías de imágenes rurales como de entornos humanos son semejantes (figura 8.6); además la gama cromática correspondiente a los entornos humanos es más reducida que la de las imágenes rurales que presentan tonalidades magentas, moradas y amarillas verdosas que no son consideradas relevantes por el algoritmo en el caso de la primera categoría.

Para comprobar si las impresiones subjetivas de los observadores de su experimento venían derivadas de un procesamiento óptimo y eficiente de las instancias

más representativas de las imágenes de las pinturas, Nieves et al [Nie21] aplicaron ICA al conjunto de pinturas del experimento. La figura 9.1 representa los histogramas de las direcciones de color de las bases obtenidas con ICA y de los colores relevantes obtenidos en el experimento psicofísico por Nieves et al. La relativamente buena concordancia existente entre los ángulos obtenidos, tanto para los colores relevantes como para las funciones base del ICA, justificaba la hipótesis de que los observadores seleccionaban el contenido cromático más relevante de las pinturas siguiendo una codificación eficiente [Nie21].



**Figura 9.1.-** Histograma de la distribución de direcciones cromáticas asociada a cada función base obtenidas al aplicar ICA al conjunto de imágenes de cuadros de la base de datos de Khan (a la derecha) y de los ángulos en el plano RG y BY de los colores relevantes de todos los observadores en el experimento psicofísico [Nie21].

Para ver si los colores relevantes seleccionados por el algoritmo, al aplicarlo a imágenes naturales, seguían también una codificación eficiente, en la sección 8.3.1. se repitió el análisis descrito utilizando los colores relevantes obtenidos al aplicar el algoritmo y los colores de las funciones base resultantes de aplicar ICA. Los histogramas conjuntos normalizados de los ángulos de ambas variables mostrados en la figura 8.9, aunque presentan una buena correspondencia para la dirección vertical ( $90^\circ$ ) en todos los casos, muestran diferencias entre las direcciones de los colores relevantes y las direcciones preferidas por las funciones base del ICA para ángulos pequeños, siendo estas menores para las imágenes de entornos humanos y para el conjunto de escenas de la base de datos de McGill. Estos resultados no nos permiten confirmar que los colores relevantes obtenidos por el algoritmo sigan una codificación eficiente.

Queda para trabajos posteriores la realización de un experimento psicofísico similar al de Nieves et al [Nie21] en el que se determinaran los colores relevantes de imágenes naturales y una revisión del algoritmo de selección de los mismos siempre que

se confirmara una codificación eficiente de los colores relevantes señalados por los observadores en dicho experimento.

---

**BIBLIOGRAFÍA.**

---

- [All04] D. Alleysson y S. Süsstrunk, “Spatio-chromatic PCA of a mosaiced color image”, 2nd European Conference on Color in Graphics, Imaging and Vision, 311 – 314 (2004).
- [Ati92a] J.J. Atick, “Could information theory provide an ecological theory of sensory processing?”, *Network: Computation in Neural Systems* 3, 213 – 251 (1992).
- [Ati92b] J. J. Atick y A. N. Redlich, “What does the retina know about natural scenes?”, *Neural Computation* 4, 196 – 210 (1992).
- [Att07] D. Attewell y R. J. Baddeley, “The distribution of reflectances within the visual environment”, *Vision Research*, 47, 548 – 554 (2007).
- [Att54] F. Attneave, “Some informational aspects of visual perception”, *Psychological Review* 61, 183 – 193 (1954).
- [Bad96] R. Baddeley, “Searching for filters with ‘interesting’ output distributions: an uninteresting direction to explore?”, *Network: Computation in Neural Systems* 7, 409 – 421 (1996).
- [Bal00] R. M. Balboa y N. M. Grzywacz, “Occlusions and their relationship with the distribution of contrasts in natural images”, *Vision Research* 40, 2661–2669 (2000).
- [Bar59] H. B. Barlow, “Sensory Mechanisms, The Reduction of Redundancy and Intelligence”, *NPL Symposium on the Mechanization of Thought Process* 10, 535-539 (1959).



- [Bar01] H.B. Barlow, “Redundancy reduction revisited”, *Network* 12, 241–253 (2001).
- [Bar12] H. B. Barlow, “Possible principles underlying the transformations of sensory messages”, en “*Sensory Communication*”, 217–234, MIT Press (2012).
- [Bel95] A. J. Bell y T. Sejnowski, “An information-maximization approach to blind separation and blind deconvolution”, *Neural Computation* 7, 1129 – 1159 (1995).
- [Bel97] A. J. Bell y T. J. Sejnowski, “The ‘Independent Components’ of Scenes are Edge Filters”, *Vision Research* 37, 3327-3338 (1997).
- [Ber69] B. Berlin y P. Kay, “Basic Color Terms: Their Universality and Evolution”, The University of California Press (1969).
- [Bez81] J. Bezdek, “Pattern Recognition with Fuzzy Objective Function Algorithms”, Kluwer Academic (1981).
- [Bla86] G.G. Blasdel y G. Salama, “Voltage-sensitive dyes reveal a modular organization in monkey striate cortex”, *Nature* 321, 579 – 585 (1986).
- [Bon89] A.B. Bonds, “Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex”, *Visual Neuroscience* 2, 41 – 55 (1989).
- [Bra00] N. Brady y D. J. Field, “Local contrast in natural images: Normalisation and coding efficiency”, *Perception* 29, 1041 – 1055 (2000).
- [Buc80] G. Buchsbaum, “A spatial processor model for object colour perception”, *Journal of the Franklin Institute*, 310, 1-26 (1980).
- [Buc83] G. Buchsbaum y A. Gottschalk, “Trichromacy, opponent colour coding and optimum colour information transmission in the retina”, *Proceedings of the Royal Society of London B* 220, 89 – 113 (1983).
- [Bur87] G. J. Burton y I. R. Moorehead, “Color and spatial structure in natural scenes”, *Applied Optics* 26, 157-170 (1987).
- [Can05] R.L. Canosa, “Modeling selective perception of complex natural scenes”, *International Journal on Artificial Intelligence Tools* 14, 233 – 260 (2005).
- [Car97] M. Carandini, D.J. Heeger y J.A. Movshon, “Linearity and normalization in simple cells of the macaque primary visual cortex”, *Journal of Neuroscience* 17, 8621 – 8644 (1997).
- [Cha65] A. Chapanis, “Color names for color space”, *American Science* 53, 327–346 (1965).

- [Chi00] C. C. Chiao, T. W. Cronin y D. Osorio, “Color signals in natural scenes: characteristics of reflectance spectra and effects of natural illuminants”, *Journal of the Optical Society of America A* 17, 218–224 (2000).
- [Cho11] E. Chorro, “Procesado cortical y uniformidad del espacio de color: hacia una métrica euclídea para las diferencias de color”, Tesis doctoral Universidad Alicante (2011). <http://hdl.handle.net/10045/20052>
- [Chu94] M. E. Churma, “Blue shadows: physical, physiological, and psychological causes,” *Applied Optics* 33, 4719–4722 (1994).
- [CIE11] International Lighting Vocabulary. Publication CIE S 017 / E : 2011. (CIE, Vienna 2011).
- [Čis10] P. Čisar y S.M. Čisar, “Skewness and Kurtosis in Function of Selection of Network Traffic Distribution”, *Acta Polytechnica Hungarica* 7, 95-106 (2010).
- [Cla03] P.L. Clatworthy, M. Chirimuuta, J.S. Lauritzen y D.J. Tolhurst, “Coding of the contrasts in natural images by populations of neurons in primary visual cortex (V1)”, *Vision Research* 43, 1983 – 2001 (2003).
- [Com94] P. Comon, “Independent component analysis—a new concept?”, *Signal Processing* 36, 287–314 (1994).
- [Con02] B.R. Conway, D.H. Hubel y M.S. Livingstone, “Color contrast in macaque V1”, *Cerebral Cortex* 12, 915–925 (2002).
- [Con07] B.R. Conway, S. Moeller y D.Y. Tsao, “Specialized color modules in macaque extrastriate cortex”, *Neuron* 56, 560 – 573 (2007).
- [Con10] B. R. Conway, S. Chatterjee, G. D. Field, G. D. Horwitz, E. N. Johnson, K. Koida, y K. Mancuso, “Advances in color science: from retina to behavior”, *Journal of Neuroscience*, 30, 14955–14963 (2010).
- [Cov91] T. M. Cover y J. A. Thomas, “Elements of Information Theory”, Wiley. (1991).
- [Dau89] J. G. Daugman, “Entropy reduction and decorrelation in visual coding by oriented neural receptive fields”, *IEEE Transactions on Biomedical Engineering* 36, 107 – 114 (1989).
- [Del05] J. Delon, A. Desolneux, J. L. Lisani, y A. B. Petro, “Automatic color palette”, *IEEE International Conference on Image Processing (ICIP)* 2, 706 – 709 (2005).
- [DeM75a] F. M. DeMonasterio y P. Gouras, “Functional properties of ganglion cells of the rhesus monkey retina”, *The Journal of Physiology* 251, 167 – 195 (1975).

- [DeM75b] F. M. DeMonasterio, P. Gouras y D. J. Tolhurst, “Trichromatic colour opponency in ganglion cells of the rhesus monkey retina”, *The Journal of Physiology* 251, 197 – 216 (1975).
- [Der95] G. Derfeldt y T. Swartling, “Colour concept retrieval by free colour naming: Identification of up to 30 colours without training”, *Displays* 16, 69 – 77 (1995).
- [Der84] A. M. Derrington, J. Krauskopf y P. Lennie, “Chromatic mechanisms in lateral geniculate nucleus of macaque”, *The Journal of Physiology* 357, 241–265 (1984).
- [Des85] R. Desimone, S. J. Schein, J. Moran y L.G. Ungerleider, “Contour, color and shape analysis beyond the striate cortex”, *Vision Research* 25, 441 – 452 (1985).
- [Dix78] R.E. Dixon, “Spectral distribution of Australian daylight”, *Journal of the Optical Society of America* 68, 437–450 (1978).
- [Doa11] D.P.Doane y L.E.Seward, “Measuring Skewness:A forgotten Statistic?”, *Journal of Statistics Education* 19 (2) : 3, 1-18 (2011).
- [Doi05] E. Doi y M. Lewicki, “Relations between the statistical regularities of natural images and the response properties of the early visual system”, *Japanese Cognitive Science Society, SIG P&P* 1, 1-8 (2005).
- [Dos11] W. Dos Santos y J. Lee, “An Adaptive Fuzzy-Based System to Simulate, Quantify and Compensate Color Blindness”, *Integrated Computer Aided Engineering* 18, 29-40 (2011).
- [Dro01] R. O. Dror, T. K. Leung, E. H. Adelson y A.S. Willsky, “Statistics of real-world Illumination”, *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 164-171 (2001).
- [Fan13] H. Fang, G. K-L. Tam, A. J. Aubrey, P. L. Rosin, C. Wallraven, D. Cunningham, D. Marshall y M. Chen, “Visualizing Natural Image Statistics”. *IEEE Transactions on Visualization and Computer Graphics* 19, 1228-1241 (2013).
- [Fie87] D.J.Field, “Relations between the statistics of natural images and the response properties of cortical cells”, *Journal of the Optical Society of America A* 4, 2379-2394 (1987).
- [Fie89] D. J. Field, "What the Statistics of Natural Images Tell us About Visual Coding", *Human Vision, Visual Processing, and Digital Display Proceedings SPIE* 1077, 269-276 (1989).

- [Fie93] D. J. Field, “Scale-invariance and self-similar ‘wavelet’ transforms: an analysis of natural scenes and mammalian visual systems”, en “Wavelets, Fractals, and Fourier Transforms”, 151–194, Clarendon (1993).
- [Fie94] D. J. Field, “What is the goal of sensory coding?”, *Neural Computation* 6, 559–601 (1994).
- [Fla90] P. Flanagan, P. Cavanagh, y O. E. Favreau, “Independent orientation-selective mechanisms for the cardinal directions of colour space”, *Vision Research* 30, 769–778 (1990).
- [Fos94] D. H. Foster y S. M. C. Nascimento, “Relational color constancy from invariant cone-excitation ratios”, *Proceedings of the Royal Society of London Series B-Biological Sciences* 257, 115–121 (1994).
- [Fos06] D.H. Foster, K. Amano, S.M.C. Nascimento y M. J. Foster, “Frequency of metamerism in natural scenes”, *Journal of the Optical Society of America A* 23, 2359-2372 (2006).
- [Fos11] D. H. Foster, “Color constancy”, *Vision Research* 51, 674–700 (2011).
- [Fos21] D.H. Foster, “Fluctuating environmental light limits number of surfaces visually recognizable by colour”, *Scientific Reports* 11, 1-10 (2021).
- [Fra06] R.A. Frazor y W.S. Geisler. “Local luminance and contrast in natural images”. *Vision Research* 46:1585–98 (2006).
- [Fri01] L.J. Frishman, “Basic Visual Processes”, en “Blackwell Handbook of Sensation and Perception”, 53 -91, Blackwell publishing Ltd. (2001).
- [Gao06] W. Gao, Q. Huang y F. Zhao, ”Image matching by normalized Cross-Correlation”, *IEEE International Conference on Acoustics Speed and Signal Processing Proceedings*, 729-732 (2006).
- [Gao13] S. Gao, K. Yang, C. Li y Y. Li, “A color constancy model with double-opponency mechanisms”, *IEEE International Conference on Computer Vision*, 929-936 (2013).
- [Geg00] K.R. Gegenfurtner y J. Rieger, “Sensory and cognitive contributions of color to the recognition of natural scenes”, *Current Biology* 10, 805–808 (2000).
- [Geg03] K. R. Gegenfurtner, “Cortical mechanisms of colour vision”, *Nature Reviews Neuroscience* 4, 563–572 (2003).
- [Gei92] W.S. Geisler y D.G. Albrecht, “Cortical neurons: isolation of contrast gain control”, *Vision Research* 8,1409 – 1410 (1992).
- [Gei08] W.S. Geisler, “Visual Perception and the Statistical Properties of Natural Scenes”, *Annual Review of Psychology* 59, 167–192 (2008).

- [Gei11] W.S. Geisler y J.S. Perry, “Statistics for optimal point prediction in natural images”, *Journal of Vision* 11 (12) : 14, 1-17 (2011)
- [Gla15] E. Gladilin y R. Eils, ”On the role of spatial phase NS phase correlation in vision, illusion and cognition”, *Frontiers in Computational Neuroscience* 9 : 45, 1-14 (2015).
- [Gou91] P. Gouras, “Vision and visual dysfunction: The Perception of colour vol. 6”, CRC Press (1991).
- [Gou09] P. Gouras, “Color Vision”, en “Webvision online book” (2009). <http://webvision.org.es/7-1-color-vision-by-peter-gouras/>
- [Gra16] D. Graham, B. Schwarz, A. Chatterjee y H. Leder, “Preference for luminance histogram regularities in natural scenes”, *Vision Research* 120, 11–21 (2016).
- [Gri19] L. D. Griffin y D. Mylonas, “Categorical colour geometry”, *PLOS ONE* 14, e0216296 (2019).
- [Gri86] A. Grinvald, E. Lieke, R.D. Frostig, C.D. Gilbert y T.N. Wiesel, “Functional architecture of cortex revealed by optical imaging of intrinsic signals”, *Nature* 324, 361–364 (1986).
- [Guo04] X. Guo y K. W. Houser, “A review of colour rendering indices and their application to commercial light sources”, *Lighting Research and Technology* 36, 183–199 (2004).
- [Han07] B. C. Hansen y R. F. Hess, “Structural sparseness and spatial phase alignment in natural scenes”, *Journal of the Optical Society of America A* 24, 1873–1885 (2007).
- [Han09] T. Hansen y K. R. Gegenfurtner, “Independence of color and luminance edges in natural scenes”, *Visual Neuroscience* 26, 35–49 (2009).
- [Har06] J. Harel, C. Koch y P. Perona, “Graph-based visual saliency”, en “Advances in Neural Information Processing Systems 19”, 545–552, MIT Press (2006).
- [Har78] F. J. Harris, “On the use of windows for harmonic analysis with the discrete Fourier transform”, *Proceedings of the IEEE* 66, 51-84 (1978).
- [Hee92] D. J. Heeger, “Normalization of cell responses in cat striate cortex”, *Visual Neuroscience* 9, 181–198 (1992).
- [Hei05] J. A. Heimel, S.D. Van Hooser y S.B. Nelson, “Laminar organization of response properties in primary visual cortex of the gray squirrel (*Sciurus carolinensis*)”, *Journal of Neurophysiology* 94, 3538 –3554 (2005).

- [Her64] E. Hering, “Outlines of a Theory of the Light Sense”, Harvard University Press (1964).
- [Her01] J. Hernández-Andrés, J. Romero, J. L. Nieves, y R. L. Lee, Jr., “Color and spectral analysis of daylight in southern Europe”, *Journal of the Optical Society of America A* 18, 1325–1335 (2001).
- [Hor07] G.D. Horwitz, E.J. Chichilnisky y T.D. Albright. “Cone inputs to simple and complex cells in V1 of awake macaque”. *Journal of Neurophysiology* 97, 3070–3081 (2007).
- [Hua99] J. Huang y D.Mumford, “Statistics of Natural Images and Models”, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 541-547 (1999).
- [Hub77] D.H. Hubel y T.N. Wiesel, "Ferrier lecture. Functional architecture of macaque monkey visual cortex", *Proceedings of the Royal Society of London B: Biological Sciences*, 198, 1-59 (1977).
- [Hyv98] A. Hyvärinen, “New approximations of differential entropy for independent component analysis and projection pursuit”, *Advances in Neural Information Processing Systems* 10, 273–279, MIT Press. (1998).
- [Hyv99] A. Hyvärinen, “Survey on independent component analysis”, *Neural Computing Surveys* 2, 94–128 (1999).
- [Hyv00] A. Hyvärinen y E. Oja, “Independent Component Analysis: Algorithms and Applications”, *Neural Networks* 13, 411-430 (2000).
- [Hyv04] A.Hyvärinen, “The FastICA package for Matlab”, <http://www.cis.hut.fi/projects/ica/fastica/> (2004).
- [Hyv09] A. Hyvärinen, J. Hurri y P.O. Hoyer, “Natural image Statistics: A probabilistic approach to early computational Vision”, Springer (2009).
- [Itt98] L. Itti, C. Koch y E. Niebur, “A model of saliency-based visual attention for rapid scene análisis”, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20,1254–1259 (1998).
- [Itt00] L. Itti L y C. Koch, “A saliency-based search mechanism for overt and covert shifts of visual attention”, *Vision Research* 40, 1489–1506 (2000).
- [Jam55] D. Jameson y L. M. Hurvich, “Some quantitative aspects of an opponent-colors theory. I. Chromatic responses and spectral saturation”, *Journal of the Optical Society of America* 45, 546–552 (1955).
- [Joh05] A. P. Johnson, F. Kingdom y C. J. Baker, “Spatiochromatic statistics of natural scenes: first- and second-order information and their correlational



- structure”, *Journal of the Optical Society of America A* 22, 2050–2059 (2005).
- [Joh08] E.N. Johnson, M.J. Hawken y R. Shapley, “The orientation selectivity of color-responsive neurons in macaque V1”, *Journal of Neuroscience* 28, 8096–8106 (2008).
- [Jud51] D. B. Judd, “Report of U.S. secretariat committee on colorimetry and artificial daylight”, *Proceedings of the 12th Session of the CIE*, 11 (1951).
- [Jud64] D. B. Judd, D.L. MacAdam y G.W. Wyszecki, “Spectral distribution of typical daylight as a function of correlated color temperatura”, *Journal of the Optical Society of America* 54, 1031–1040 (1964).
- [Jut91] C. Jutten y J. Héroult. “Blind separation of sources, part I: An adaptive algorithm based on neuromimetic Architecture”, *Signal Processing* 24, 1–10 (1991).
- [Kel13] C. J. Kellner y T. Wachtler, “A distributed code for color in natural scenes derived from center-surround filtered cone signals”, *Frontiers in Psychology* 4, 1-11 (2013).
- [Kha14] F. S. Khan, S. Beigpour, J. van de Weijer y M. Felsberg, “Painting-91: A large scale database for computational painting categorization”, *Machine Vision and Applications* 25, 1385– 1397 (2014).
- [Khi34] A. Khintchine, "Korrelationstheorie der stationären stochastischen Prozesse". *Mathematische Annalen* 109, 604–615 (1934).
- [Koc85] C. Koch y S. Ullman, “Shifts in selective visual attention: towards the underlying neural circuitry”, *Human Neurobiology* 4, 219– 227 (1985).
- [Kom98] H. Komatsu, “Mechanisms of central color vision”, *Current Opinion in Neurobiology* 8, 503–508 (1998).
- [Kra82] J. Krauskopf, D. R. Williams y D. W. Heeley, “Cardinal directions of color space”, *Vision Research* 22, 1123–1131 (1982).
- [Laj09] S. Lajvardi y Z. Hussain, “Facial expression recognition using log-Gabor filters and local binary pattern operators”, *Proceedings of the International Conference on Communication, Computer and Power (ICCCP'09)*, 349-353 (2009).
- [Lau81] S. B. Laughlin, “A simple coding procedure enhances a neuron’s information capacity”, *Zeitschrift für Naturforschung* 36, 910–912 (1981).
- [Lee05] H-C. Lee, “Introduction to color imaging science”, Cambridge University Press (2005).

- [Lee02] T-W. Lee, T. Wachtler y T.J. Sejnowski, “Color opponency is an efficient representation of spectral properties in natural scenes”, *Vision Research* 42, 2095–2103 (2002).
- [Lew99] M.S. Lewicki y B.A. Olshausen, “A probabilistic framework for the adaptation and comparison of image codes”, *Journal of the Optical Society of America A: Optics, Image Science, and Vision* 16, 1587–1601 (1999).
- [Lin06] D. T. Lindsey y A. M. Brown, “Universality of color names”, *Proceedings of the National Academy of Sciences USA* 103, 16608–16613 (2006).
- [Lin09] D. T. Lindsey y A. M. Brown, “World Color Survey color naming reveals universal motifs and their within-language diversity”, *Proceedings of the National Academy of Sciences USA* 106, 19785–19790 (2009).
- [Lin14] D. T. Lindsey y A. M. Brown, “The color lexicon of American English”, *Journal of Vision* 14 (2) : 17, 1-25, (2014).
- [Lin08] J.M. Linhares, P.D. Pinto y S. M. Nascimento, “The number of discernible colors in natural scenes”, *Journal of the Optical Society of America A* 25, 2918–2924 (2008).
- [Liv84] M.S. Livingstone y D.H. Hubel, “Anatomy and Physiology of a Color System in the Primate Visual-Cortex”, *The Journal of Neuroscience* 4, 309- 356 (1984).
- [Lov05] P. G. Lovell, D. J. Tolhurst, C. A. Párraga, R. Baddeley, U. Leonards, J. Troscianko y T. Troscianko, “Stability of the color-opponent signals under changes of illuminant in natural scenes”, *Journal of the Optical Society of America A* 22, 2060-2071 (2005).
- [Mac79] D. I. MacLeod y R. M. Boynton, “Chromaticity diagram showing cone excitation by stimuli of equal luminance”, *Journal of the Optical Society of America* 69, 1183–1186 (1979).
- [Mal86a] L.T. Maloney, “Evaluation of linear models of surface spectral reflectance with small numbers of parameters”, *Journal of the Optical Society of America A* 3, 1673–83 (1986).
- [Mal86b] L.T. Maloney y B.A. Wandell, “Color constancy: a method for recovering surface spectral Reflectance”, *Journal of the Optical Society of America A* 3, 29–33 (1986).
- [Mar10] I. Marin-Franch y D. H. Foster, “Number of perceptually distinct surface colors in natural scenes”, *Journal of Vision* 10 (9) : 9, 1-7 (2010).
- [Mar07] F. Martínez-Verdu, E. Perales, E. Chorro, D. de Fez, V. Viqueira y E. Gilabert, “Computation and visualization of the MacAdam limits for any

- lightness, hue angle and light source”, *Journal of the Optical Society of America A* 24, 1501–1515 (2007).
- [Mas13] K. Masaoka, R. S. Berns, M. D. Fairchild y F. M. Abed, “Number of discernible object colors is a conundrum”, *Journal of the Optical Society of America A* 30, 264–277 (2013).
- [Mas12] D. Massaro, F. Savazzi, C. Di Dio, D. Freedberg, V. Gallese, G. Gilli y A. Marchetti, “When art moves the eyes: a behavioral and eye-tracking study”. *PLOS ONE* 7: e37285–e37216 (2012).
- [Ma13] O. Masuda y S. M. C. Nascimento, “Best lighting for naturalness and preference”, *Journal of Vision* 13 (7) : 4, 1–14 (2013).
- [McD12] K. C. McDermott y M. A. Webster, “Uniform color spaces and natural image statistics”, *Journal of the Optical Society of America A* 29, 182–187 (2012).
- [Mdp21] <https://www.museodelprado.es/coleccion/obras-de-arte> (2021).
- [Mil18] Z. Milojevic, R. Ennis, M. Toscani y K. R. Gegenfurtner, “Categorizing natural color distributions”, *Vision Research* 151, 18–30 (2018).
- [Mon16] C. Montagner, J. M. M. Linhares, M. Vilarigues y S. M. C. Nascimento, “Statistics of colors in paintings and natural scenes”, *Journal of the Optical Society of America A* 33, 170–177 (2016).
- [Mor98] M. C. Morrone y D. C. Burr, “Feature detection in human vision: a phase-dependent energy model”, *Proceedings of the Royal Society of London, Serie B* 235, 221–224 (1988).
- [Mor07] B. Morse, D. Thornton, Q. Xia y J. Uibel, “Image-based color schemes”, *IEEE International Conference on Image Processing* 3, 497–500. (2007)
- [Nad94] J-P. Nadal y N. Parga, “Non-linear neurons in the low noise limit: a factorial code maximizes information transfer”, *Network* 5, 565–581 (1994).
- [Nas02a] S. M. C. Nascimento, F. P. Ferreira y D.H. Foster, “Statistics of spatial cone excitation ratios in natural scenes”, *Journal of the Optical Society of America A* 19, 1484–1490. (2002).
- [Nas02b] S.M.C. Nascimento, F.P. Ferreira y D.H. Foster, “Statistics of spatial cone-excitation ratios in natural scenes”, *Journal of the Optical Society of America A* 19, 1484-1490. (2002).
- [Nas12] S. M. C. Nascimento y O. Masuda, “Psychophysical optimization of lighting spectra for naturalness, preference and chromatic diversity”, *Journal of the Optical Society of America A* 29, 144–150 (2012).

- 
- [Nas16] S. M. C. Nascimento, K. Amano y D. H. Foster, “Spatial distributions of local illumination color in natural scenes”, *Vision Research* 120, 39–44 (2016).
- [Nie12] J. L. Nieves, S. M. C. Nascimento y J. Romero, “Contrast Edge colors under different natural illuminations”, *Journal of the Optical Society of America A* 29, 240–246 (2012).
- [Nie18] J.L. Nieves y J. Romero, “Heuristic analysis influence of saliency in the color diversity of natural images”, *Color Research and Application*, 1–13 (2018).
- [Nie20] J. L. Nieves, L. Gomez-Robledo, Yu-Jung Chen y J. Romero, “Computing the relevant colors that describe the color palette of paintings”, *Applied Optics* 59, 1732-1740 (2020).
- [Nie21] J.L. Nieves, J. Ojeda, L. Gómez-Robledo y J. Romero, “Psychophysical Determination of the Relevant Colours That Describe the Colour Palette of Paintings”, *Journal of Imaging*, 7(4), 72. (2021).
- [Nod08] C. Nodine, C. Mello-Thoms, E. Krupinski y P Locher, “Visual interest in pictorial art during an aesthetic experience”. *Spatial Vision* 21, 55–77 (2008).
- [Ohk06] K. Ohki, S. Chung, P. Kara, M. Hübener, T. Bonhoeffer y R.C. Reid, “Highly ordered arrangement of single neurons in orientation pinwheels”, *Nature* 442, 925–928 (2006).
- [Oje17] J. Ojeda, J. L. Nieves y J. Romero, “How daylight influences high-order chromatic descriptors in natural images”, *Applied Optics* 56, 120-127 (2017).
- [Oli02] A. Oliva, A. Torralba, A. Guerin-Dugue y J. Herault, “Global Semantic Classification of Scenes using Power Spectrum Templates”, *Proceedings of The Challenge of Image Retrieval, Electronic Workshops in Computing Series*, 1-12 (2002).
- [Olm04] A. Olmos y F. A. A. Kingdom, “A biologically inspired algorithm for the recovery of shading and reflectance images”, *Perception* 33, 1463-1473 (2004).
- [Ols96] B. Olshausen y D. Field, “Emergence of simple-cell receptive field properties by learning a sparse code for natural images”, *Nature* 381, 607–609 (1996).
- [Oom08] T. Oommen, D. Misra, N.K.C. Twarakavi, A. Prakash, B. Sahoo y S. Bandopadhyay, “ An Objective Analysis of Support Vector Machine

- Based Classification for Remote Sensing”, *Mathematical Geosciences* 40, 409-424 (2008).
- [Ori17] E.A.T. Orihuela, J.S. Martínez y E. J. S. Villaseñor (2017).  
[http://ocw.uc3m.es/matemáticas/algebra-lineal/teoría/algebra\\_teoría\\_11.pdf](http://ocw.uc3m.es/matemáticas/algebra-lineal/teoría/algebra_teoría_11.pdf).
- [Pap84] A. Papoulis, “Probability, Random Variables and Stochastic Processes”, McGraw-Hill (1984).
- [Par09] C.A. Parraga, J. Vazquez-Corral y M. Vanrell, “A new cone activation-based natural images dataset”, *Perception* 36 (Suppl), 180 (2009).
- [Par10] C.A. Parraga, R. Baldrich y M. Vanrell, “Accurate Mapping of Natural Scenes Radiance to Cone Activation Space: A New Image Dataset”, *Proceedings of Conference on Colour in Graphics, Imaging, and Vision*, 50-57 (2010).
- [Pey16] S. Peyvandi, J. Hernández-Andrés, F. J. Olmo, J. L. Nieves y J. Romero, “Colorimetric analysis of outdoor illumination across varieties of atmospheric conditions”, *Journal of the Optical Society of America A* 33, 1049-1059 (2016).
- [Pha92] D-T. Pham, P. Garrat y C. Jutten, “Separation of a mixture of independent sources through a maximum likelihood approach”, *Proceedings of European Signal Processing Conference*, 771–774 (1992).
- [Pit38] F H G Pitt y E W H Selwyn, “Colour of outdoor photographic objects”, *The Photographic Journal* 78, 115–121 (1938).
- [Pla12].- <https://es.m.wikipedia.org/wiki/Archivo:PlanckianLocus.png>
- [Poi98] M. R. Pointer y G. G. Attridge, “The number of discernible colours”, *Color Research and Application* 23, 52–54 (1998).
- [Pou13] T. Pouli, E. Reinhard y D. Cunningham, “Image Statistics in Visual Computing”, CRC Press (2013).
- [Psy20] <http://psychtoolbox.org/>
- [Rei79] W. Reichardt y T. Poggio, “Figure-ground discrimination by relative movement in the visual system of the fly”, *Biological Cybernetics* 35, 1–100 (1979).
- [Rei92] R. C. Reid y R. M. Shapley, “Spatial structure of cone inputs to receptive fields in primate lateral geniculate nucleus”, *Nature* 356, 716–718 (1992).
- [Rei01] E. Reinhard, P. Shirley y T. Troscianko, “Natural image statistics for computer graphics”, University of Utah School of Computing Technical report UUCS-01-002, 1-9, (2001).

- [Ric98] P. Ricchiazzi, S. Yang, C. Gautier y D. Sowle, “SBDART: a research and teaching software tool for plane-parallel radiative transfer in the earth’s atmosphere”, *Bulletin of the American Meteorological Society* 79, 2101–2114 (1998).
- [Rod65] R. W. Rodieck, “Quantitative analysis of cat retinal ganglion cell response to visual stimuli”, *Vision research*, 5, 583– 601 (1965).
- [Ros76] E. Rosch, C. B. Mervis, W. D. Gray, W. D. Johnson y P. Boyes-Braem, “Basic objects in natural categories”, *Cognitive Psychology* 8, 382-439 (1976).
- [Rub82] J. M. Rubin y W. A. Richards, “Color vision and image intensities: when are changes material?” *Biological Cybernetics* 45, 215–226 (1982).
- [Rud94] D. L. Ruderman y W. Bialek, “Statistics of natural images: scaling in the woods”, *Physical Review Letters* 73: 814–817 (1994).
- [Rud97] D.L. Ruderman, ” Origins of scaling in natural images”. *Vision Research* 37, 3385–3398 (1997).
- [Rud98] D.L. Ruderman, T.W. Cronin y C-C Chiao, “Statistics of cone responses to natural images: implications for visual coding”, *Journal of the Optical Society of America A* 15, 2036-2045 (1998).
- [Ruz20] M. Ruzon (2020).  
<https://es.mathworks.com/matlabcentral/fileexchange/24009-rgb2lab>
- [Sch07] J. Schanda, “Colorimetry: Understanding the CIE system”, John Wiley & Sons (2007).
- [Sha49] C.E. Shannon y W. Weaver, “The mathematical theory of communication”, University of Illinois Press (1949).
- [Sha11] R. Shapley y M. J. Hawken, “Color in the cortex: single and double-opponent cells”, *Vision research* 51, 701–717 (2011).
- [She08] S. K. Shevell y F. A. A. Kingdom, “Color in complex scenes”, *Annual Review of Psychology* 59, 143–166 (2008).
- [Sim97] E.P. Simoncelli, “Statistical Models for Images: Compression, Restoration and Synthesis”, *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems and Computers* 1, 673-678 (1997).
- [Sim99] E.P. Simoncelli y O. Schwartz, “Image statistics and cortical normalization models”, *Advances in Neural Information Processing* 11, 153–59 (1999).



- [Sim01] E.P. Simoncelli y B.A. Olshausen, “Natural image statistics and neural representation”, *Annual Review of Neuroscience*. 24, 1193–1216 (2001).
- [Sri82] M.V. Srinivasan, S.B. Laughlin y A. Dubs, “Predictive coding: A fresh view of inhibition in the retina”, *Journal of the Royal Society of London B* 216, 427–459 (1982).
- [Süs00] S. Süsstrunk, J. Holm y G. D. Finlayson, “Chromatic adaptation performance of different RGB sensors”, *Proceedings of SPIE Color Imaging* 4300, 172–183 (2000).
- [Tad00] Y. Tadmor y D.J. Tolhurst, “Calculating the contrasts that retinal ganglion cells and LGN neurones encounter in natural scenes”. *Vision Research* 40, 3145–3157 (2000).
- [Tai00] D.R. Taylor, L.H. Finkel y G. Buchsbaum, “Color-opponent receptive fields derived from independent component analysis of natural images”, *Vision Research* 40, 2671-2676 (2000).
- [Tap02] M. F. Tappen, W. T. Freeman y E. H. Adelson, “Recovering intrinsic images from a single image”, *Advances in Neural Information Processing Systems* 15, 1459–1472 (2002).
- [Tho99a] M. G. A. Thomson, “Higher-order structure in natural scenes”, *Journal of the Optical Society of America* 16, 1549-1553 (1999).
- [Tho99b] M. G. A. Thomson, “Visual coding and the phase structure of natural scenes”, *Network: Computation in Neural Systems* 10, 123–132 (1999) .
- [Tho00] M. G. A. Thomson, D. H. Foster y R. J. Summers, “Human sensitivity to phase perturbations in natural images: a statistical framework”, *Perception* 29, 1057- 1069 (2000).
- [Tho84] L.G. Thorell, R.L. De Valois y D.G Albrecht, “Spatial mapping of monkey V1 cells with pure color and luminance stimuli”. *Vision Research* 24, 751–769 (1984).
- [Tia10] M. Tian, S. Wan y L. Yue, “A Color Saliency Model for Salient Objects Detection in Natural Scenes”, *Advances in Multimedia Modeling*, vol. 5916, *Lecture Notes in Computer Science*, 240–250 (2010).
- [Tka11] G. Tkacik, P. Garrigan, C. Ratliff, G. Milčinski, J. M. Klein, L. H. Seyfarth, P. Sterling, D. H. Brainard y V. Balasubramanian, “Natural images from the birthplace of the human eye”, *PLOS ONE* 6: e20409 (2011).
- [Tol92] D. J. Tolhurst, Y. Tadmor y Tang Chao, “Amplitude spectra of natural images”, *Ophthalmic and Physiological Optics* 12, 229-232 (1992).

- 
- [Tor02] A. Torralba y A. Oliva, “Depth estimation from image structure”, IEEE Transactions on Pattern Analysis and Machine Intelligence 24, 1-13 (2002).
- [Tor03] A. Torralba y A. Oliva, “Statistics of natural image categories”, Network: Computation Neural Systems 14, 391-412 (2003).
- [Ugr] <http://wpd.ugr.es/~bioestad/guia-spss/practica-6/>
- [Uv10] <https://www.uv.es/innomide/spss/>
- [VdS98] A. van der Schaaf, “Natural image statistics and visual processing”, [University Library Groningen][Host] (1998).
- [VGe20] E. Van Geert y J. Wagemans, “Order, complexity, and aesthetic appreciation”, Psychology of Aesthetics, Creativity and the Arts 14, 135 - 152 (2020).
- [vHa92] J.H. van Hateren, “A theory of maximizing sensory information”, Biological Cybernetics 68, 23–29 (1992).
- [vHa98] J. H. van Hateren y A. van der Schaaf, “Independent component filters of natural images compared with simple cells in primary visual cortex”. Proceedings of the Royal Society London B 265, 359– 366 (1998).
- [Vin00] W.E. Vinje y J.L. Gallant, “Sparse coding and decorrelation in primary visual cortex during natural vision”, Science 287, 1273–1276 (2000).
- [Vla07] T. Vladusich, “Chromatic aberration and the roles of double-opponent and color-luminance neurons in color vision”, Neural Network 20,153–155 (2007).
- [Wac01] T. Wachtler, T-W. Lee y T.J. Sejnowski, “Chromatic structure of natural scenes”, Journal of the Optical Society of America A 18, 65-77 (2001).
- [Wie66] T.N. Wiesel y D.H. Hubel, “Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey”, Journal of Neurophysiology 29, 1115-1156 (1966).
- [Wil01] B. Willmore y D.J. Tolhurst, “Characterizing the sparseness of neural codes”, Network 12, 255-270 (2001).
- [Xia10] J. Xiao, J. Hays, K. Ehinger, A. Oliva y A. Torralba, “SUN database: large-scale scene recognition from Abbey to Zoo”, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 3485-3492 (2010).
- [Yas10] M. Yasuda, T. Banno y H. Komatsu, “Color selectivity of neurons in the posterior inferior temporal cortex of the macaque monkey”, Cerebral Cortex 20,1630 –1646 (2010).

- [Yoo08].-Yoonessi, F. A. A. Kingdom, y S. Alqawlaq, “Is color patchy?”, *Journal of the Optical Society of America A* 25, 1330–1338 (2008).
- [Zek98] S. Zeki y L. Marini. “Three cortical stages of colour processing in the human brain”, *Brain* 121,1669 –1685 (1998).
- [Zha97].- X. Zhang y B. A. Wandell, "A spatial extension of CIELAB for digital color-image reproduction", *Journal of Society for Information Display* 5, 61-63 (1997).
- [Zho08] C. Zhou y B. W. Mel, “Cue combination and color edge detection in natural scenes”, *Journal of Vision* 8 (4) : 4, 1–25 (2008).

---

**ANEXOS.**

**Anexo I: Procedimiento para comparar las medias de  
múltiples distribuciones de un descriptor estadístico**

**Anexo II: Lista de figuras**

**Anexo III: Lista de tablas**

---



---

## **ANEXO I: PROCEDIMIENTO PARA COMPARAR LAS MEDIAS DE MÚLTIPLES DEISTRIBUCIONES DE UN DESCRIPTOR ESTADÍSTICO**

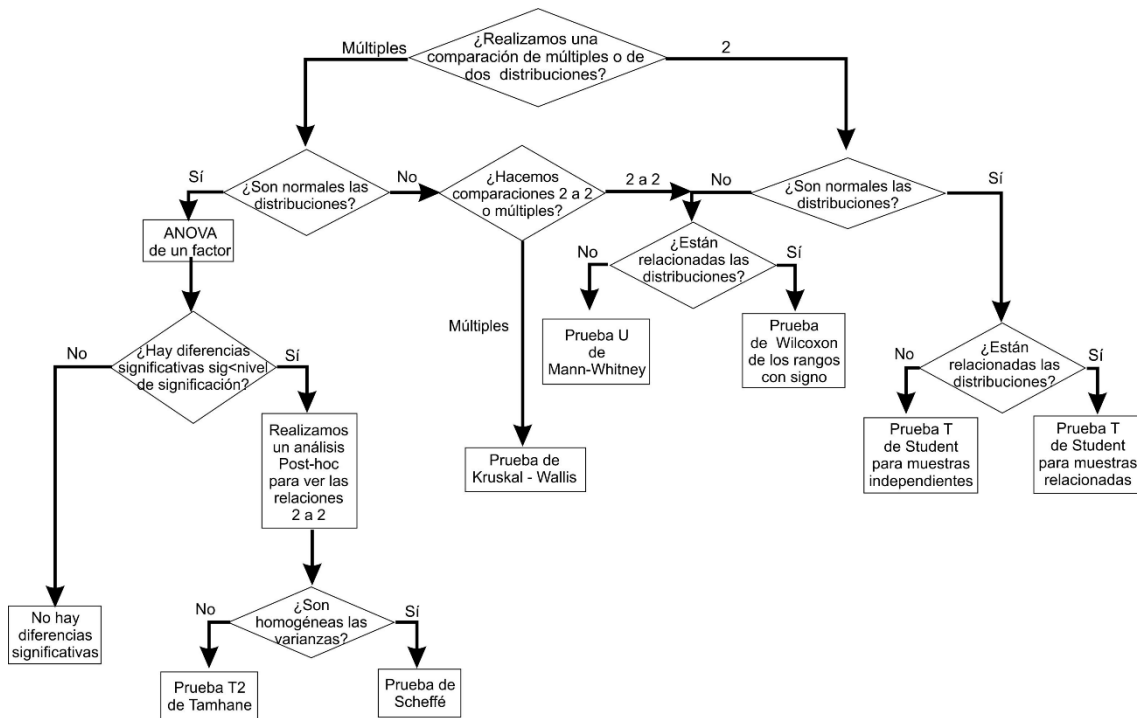
---

Para comparar las medias de múltiples distribuciones de un descriptor dado hemos utilizado el esquema de la figura A.1 basado en el propuesto por el Grupo de innovación educativa de la Universitat de Valencia [Uv10] y en la implementación en SPSS descrita en la guía elaborada por la universidad de Granada [Ugr].

La normalidad de una distribución se analiza con el Test Kolmogorov-Smirnov (K-S) y la homogeneidad de las varianzas con la prueba de Levene. En la práctica se comenzaba estudiando la normalidad de todas las muestras que había que comparar; de todas aquellas que tenían una distribución normal se realizaba el ANOVA de un factor y de aquellas que no tenían una distribución normal se analizaba su relación con las demás mediante la prueba U de Mann-Whitney (pues se trata de muestras independientes). En el caso de que las distribuciones no fueran normales se utilizó un test no paramétrico como el de Kruskal-Wallis; no obstante, dado que ANOVA se comporta con robustez si el tamaño de la muestra es grande ante la falta de normalidad, se obtuvieron en general resultados parecidos en estos casos aplicando ambas pruebas.

Cuando para el intervalo de confianza elegido (en este trabajo el 95%) había diferencias significativas entre todas las distribuciones del análisis de ANOVA de un factor se realizaba uno de los análisis Post-hoc descritos según las varianzas de las muestras fueran o no homogéneas, para poder determinar los posibles distintos grupos homogéneos (con medias estadísticamente iguales entre sí).





*Figura A.1: Esquema del proceso seguido en la comparación de los valores medios de múltiples distribuciones de un descriptor estadístico.*

---

## ANEXO II: LISTA DE FIGURAS

---

|   |    |
|---|----|
| Figura 1.1: Relación de los descriptores estadísticos en el dominio espacio-cromático analizados en el presente trabajo. ....   | 2  |
| Figura 2.1: Esquema de los elementos y componentes de la luz día.....   | 8  |
| Figura 2.2: Diagrama de cromaticidad CIE xy 1931, lugar geométrico de las cromaticidades del cuerpo negro y varias rectas de temperatura de color correlacionada constante .....  | 8  |
| Figura 2.3: Esquema del sistema visual .....  | 9  |
| Figura 2.4: Estructura celular de la retina humana .....  | 10 |
| Figura 2.5: Campos receptivos y sus funciones correspondientes de las células ganglionares y del LGN.....   | 11 |
| Figura 2.6: Esquema de la estructura del córtex visual estriado .....   | 13 |
| Figura 2.7: Campos receptivos y sus funciones correspondientes para una célula oponente doble rojo-verde concéntrica (primera fila) y orientada (segunda fila). La tercera fila muestra las posibles topologías de las células oponentes dobles concéntricas cromaticas ..... | 14 |
| Figura 3.1 .- Campos receptivos de células oponentes simples orientadas de un macaco (a), ejemplos de tres vectores peso encontrados en la maximización de la dispersión en   |    |

|   |    |
|---|----|
| imágenes naturales (b) y algunas de las características obtenidas al aplicar ICA a trozos de imágenes naturales monocromáticas (c).....   | 21 |
| Figura 3.2 .- Ejemplo de una imagen natural y de su correspondiente versión blanqueada en la que todavía se aprecian estructuras claras. ....   | 22 |
| Figura 4.1: Ejemplos de imágenes pertenecientes a las doce categorías semánticas utilizadas en este trabajo. ....   | 28 |
| Figura 4.2: Espacio cromático CIELAB.<br>( <a href="http://sensing.konicaminolta.com.mx/images/blogImages/ColorSphere.jpg">http://sensing.konicaminolta.com.mx/images/blogImages/ColorSphere.jpg</a> ) .....  | 29 |
| Figura 4.3 : Histogramas correspondientes a los tres canales $L^*$ , $a^*$ y $b^*$ de las imágenes rurales (fila superior) y de entornos humanos (fila inferior). Superpuesta en rojo aparece la distribución normal y en verde la distribución Laplaciana correspondientes a cada conjunto de datos .....  | 30 |
| Figura 4.4 : Una imagen original (a), su gama de color con la elipse de ajuste de mínimos cuadrados (b) y volumen de color encerrado en un sólido convexo. ....   | 33 |
| Figura 4.5: Gráfica de coordenadas paralelas de cinco descriptores cromáticos normalizados al intervalo [0,1] (Ratio de los ejes, ángulo del eje mayor con respecto a $a^*$ positivo y el área de las elipses de ajuste de la gama de color, volumen de color y número de colores discernibles). La imagen superior representa las siete clases de imágenes rurales, la central las de entornos humanos y la inferior los dos tipos rurales y entornos humanos entre si. .... | 37 |
| Figura 5.1 : La imagen de la derecha se obtiene distribuyendo de manera aleatoria los píxeles de la imagen natural de la izquierda de modo que las dos presentarían la misma estadística de primer orden. ....  | 39 |
| Figura 5.2 : Histogramas del gradiente correspondientes a las tres componentes ( $L^*$ , $a^*$ y $b^*$ ) de imágenes rurales (fila superior) y de entornos humanos (fila inferior). ....  | 41 |
| Figura 5.3 : En la fila superior, imagen original. La segunda fila representa, en falso color, sus componentes $L^*$ , $a^*$ y $b^*$ . Tercera fila, los bordes correspondientes a los tres planos. Fila inferior, histogramas conjuntos de los distintos canales. ....   | 42 |
| Figura 5.4 : Histogramas conjuntos de los gradientes entre los canales $L^*$ y $a^*$ (primera columna), $L^*$ y $b^*$ (segunda columna) y $a^*$ y $b^*$ (tercera columna) correspondientes a 4 categorías rurales (bosques y parques, flores y frutas, montaña y playa) y 2 de entornos humanos (interiores y calles). ....   | 44 |
| Figura 5.5 : En la fila superior una imagen original y una representación conjunta en RGB de la autocorrelación calculada por separado para cada uno de los planos $L^*$ , $a^*$ y $b^*$ . En la fila inferior la función de autocorrelación de cada uno de los tres canales. ( $L^*$ izquierda, $a^*$ centro y $b^*$ derecha).....   | 47 |
| Figura 5.6 : Rectas de ajuste entre el logaritmo del espectro de potencia para los tres canales $L^*$ , $a^*$ y $b^*$ en tres direcciones (horizontal $\theta=0^\circ$ ,diagonal $\theta=45^\circ$ y vertical $\theta= 90^\circ$ )  |    |

|   |    |
|---|----|
| y el logaritmo de la frecuencia espacial en ciclos/imagen, para las tres imágenes perteneciendo las dos primeras a la misma categoría semántica. ....   | 49 |
| Figura 5.7: Espectros de potencia de la imagen superior para los tres canales $L^*$ , $a^*$ y $b^*$ (primera columna); intersección de planos horizontales con el espectro de potencia que nos proporcionan las regiones que encierran el 60% (plano azul), 80% (plano rojo) y 90%(plano verde)de la energía total (segunda columna) y las firmas espectrales correspondientes (60% verde claro, 80% verde oscuro y 90% azul)(tercera columna).....   | 52 |
| Figura 5.8: Firma espectral de los planos $L^*$ , $a^*$ y $b^*$ de las distintas clases semánticas correspondientes al 60% (verde claro), 80% (verde oscuro) y 90% (azul). ....   | 54 |
| Figura 5.9: Mapas de distribución del logaritmo neperiano del espectro de potencia de los planos $L^*$ , $a^*$ y $b^*$ de las distintas clases semánticas. ....   | 55 |
| Figura 5.10: Representación de la nube de puntos de nuestras imágenes rurales en tres espacios de color LMS, RGB y CIELAB en tres sistemas de referencia diferentes. En la primera fila se representan en sus valores originales (en el caso del espacio LMS está representado en el espacio logarítmico). En la segunda fila en las tres componentes principales obtenidas al aplicar el PCA. En la tercera fila en los tres canales de una representación color oponente (canales luminancia (Lum), oponente rojo-verde (RG) y oponente amarillo-azul (BY)). .... | 58 |
| Figura 5.11: Representación de la nube de puntos de nuestras imágenes de entornos humanos en tres espacios de color logaritmo decimal de LMS, RGB y CIELAB. En la primera fila se representan en las tres componentes principales obtenidas al aplicar el PCA. En la segunda fila en los tres canales de una representación color oponente (canales luminancia (Lum), oponente rojo-verde (RG) y oponente amarillo-azul (BY)). ....   | 59 |
| Figura 5.12: Ejemplo de un trozo cuadrado de una de las imágenes de trabajo de dimensiones 8 x 8 (a). Representación de los valores de cada píxel del trozo seleccionado en los tres planos cromáticos de definición (b). Ordenación de los datos de cada trozo como un vector columna de dimensión 192 x 1 (c) de la matriz de datos a la que le aplicamos el PCA. ....  | 60 |
| Figura 5.13: (a) 192 componentes principales obtenidas de nuestras imágenes rurales en los espacios de color log LMS, RGB y CIELAB dispuestas de izquierda a derecha y de arriba abajo según el orden decreciente de sus autovalores. (b) Cromaticidades de los valores de los 192 píxeles de las componentes indicadas en un plano isoiluminante. El eje horizontal corresponde a variaciones en el canal RG mientras que el eje vertical corresponde con cambios en el canal BY. La luminancia puede deducirse de la luminosidad de los puntos. ....              | 61 |
| Figura 5.14: Histogramas circulares que muestran los ángulos de alineación (entre $0^\circ$ y $180^\circ$ ) de los píxeles de las componentes principales obtenidas al aplicar PCA a los trozos de imágenes rurales en cada uno de los espacios de color. ....  | 63 |

|  |    |
|--|----|
| Figura 5.15: 192 componentes principales obtenidas de nuestras imágenes de entornos humanos en los espacios de color log LMS, RGB y CIELAB dispuestas de izquierda a derecha y de arriba abajo según el orden decreciente de sus autovalores. ....   | 64 |
| Figura 5.16: Autovalores correspondientes a las componentes principales de los trozos de imágenes rurales en el espacio CIELAB .....   | 64 |
| Figura 5.17: Una imagen natural y tres reconstrucciones de la misma usando las 5 primeras componentes principales (78,6% del total de la varianza de los datos), las 15 primeras (90,8% de la varianza total) y finalmente las 25 primeras (95% de la varianza total) de los trozos de las imágenes de estudio. ....   | 65 |
| Figura 5.18: (Primera columna) 2 ejemplos de imágenes naturales (Segunda columna) Reconstrucción de la imágenes originales usando sus respectivos espectros de amplitud y un espectro de fase de 0°. (Tercera columna) Reconstrucción de las imágenes originales usando un espectro de amplitud normalizado y su espectro de fase original. (Última columna) Imágenes resultantes de combinar los espectros de fase originales intercambiando los espectros de amplitud. ....  | 66 |
| Figura 6.1: Nube de puntos de los valores originales (izquierda) de los píxeles de la imagen en el espacio de color CIELAB y la nube de puntos de los mismos datos blanqueados (derecha). ....   | 69 |
| Figura 6.2.- A la izquierda se representan tres distribuciones con asimetrías negativa, cero (función simétrica) y positiva respectivamente. Las funciones asimétricas presentan una cola más larga en el sentido de la asimetría [Doa11]. A la derecha se presentan tres nuevas distribuciones simétricas con curtosis menor que 3 (distribución platicúrtica), 3 (distribución mesocúrtica) y mayor que 3 (distribución leptocúrtica) respectivamente [Čis10] .....  | 71 |
| Figura 6.3: (a) 192 características (filas de la matriz <b>A</b> ) obtenidas de aplicar ICA a nuestras imágenes rurales en los espacios de color log LMS, RGB y CIELAB dispuestas de izquierda a derecha y de arriba abajo según el orden decreciente de su norma L2. (b) Cromaticidades de los valores de los 192 píxeles de las características indicadas en un plano isoiluminante. El eje horizontal corresponde a variaciones en el canal RG mientras que el eje vertical corresponde con cambios en el canal BY. La luminancia puede deducirse de la luminosidad de los puntos. .... | 78 |
| Figura 6.4: Histogramas circulares que muestran los ángulos de alineación (entre 0° y 180°) de los píxeles de las características obtenidas al aplicar ICA a los trozos de imágenes rurales en cada uno de los espacios de color .....   | 80 |
| Figura 6.5: 192 características (filas de la matriz <b>A</b> ) obtenidas de aplicar ICA nuestras imágenes de entornos humanos en los espacios de color log LMS, RGB y CIELAB dispuestas de izquierda a derecha y de arriba abajo según el orden decreciente de su norma L2. ....   | 81 |

|  |    |
|--|----|
| Figura 6.6: Los 4 filtros log Gabor de frecuencias 3, 6, 12 y 24 ciclos por grado utilizados en el estudio del índice de agrupamiento.....   | 82 |
| Figura 6.7: Muestra el método utilizado para obtener el índice de agrupamiento (última columna). La primera columna muestra la imagen original utilizada en el ejemplo. La segunda columna sus canales L* (luminancia), a* (rojo-verde) y b* (amarillo-azul) en el espacio CIELAB. En las cuatro columnas siguientes (correspondientes a los 4 filtros log Gabor de frecuencias 3, 6, 12 y 24 ciclos por grado utilizados) bajo la imagen filtrada aparece la correspondiente imagen binarizada aplicando los umbrales fijos. .... | 83 |
| Figura 7.1: Ejemplo de una imagen natural y una de entonos humanos simuladas bajo tres iluminantes diferentes caracterizados por CCTs de 2735, 6478 y 25889 K.....   | 87 |
| Figura 7.2: Imágenes hiperespectrales utilizadas a la hora de evaluar la transformación cromática de Bradford para simular escenas bajo distintos iluminantes. Las imágenes de la primera fila pertenecen a la colección de Foster, Nascimento y Amano (2004) y el resto a la colección de Nascimento, Ferreira y Foster (2002) [Fos06,Nas02b]. ....   | 89 |
| Figura 7.3: En la columna de la izquierda aparecen dos imágenes simuladas con la transformación cromática de Bradford bajo los iluminantes de CCTs de 2735 K y 25889 K. En la columna de la derecha se muestra el histograma de las diferencias de color s-CIELAB entre la imagen original y la simulada bajo la transformación cromática de Bradford utilizada. ....  | 91 |
| Figura 7.4: Valor medio del primer momento (media) de la intensidad de las componentes L*, a* y b* para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante.....  | 92 |
| Figura 7.5: Valor medio del segundo momento (desviación típica) de la intensidad de las componentes L*,a* y b* para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante. ....   | 94 |
| Figura 7.6: Elipses de ajuste de mínimos cuadrados de las gamas de color media de cada una de las categorías semánticas para las temperaturas de color de los iluminantes 2735 K (líneas continuas), 6478 K (líneas discontinuas) y 25889 K (líneas de puntos). ....   | 95 |
| Figura 7.7: Variación con la CCT del iluminante de los descriptores de la elipse que mejor ajusta la gama de color: área en unidades CIELAB (primera fila), ángulo del eje mayor con el semieje positivo del eje a*(segunda fila) y razón entre los semiejes de la elipse (tercera fila); para imágenes rurales (primera columna) y de entornos humanos (segunda columna). ....  | 96 |
| Figura 7.8: Valor medio del número de colores discernibles de cada categoría semántica para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante.....  | 97 |
| Figura 7.9: Volumen de color en unidades CIELAB de cada categoría semántica para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante.....   | 99 |

|  |     |
|--|-----|
| Figura 7.10: Histogramas conjuntos en escala logarítmica para los tres canales $L^*$ , $a^*$ y $b^*$ de las imágenes rurales y las imágenes de entornos humanos bajo tres iluminantes caracterizados por temperaturas de 2735 K, 6478 K y 25889 K.....   | 101 |
| Figura 7.11: Información mutua entre los gradientes de las componentes $L^*$ , $a^*$ y $b^*$ para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante.....  | 103 |
| Figura 7.12: Pendiente espectral de las componentes cromáticas ( $a^*$ y $b^*$ ) para tres orientaciones horizontal ( $0^\circ$ ), diagonal ( $45^\circ$ ) y vertical ( $90^\circ$ ) para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante. ....   | 105 |
| Figura 7.13: Diagramas de cajas y bigotes de las pendientes espectrales del canal luminancia ( $L^*$ ) en función de la temperatura de color del iluminante para las distintas categorías semánticas (bp: bosques y parques, ca: campo, co: costa, ff: flores y frutas, mo: montaña, pl: playa, rc: ríos y cascadas, rur: imágenes rurales, au: autopistas, ci: ciudades, ed:edificios, in: interiores, st: calles y art: imágenes de entornos humanos) y tres orientaciones $0^\circ$ (horizontal), $45^\circ$ (diagonal) y $90^\circ$ (vertical). .... | 106 |
| Figura 7.14: Firma espectral (las tres primeras filas de cada figura) y mapas de distribución (las tres últimas filas de cada figura) para los tres canales $L^*$ , $a^*$ y $b^*$ y las diferentes categorías semánticas para tres iluminantes caracterizados por las temperaturas de color 2735 K (la primera fila y la cuarta fila), 6478 K (la segunda fila y la quinta fila) y 25889 K (la tercera fila y la sexta fila) .....   | 107 |
| Figura 7.15: Las 192 componentes principales correspondientes a 17500 trozos de imágenes rurales (primera fila) y de entornos humanos (segunda fila) simuladas en el espacio de color CIELAB bajo tres iluminantes: 2735 K, 6478 K y 25889 K, histogramas circulares correspondientes al ángulo (entre $0^\circ$ y $180^\circ$ ) que forman dichas componentes con el semieje horizontal positivo para las imágenes rurales(tercera fila) y de entornos humanos (cuarta fila).....   | 109 |
| Figura 7.16: Autovalores normalizados de las 25 primeras componentes principales correspondientes al PCA de trozos de imágenes rurales y de entornos humanos bajo tres iluminantes: 2735 K, 6478 K y 25889 K .....   | 110 |
| Figura 7.17: Valor medio de la asimetría de la intensidad de las componentes $L^*$ , $a^*$ y $b^*$ para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante.....  | 111 |
| Figura 7.18: Valor medio de la curtosis de la intensidad de las componentes $L^*$ , $a^*$ y $b^*$ para imágenes rurales (primera columna) y de entornos humanos (segunda columna) en función de la CCT del iluminante.....   | 112 |
| Figura 7.19: Las 192 características de las componentes independientes correspondientes a 17500 trozos de imágenes rurales (primera fila) y de entornos humanos (segunda fila) simuladas en el espacio de color CIELAB bajo tres iluminantes: 2735 K, 6478 K y 25889   |     |



|   |     |
|---|-----|
| K (parte superior) e histogramas normalizados comparativos para los tres iluminantes de los ángulos que forman dichas componentes con el semieje horizontal positivo (imágenes rurales a la izquierda, imágenes de entornos humanos a la derecha). .....  | 114 |
| Figura 7.20.-. Diagramas de cajas y bigotes del valor medio del índice de agrupamiento en función de la CCT del iluminante para cada uno de los canales luminancia $L^*$ y cromáticos $a^*$ y $b^*$ para las distintas categorías semánticas (bp: bosques y parques, ca: campo, co: costa, ff: flores y frutas, mo: montaña, pl: playa, rc: ríos y cascadas, rur: imágenes rurales, au: autopistas, ci: ciudades, ed:edificios, in: interiores, st: calles y art: imágenes de entornos humanos) ..... | 115 |
| Figura 8.1: Esquema general del modelo de saliencia clásico bottom-up.....  | 118 |
| Figura 8.2: Ejemplo procedente de Nieves y Romero [Nie18] de localización sobre la imagen original de los colores salientes destacables (puntos amarillos) obtenidos por el algoritmo de Harel et al. [Har06] (izquierda) y la imagen correspondiente segmentada según esos colores salientes destacables (derecha). .....  | 119 |
| Figura 8.3.- Ejemplo en una de las imágenes de la base de datos de Khan et al. [Kha14] de los resultados obtenidos al aplicar el algoritmo de Nieves et al. [Nie20]. En la fila superior izquierda la imagen original, en la fila superior derecha la imagen segmentada usando los colores relevantes obtenidos que aparecen en la fila inferior (paleta de color de la pintura).....   | 122 |
| Figura 8.4: Análisis de la dependencia de los parámetros del algoritmo de colores relevantes. Arriba: NRC en función del tamaño del cubo del espacio de color, central: NRC en función del umbral general (porcentaje del total de los píxeles en un cubo) y abajo: NRC para un tamaño del cubo de 20 unidades y un umbral del 4% en función del valor luminancia y percentil del croma .....   | 125 |
| Figura 8.5: Esquema del algoritmo del cálculo del número de colores relevantes de Nieves et al. [Nie20] que hemos aplicado para determinar el NRC de cada una de nuestras imágenes naturales.....   | 126 |
| Figura 8.6: Total de colores relevantes en el espacio de color CIELAB de los dos conjuntos de imágenes (puntos azules para las imágenes de este trabajo y puntos magenta para las imágenes de la base de datos McGill) rurales y de entornos humanos. ....  | 128 |
| Figura 8.7 : Histogramas circulares de los ángulos de las direcciones de los colores relevantes en el plano RG y BY para los dos conjuntos de imágenes y las categorías rurales y de entornos humanos .....   | 130 |
| Figura 8.8 : (Izquierda) Las 192 características (filas de la matriz A) obtenidas de aplicar ICA a las imágenes de la base de datos McGill en el espacio de color RGB (imágenes rurales en la fila superior y de entornos humanos en la fila inferior). (Derecha) Cromaticidades de los valores de los 192 píxeles de cada una de las características indicadas en un plano isoiluminante.....  | 131 |

Figura 8.9: Histogramas conjuntos normalizados de los ángulos (entre 0° y 180 °) de las direcciones de los colores relevantes y de la primera componente principal de cada una de las características obtenidas con el ICA para nuestras imágenes (fila superior) y de la base de datos de McGill (fila inferior) para imágenes rurales (izquierda) y de entornos humanos (derecha).....132

Figura 9.1.- Histograma de la distribución de direcciones cromáticas asociada a cada función base obtenidas al aplicar ICA al conjunto de imágenes de cuadros de la base de datos de Khan (a la derecha) y de los ángulos en el plano RG y BY de los colores relevantes de todos los observadores en el experimento psicofísico [Nie21]. .....141

Figura A.1: Esquema del proceso seguido en la comparación de los valores medios de múltiples distribuciones de un descriptor estadístico. ....162

---

### ANEXO III: LISTA DE TABLAS

---

|   |    |
|---|----|
| Tabla 4.1: Promedio para las distintas clases utilizadas de la media ( $m_1$ ) y la desviación típica ( $\sqrt{m_2}$ ). En la última columna se muestra el color en el espacio RGB que le correspondería a dichos valores medios. ....  | 32 |
| Tabla 4.2: Valores medios del área (en unidades CIELAB), ratio entre los ejes y ángulo (en grados) del eje mayor con respecto al eje positivo de la componente $a^*$ de las elipses de mejor ajuste cuadrático a la gama cromática de las distintas categorías semánticas de imágenes. Representación de dichas elipses para cada categoría.....    | 34 |
| Tabla 4.3: Rango del número de colores discernibles correspondientes a las distintas categorías semánticas. ....  | 36 |
| Tabla 4.4: Fila superior: Una imagen ejemplo y el diagrama de puntos junto con la recta de ajuste correspondiente al representar el logaritmo de la ratio del número de colores discernibles frente al logaritmo de la frecuencia absoluta. Fila inferior: Pendiente media de la recta de ajuste para las distintas categorías semánticas. ....     | 36 |
| Tabla 5.1 : Información mutua de los gradientes entre los distintos canales para las diferentes categorías semánticas.....  | 45 |
| Tabla 5.2 : Pendientes medias de las rectas de ajuste entre el logaritmo del espectro de potencia y el logaritmo de la frecuencia espacial para los tres canales $L^*$ , $a^*$ y $b^*$ en tres direcciones (horizontal $\theta=0^\circ$ ,diagonal $\theta=45^\circ$ y vertical $\theta= 90^\circ$ ) para las diferentes categorías semánticas. .... | 50 |
| Tabla 6.1 : Valores de la asimetría (S) en los tres canales del espacio cromático CIELAB para las distintas categorías semánticas de nuestro estudio.....   | 71 |

|   |     |
|---|-----|
| Tabla 6.2: Valores de la curtosis (K) en los tres canales de los espacios cromáticos RGB (entradas a los fotorreceptores) y CIELAB (salidas de las células ganglionares retinianas) para las distintas categorías semánticas de nuestro estudio.....  | 73  |
| Tabla 6.3: muestra el índice de agrupamiento de los tres canales L*, a*,b* de las distintas categorías de nuestras imágenes ordenados en orden decreciente del índice en el canal luminancia L*. .....  | 85  |
| Tabla 7.1: Diferencia de color CIELAB ( $\Delta E$ ) media para cada una de las imágenes hiperespectrales y cada una de las categorías semánticas (imágenes rurales y de entornos humanos), entre la imagen original y la simulada por la adaptación cromática de Bradford para los iluminantes extremos de nuestro rango de trabajo de CCTs 2735 K y 25889 K. .... | 90  |
| Tabla 8.1: NRC obtenidos para los distintos conjuntos de imágenes naturales. ....   | 127 |
| Tabla 8.2 : Coeficientes de correlación para cada plano del espacio RGB entre las imágenes originales y las imágenes segmentadas de los dos conjuntos de imágenes. ....   | 129 |