

Procesamiento de imágenes nocturnas en espacios de color perceptuales.

Cepeda-Negrete, J.

Cita:

Cepeda-Negrete, J. (2016). *Procesamiento de imágenes nocturnas en espacios de color perceptuales* (Tesis de Doctorado). Universidad de Guanajuato, Salamanca, México.

Dirección estable: <https://www.aacademica.org/jcepedanegrete/15>

ARK: <https://n2t.net/ark:/13683/pa8v/3mB>



Esta obra está bajo una licencia de Creative Commons.
Para ver una copia de esta licencia, visite
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>.

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.



UNIVERSIDAD DE GUANAJUATO

Campus Irapuato-Salamanca

División de Ingenierías

**PROCESAMIENTO DE IMÁGENES
NOCTURNAS EN ESPACIOS DE COLOR
PERCEPTUALES**

TESIS PROFESIONAL

QUE PARA OBTENER EL GRADO DE

DOCTOR EN INGENIERÍA ELÉCTRICA

PRESENTA:

M.I. Jonathan Cepeda Negrete

Director:

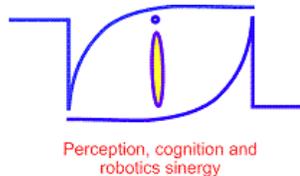
Dr. Raúl Enrique Sánchez Yáñez

*“Para mi hija Allison
y mi esposa Pamela”.*

Agradecimientos

Agradezco a las personas e instituciones que hicieron posible que este proyecto iniciara y concluyera en tiempo y forma.

Al Dr. Raúl E. Sánchez Yáñez por confiar en mis capacidades para poder desarrollar un proyecto de doctorado. Por sus valiosas ideas y aportaciones a este trabajo de tesis. Por su apoyo moral, y por gestionar los recursos necesarios para la realización del proyecto y de una estancia de investigación. Por su invaluable amistad que día tras día se hacía presente y que sin duda deja marca en cada una de estas páginas.



A todo el equipo de trabajo del Laboratorio de Visión, Robótica e Inteligencia Artificial (LaViRIA, www.laviria.org) dirigido por el Dr. Victor Ayala Ramirez. Gracias a los conocimientos generados y compartidos por cada uno de los integrantes, que de una u otra forma influyen en la generación de ideas que terminan reflejándose en este trabajo.



Al director de la DICIS, Dr. Roberto Rojas Laguna y al Rector del Campus Irapuato-Salamanca, Dr. Ernesto A. Camarena Aguilar, por el apoyo económico brindado para gastos de transporte en la estancia de investigación. A la Dirección de Apoyo a la Investigación y al Posgrado de la Universidad de Guanajuato (DAIP) y a la Dirección de Cooperación Académica, por los apoyos económicos brindados para la realización de

la estancia. Y también al Consejo Estatal de Ciencia y Tecnología (CONCYTEG) por el apoyo en gastos de instalación.



Deseo agradecer enormemente al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca 290747 (388681/254884) para el programa de doctorado. De igual manera por su beca Mixta como apoyo adicional para la estancia de tres meses.

A los doctores Juan Gabriel Aviña Cervantes, Mario Alberto Ibarra Manzano, José Francisco Martínez Trinidad y Jesús Ariel Carrasco Ochoa, por sus valiosas aportaciones para el mejoramiento de este documento.

Finalmente, quiero señalar que esta tesis no hubiera sido posible sin el apoyo de mi familia, esposa e hija, quienes con energía y cariño, me motivaron a concluir este trabajo de manera exitosa.

*"Para las personas creyentes, Dios está al principio.
Para los científicos está al final de todas sus reflexiones".*

Max Planck (1858-1947) Físico alemán.

Índice general

Índice general	III
Índice de figuras	V
Índice de tablas	VIII
Prefacio	X
1. Visión nocturna y procesamiento de imágenes oscuras	1
1.1. Constancia de color para el mejoramiento de imágenes oscuras	3
1.2. Transferencia de color en la visión nocturna	6
1.2.1. Transferencia de color usando estadísticas globales	6
1.2.2. Mapeo de color basado en muestras	8
1.3. Sistemas basados en reglas de conocimiento	10
1.3.1. Sistemas basados en reglas difusas	10
2. Constancia de color en espacios de color perceptuales	13
2.1. Algoritmos de constancia de color	13
2.1.1. White-Patch	14
2.1.2. Gray-World	16
2.1.3. Gray-Edge	17
2.2. La teoría del Gray-World en espacios de color perceptuales	17
2.2.1. Marco de trabajo	18
2.2.2. Implementación del algoritmo en los espacios perceptuales	19
2.2.3. Resultados experimentales	20
3. Selección automática de algoritmos de constancia de color para el mejoramiento de imágenes oscuras usando razonamiento basado en reglas difusas	24
3.1. Selección de un algoritmo de constancia de color	25
3.1.1. Métrica de comparación	27

ÍNDICE GENERAL

3.2. Extracción de características	28
3.2.1. Color	29
3.2.2. Textura	31
3.2.3. Iluminación	32
3.3. Aprendizaje de las reglas difusas	32
3.3.1. Definición de los términos lingüísticos	33
3.3.2. Reglas y selección de las características	34
3.3.3. Sintonización de las reglas	36
3.3.4. Evaluación del desempeño	37
3.4. Resultados del sistema	38
3.4.1. Comparación con otros métodos de selección	39
4. Transferencia de color perceptual para mejorar imágenes de noche	43
4.1. Marco de trabajo	44
4.1.1. Espacios de color perceptuales usados en la transferencia de color .	45
4.1.2. Valoración de la transferencia de color	46
4.2. Resultados de la comparación entre enfoques	47
5. Iluminando imágenes de noche para mejorar la segmentación por color	52
5.1. Marco de trabajo	54
5.1.1. Creación de base de datos nocturna para segmentación	55
5.1.2. Métodos de mejoramiento en imágenes nocturnas	56
5.2. Pruebas y Resultados	56
5.2.1. Pruebas estadísticas	58
Conclusiones y productos generados	62
A. Ecuaciones de transformación entre espacios de color	66
A.1. LMS	66
A.2. $l\alpha\beta$	66
A.3. CIECAM97s	67
A.4. HSI	68
A.5. YIQ	69
A.6. YCbCr	69
A.7. XYZ	70
A.8. CIELUV	71
A.9. CIELAB	72
A.10. RLAB	72
Bibliografía	74

Índice de figuras

1.1. Un ejemplo de la mejora producida por algoritmos de constancia de color. (a) Imagen oscurecida artificialmente. (b) Resultado usando el algoritmo White-Patch. (c) Resultado usando el algoritmo Gray-World.	5
1.2. Transferencia de color. (a) imagen de origen [Premoze y Ashikhmin, 2001], (b) imagen de referencia (cortesía de <i>National Oceanic and Atmospheric Administration Photo Library</i>), (c) resultante de la transferencia de color	7
1.3. Dos ejemplos de transferencia de color. (a) Una imagen con colores falsos, (b) Una imagen real de referencia con luz de día [Hogervorst y Toet, 2010] y, (c) el resultado obtenido después de la transferencia de color usando la imagen de referencia (b) sobre la imagen (a). (d) Una imagen oscurecida artificialmente, y (e) es su salida correspondiente usando la imagen (b) como referencia.	8
1.4. Mapeo de color basado en muestras. (a) Imagen visible y (b) NIR de la misma escena. (c) Combinación de las imágenes (a) y (b) en una representación RG de color falso, obtenida de asignar (a) al canal <i>R</i> y (b) al canal <i>G</i> de una imagen a color RGB (el canal <i>B</i> se establece en cero); este mapa de color se muestra en (d). (e) El mapeo de color derivado de los pares correspondientes en (c); y (f) es la imagen de referencia con colores de día reales [Hogervorst y Toet, 2010]. (g) Es el resultante de la aplicación del esquema de mapeo en (e) a las imagen de color falso en (c).	9
2.1. Imagen corregida por el algoritmo WP. a) Imagen original. b) Imagen procesada con el algoritmo WP original. c) Imagen procesada con la variante de Finlayson.	15
2.2. Imagen corregida por el algoritmo GW. a) Imagen original. b) Imagen procesada con el algoritmo WP. c) Imagen procesada con GW.	17
2.3. Metodología seguida en las pruebas experimentales para el algoritmo GW en espacios perceptuales.	18

2.4.	Ajuste a la curva para cada enfoque mostrando la tendencia de los errores angulares dados por los resultantes. A menor error angular, la estimación del iluminantes es mejor.	21
2.5.	Tres ejemplos de imágenes donde el error angular se muestra en la esfera gris. La imagen ideal también es incluida. a) ApacheTrial frame no. 01520, b) CIC2002 frame no. 10150, c) DeerLake frame no. 00140.	22
3.1.	Diagrama de flujo de la metodología seguida en este estudio.	26
3.2.	Imagen “vase_syl-cwf” del conjunto de datos SFU Laboratory. La escena original se procesa usando los tres algoritmos de constancia de color. . . .	27
3.3.	Distribución de los conjuntos difusos Gaussianos en el dominio de una característica.	34
3.4.	Cinco ejemplo de los 429 usados en los experimentos. Éstos fueron seleccionados correctamente. De arriba a abajo: imágenes “ball_solux-4100”, “weave_syl-50MR16Q”, “clr_cubes3-solux-4700”, “elephant_solux-4100+3202” y “fcooper_syl-wwf”. De izquierda a derecha: Imagen original, resultantes del WP, resultantes del GW, resultantes del GE1 y la imagen seleccionada por el sistema.	41
4.1.	Proceso de la transferencia de color para una imagen oscura.	44
4.2.	Diagrama a bloques de la metodología seguida para la comparación de enfoques y transferencia de color en distintos espacios.	45
4.3.	Una imagen oscura y sus resultantes obtenidas usando diferentes métodos. (a) Entrada no. 92 (paisaje); (b) imagen de referencia no. 19; resultantes usando (c) WP, (d) GW, y (e) HE; resultantes usando (b) como objetivo y (f) CT en RGB, (g) CT en $l\alpha\beta$, (h) CT en CIELUV, (i) CT en CIELAB, y (j) CT en RLAB.	48
4.4.	Resultados para la Imagen no. 92 (paisaje) mostrado en la Figura 4.3. Las medidas usando una métrica específica para los resultantes de diferentes métodos. Detalles de esos resultados son reportados en la Tabla 4.2. . . .	48
4.5.	Ejemplos de imágenes objetivo usadas en la transferencia de color para la Imagen no. 92, y sus resultados correspondientes.	50
4.6.	Valores promedios de las 2000 resultantes en cada metodología usando las cuatro medidas. Detalles de los resultados son reportados en la Tabla 4.3.	50
5.1.	Diagrama a bloques de la metodología seguida para la evaluación de la segmentación por un método sobre diferentes enfoques.	54
5.2.	Tres ejemplos de la nueva base de datos BSDS300 de noche y sus respectivas imágenes originales con luz natural.	55

ÍNDICE DE FIGURAS

5.3. Un ejemplo mostrando los resultantes por los métodos de mejoramiento. (a) Imagen original, (b) imagen de noche, (c) resultante por WP, (d) resultante por GW, (e) resultante por HE y (f) resultante por CT.	56
5.4. Un catálogo mostrando las imágenes señaladas en la Figura 5.3 y sus respectivas segmentaciones por cuatro métodos diferentes.	57

Índice de tablas

2.1. Errores angulares estadísticos para los diferentes enfoques.	22
2.2. Tiempo de cómputo para cada enfoque.	23
3.1. Ejemplos de la asignación de etiqueta en las imágenes de entrenamiento pre-seleccionadas de acuerdo al Promedio de Cromo. Los mejores resultados son mostrados en negrita.	28
3.2. Características y su rango calculado en la etapa de entrenamiento. El vector de características en el protocolo de prueba es normalizado con base en estos valores.	33
3.3. Valores máximos de membresía ($A_{(i,m)}$).	35
3.4. Una manera práctica de representar las 15 reglas. Todas las premisas se incluyen en cada antecedente.	36
3.5. Una manera práctica de representar las 15 reglas sintonizadas. Algunas premisas fueron excluidas para incrementar la cobertura del espacio.	37
3.6. Desempeño usando las cinco reglas “no optimizadas” por algoritmo. Usando la inferencia de Larsen se obtiene una mejor tasa de selección.	38
3.7. Matriz de confusión de los resultados de selección usando la inferencia de Mamdani en el conjunto de datos de prueba.	38
3.8. Matriz de confusión de los resultados de selección usando la inferencia de Larsen en el conjunto de datos de prueba.	39
3.9. Comparación con otros métodos de selección.	40
4.1. Las componentes correspondientes al índice i de acuerdo a cada espacio de color.	46
4.2. Medidas de distancia entre la imagen objetivo (Figure 4.3b) y los resultados después de procesar la entrada (Figure 4.3a) usando los diferentes enfoques (Figure 4.3c-j).	49
4.3. Valores promedio de las 2000 medidas de cada método bajo análisis. Se dan los datos para las cuatro métricas de distancia entre los histogramas	51

ÍNDICE DE TABLAS

5.1. Valores respectivos de la medición por el índice NPR para las imágenes mostradas en la Figura 5.4.	58
5.2. Resultados del índice NPR usando el método de segmentación Kmeans por color.	58
5.3. Resultados del índice NPR usando el método de segmentación Kmeans por color e información espacial.	59
5.4. Resultados del índice NPR usando el método de segmentación Mean shift por color.	59
5.5. Resultados del índice NPR usando el método de segmentación Mean Shift por color e información espacial.	60

Prefacio

La información visual es nuestra mayor fuente natural de comunicación. Aparte de la visión humana, la información visual juega un papel importante e indispensable en la sociedad y es el núcleo de los medios actuales de comunicación, tales como el amplio mundo del Internet y los teléfonos celulares. Debido a la creciente producción y uso de información visual digital (documentos, sitios web, imágenes, vídeos), existe una sobrecarga de información visual y por lo tanto se requiere la automatización en el entendimiento de esa información. Además, como la información digital actualmente está disponible en formato a color, hay una urgente necesidad de entender esa información usando la propiedad de color. La visión por computadora es nuestra mejor herramienta para tratar de entender esa información visual. Aunque el color ha sido un tema central en varias disciplinas por siglos (que van desde las matemáticas y la física, hasta humanidades y arte), lleva relativamente poco tiempo siendo abordado en la visión por computadora.

El ojo humano no puede distinguir todos los tonos de gris en un mismo instante, por lo que la información visual nocturna es usualmente difícil de interpretar. Por otra parte, el sistema visual humano puede discriminar varios miles de colores definidos variando el matiz, la saturación y el brillo. Por esta razón, las imágenes a color tienen varios beneficios para aplicaciones de vigilancia y seguridad, y una representación por color podría facilitar el reconocimiento de imágenes nocturnas y su representación.

En esta tesis proporcionamos un conjunto de estudios y métodos que contribuyen a la tarea de la interpretación de imágenes, especialmente aquellas que representan escenas nocturnas, a partir de una perspectiva de color. Ofreciendo alternativas útiles y, hasta cierto punto rápidas, para el campo de la visión nocturna. Los métodos propuestos se pueden agrupar dentro de tres categorías principales del procesamiento de imágenes:

- Basados en constancia de color;
- Basado en transferencia de color;
- Utilizando métodos inteligentes.

En el Capítulo 1 se establece la importancia del procesamiento de imágenes en el

campo de la visión nocturna. A su vez, se introducen las áreas del procesamiento de imágenes que han sido abordadas para el mejoramiento de imágenes nocturna. Estos temas son desarrollados posteriormente con más detalle en cada uno de los capítulos. Algunos temas principales discutidos en este capítulo son: la constancia de color, la transferencia de color y los sistemas basados en reglas.

En el Capítulo 2 se propone la estimación del iluminante en una escena directamente en un espacio de color perceptual. Primero, se desarrollan los algoritmos principales que son utilizados a lo largo de esta tesis. Y posteriormente, nos enfocamos en resolver el problema de la constancia de color para una escena usando uno de los algoritmos de constancia de color en espacios perceptuales. El Gray-World es analizado en dos espacios de color perceptuales. Específicamente, se usan los ampliamente conocidos CIELAB y CIELUV ya que sus transformaciones son simples y rápidas. El desempeño del método propuesto es evaluado calculando la exactitud entre el iluminante estimado y el iluminante real. También se incluyen los resultados usando el algoritmo Gray-World (en RGB), y el algoritmo Gray-Edge con fines comparativos. Este trabajo aunque no está orientado a imágenes oscuras, resultó en un ejercicio que sirvió para comprender los fundamentos de la constancia de color.

En el Capítulo 3 se presenta un sistema difuso basado en reglas, el cual opera como un selector de algoritmos de constancia de color para la mejora de imágenes oscuras. De acuerdo con el contenido en color de una imagen, el sistema selecciona entre tres algoritmos de constancia de color, el White-Patch, el Gray-World y el Gray-Edge. Estos algoritmos han sido considerados debido a la remoción precisa del color del iluminante en la escena, además de mostrar una excelente mejora del contenido cromático en las imágenes. Nuestra propuesta consiste en un sistema difuso, modelando el proceso de decisión usando reglas simples. La metodología se basa en dos etapas principales. Por un lado, un protocolo de entrenamiento determina las reglas difusas, de acuerdo a las características calculadas a partir de un subconjunto de imágenes de entrenamiento, las cuales se tomaron del conjunto de datos SFU Laboratory. Después, se eligieron cuidadosamente doce características de la imagen para la formulación de las reglas: siete características de color, tres descriptores de textura, y dos descriptores de iluminación. En las reglas, se modelan los conjuntos difusos mediante funciones de membresía Gaussianas. Por otro lado, los experimentos se llevaron a cabo utilizando las inferencias difusas de Mamdani y Larsen. Para una imagen de prueba, se selecciona un algoritmo de constancia de color por medio del proceso de inferencia y de las reglas formuladas. Los resultados muestran que el método propuesto alcanza una alta tasa de selección correcta de acuerdo al algoritmo más adecuado para una escena particular.

En el Capítulo 4 se extiende la aplicabilidad de una técnica de transferencia de color usando estadísticas globales de las imágenes, para transformar una imagen oscura a

otra con una apariencia de iluminación natural. Este enfoque transfiere los colores desde una imagen deseada hasta una imagen oscura, obtenida a partir de un sensor RGB común. En nuestro enfoque, para obtener un resultado que muestra colores similares a los de la imagen de destino, sólo se requieren estadísticas a partir de una sola imagen. En un espacio de color perceptual, la representación de color de la imagen se ajusta para emular la forma en que el ojo humano percibe el color. Los espacios de color perceptuales se utilizan para conseguir distancias precisas de entre colores, porque la transformación de una imagen RGB a tal espacio de color elimina los errores producidos por la percepción no lineal inducida en las zonas oscuras. El desempeño de la transferencia de color se compara en cuatro espacios de color perceptuales, y adicionalmente con los generados por métodos de referencia. Debido a que no existe una base de datos de referencia para este fin, proporcionamos un conjunto de datos que consiste en imágenes oscuras en RGB, que se obtiene utilizando una cámara de uso común. Las pruebas realizadas apoyan nuestra hipótesis: los resultados de la transferencia de color son mejores cuando se utilizan espacios de color perceptuales. Entre los espacios de color evaluados, el espacio RLAB muestra los mejores resultados.

En el Capítulo 5 se propone utilizar métodos de mejoramiento, para iluminar imágenes nocturnas y aplicarles un proceso de segmentación. La idea consiste en mostrar que se produce una mejor segmentación sobre una imagen mejorada, que sobre aquella que sea oscura o de noche. Aunque los métodos de segmentación funcionan bien, incluso en condiciones de baja iluminación, al aplicar un pre-procesamiento a una imagen nocturna es posible obtener una segmentación equivalente a la que se obtendría a partir de la escena diurna. Para esto, se utilizan imágenes de una base de datos muy conocida en este campo. Sin embargo, en esta ocasión a toda la base de datos se le aplicará una etapa de oscurecimiento artificial para emular un efecto de noche. En este estudio, se comparan los resultados obtenidos por cuatro métodos de segmentación. Cada método se aplica sobre las imágenes resultantes por los métodos de mejoramiento, sobre la imagen oscura y sobre la imagen original (sin ser oscurecida). Los resultados se miden utilizando el índice NPR que evalúa la calidad de la imagen segmentada. Además, en las pruebas realizadas los resultados muestran que con la aplicación de métodos de segmentación después de la aplicación de un pre-procesamiento, se obtiene una calidad de segmentación similar a la segmentación realizada sobre las imágenes originales.

Una vez presentada la estructura del documento, cabe señalar que en la organización de los capítulos se ha procurado, por una parte, dar continuidad a la obra, y por otra parte, que cada uno de los Capítulos 2 a 5 puedan leerse de manera independiente. Aunque, podría requerirse la previa lectura del Capítulo 1, el cual presenta los fundamentos de los temas tratados.

Visión nocturna y procesamiento de imágenes oscuras

La visión nocturna es un área importante para un amplio número de aplicaciones militares relacionadas a la vigilancia [Maines, 1984; Rao et al., 2010], reconocimiento de escenas [Hogervorst y Toet, 2010; Qian et al., 2012, 2013], y seguridad [Yamasaki et al., 2008]. El término visión nocturna se refiere a la acción de observar durante la noche o condiciones de baja iluminación. Tradicionalmente, los sistemas de visión nocturna requieren el uso de dispositivos especiales, tales como cámaras de bajo nivel de luz (intensificadores de imágenes) y las cámaras infrarrojas térmicas (infrared, IR) [Hogervorst y Toet, 2010]. Las cámaras pertenecientes al primer tipo incrementan la poca luz visible y la cercana al infrarrojo (visible near-infrared VNIR). Las del segundo tipo transforman la energía térmica del espectro infrarrojo ($3 - 12\mu m$) a una representación de imagen digital. Recientemente, una tercer familia de dispositivos está creciendo en disponibilidad, los cuales son sensores y sistemas multiespectrales. Las metodologías de visión nocturna que usan sensores espectrales, comúnmente mapean la imagen espectral a una representación digital en escala de grises.

El ojo humano no puede distinguir todos los tonos de gris en un mismo instante, por lo que las imágenes de visión nocturna son usualmente difíciles de interpretar. A pesar de ello, la representación monocromática ha sido un estándar por décadas en el campo de la visión nocturna. Por otra parte, el sistema visual humano puede discriminar varios miles de colores definidos variando el matiz, la saturación y el brillo. Por esta razón, las imágenes a color tienen varios beneficios sobre las imágenes monocromáticas para aplicaciones de vigilancia y seguridad, y una representación por color podría facilitar el reconocimiento de imágenes nocturnas y su representación.

En las últimas décadas, la técnica de mapeo de colores ha sido utilizada para un conjunto de métodos que combinan imágenes espectrales [Toet y Walraven, 1996; Goffaux et al., 2005; Hogervorst y Toet, 2010]. Estas metodologías de mapeo también mejoran

el rango dinámico de la imagen resultante por la aplicación de diferentes técnicas. Una imagen con un rango dinámico estrecho puede ser ajustada usando técnicas de procesamiento de imágenes para mejorar las características visuales como la luminancia, sin comprometer la calidad.

Enfoques de constancia de color son usados para incrementar la luminancia global en la imagen. Aunque originalmente los algoritmos de constancia de color han sido desarrollados para estimar el color de la fuente de iluminación con el fin de eliminar el efecto ocasionado por ella en la escena, también mejoran el contenido cromático de la imagen. Esta mejora puede ser usada para el procesamiento de imágenes oscuras [Cepeda-Negrete y Sanchez-Yanez, 2015].

La fusión de imágenes es otro enfoque usado para el mejoramiento de imágenes oscuras. Esta técnica incrementa la información visual en una imagen, combinando bandas espectrales al espacio RGB, para expandir el rango dinámico de un sensor [Goffaux et al., 2005]. Generalmente, se usan dos imágenes monocromáticas derivadas de diferentes bandas espectrales. Una imagen visible (o cercana al infrarrojo) es considerada como la componente *R*, y una imagen térmica como la componente *G* [Toet y Walraven, 1996; Hogervorst y Toet, 2010]. Esta fusión se usa para construir una tabla de búsqueda (LUT) para transferir colores a otras imágenes. Sin embargo, este esquema puede producir imágenes que contengan colores falsos sin apariencia natural, lo que podría disminuir la comprensión de la escena [Toet y Franken, 2003; Toet, 2003]. Otros métodos han emergido como una opción viable con el fin de evitar la producción de colores falsos.

Otra forma de mejorar una imagen con condiciones de baja iluminación o corregir colores falsos, consiste en usar un enfoque de transferencia de color (también llamado mapeo de colores). Esta técnica ayuda a recolorar una imagen dada, transfiriendo estadísticas a partir de otra usada como referencia. Se ha demostrado que si se diseña apropiadamente un método de transferencia de color aplicado a imágenes de noche, esto ayuda a mejorar la habilidad de un observador para entender una escena [Toet y Hogervorst, 2012]. Varios estudios han abordado la corrección de colores no reales usando estas técnicas de transferencia de color [Hogervorst y Toet, 2010; Qian et al., 2012, 2013; Toet, 2003]. Estos trabajos han sido específicamente orientados para tareas de segmentación de escenas y clasificación [Toet y Hogervorst, 2012; Essock et al., 1999]. Sin embargo, en estos estudios aún se siguen usando dos imágenes monocromáticas obtenidas de diferentes dispositivos de captura.

Si no se cuenta con más información que aquella proporcionada por una única imagen tomada por un sensor de luz visible, deberían ser usadas otras metodologías de procesamiento de imágenes. Ya que para imágenes tomadas bajo condiciones de baja iluminación, donde el rango dinámico es muy estrecho, el sensor introduce cantidades significativas de ruido, y el contenido de color no es preciso. Hay diferentes metodologías

que pueden ser usadas para mejorar la representación visual de imágenes oscuras bajo las restricciones antes mencionadas.

En este capítulo se introducen los conceptos utilizados para el procesamiento de imágenes a color (constancia de color, transferencia de color, sistemas basados en reglas). Especialmente orientados al procesamiento de imágenes oscuras y usando espacios de color perceptuales. A continuación se presentan los temas de: constancia de color, transferencia de color y sistemas basados en reglas.

1.1. Constancia de color para el mejoramiento de imágenes oscuras

El color es una propiedad importante en la visión por computadora y reconocimiento de patrones. Algunas áreas típicamente abordadas en la visión por computadora son la interacción hombre-máquina [Yang et al., 1998], la extracción de características [Gevers y Smeulders, 2000], clasificación de imágenes [Schroder y Moser, 2001], reconocimiento de objetos [Gasparini y Schettini, 2004; van de Weijer et al., 2007b], categorización de escenas [Bianco et al., 2008; Gijssenij y Gevers, 2007], y los modelos en apariencia de color [Fairchild, 2005]. Los colores que están presentes en imágenes son determinados por propiedades intrínsecas de objetos y superficies, así como el color de la fuente de luz. Para un sistema robusto basado en color, estos efectos de la fuente de luz deberían ser filtrados.

La habilidad de distinguir el mismo color de una o varias superficies, independientemente del color de la fuente de luz presente en la escena, se conoce como *Constancia de Color* [Zeki, 1993]. La constancia de color permite la interpretación de los colores en el contexto de su entorno. El sistema visual humano aprende de manera natural esta habilidad [Fairchild, 2005], hasta cierto punto. Sin embargo, los mecanismos que están involucrados en esta habilidad no están del todo entendidos aún [Gijssenij et al., 2011]. Por lo contrario, los sistemas visuales mecánicos y electrónicos, solamente reciben luz del entorno, por lo cual resulta complicado entender la información correcta de la escena proporcionada [Agarwal et al., 2006].

Desde el punto de vista computacional, la constancia de color se define como la transformación de una imagen de entrada, tomada bajo una iluminación desconocida, a una imagen que aparenta ser tomada bajo una luz conocida, comúnmente blanca. Para esto, se necesita estimar el color de la fuente de luz en la imagen. Estos valores del color de la fuente luminosa son usados para transformar la imagen de entrada a una de salida. La imagen resultante representa la misma escena que la imagen de entrada, pero ahora aparentando haber sido tomada bajo una fuente de luz conocida (blanca).

Por décadas, varios investigadores han intentado resolver el problema de la cons-

tancia de color, proponiendo varios algoritmos y métodos. Sin embargo, no se ha podido establecer una única solución que resuelva el problema. Debido al amplio rango de aplicaciones de visión por computadora que requieren de la constancia de color, es imposible tener una solución única para todos los casos. Esto ha llevado a las investigaciones en identificar un conjunto de posibles enfoques que puedan aplicarse a problemas particulares. Principalmente, los esfuerzos se concentran en identificar enfoques de constancia de color para el seguimiento por video en tiempo real [Agarwal et al., 2006].

Las imágenes se componen de una combinación de reflectancia del objeto, la posición geométrica en que se encuentra y el color de la fuente de luz. Esto hace que el problema de la estimación de la fuente de luz sea mal planteado, y no pueda ser recuperada en base sólo a los colores de las imágenes. Debido a esto, todos los algoritmos hacen uso de suposiciones para simplificar el problema. Una de las suposiciones más conocidas y utilizadas, es la suposición del Mundo Gris o *Gray-World* [Buchsbaum, 1980]. Ésta asume que la iluminación reflejada por los objetos (reflectancia) promedio en una escena, bajo una fuente de luz blanca, es acromática. Otro algoritmo muy conocido se basa en la suposición del Parche Blanco o *White-Patch*, es decir, la suposición de que la respuesta máxima en los canales del espacio RGB es causada por una reflectancia perfecta [Land y McCann, 1971]. Otros métodos que se basan en estadísticas simples de las imágenes son los algoritmos de Tonos de Gris o *Shades-of-Gray* [Finlayson y Trezzi, 2004], Borde Gris o *Gray-Edge* [van de Weijer et al., 2007a] y el algoritmo de *Local Space Average Color* [Ebner, 2009].

Métodos más complejos se desarrollan con base en la información adquirida en una fase de aprendizaje. Las posibles fuentes de luz, la distribución de los colores en posibles reflectancias y las previas probabilidades sobre la combinación de colores, se aprenden y se utilizan para estimar el color de la fuente de luz. Uno de los primeros algoritmos de este tipo es el algoritmo *Gamut-Mapping* [Forsyth, 1990]. Este algoritmo, se basa en la suposición de que en imágenes del mundo real, para una iluminación dada, sólo un número limitado de colores pueden ser observados. Al hacer uso de esta suposición, la fuente luminosa se puede estimar comparando la distribución de los colores en la imagen actual y una distribución pre-aprendida de los colores (llamada gama canónica). Muchos algoritmos se han derivado del algoritmo original incluyendo el *Color-by-Correlation* [Finlayson et al., 2001], la estimación del *Gamut constrained* [Finlayson et al., 2006b] y el *Generalized Gamut-Mapping* [Gijssen et al., 2010]. Otros enfoques están basados en métodos probabilísticos [Brainard y Freeman, 1997; D'Zmura et al., 1995] y en computación flexible [Ebner, 2006]. Sin embargo, ninguna de estas extensiones consideran estadísticas de orden superior, descartando mucha información presente en las imágenes.

El mejoramiento de imágenes por color es un campo de investigación muy retador, en comparación con la mejora de una imagen en escala de grises. Una revisión detallada

1.1 Constancia de color para el mejoramiento de imágenes oscuras

y clara sobre asuntos del mejoramiento de imágenes es aquel hecho por Lucchese y Mitra [2004]. Es importante mencionar que los algoritmos de constancia de color mejoran el contenido cromático de las imágenes, aunque ellos fueron originalmente desarrollados sólo para la estimación de color de la fuente luminosa. Provenzi et al. [2008] exploraron el uso de dos algoritmos de constancia de color para propósitos de mejora. En particular, su trabajo fue orientado al mejoramiento de contraste local usando los algoritmos White-Patch y Gray-World en combinación con una técnica automática de igualación. El mejoramiento de imágenes usando algoritmos de constancia de color parece ser más significativo cuando son aplicados sobre imágenes oscuras, una premisa tomada en cuenta en nuestro estudio.

Tres algoritmos de constancia de color se han usado para el mejoramiento de imágenes oscuras: White-Patch (WP), Gray-World (GW) y Gray-Edge (GE). Estos métodos se explican detalladamente en el Capítulo 2.

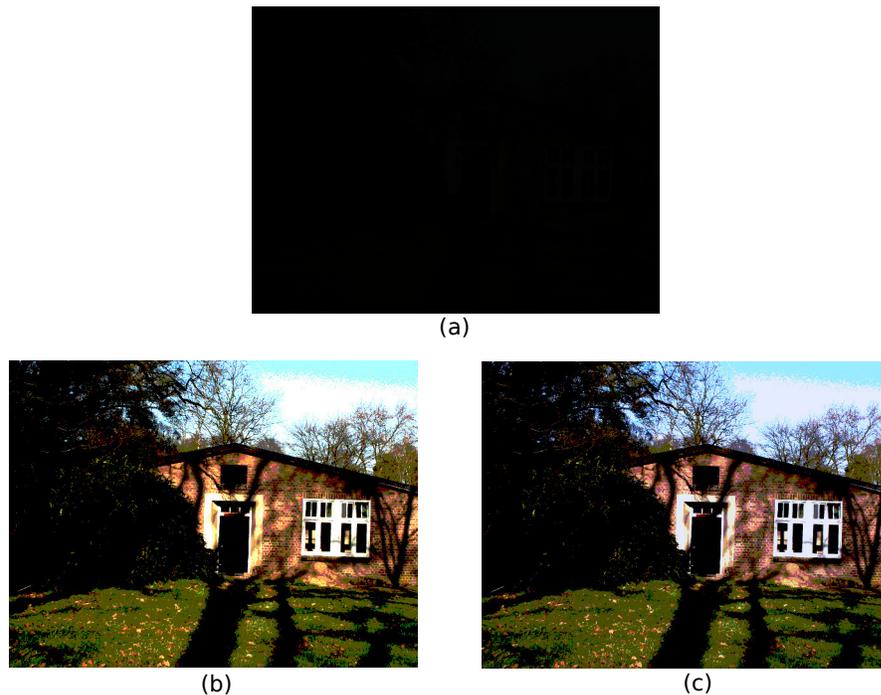


Figura 1.1: Un ejemplo de la mejora producida por algoritmos de constancia de color. (a) Imagen oscurecida artificialmente. (b) Resultado usando el algoritmo White-Patch. (c) Resultado usando el algoritmo Gray-World.

En la Figura 1.1 se muestra un ejemplo con los resultados proporcionados sólo por dos algoritmos sobre una imagen oscurecida artificialmente. El algoritmo GW es útil para el mejoramiento de imágenes de noche, iluminando áreas oscuras aunque pueden surgir saturaciones en zonas claras. Con el algoritmo WP, por el contrario, se evitan saturaciones aunque puede que algunas zonas permanezcan aún oscuras.

1.2. Transferencia de color en la visión nocturna

La transferencia de color es una técnica de procesamiento de imágenes en muchas aplicaciones, desde la mejora de una imagen personal hasta el pos-procesamiento de películas. Todos los colores dominantes en una imagen objetivo o de referencia deben ser transferidos a una de origen (por transferencia de estadísticas), siempre y cuando los colores originales de la imagen origen no prevalezcan al final de la transferencia. Un algoritmo ideal de transferencia de color debe mantener la escena de la imagen de origen pero con un estilo de color similar al de la imagen de referencia.

Existen tres estrategias principales para transferir colores entre imágenes: métodos basados en geometría [HaCohen et al., 2011], soluciones asistidas por el usuario [Levin et al., 2004; Cohen-Or et al., 2006; Sauvaget et al., 2010] y enfoques estadísticos [Reinhard et al., 2001; Xiao y Ma, 2006; Abadpour y Kasaei, 2007; Senanayake y Alexander, 2007; Pouli y Reinhard, 2011; Papadakis et al., 2011; Liu et al., 2012].

1.2.1. Transferencia de color usando estadísticas globales

Un método de transferencia de color tiene el propósito de modificar el contenido cromático de una imagen dada por la transferencia de estadísticas a partir de una imagen de referencia. En esta tesis nosotros usamos el método propuesto por Reinhard et al. [2001]. En este método, se calculan solamente estadísticas globales, lo que lo hace bastante rápido y simple. El objetivo del método es transformar una imagen de entrada (origen) en otra con apariencia similar a una imagen de referencia (objetivo). La meta es modificar el contenido de color de la imagen de origen aproximando sus estadísticas a las de la imagen de referencia. Este proceso puede ser mejorado usando un espacio de color perceptual específico (CIELAB, CIELUV, RLAB). Tomemos en cuenta que el método originalmente fue propuesto en el espacio $l\alpha\beta$. Las estadísticas que se usan son el promedio y la desviación estándar. Estas estadísticas se calculan para cada componente de color, y para ambas imágenes (origen y referencia). La Figura 1.2 muestra un clásico ejemplo de

1.2 Transferencia de color en la visión nocturna

la transferencia de color. Las estadísticas se calculan como sigue:

$$\mu_i^O = \frac{1}{M_O N_O} \sum_{x=1}^{M_O} \sum_{y=1}^{N_O} O_i(x, y), \quad (1.1)$$

$$\mu_i^R = \frac{1}{M_R N_R} \sum_{x=1}^{M_R} \sum_{y=1}^{N_R} R_i(x, y), \quad (1.2)$$

$$\sigma_i^O = \sqrt{\frac{1}{M_O N_O} \sum_{x=1}^{M_O} \sum_{y=1}^{N_O} (O_i(x, y) - \mu_i^O)^2}, \quad (1.3)$$

$$\sigma_i^R = \sqrt{\frac{1}{M_R N_R} \sum_{x=1}^{M_R} \sum_{y=1}^{N_R} (R_i(x, y) - \mu_i^R)^2}, \quad (1.4)$$

donde μ_i y σ_i son la media y la desviación estándar, i es el índice del canal, M es el número de filas, y N es el número de columnas de la imagen. Aquí, las señales R y O corresponden a las imágenes de referencia y origen, respectivamente.

La transferencia de color entre la imagen de origen y la referencia para el canal i se realiza usando la Ecuación (1.5)

$$S_i(x, y) = \frac{\sigma_i^R}{\sigma_i^O} (O_i(x, y) - \mu_i^O) + \mu_i^R, \quad (1.5)$$

donde S representa la imagen de salida en la transferencia. Finalmente, transformamos de vuelta la imagen O a RGB usando la transformación inversa.

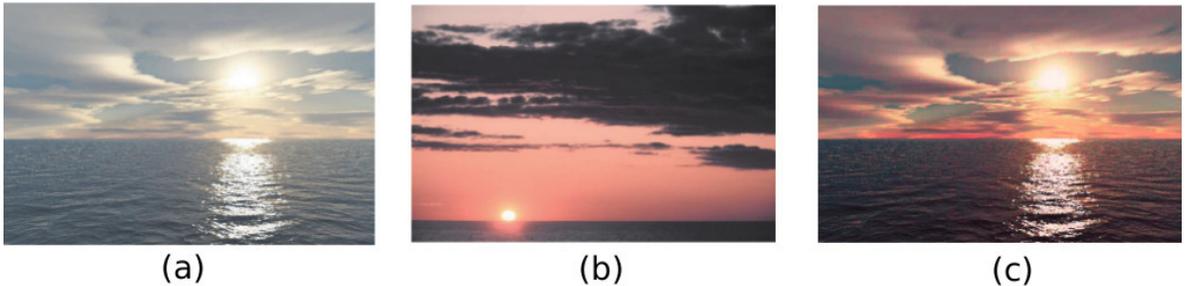


Figura 1.2: Transferencia de color. (a) imagen de origen [Premoze y Ashikhmin, 2001], (b) imagen de referencia (cortesía de *National Oceanic and Atmospheric Administration Photo Library*), (c) resultante de la transferencia de color .

Es importante mencionar que para las imágenes dadas, sus componentes se procesan separadamente. Si se usa otro espacio de color, la imagen RGB debe ser transformada primeramente a ese espacio para poder realizar la transferencia de color. Después de haber efectuado la transferencia, la imagen se transforma de vuelta a RGB para visualizar el resultado.

1.2 Transferencia de color en la visión nocturna

El algoritmo de transferencia de color usado en imágenes oscuras es el propuesto por Reinhard et al. [2001] usando estadísticas globales. La Figura 1.3 muestra dos ejemplos de la transferencia de color. En el primer caso, la aplicación sobre una imagen con colores falsos, y en la segunda un resultados después de aplicar la transferencia sobre una escena oscura (artificialmente oscurecida).

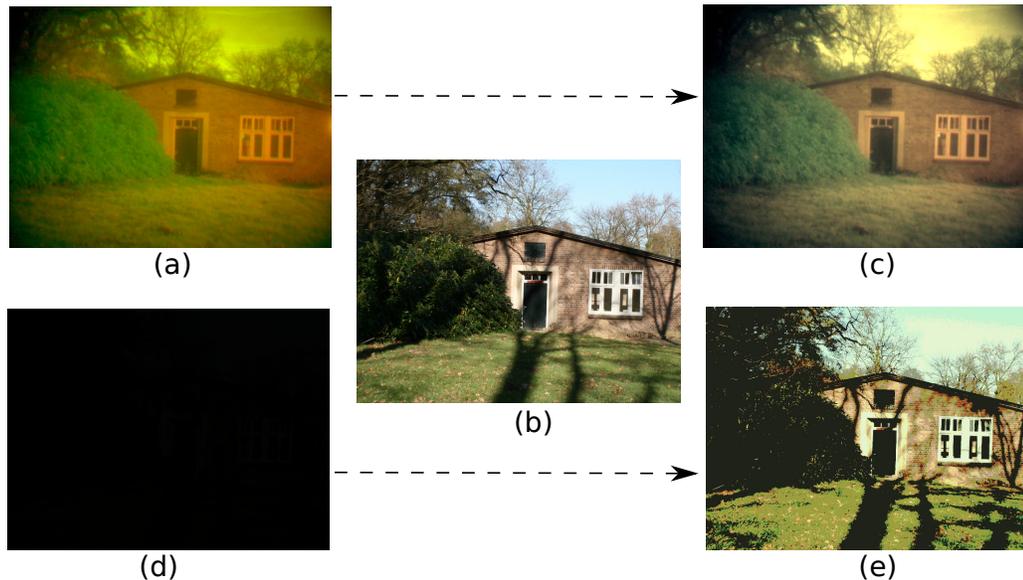


Figura 1.3: Dos ejemplos de transferencia de color. (a) Una imagen con colores falsos, (b) Una imagen real de referencia con luz de día [Hogervorst y Toet, 2010] y, (c) el resultado obtenido después de la transferencia de color usando la imagen de referencia (b) sobre la imagen (a). (d) Una imagen oscurecida artificialmente, y (e) es su salida correspondiente usando la imagen (b) como referencia.

1.2.2. Mapeo de color basado en muestras

Un método alternativo basado en una tabla de búsqueda (LUT) fue propuesto por Hogervorst y Toet [2010] para aplicar colores reales a imágenes multibanda (comúnmente visibles, cercanas al infrarrojo o térmicas). La transformación del color se deriva de un conjunto correspondiente de muestras para las que se conocen tanto los valores de los sensores multibanda como los de la imagen real de día. Se puede apreciar que este método resulta en la reproducción de imágenes multibanda con colores que correspondan a aquellos de día de manera más similar que el resultado del enfoque estadístico. Sin embargo, en contraste con el método estadístico, la derivación del mapeo de colores requiere la captura de un par de imágenes de distintos sensores y una imagen de referencia de día de la misma escena. Esto, porque los conjuntos de píxeles correspondientes se utilizan para definir la transformación de color por pares en este enfoque. Una vez que se ha generado el mapeo de colores a una tabla de búsqueda, estos colores se pueden

1.2 Transferencia de color en la visión nocturna

aplicar a diferentes imágenes nocturnas multibanda. Y así, la transformación de color se realiza utilizando solamente la tabla de búsqueda, lo que permite que la aplicación sea en tiempo real.

El método funciona de la siguiente manera. Dado un conjunto de muestras (píxeles) para los que se conocen los colores tanto en la salida de los sensores multibanda y la correspondiente de día, el problema de obtener la transformación de color óptima es encontrar una transformación que asigne de forma óptima la dimensión N (en nuestros ejemplos, $N = 2$) los vectores de salida del sensor multibanda (una para cada muestra) a los vectores de 3-D correspondientes a los colores durante el día (RGB).

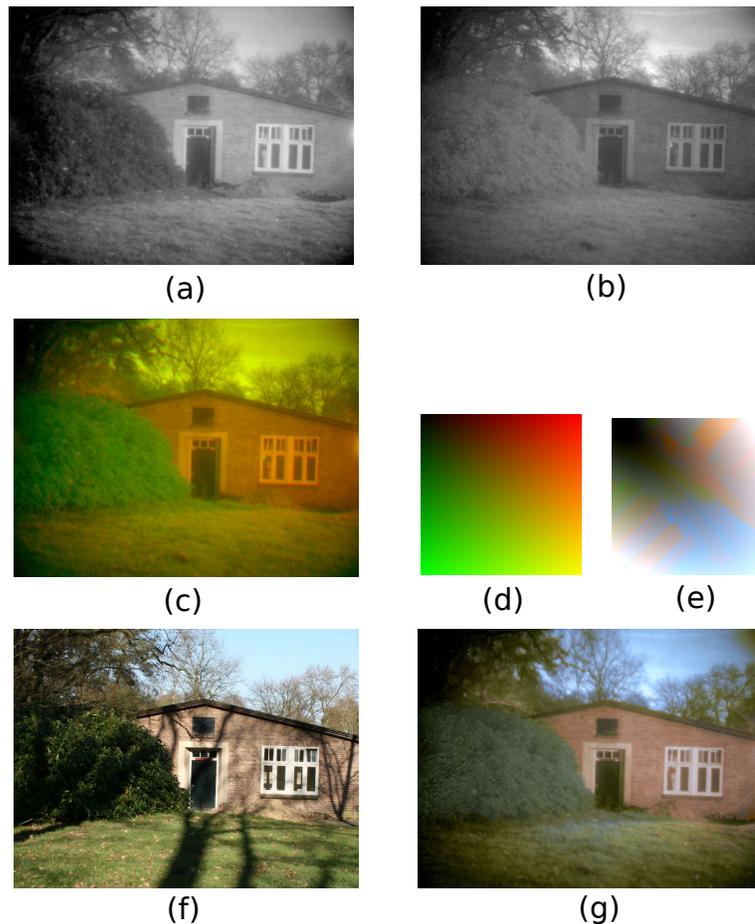


Figura 1.4: Mapeo de color basado en muestras. (a) Imagen visible y (b) NIR de la misma escena. (c) Combinación de las imágenes (a) y (b) en una representación RG de color falso, obtenida de asignar (a) al canal R y (b) al canal G de una imagen a color RGB (el canal B se establece en cero); este mapa de color se muestra en (d). (e) El mapeo de color derivado de los pares correspondientes en (c); y (f) es la imagen de referencia con colores de día reales [Hogervorst y Toet, 2010]. (g) Es el resultante de la aplicación del esquema de mapeo en (e) a las imagen de color falso en (c).

1.3 Sistemas basados en reglas de conocimiento

En el método de mapeo, primero se transforma la imagen del sensor multibanda a una imagen de colores falsos (aquellos no pertenecientes a los colores reales de la escena), tomando la imagen visual como la entrada en el canal R y la imagen NIR como el canal G (Figuras 1.4a y 1.4b, respectivamente), se hace referencia como la imagen RG (Figura 1.4c). El mapeo de las dos bandas a una imagen RGB de colores falsos nos permite usar técnicas de conversión estándares tales como la indización. En el siguiente paso, la resultante de color falso (imagen RG, Figura 1.4c), se convierte a una imagen indizada. Cada píxel en cada una de las imágenes contiene un único índice. Éste refiere a un valor RGB en la tabla de búsqueda. Cada índice representa una combinación R-G, donde el color real correspondiente se obtiene de localizar los píxeles en la imagen objetivo con este índice y encontrando los píxeles correspondientes en la imagen de referencia (Figura 1.4f). En tal caso, todos los píxeles se localizan en la imagen objetivo de color con un índice de color i . Así, todos los píxeles correspondientes (píxeles con las mismas coordenadas) se registran con base en la imagen de color de referencia, para después calcular sus promedios. Finalmente, este promedio en valor RGB se asigna a otro índice i en una nueva tabla de búsqueda. Todos los pasos se llevan a cabo sucesivamente para todos los índices de color. Este proceso produce otra tabla de búsqueda que contiene colores reales asociados con las diferentes combinaciones multibandas, en la tabla de búsqueda de color falso RG. Sustituyendo la tabla de búsqueda RG (Figura 1.4d) con la tabla de búsqueda con colores reales (Figura 1.4e), se produce una imagen con una apariencia natural y real, en la que los colores son optimizados para un conjunto de muestras particulares.

1.3. Sistemas basados en reglas de conocimiento

Los sistemas basados en reglas permiten representar conocimiento, basado en la experiencia de un experto, en un conjunto de reglas tipo SI-ENTONCES. En las reglas, un grupo de premisas son evaluadas para concluir un resultado. Estos sistemas han demostrado resultados satisfactorios en una gran cantidad de aplicaciones como son los sistemas de control [Bernard, 1988; Nilashi y Ibrahim, 2013; Shobha et al., 2013]. Además, han incrementado su flexibilidad, robustez y su interpretabilidad, cuando son fusionados con otras técnicas de computación flexible como la lógica difusa [Cordon et al., 1999] la cual fue presentada como una extensión de los conjuntos clásicos [Zadeh, 1973, 1965].

1.3.1. Sistemas basados en reglas difusas

Actualmente, la lógica difusa (fuzzy logic en Inglés) ha sido apreciada como una herramienta para el manejo de la información en los sistemas basados en reglas [Ishibuchi et al., 2004; Trawinski et al., 2013], debido a su tolerancia a la imprecisión, a la ambigüedad y a la falta de información, como ocurre en tareas de extracción visual de

alto nivel. En el Capítulo 3 se presenta un trabajo que usa reglas difusas para la selección de algoritmos de constancia de color.

Formulación de reglas difusas

Un paso importante en el proceso de aprendizaje es la adquisición del conocimiento, es decir, la formulación de las reglas. En este enfoque, las entradas y salidas están representadas por términos lingüísticos relacionados por una regla *SI-ENTONCES*. Estas reglas son expresiones usadas en un proceso de inferencia tal que si el conjunto de hechos es conocido (antecedentes), un resultado (consecuente) puede ser inferido, por un algoritmo. La forma genérica de las reglas es:

$$R^n : \text{SI } x_1 \text{ es } A_{(1,m)} \text{ y } x_2 \text{ es } A_{(2,m)} \text{ y } \dots \text{ y } x_i \text{ es } A_{(i,m)} \text{ ENTONCES } \Omega \text{ es } \omega_p,$$

donde, $A_{(i,m)}$ es el conjunto difuso m_i en la característica i para la regla n , y ω_p representa el consecuente conjunto difuso p en la regla n . R^n es la regla n que proviene el conjunto de reglas \mathbf{R} .

La metodología para la formulación de la regla n es la siguiente. Primero, son identificadas las imágenes correspondientes en la clase p del conjunto de entrenamiento. Después, se calcula el promedio de cada característica i del subconjunto de imágenes identificadas previamente. La definición del conjunto difuso $A_{(i,m)}$ considerado en el antecedente de las reglas está dado por

$$A_{(i,m)} = \max \left[\mu_{(i,1)}(x_i), \mu_{(i,2)}(x_i), \dots, \mu_{(i,m)}(x_i) \right] \quad (1.6)$$

donde $\mu_{(i,m)}(x_i)$ es el grado de membresía o pertenencia de x_i en el conjunto difuso. Entonces, la regla es evaluada en el conjunto difuso m correspondiente. Como un ejemplo, una regla para el algoritmo WP (ω_1) es establecida como:

$$R^1 : \text{SI } x_1 \text{ es } A_{(1,1)} \text{ y } x_2 \text{ es } A_{(2,1)} \text{ y } \dots \text{ y } x_{12} \text{ es } A_{(12,2)}, \text{ ENTONCES } \Omega \text{ es } \omega_1,$$

donde ω_1 es el conjunto difuso consecuente en la regla R^1 .

Modelos de inferencia difusa

Este proceso comienza con la estimación de un nivel de activación $\tau_n(\mathbf{x})$ en cada regla generada. Así, el nivel de activación representa qué tanto \mathbf{x} satisface el conjunto completo de antecedentes.

Usando el modelo de inferencia Mamdani [Mamdani, 1976], el nivel de activación $\tau_n(\mathbf{x})$ de una regla n se calcula usando

$$\tau_n(\mathbf{x}) = \min \left[\mu_{(1,m)}^n(x_1), \mu_{(2,m)}^n(x_2), \dots, \mu_{(i,m)}^n(x_i) \right] \quad (1.7)$$

o, si se usa la inferencia del producto Larsen [Vadiee, 1993], el nivel de activación es $\tau_n(\mathbf{x})$ es calculado usando

$$\tau_n(\mathbf{x}) = \prod_i \mu_{(i,m)}^n(x_i). \quad (1.8)$$

Finalmente, la clase asignada a la salida correspondiente Ω es aquella con el máximo nivel de activación de todas las reglas

$$\Omega = \max[\tau_1(\mathbf{x}), \tau_2(\mathbf{x}), \dots, \tau_n(\mathbf{x})]. \quad (1.9)$$

Cada modelo de inferencia es considerado independientemente en los experimentos, con el fin de probar su desempeño en la selección, y por tanto considerar aquel modelo de inferencia que sea mejor para esta tarea.

Resumen

Se han establecido los fundamentos de esta tesis y los temas principales relacionado con el problema de la visión nocturna y el procesamiento de imágenes nocturnas. Tanto la constancia de color, la transferencia de color y los sistemas basados en reglas son herramientas que nos ayudan a la solución de los problemas que surgen en la visión nocturna. En capítulos posteriores se estarán desarrollando con más detalle cada una de estas tres áreas, enfocadas a la solución de un problema en particular. El Capítulo 2 describe la constancia de color, los algoritmos básicos a usar y un enfoque propuesto. En el Capítulo 3 se desarrolla un sistema inteligente para la selección de algoritmos de constancia de color en imágenes oscuras. El Capítulo 4 estudia la transferencia de color para la mejora de imágenes de noche usando diferentes espacios de color. Especialmente aquellos que son considerados perceptuales. Finalmente, en el Capítulo 5 se usan métodos para iluminar imágenes nocturnas con el objetivo de mejorar la tarea de segmentación.

Constancia de color en espacios de color perceptuales

La habilidad de distinguir el mismo color de una o varias superficies, independiente del color de la fuente de luz presente en la escena, se conoce como *Constancia de Color* [Zeki, 1993]. La constancia de color permite la interpretación de los colores en el contexto de su entorno. El sistema visual humano aprende de manera natural esta habilidad [Fairchild, 2005], hasta cierto punto. Sin embargo, los mecanismos que están involucrados en esta habilidad no están del todo entendidos aún [Gijssen et al., 2011]. Por lo contrario, los sistemas visuales mecánicos y electrónicos, solamente reciben luz del entorno, por lo cual resulta complicado entender la información correcta de la escena proporcionada [Agarwal et al., 2006].

En este capítulo se propone la estimación del iluminante directamente en un espacio de color perceptual. Así, nos enfocamos en resolver el problema real de la constancia de color. El Gray-World es analizado en dos espacios de color perceptuales. Específicamente, los ampliamente conocidos CIELAB y CIELUV debido a sus transformaciones simples y rápidas. Para estos dos espacios, la distancia Euclidiana entre dos puntos es proporcionalmente uniforme a la diferencia perceptual de los colores correspondientes a los puntos [Fairchild, 2005]. El desempeño del método es evaluado calculando la exactitud entre el iluminante estimado y el valor real. También se incluyen los resultados usando el algoritmo Gray-World (en RGB) y el algoritmo Gray-Edges con fines comparativos.

2.1. Algoritmos de constancia de color

Una vez que el color de la fuente de luz se ha estimado, ésta es usada para transformar la imagen de entrada de modo que produzca una imagen de salida, la cual aparenta haber sido tomada bajo una fuente de luz canónica, este proceso es llamado adaptación

cromática. La adaptación cromática usualmente se modela utilizando una transformación lineal, que a su vez se puede simplificar a una transformación en diagonal, cuando se cumplan ciertas condiciones [Finlayson et al., 1994], [Funt y Lewis, 2000], [West y Brill, 1982]. Este tipo de transformación es utilizada en esta tesis. Este modelo escala los canales de color de cada píxel de forma independiente, de acuerdo con la fuente luminosa que estima, de manera que el color original es transformado para que aparente haber sido capturado bajo una fuente de luz blanca. Existen otros métodos de adaptación cromática, incluyendo correlación de diagonal o modelo *Von Kries* [Von Kries, 1970], linealizado *Bradford* [Lam, 1985] y *CIECAT02* [Moroney et al., 2002], pero todos estos métodos suponen que el color de la fuente de luz en la imagen original es conocido.

La Ecuación (2.1) nos indica la relación que existe para la formación del color en una imagen

$$f_i(x, y) = G(x, y)R_i(x, y)I_i \quad (2.1)$$

donde, la $f_i(x, y)$ es la intensidad del píxel en una posición de una imagen o un cuadro de video, $G(x, y)$ es un factor que depende de la geometría de la escena, $R_i(x, y)$ es la reflectancia de un punto del objeto mostrado en la misma posición, I_i es el iluminante en la escena. El índice i corresponde al canal de color en la imagen.

Para escenas con iluminación uniforme, los algoritmos de constancia de color asumen que $G(x, y) = 1$ y $R_i(x, y) = 1$, permitiendo que el iluminante dependa totalmente de la imagen de entrada. Por otra parte, una imagen de salida ideal es aquella donde no existe un iluminante que altere la escena, o mejor dicho, este iluminante es el blanco perfecto,

$$o_i(x, y) = G(x, y)R_i(x, y). \quad (2.2)$$

La imagen de salida resultante de la sustitución de las variables en las ecuaciones está dada finalmente por

$$o_i(x, y) = \frac{f_i(x, y)}{I_i}. \quad (2.3)$$

2.1.1. White-Patch

El algoritmo Retinex fue propuesto por Land y McCann [1971]. El algoritmo Retinex en su forma más simple, el White-Patch (WP), o Parche Blanco en Español, toma en cuenta el valor más grande en cada componente de color como una representación del blanco de la imagen [Land, 1977]. Computacionalmente, este parche blanco es encontrado al buscar la máxima intensidad en cada canal, y está dado por

$$I_{i_{max}} = \max\{f_i(x, y)\}. \quad (2.4)$$

2.1 Algoritmos de constancia de color

Después, todas las intensidades de los píxeles son escalados usando (2.5), con base en el iluminante calculado,

$$o_i(x, y) = \frac{f_i(x, y)}{I_{i_{max}}}. \quad (2.5)$$

El algoritmo WP resulta más robusto si se calcula un histograma H_i para cada componente de color. El histograma nos muestra cuántos píxeles de la imagen tienen una intensidad particular por componente, donde n_b es el número de casillas del histograma y $H_i(j)$ es el histograma de un canal, j es el índice para una casilla de la intensidad correspondiente en el histograma. En lugar de elegir el píxel con la intensidad máxima para cada canal de color, se puede elegir la intensidad de tal manera que todos los píxeles con intensidades altas no influyan en la estimación. Este método es usado en [Finlayson et al., 2006a] para remover sombras. Siendo $f_i(j)$ la intensidad de cada canal i representada por la casilla j del histograma H_i . Entonces la iluminación estimada está dada por

$$I_i = f_i(j_i), \quad (2.6)$$

Para la selección de j_i deben cumplirse las siguientes dos condiciones:

$$pn \leq \sum_{k=j_i}^{n_b} H_i(k) \quad \text{y} \quad pn \geq \sum_{k=j_i+1}^{n_b} H_i(k), \quad (2.7)$$

donde pn es un porcentaje (usualmente cerca del 1%) del total de los píxeles en la imagen [Ebner, 2007]. De ahora en adelante, al referirnos al algoritmo WP es tomando en cuenta esta consideración de Finlayson y no la forma clásica inicialmente propuesta. La Figura 2.1 muestra un ejemplo de la diferencia al aplicar el WP clásico contra la mejora propuesta por Finlayson.

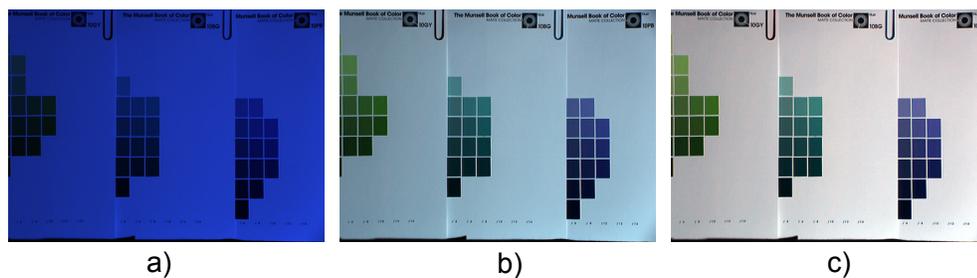


Figura 2.1: Imagen corregida por el algoritmo WP. a) Imagen original. b) Imagen procesada con el algoritmo WP original. c) Imagen procesada con la variante de Finlayson.

2.1.2. Gray-World

La suposición del Gray-World (GW), Mundo Gris en español, es el algoritmo más conocido de constancia de color y fue propuesto por Buchsbaum [1980]. El GW está basado en la suposición de que, en promedio, el mundo es gris, y estima el iluminante usando el color promedio de todos los píxeles. Se asume que la información dada por el promedio de cada canal de la imagen representa al tono gris.

El primer paso a realizar es calcular el color promedio en la imagen, como se indica en (2.8),

$$a_i = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N \{f_i(x, y)\}, \quad (2.8)$$

donde M y N son el número de columnas y filas, respectivamente. Igualmente, a_i puede ser representado por

$$a_i = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N G(x, y) R_i(x, y) I_i, \quad (2.9)$$

$$a_i = I_i \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N G(x, y) R_i(x, y), \quad (2.10)$$

$$a_i \approx I_i E[GR_i] = I_i E[G] E[R_i]. \quad (2.11)$$

La función $E[GR_i]$ es el valor esperado para el factor de geometría G multiplicado por la reflectancia R_i . Ambos pueden ser considerados como variables independientes, ya que no hay correlación en la forma y el color de un objeto. Suponiendo que muchos colores diferentes están presentes en la escena y cada color está igualmente equilibrado, la reflectancia puede ser considerada una variable aleatoria con rango $[0, 1]$, y

$$E[G]E[R_i] = E[G] \left(\int_0^1 x dx \right) = E[G] \frac{1}{2}, \quad (2.12)$$

$$a_i \approx I_i E[G] \frac{1}{2}. \quad (2.13)$$

Una vez que este valor global es conocido, se calcula el iluminante I_i . Suponiendo que hay una orientación perpendicular entre el objeto y la cámara, $E[G] = 1$. La Ecuación (2.14) describe el cálculo del factor ϕ para ser usado después en (2.15).

$$\phi = \frac{2}{E[G]} = 2 \quad (2.14)$$

$$I_i \approx \frac{2}{E[G]} a_i = \phi a_i \quad (2.15)$$

Ya que $f_i(x, y) = G(x, y) R_i(x, y) I_i$, el valor de salida está dado por

$$o_i(x, y) = \frac{f_i(x, y)}{I_i} = \frac{f_i(x, y)}{\phi a_i}. \quad (2.16)$$

2.2 La teoría del Gray-World en espacios de color perceptuales

Un ejemplo se muestra en la Figura 2.2, donde presenta el resultado de aplicar el GW y la diferencia con el obtenido con el WP.

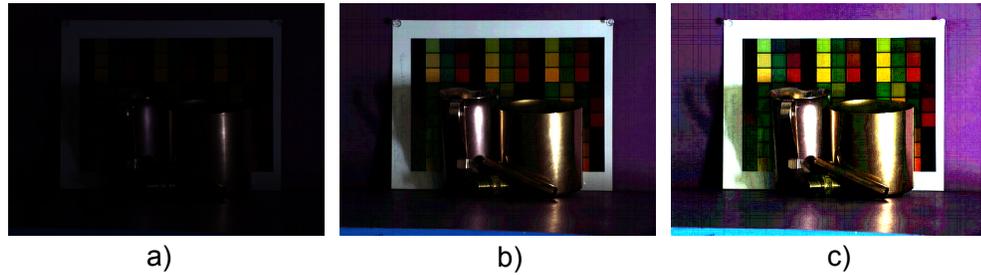


Figura 2.2: Imagen corregida por el algoritmo GW. a) Imagen original. b) Imagen procesada con el algoritmo WP. c) Imagen procesada con GW.

2.1.3. Gray-Edge

Este algoritmo surge de una teoría alternativa a las anteriores, la cual denominan van de Weijer et al. [2007a] como teoría del Gray-Edge (GE), Borde Gris en Español. Esta teoría supone que el iluminante es el promedio en las diferencias de las reflectancias y está dado por

$$\mathbf{I} = \left(\int \left| \frac{\partial^n \mathbf{f}^\sigma(\mathbf{x})}{\delta \mathbf{x}^n} \right|^p d\mathbf{x} \right)^{1/p}, \quad (2.17)$$

$$I_i = \frac{\left(\sum_{x=1}^m \sum_{y=1}^n h_i^\sigma(x, y)^p \right)^{1/p}}{(m \cdot n)^{1/p}}. \quad (2.18)$$

donde $\mathbf{h}^\sigma = \left(\sqrt{(\partial \mathbf{f}_x)^2 + (\partial \mathbf{f}_y)^2} \right) * \mathbf{G}^\sigma$.

Esta ecuación describe un marco de trabajo el cual produce diferentes estimaciones del color de la iluminación dependiente de tres variables. i) El orden n de la derivada en la imagen, ii) la norma Minkowski p , y, iii) el tamaño de σ para el difuminado local. Para este trabajo, solo se usa este algoritmo con $n = 1$ (primera derivada de la imagen), caso denominado GE1.

2.2. La teoría del Gray-World en espacios de color perceptuales

El enfoque del Gray-World fue propuesto en el espacio de color RGB el cual no es perceptual, a pesar de que el color es una propiedad de la visión humana. La mayoría de los algoritmos de constancia de color han sido propuestos e implementados en el

2.2 La teoría del Gray-World en espacios de color perceptuales

espacio RGB [Ebner, 2007], y a pesar de la existencia de una gran cantidad de enfoques de constancia de color, sigue sin existir una solución general a este problema.

Entre los pocos trabajos de investigación dirigidos a la estimación del iluminante en espacios de color perceptuales, podemos mencionar los estudios realizados en [Kloss, 2009; Cepeda-Negrete y Sanchez-Yanez, 2012]. Aunque estas propuestas tuvieron éxito en cuanto a la adaptación cromática, el desempeño de los métodos no fue evaluado usando el iluminante de referencia, dejando la medición de la constancia de color como tal sin verificar. A continuación, se presenta el marco de trabajo llevado a cabo para el enfoque propuesto y el cual sí presentará una evaluación con respecto al iluminante estimado y el de referencia.

2.2.1. Marco de trabajo

En este estudio, los resultados del enfoque GW en espacios de color perceptuales son comparados con aquellos obtenidos usando el clásico GW en RGB, como se muestra en la Figura 2.3.

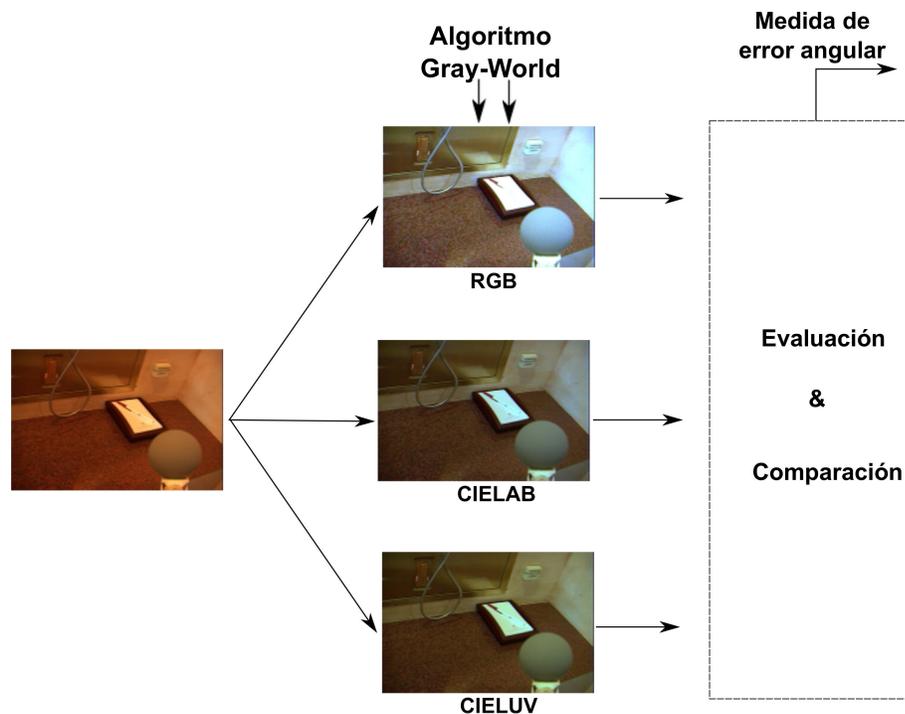


Figura 2.3: Metodología seguida en las pruebas experimentales para el algoritmo GW en espacios perceptuales.

Básicamente la metodología consiste en los siguientes módulos:

- Estimación del iluminante usando el algoritmo GW, para todas la imágenes en RGB.

- Estimación del iluminante usando el algoritmo GW propuesto, para todas la imágenes en CIELAB.
- Estimación del iluminante usando el algoritmo GW propuesto, para todas la imágenes en CIELUV.
- Evaluación de la exactitud para cada enfoque, usando por el error angular entre el iluminante estimado y su valor de referencia correspondiente.
- Comparar los resultados con las medidas obtenidas.

2.2.2. Implementación del algoritmo en los espacios perceptuales

La teoría del Gray-World supone que el color promedio en una imagen tiende al tono gris. En otras palabras, la media de cada componente cromática tiende al gris. Esta suposición fue originalmente aplicada al espacio de color RGB, el cual tiene tres componentes de color. Pero en este estudio aplicamos esta suposición en espacios perceptuales. En el Apéndice A se han incluido las ecuaciones de transformación entre los espacios de color considerados en este capítulo, además de otras transformaciones a utilizar posteriormente.

Los espacios perceptuales analizados constan de dos componentes cromáticas y una de luminancia. Estas dos componentes de color son consideradas en nuestro enfoque mientras la información de iluminación es omitida. El espacio CIELAB es el primero en ser considerado. Este espacio es perceptualmente uniforme y las componentes de color a^* y b^* no están correlacionadas con la componente de iluminación L^* . Esta última componente corresponde a una escala perceptual de brillo y excluye cualquier información de color. Siguiendo la suposición del GW, el valor promedio de cada componente cromática es calculada como

$$I_{a^*} = \frac{1}{MN} \sum_{y=0}^{M-1} \sum_{x=0}^{N-1} a^*(x, y), \quad (2.19)$$

$$I_{b^*} = \frac{1}{MN} \sum_{y=0}^{M-1} \sum_{x=0}^{N-1} b^*(x, y), \quad (2.20)$$

donde M y N son el número de columnas y filas en la imagen, respectivamente. Estos dos valores son muy importantes porque representan el color del iluminante en este espacio perceptual.

Para incrementar la iluminación en la imagen de salida, y para completar la estimación del iluminante, se sugiere el uso de $I_{L^*} = \max\{L^*(x, y)\}$ como una tercer componente. Esto es, se asume que el valor de iluminación correspondiente al gris promedio es el más alto posible en la imagen. Pruebas experimentales apoyan esta suposición. Por lo

tanto, se puede considerar este enfoque como una combinación de los algoritmos WP y GW.

Ahora, la imagen de salida es aquella obtenida usando el iluminante calculado. Pero para esto, el iluminante debe estar en el espacio RGB. Para el cálculo de $I_{i \in \{R,G,B\}}$, el iluminante $I_{j \in \{L^*, a^*, b^*\}}$ es llevado a RGB usando las ecuaciones conocidas de la transformación inversa. Finalmente, la exactitud del iluminante se evalúa usando la métrica del error angular y la imagen de salida se calcula usando la Eq. (2.3).

El espacio CIELUV es el segundo espacio perceptual considerado en este estudio. Aunque tiene propiedades similares que el CIELAB (dos componentes cromáticas y una de luminancia), CIELUV incorpora una formulación más simple para la adaptación cromática. El cálculo del iluminante se realiza exactamente igual que como fue explicado con CIELAB.

2.2.3. Resultados experimentales

Para probar nuestro enfoque, se compara el desempeño con los resultados correspondientes usando el GW clásico en RGB. Usamos la base de datos de referencia Gray-Ball [Ciurea y Funt, 2003] en los experimentos. Este conjunto de imágenes es comúnmente usado para la evaluación de los algoritmos de constancia de color ya que cada imagen posee su iluminante real como referencia. Más de 11000 imágenes se incluyen en este conjunto, sin embargo la mayoría de éstas están altamente correlacionadas. Bianco et al. [2008] propusieron el uso de un subconjunto que contiene el 10 % del conjunto total. En este trabajo, se usa dicho subconjunto, que incluye 1135 imágenes.

Para evaluar el desempeño de los algoritmos de constancia de color, una métrica de error debe ser considerada. La constancia de color debe ser evaluada de acuerdo al color de la iluminación estimada, independientemente de la intensidad de la misma. Hordley y Finlayson [2004] mostraron que una buena medida para evaluar la constancia de color es el error angular. En tal estudio, se discuten tres descriptores para la distribución del error angular, que son los valores de media, mediana y máximo. Esta métrica está dada por

$$e_{ang} = \cos^{-1} \left(\frac{\mathbf{I}_r \mathbf{I}_e}{|\mathbf{I}_r| |\mathbf{I}_e|} \right) \quad (2.21)$$

donde \mathbf{I}_r es el color de la iluminación real de la escena (verdad de referencia) e \mathbf{I}_e es el color de la iluminación estimada por un algoritmo. Nótese que, para evaluar la constancia de color, es necesario contar con la iluminación real, comúnmente proporcionada por los autores de la base de datos. Para esta medición, entre menor sea el valor del error angular, mejor es la estimación del iluminante.

El experimento se realizó con un análisis comparativo entre los resultados de nues-

2.2 La teoría del Gray-World en espacios de color perceptuales

tros enfoques y los del GW clásico en RGB. Se obtuvieron 1135 iluminantes y se calculó el error angular para cada uno de ellos. Con el fin de aclarar el comportamiento de los resultados, los errores angulares obtenidos fueron aproximados a un ajuste a la curva. Esta aproximación se hizo usando tres términos de la serie de Fourier, así que la curva ajustada describe la tendencia de un enfoque particular, como se muestra en la Figura 2.4.

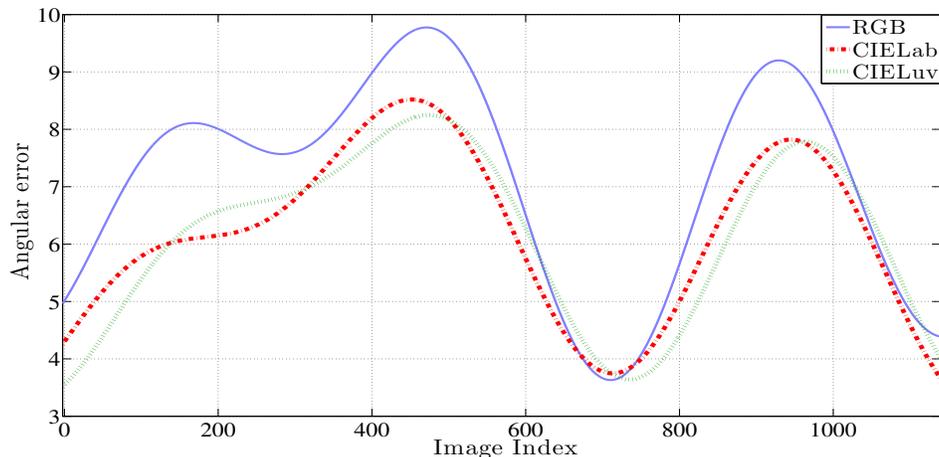


Figura 2.4: Ajuste a la curva para cada enfoque mostrando la tendencia de los errores angulares dados por los resultantes. A menor error angular, la estimación del iluminantes es mejor.

En la Figura 2.5 se muestran tres ejemplos de imágenes resultantes usando los algoritmos bajo evaluación. Cada imagen es presentada bajo diversas condiciones de iluminación y es procesada usando los tres enfoques. Además, también se muestra la imagen ideal correspondiente. Una imagen ideal es aquella que es corregida usando el iluminante proporcionado como verdad de referencia. El GW en RGB generalmente introduce una saturación que no se encuentra en las imágenes de entrada. Por el contrario, el GW aplicado en espacios de color perceptuales preserva mejor la relación de los colores. Se puede apreciar esto en la Figura 2.5.

La Tabla 2.1 muestra, para los tres descriptores, la distribución del error angular sugerida anteriormente. De acuerdo a estos datos, la aplicación del GW en cualquier espacio perceptual es significativamente mejor que aplicado en RGB. Además, el algoritmo GW aplicado en CIEluv es ligeramente mejor que aplicado en CIELab. El algoritmo Gray-Edge (GE) se incluye en los experimentos con propósitos comparativos. Este algoritmo ha mostrado tener el mejor desempeño para este conjunto de imágenes [Bianco et al., 2008]. Sin embargo, tiene una particular desventaja ya que requiere un proceso de sintonización usando imágenes de entrenamiento. Particularmente, el método depende de tres parámetros, los cuales deben de ser elegidos correctamente para mejorar el desem-

2.2 La teoría del Gray-World en espacios de color perceptuales

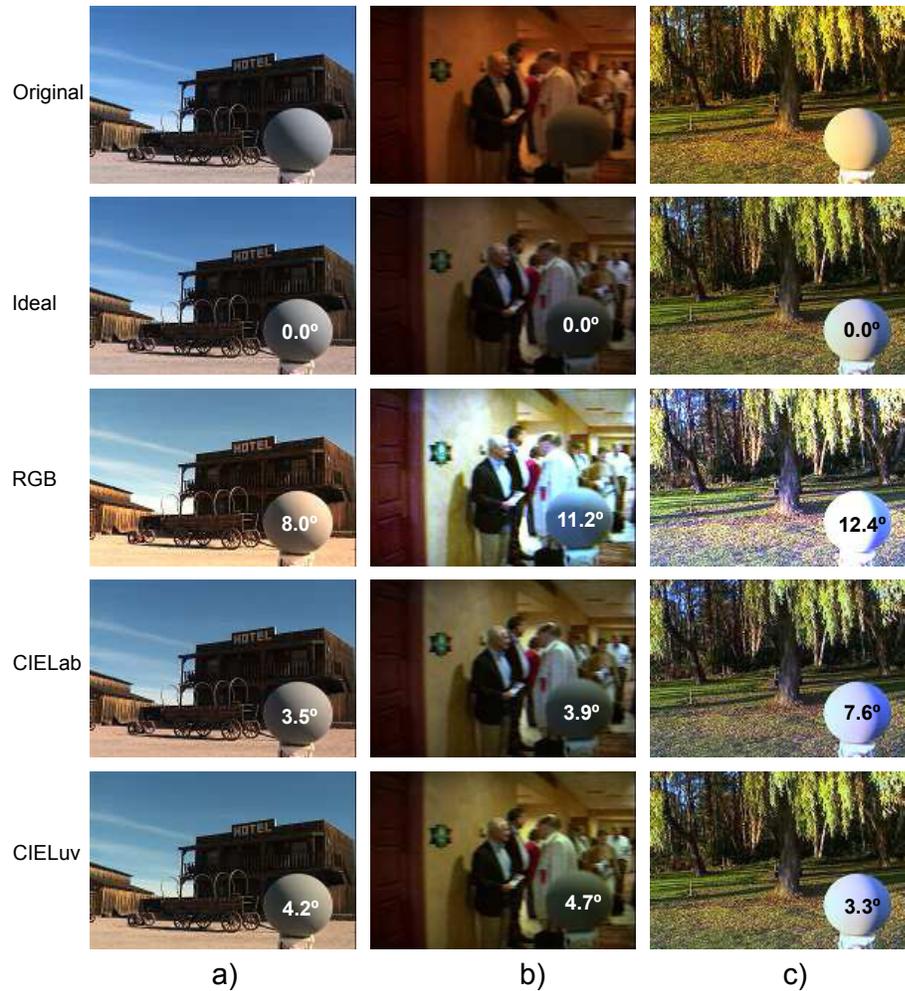


Figura 2.5: Tres ejemplos de imágenes donde el error angular se muestra en la esfera gris. La imagen ideal también es incluida. a) ApacheTrial frame no. 01520, b) CIC2002 frame no. 10150, c) DeerLake frame no. 00140.

peño. Para este estudio, GE opera usando los parámetros propuestos en [Bianco et al., 2010]. No obstante, la diferencia en desempeño entre el GW en espacios perceptuales y el GE es muy pequeña. La principal diferencia es que el algoritmo GW no requiere ningún proceso de sintonización.

Tabla 2.1: Errores angulares estadísticos para los diferentes enfoques.

Algoritmo	Media	Mediana	Max
Gray-World (RGB)	7.2	6.0	33.3
Gray-World (CIELAB)	6.5	5.7	25.9
Gray-World (CIELUV)	6.2	5.5	27.1
Gray-Edge 2° orden	6.1	5.1	24.0
Gray-Edge 1 ^{er} orden	6.0	4.9	23.7

2.2 La teoría del Gray-World en espacios de color perceptuales

La Tabla 2.2 muestra los tiempos de procesamiento de cada enfoque. Las especificaciones de la computadora usada son las siguientes: iMac Apple, 2.5 GHz Intel Core i5, 4GB ram 1333MHz DDR3. Se puede apreciar que la diferencia en el tiempo de cómputo entre el GE y los enfoques de GW es significativa. El algoritmo GW usando CIELUV fue el mejor en desempeño, excepto contra el GE. Sin embargo, el GW es 8.3 veces más rápido que el GE de 1^{er} orden, y 13.7 más rápido que el GE de 2^o orden. Cabe señalar que los resultados de este estudio ya han sido publicados [Cepeda-Negrete y Sanchez-Yanez, 2014].

Tabla 2.2: Tiempo de cómputo para cada enfoque.

Algoritmo	Tiempo (ms)
Gray-World (RGB)	0.57
Gray-World (CIELUV)	15.73
Gray-World (CIELAB)	31.22
Gray-Edge 1 ^{er} orden	130.44
Gray-Edge 2 ^o orden	216.20

Resumen

A lo largo de este capítulo, se propuso una variación del método utilizando la hipótesis de GW para constancia del color en los espacios de color CIELAB y CIELUV. Dada una imagen de prueba, se estima el iluminante y después se compara con un valor de referencia. Las pruebas experimentales se realizaron comparando los resultados de los métodos propuestos con los resultados utilizando el GW estándar en RGB y el algoritmo de GE. Según las pruebas, se concluye que los resultados de nuestro enfoque, los resultados del GW en un espacio de color perceptual, son mejores que los obtenidos utilizando el procedimiento estándar en RGB. A pesar de que los resultados utilizando el algoritmo de GE son ligeramente mejores que los que utilizan nuestro enfoque, para aplicaciones prácticas podemos optar por el nuestro, porque es mucho más rápido y no requiere un proceso de sintonía. También, podemos apreciar que GW en CIELUV es marginalmente mejor que en CIELAB de acuerdo a la precisión de la fuente luminosa estimada. Por otra parte, el tiempo de proceso es considerablemente más rápido en CIELUV. Por estas razones, nuestros enfoques, en particular el uso del GW en el espacio CIELUV, pueden considerarse métodos de pre-procesamiento muy adecuados para aplicaciones en tiempo real en los campos de reconocimiento de patrones y visión por computadora.

Selección automática de algoritmos de constancia de color para el mejoramiento de imágenes oscuras usando razonamiento basado en reglas difusas

Algunos trabajos de investigación se han orientado a la selección de algoritmos de constancia de color de acuerdo a varias características. El uso de análisis de imágenes basado en contenido para la corrección automática del color fue originalmente propuesto por Schroder y Moser [2001]. Este estudio clasificó las imágenes en un número de clases (e.g. interiores, exteriores, vegetación, montañas), y las asoció a un algoritmo particular (White-Patch y Gray-World). Gasparini y Schettini [2004] propusieron un método para el análisis de índices de influencia de color en las imágenes y su clasificación. Esta clasificación (e.g. piel, cielo, mar, vegetación) también permite detectar la presencia de un posible color dominante. El trabajo de van de Weijer et al. [2007b] usa información visual de alto nivel para modelar imágenes como una mezcla de clases semánticas (e.g. piel, camino, construcciones, pasto). Este último estudio utiliza la información visual para la selección del algoritmo de color más adecuado. Bianco et al. [2008] propusieron clasificar las imágenes en categorías de exteriores e interiores con el fin de seleccionar el mejor algoritmo para cada escenario. Mas tarde, ellos implementaron un selector automático de algoritmos tomando en cuenta propiedades de bajo nivel [Bianco et al., 2010]. Recientemente, Faghieh y Moghaddam [2014] usaron un clasificador para determinar el mejor grupo de algoritmos de constancia de color para una imagen de entrada, y después, algunos de los algoritmos de este grupo son combinados usando Optimización por Enjambre de Par-

tículas (PSO). Es importante señalar que todos estos trabajos de investigación han sido únicamente dirigidos a la estimación del iluminante. Según nuestro conocimiento no hay sistemas de selección de algoritmos de constancia de color enfocados a la mejora del contenido cromático.

En este capítulo, se propone un sistema basado en reglas difusas para la selección de uno de tres algoritmos básicos de constancia de color: El Gray-World (GW), el White-Patch (WP) y el Gray-Edge (GE). Este marco de trabajo es un enfoque de tres partes para resolver el problema. (1) El estudio se enfoca en la mejora del color usando algoritmos de constancia de color, al mismo tiempo que se remueve la influencia del iluminante. Además, éste es particularmente enfocado en el procesamiento de imágenes oscuras. (2) Un importante problema para desarrollar el sistema basado en reglas está en la elección correcta de las características de la imagen. (3) Se usa un conjunto de reglas difusas, recopilando el conocimiento necesario para tomar una decisión acerca del algoritmo más adecuado a ser aplicado a una imagen en consideración. Para realizar la selección, una imagen de prueba es sometida a una proceso de inferencia, donde el mejor algoritmo es elegido de acuerdo al nivel de activación correspondiente a las reglas.

3.1. Selección de un algoritmo de constancia de color

Nuestro enfoque consiste en un sistema MISO (Multiple Input, Single Output), incluyendo doce entradas calculadas de la escena bajo análisis y una sola salida, la etiqueta del algoritmo correspondiente. Las entradas y salidas son términos lingüísticos, enlazados por un conjunto de reglas *SI-ENTONCES*.

Básicamente, el marco de trabajo ha sido dividido en dos etapas. Por un lado, un protocolo de entrenamiento determina las reglas difusas, de acuerdo a las características calculadas de un conjunto de imágenes de entrenamiento. Como se mencionó, se usan doce características de la imagen de bajo nivel para el proceso de selección: siete características de color, tres de descriptores de textura y dos descriptores de iluminación. Y así aprovechar la ventaja de modelar intuitivamente pocas características. Por otro lado, dada una imagen de prueba, se elige el mejor algoritmo de acuerdo a la evaluación de las reglas en dos modelos de inferencia. Mamdani [Mamdani, 1976] y Larsen [Vadiee, 1993]. La Figura 3.1 muestra un diagrama de flujo de la metodología realizada en este estudio.

Para desarrollar el sistema basado en reglas usamos las imágenes de la base de referencia *SFU Laboratory* [Barnard et al., 2002]. Este conjunto de imágenes esta conformado por 529 imágenes oscuras con iluminaciones controladas. Es importante mencionar que este conjunto de datos fue creado para propósitos del problema real de la constancia de color. Sin embargo, estas imágenes se adaptan muy bien para el campo del mejoramiento de imágenes oscuras debido a su escasa iluminación.

3.1 Selección de un algoritmo de constancia de color

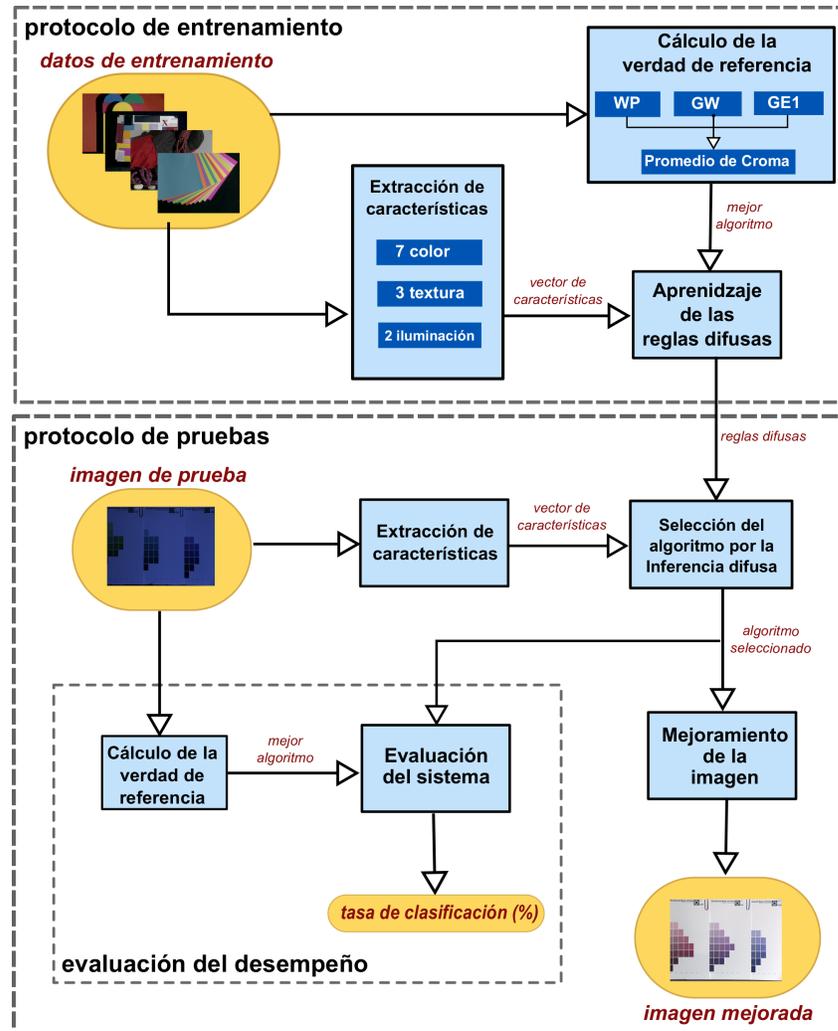


Figura 3.1: Diagrama de flujo de la metodología seguida en este estudio.

Los tres algoritmos que se consideran en la selección son:

- White-Patch (WP),
- Gray-World (GW),
- Gray-Edge de 1er orden (GE1).

Estos tres algoritmos son considerados por su exactitud en la estimación del iluminante y por su sobresaliente mejoramiento del color en imágenes oscuras. Además, ya habíamos llevado a cabo un estudio acerca del desempeño de algoritmos de constancia de color en tareas de mejoramiento de imágenes. Seis algoritmos de constancia de color fueron comparados y donde WP, GW y GE1 resultaron los mejores algoritmos en más del

3.1 Selección de un algoritmo de constancia de color

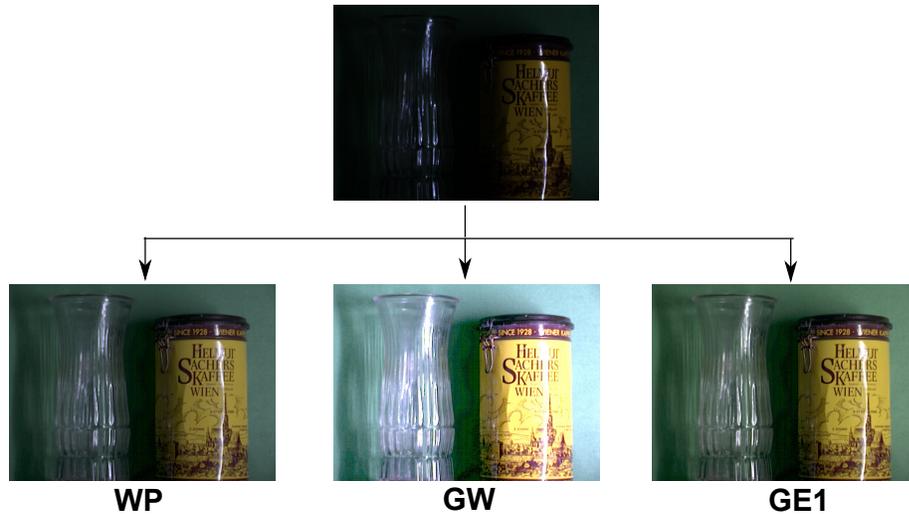


Figura 3.2: Imagen “vase_syl-cwf” del conjunto de datos SFU Laboratory. La escena original se procesa usando los tres algoritmos de constancia de color.

90 % de los casos de prueba usando la base de datos SFU Laboratory. Como consecuencia, el resto de los algoritmos se excluyeron de este estudio. Como ejemplo, la Figura 3.2 muestra una imagen perteneciente a la base de datos; además de la escena original y las salidas después de aplicar los tres algoritmos de constancia de color.

3.1.1. Métrica de comparación

Una medida de calidad es fundamental para preseleccionar las imágenes en la etapa de entrenamiento, para el aprendizaje de las reglas, y en el protocolo de pruebas, con el fin de comparar y evaluar el desempeño de nuestro sistema. La medida usada aquí, y descrita a continuación, permite encontrar la imagen resultante con la mejor adaptación cromática. Justo después, podremos concluir el algoritmo que logra el mejor realce en color para una imagen particular.

No existe una medida sin referencia establecida para evaluar el mejoramiento de imágenes a color. Los seres humanos tienen una predisposición natural a calificar positivamente una imagen con colores intensos [Joshi et al., 2011]. Para entender esto, necesitan ser estudiados aspectos filosóficos y psicológicos de la mente humana. Por ejemplo, Joshi et al. [2011] describieron aspectos acerca de estéticas y emociones presentes en imágenes, incluyendo referencias para entender la mente humana. Algunos trabajos de investigación han tratado de establecer medidas simples en este campo. También, llevamos a cabo pruebas experimentales y se observó que la medida de *Promedio de Cromo* [Tsagaris et al., 2005] proporciona información importante acerca de la calidad del color de las imágenes. Esta medida es también una característica de imágenes usada como entrada para el

3.2 Extracción de características

sistema de las reglas de este estudio.

En este trabajo son usadas 100 imágenes, elegidas aleatoriamente del conjunto de datos SFU Laboratory, como información de entrenamiento y las 429 restantes como información para prueba. Los tres algoritmos en consideración se aplican a las imágenes de entrenamiento. Después, se calcula la métrica de calidad para los resultantes de cada imagen para después compararlas entre sí. La medida con el valor más alto corresponde al mejor algoritmo para una imagen en particular. Así, es posible saber cuál es el mejor algoritmo para cada imagen. Esta pre-selección será nuestra verdad de referencia que se usará en el proceso de aprendizaje del selector. La Tabla 3.1 muestra con ejemplos de la clase correcta determinada en el proceso de entrenamiento.

Tabla 3.1: Ejemplos de la asignación de etiqueta en las imágenes de entrenamiento pre-seleccionadas de acuerdo al Promedio de Cromo. Los mejores resultados son mostrados en negrita.

Índice	WP	GW	GE1	Clase
0	22.24	23.43	21.79	2
1	21.03	22.04	19.83	2
2	30.12	22.39	25.86	1
3	33.25	23.84	27.56	1
4	14.43	14.79	12.17	2
5	13.53	13.88	12.11	2
⋮		⋮		
97	27.29	19.16	25.36	1
98	23.85	12.75	22.30	1
99	25.11	13.90	22.37	1

Para cada algoritmo se estimó la probabilidad a priori de ser la mejor elección. En el 48 % de los casos, la suposición del GW fue la mejor opción. En poco más de la tercera parte, el 36 % de las imágenes, el algoritmo WP corresponde a la mejor opción. El restante 16 % de las imágenes fueron mejor procesadas por el algoritmo GE1.

3.2. Extracción de características

Las características son usadas para describir imágenes numéricamente y deberían proporcionar información para representar una escena. Las características consideradas en este capítulo están principalmente relacionadas a la descripción del color. Algunas de estas características han sido usadas en trabajos similares [Bianco et al., 2010, 2008]. Además, también son consideradas otras características relacionadas a la descripción de textura e iluminación de la escena. Usamos doce características: siete de color, tres de

textura y dos de iluminación.

3.2.1. Color

Las características de color usadas en el sistema son las siguientes:

- Número de colores.
- Valor promedio del espectro de potencia (APSV).
- Índices de influencia del color (σ , D and D_σ).
- Promedio de croma (μ_{c^R}).
- Probabilidad de color dominante (PDC).

Además, es importante señalar que se usa el espacio de color RLAB para la extracción de cinco de las siete características de color. Este espacio es una extensión del espacio perceptual CIELAB. Los tres índices de influencia de color, el promedio de croma y la probabilidad de color dominante, son las características calculadas en el espacio RLAB. Para las dos restantes se usa el espacio RGB.

Las coordenadas cilíndricas del espacio RLAB están dadas por

$$h^R = \tan^{-1} \left(b^R / a^R \right), \quad (3.1)$$

$$c^R = \sqrt{(a^R)^2 + (b^R)^2}, \quad (3.2)$$

donde a^R y b^R son las componentes cromáticas del RLAB. También, h^R y c^R son el Matiz y la Croma, respectivamente. Cinco características son calculadas usando estas componentes cilíndricas.

Número de colores

La primer característica consiste en el número de colores y está relacionada al rango de color de la imagen. Ya que varios algoritmos de constancia de color están basados en la suposición del GW, el número de colores es un referente de si esta teoría es verdadera o no. Los colores reales de los píxeles pueden cancelar esta suposición, pero si una imagen contiene una gran variedad de colores, entonces el promedio es probable que sea un valor gris. En el espacio RGB, y considerando 8 bits por canal, es posible cuantificar 16 millones de colores ($2^8 \times 2^8 \times 2^8$). Aquí se usa un enfoque similar al de Bianco et al. [2008] para obtener el número de colores en la imagen.

Valor promedio del espectro de potencia

La medida del espectro de potencia [Zhang et al., 2011] ha sido usada para evaluar la calidad de la imagen y determinar la distribución luminosa y de color en la imagen. Esta característica permite el análisis de la información que no es evidente en el dominio espacial.

El Valor Promedio del Espectro de Potencia (APSV) está dado por

$$APSV = (1/3) \sum_i |F_i|^2, \quad (3.3)$$

donde i corresponde a la componente de color en el espacio RGB y $|F|^2$ está dado por $|F|^2 = (1/MN) \sum_u \sum_v |F(u, v)|^2$, donde M y N son las columnas y renglones, respectivamente. $F(u, v)$ es la imagen en el dominio de Fourier. El APSV tiende a ser más alto cuando la imagen muestra mayor contenido cromático bajo iluminaciones naturales.

Índices de influencia del color

Estos índices son características para el análisis de la distribución del color en la imagen. Fueron propuestos por Gasparini y Schettini [2004] para clasificación de imágenes (e.g. piel, cielo, mar, vegetación). También, estos índices han sido usados en trabajos similares [Bianco et al., 2010, 2008].

El cómputo de estos índices de influencia del color debería realizarse en espacios descorrelacionados donde la luminancia se encuentra separada de las componentes cromáticas. Primero, la imagen se transforma al espacio RLAB y se lleva a cabo el procedimiento propuesto en [Gasparini y Schettini, 2004]. Así, las medias y varianzas de las componentes a^R y b^R en RLAB (μ_{a^R} , μ_{b^R} , $\sigma_{a^R}^2$ y $\sigma_{b^R}^2$) son usadas para el cálculo del primer índice

$$\sigma = \sqrt{\sigma_{a^R}^2 + \sigma_{b^R}^2}, \quad (3.4)$$

que corresponde al radio de un Círculo Equivalente (EC) con centro en $C = (\mu_{a^R}, \mu_{b^R})$. Los otros dos índices están dados por

$$D = \mu - \sigma, \quad (3.5)$$

$$D_\sigma = D/\sigma, \quad (3.6)$$

donde $\mu = \sqrt{\mu_{a^R}^2 + \mu_{b^R}^2}$. D es una medida de qué tan lejos está la distribución de color del EC con respecto al origen ($a^R = 0$, $b^R = 0$). D_σ cuantifica la fuerza de la influencia de la distribución.

Promedio de croma

La croma es un atributo relacionado a la intensidad de los colores. Esta señal corresponde a una componente del espacio RLAB en coordenadas cilíndricas. También, la croma es un vector que representa la magnitud entre las componentes a^R y b^R , la cual se calcula como $c^R = \sqrt{(a^R)^2 + (b^R)^2}$. Entre más alto es el valor de la croma, más intenso es el color. Un estudio acerca de la croma como medida de calidad fue hecho por Tsagaris et al. [2005]. Nosotros usamos esta medida como un indicador de la calidad percibida por humanos y también como un descriptor de color. El promedio de croma es usado para la generación de verdad de referencia

$$\mu_{c^R} = \text{mean}\{c^R(x, y)\}. \quad (3.7)$$

Probabilidad de color dominante

La Probabilidad de Color Dominante (PDC) puede ser extraída de un histograma de color. Nosotros usamos un histograma de la componente de matiz del RLAB. Este histograma tiene 12 celdas, donde cada celda registra 30 grados sucesivos del matiz. Posteriormente, se calcula la probabilidad de cada celda como la razón entre el número de ocurrencias y el número total de píxeles en la imagen. P_{HR} corresponde a la función de distribución de probabilidad del matiz. La máxima probabilidad registrada en una celda está dada por

$$\text{PDC} = \max\{P_{HR}\}. \quad (3.8)$$

3.2.2. Textura

El uso de descriptores de textura está inspirado por el estudio de Bianco et al. [2010]. Ellos usaron algunas características relacionadas al análisis de textura y consideraron que ellas podrían describir la composición de la imagen, independientemente del estudio de color realizado. No obstante, nosotros tomamos en cuenta otras características de textura, las cuales fueron propuestas por Haralick et al. [1973], pero calculadas usando histogramas de Sumas y Diferencias (SDH), como una alternativa más rápida propuesta por Unser [1986]. El desplazamiento relativo del vector (\mathbf{V}) entre dos elementos de la imagen es un parámetro del SDH y en nuestro estudio está definido como una composición del producto cartesiano $R \times \theta$, donde $R = \{1, 2\}$ y $\theta = \{0, \pi/4, \pi/2, 3\pi/4\}$.

Las tres características de textura consideradas son las siguientes

$$\begin{aligned} \text{Entropía} = & - \sum_i P_s(i) \cdot \log(P_s(i)) \\ & - \sum_j P_d(j) \cdot \log(P_d(j)) \end{aligned} \quad (3.9)$$

$$\text{Contraste} = \sum_j j^2 \cdot P_d(j) \quad (3.10)$$

$$\text{Homogeneidad} = \sum_j (1 + j^2)^{-1} \cdot P_d(j) \quad (3.11)$$

donde P_s y P_d son los SDH normalizados. Es importante mencionar que estas características son calculadas usando sólo la componente de luminancia L^* del espacio RLAB.

3.2.3. Iluminación

Dos componentes de iluminación son consideradas en este estudio,

- promedio de iluminación (μ_{LR}).
- Probabilidad de especularidad (PS).

Estas características también se obtienen del espacio de color RLAB. Es importante señalar que la componente de luminancia mide la intensidad luminosa o de iluminación en la imagen. El promedio de iluminación se calcula como

$$\mu_{LR} = \text{mean}\{L^R(x, y)\} \quad (3.12)$$

La especularidad de refiere a la cantidad de luz que se emite por la reflexión de un objeto. Comúnmente, una imagen oscura tiene baja probabilidad de especularidad. Ésta se obtiene a partir de un histograma de luminancia (0 a 100 en RLAB) y se normaliza. Y así, se obtiene la función de densidad de probabilidad (pdf) de la luminancia. Consideramos la función de distribución acumulada (cdf) entre las celdas 91 y 100, ya que consideramos que esta información esta contenida en esta parte del histograma,

$$\text{PS} = \sum_{i=91}^{100} P_{LR}(i), \quad (3.13)$$

donde P_{LR} es el histograma de luminancia, e i es el índice de la celda del histograma.

3.3. Aprendizaje de las reglas difusas

La determinación de las reglas difusas es un aspecto clave del sistema propuesto. Estas reglas son significativamente importantes porque contienen la información necesaria para tomar una decisión sobre el algoritmo correcto a aplicar. En el protocolo de entrenamiento, necesitamos definir los términos lingüísticos y conjuntos difusos. Después, se

formulan y sintonizan las reglas difusas para mejorar la selección. Posteriormente, para una imagen de prueba se usan las reglas en un modelo de inferencia para elegir el mejor algoritmo.

El primer paso para el desarrollo de nuestro sistema consiste en la definición de las características funcionales y operacionales. Tales características incluyen los términos lingüísticos, la formulación de las reglas y el modelo de inferencia. Se considera un conjunto de i entradas $\mathbf{X} = [x_1, x_2, \dots, x_i]^T \in \mathbb{R}^i$ para la selección, correspondiente a un vector de características con i elementos, donde un elemento dado es una característica cuantitativa. Doce características se incluyen en el vector \mathbf{X} : Número de colores (x_1), APSV (x_2), σ del EC (x_3), D del EC (x_4), D_σ del EC (x_5), Promedio de croma (x_6), Probabilidad de un color dominante (x_7), Entropía (x_8), Contraste (x_9), Homogeneidad (x_{10}), Promedio de iluminación (x_{11}), y Probabilidad de especularidad (x_{12}). La salida del sistema (Ω) toma un valor de un conjunto $\Omega = \{\omega_1, \omega_2, \omega_3\}$ correspondiente a las etiquetas de 3 clases conocidas. Las clases de salida representan el algoritmo de constancia de color considerado en la selección: WP (ω_1), GW (ω_2) y GE1 (ω_3).

Para cada imagen de entrenamiento se extrae su vector de características. Una vez que todas las características se extraen para todo el conjunto de entrenamiento, éstas son normalizadas entre 0 y 1, de acuerdo a los valores mínimos y máximos registrados en cada característica registrada en la etapa de entrenamiento (ver la Tabla 3.2).

Tabla 3.2: Características y su rango calculado en la etapa de entrenamiento. El vector de características en el protocolo de prueba es normalizado con base en estos valores.

Característica	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
min	1083.00	0.03	5.96	22.78	0.95	1.79	0.08	2.73	1.31	0.26	11.79	0.00
max	62166.00	60.19	32.46	38.97	4.99	49.72	0.88	12.31	434.34	0.88	71.71	0.01

3.3.1. Definición de los términos lingüísticos

El siguiente paso en el proceso de entrenamiento es la definición de los términos lingüísticos y sus correspondientes conjuntos difusos. A partir de cada vector \mathbf{X} , cada característica de entrada x_i es un valor en el dominio de términos lingüísticos, particionados en un número de conjuntos difusos.

La definición de conjuntos difusos representa uno de los aspectos más importantes en el proceso de diseño. De acuerdo a nuestra experiencia, los términos lingüísticos deben ser separados en 4 conjuntos difusos con la intención de obtener buena resolución y alcanzar una capacidad de decisión adecuada evitando un término lingüístico intermedio. La Figura 3.3 muestra como cuatro conjuntos difusos Gaussianos son distribuidos a través del dominio de una característica, $x_i \rightarrow \{0, 1\}$, considerando una desviación

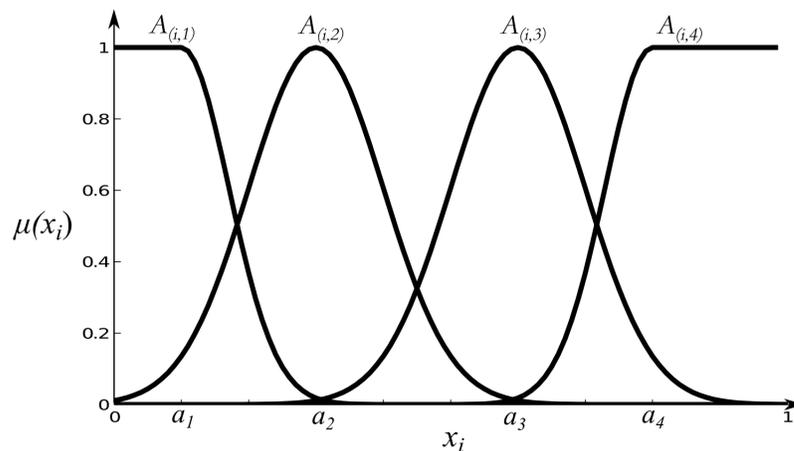


Figura 3.3: Distribución de los conjuntos difusos Gaussianos en el dominio de una característica.

estándar de $\sigma = 0.1$.

Usamos funciones Gaussianas, con media en a_m y $\sigma = 0.1$ donde $m = \{1, 2, 3, 4\}$. La conformación del conjunto difuso se hace usando 4 agrupaciones por característica sobre cada variable lingüística. El centro de una agrupación a_m , es el máximo grado de membresía en el conjunto difuso $A_{(i,m)}$. Cabe mencionar, que el término membresía es equivalente a pertenencia, el cual también se encuentra en literatura de lógica difusa. Para casos específicos, las etiquetas $A_{(i,m)}$ son remplazadas por términos lingüísticos como $A_{(i,1)} = \text{"MUY BAJO"}$, $A_{(i,2)} = \text{"BAJO"}$, $A_{(i,3)} = \text{"ALTO"}$ and $A_{(i,4)} = \text{"MUY ALTO"}$. Para la primera función Gaussiana $A_{(i,1)}$, su lado izquierdo permanece con un valor de membresía de 1 para valores más bajos en el dominio de x_i . De manera similar ocurre para la última función $A_{(i,4)}$, pero ahora hacia el lado derecho. La Tabla 3.3 muestra el núcleo con máximo grado de membresía (centro de las Gaussianas) para cada característica.

3.3.2. Reglas y selección de las características

Al principio del proceso de entrenamiento, el selector automático fue diseñado usando tres reglas, una para cada algoritmo de salida. Las imágenes pertenecientes al algoritmo (aquellas donde dicho algoritmo fue mejor) se incluyeron para generar la regla correspondiente. Sin embargo, los resultados mostraron un rendimiento de selección muy bajo, logrando un valor cercano a 30 %. Este resultado tiene dos explicaciones posibles. En primer lugar, las características seleccionadas no son lo suficientemente discriminantes. En segundo lugar, hay imágenes con diferentes características que corresponden al mismo algoritmo de constancia del color, es decir, un algoritmo específico es adecuado para imágenes que no son similares entre sí. Hemos experimentado con un gran número de características, sin embargo, las doce descritas anteriormente fueron escogidas debido

Tabla 3.3: Valores máximos de membresía ($A_{(i,m)}$).

Característica	$A_{(i,1)}$	$A_{(i,2)}$	$A_{(i,3)}$	$A_{(i,4)}$
x_1	0.09	0.20	0.33	0.52
x_2	0.05	0.20	0.37	0.80
x_3	0.06	0.25	0.38	0.49
x_4	0.20	0.25	0.35	0.50
x_5	0.06	0.15	0.25	0.50
x_6	0.05	0.27	0.40	0.50
x_7	0.03	0.18	0.35	0.70
x_8	0.12	0.47	0.65	0.80
x_9	0.06	0.12	0.20	0.39
x_{10}	0.30	0.45	0.60	0.92
x_{11}	0.05	0.30	0.50	0.75
x_{12}	0.00	0.01	0.04	0.86

a su mayor capacidad de discriminación. Por esta razón, se consideró que la segunda hipótesis puede ser correcta.

Numerosas pruebas experimentales se realizaron utilizando una búsqueda exhaustiva para determinar el número apropiado de reglas. Como resultado, se encontró que, por un lado, el uso de menos de cinco reglas por algoritmo proporciona un rendimiento deficiente. Por otra parte, la inclusión de más de cinco reglas resultó en una mejora insignificante en el rendimiento. En consecuencia, consideramos que cinco reglas por algoritmo son adecuadas, lo que nos permite tener un control razonablemente intuitivo de la información, aprovechando la ventaja de manejar grandes cantidades de datos en unas pocas reglas.

El algoritmo k -means es usado sobre las características para generar cinco grupos por clase en un espacio de 12 dimensiones. Así, una regla de cada subconjunto se genera usando el proceso de generación de reglas. Como resultado, se formula un total de 15 reglas en el proceso de aprendizaje.

Una muestra de las 15 reglas formuladas es

$$R^1 : \text{SI } x_1 \text{ es } A_{(1,1)} \text{ y } x_2 \text{ es } A_{(2,1)} \text{ y } \dots \text{ y } x_{12} \text{ es } A_{(12,3)}, \text{ ENTONCES } \Omega \text{ es } \omega_1.$$

Analizar las 15 reglas con esta notación podría resultar una tarea complicada. La Tabla 3.4 muestra una manera práctica de representar las 15 reglas, donde se muestra $A_{(i,j)}$ únicamente con un índice j . Cada etiqueta es representada por un número que puede ser asociado a una etiqueta intuitiva: 1 a "MUY BAJO", 2 a "BAJO", 3 a "ALTO", y finalmente, 4 a "MUY ALTO". El conjunto de salida es representado usando el acrónimo del algoritmo.

3.3 Aprendizaje de las reglas difusas

Tabla 3.4: Una manera práctica de representar las 15 reglas. Todas las premisas se incluyen en cada antecedente.

R^n	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	Ω
1	1	1	2	3	2	2	2	2	1	3	2	3	WP
2	1	1	2	4	3	3	4	2	1	2	2	1	WP
3	2	2	4	2	2	4	3	3	1	3	3	1	WP
4	3	4	1	4	4	4	4	4	4	2	4	2	WP
5	4	1	4	2	2	3	2	4	2	1	3	2	WP
6	3	1	4	2	1	2	2	3	2	2	2	1	GW
7	1	1	3	1	1	2	1	2	3	3	2	4	GW
8	1	3	3	3	2	2	2	2	1	3	2	1	GW
9	1	1	2	2	1	1	3	1	1	4	1	2	GW
10	2	1	4	1	1	3	2	3	2	2	2	2	GW
11	2	4	2	4	3	4	4	4	2	3	4	1	GE1
12	3	2	3	4	3	4	2	3	1	2	3	1	GE1
13	1	1	4	3	2	3	3	2	1	3	2	2	GE1
14	4	3	3	1	1	2	2	4	3	1	4	1	GE1
15	3	1	1	4	4	3	3	4	1	1	3	1	GE1

3.3.3. Sintonización de las reglas

El espacio de características es muy grande y el número posible de reglas difusas que cubren todo ese espacio es de 4^{12} , aproximadamente diecisiete millones de reglas. Esto significa que si sólo se consideran 15 reglas por el sistema experto, no se toma en cuenta una gran parte del espacio de características. Sin embargo, si cada regla es sintonizada con la intención de cubrir una sección más grande del espacio, el rendimiento de la selección debe aumentar. Con el fin de ampliar la cobertura del espacio por cada regla es necesario excluir una o varias premisas en el antecedente. Por lo tanto, la regla formulada es más robusta e implícitamente incluye otras premisas en el antecedente.

El proceso de sintonización se realiza a través de una experimentación exhaustiva. Los conjuntos difusos ya están establecidos en cada premisa para toda regla. Ahora se excluye (considera '-') de una por una cada premisa en todas las reglas, de manera que al probar con cada exclusión podamos probar el sistema y medir su tasa de selección global. Todas las combinaciones posibles se realizaron en la base de conocimiento generada por las 15 reglas (15^{12}). Finalmente, se adoptó la mejor combinación y las 15 reglas con las premisas sintonizadas se muestran en la Tabla 3.5. Además, como se puede apreciar en la Tabla 3.5, las reglas incluidas muestran el símbolo '-' que representa a las premisas excluidas. Un guión indica la exclusión de este conjunto difuso de las premisas de la regla. Esta situación de "no importa" resulta en la cobertura de todo el espacio de la característica.

3.3 Aprendizaje de las reglas difusas

Es decir, cualquier valor que tenga la característica, la regla no se ve afectada. Un ejemplo de las 15 reglas sintonizadas es el siguiente

$$R^2 : \text{SI } x_3 \text{ es } A_{(3,2)} \text{ y } x_6 \text{ es } A_{(6,3)} \text{ y } x_{11} \text{ es } A_{(11,2)}, \text{ ENTONCES } \Omega \text{ es } \omega_1$$

Tabla 3.5: Una manera práctica de representar las 15 reglas sintonizadas. Algunas premisas fueron excluidas para incrementar la cobertura del espacio.

R^n	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	Ω
1	1	-	2	3	-	2	2	2	-	3	2	3	WP
2	-	-	2	-	-	3	-	-	-	-	2	-	WP
3	-	2	-	2	-	4	-	-	-	-	-	-	WP
4	-	4	-	4	4	-	-	-	4	-	-	-	WP
5	-	1	-	2	2	-	2	4	2	1	3	-	WP
6	3	-	4	2	-	-	-	3	-	2	2	-	GW
7	-	1	3	1	-	2	-	2	-	-	-	-	GW
8	-	3	3	-	-	2	-	2	-	-	2	-	GW
9	-	-	-	-	-	-	-	1	-	4	1	-	GW
10	-	1	4	1	-	-	-	-	-	-	2	-	GW
11	2	4	-	4	3	4	-	4	-	-	-	-	GE1
12	3	-	-	4	-	4	2	3	-	-	3	-	GE1
13	1	1	4	3	-	-	3	2	-	3	2	-	GE1
14	4	3	-	-	-	2	-	-	-	1	-	-	GE1
15	-	1	1	-	4	-	-	-	-	-	3	-	GE1

3.3.4. Evaluación del desempeño

En la evaluación se utilizaron matrices de confusión. La matriz de confusión es una herramienta de visualización que se utiliza ampliamente en el aprendizaje supervisado. Cada columna de la matriz representa el número o porcentaje de predicciones por clase, mientras que cada fila representa las instancias de la clase real. La principal ventaja de esta herramienta es que muestra la confusión entre clases. Una matriz de confusión ideal debe mantener toda la información en la diagonal principal.

Un "éxito" es obtenido cuando el sistema ha seleccionado correctamente el algoritmo que le corresponde a una imagen. Además, este resultado debe ser registrado en la diagonal principal de la matriz de confusión. Un "error" es aquel en el que el sistema elige incorrectamente el algoritmo. Por lo tanto, esta selección es registrada en la fila del algoritmo correcto y la columna del algoritmo seleccionado por el sistema. La tasa de éxito está dada por $(N_{correcto} / N) \times 100$, donde $N_{correcto}$ es el número de veces que un algoritmo se ha elegido correctamente y N es el número total de muestras.

3.4. Resultados del sistema

En esta sección se presentan los resultados obtenidos de nuestro sistema basado en reglas difusas. En el protocolo de prueba se utilizan 429 imágenes de la base de datos SFU Laboratory. En las secciones anteriores se mencionó que se tomaron de forma aleatoria 100 imágenes de todo el conjunto para el proceso de aprendizaje. Es importante tener en cuenta que los experimentos se llevaron a cabo utilizando dos modelos de inferencia: Mamdani y Larsen. La tasa de selección obtenida en los resultados utilizando las 15 reglas “no optimizadas”, se muestra en la Tabla 3.6. Usando la inferencia Mamdani el sistema obtiene el 60.8% como tasa de selección correcta. La inferencia de Larsen proporciona una mejora marginal de 62.4%.

Tabla 3.6: Desempeño usando las cinco reglas “no optimizadas” por algoritmo. Usando la inferencia de Larsen se obtiene una mejor tasa de selección.

Inferencia	Selección (%)
Mamdani	60.8
Larsen	62.4

Más tarde, los experimentos se realizaron con las cinco reglas “ optimizadas” por algoritmo. La Tabla 3.7 presenta la matriz de confusión obtenida usando la inferencia de Mamdani. La matriz de confusión utilizando la inferencia de Larsen se presenta en la Tabla 3.8. La tasa de selección de nuestro enfoque aumentó a 69.2% utilizando la inferencia de Mamdani y a 77.5% utilizando la inferencia de Larsen.

Tabla 3.7: Matriz de confusión de los resultados de selección usando la inferencia de Mamdani en el conjunto de datos de prueba.

Correcto	Seleccionado		
	WP	GW	GE1
WP	70.1	15.3	14.6
GW	17.5	74.1	8.5
GE1	36.4	7.8	55.8
Total	69.2%		

El traslape de las características es un indicador de una alta similitud en las imágenes, incluso entre aquellos que pertenecen a diferentes clases. En la evaluación, varias imágenes de prueba proporcionan al menos una característica que no está incluida en el espacio de características creado por las reglas. Es importante señalar que, si una característica está fuera del rango establecido por la regla, el grado de membresía tiende a cero en la premisa. Por lo tanto, el nivel de activación final, también es casi nulo. Aplicando el mismo proceso para cada regla, el nivel máximo de activación se mantiene cercano a

Tabla 3.8: Matriz de confusión de los resultados de selección usando la inferencia de Larsen en el conjunto de datos de prueba.

Correcto	Seleccionado		
	WP	GW	GE1
WP	79.6	14.6	5.7
GW	13.2	83.1	3.7
GE1	31.2	9.1	59.7
Total	77.5 %		

cero. En consecuencia, el selector no puede tomar una decisión adecuada para la elección de un algoritmo. Para este caso de incertidumbre, el algoritmo más fuerte (GW) podría haber sido elegido de acuerdo con las probabilidades en el protocolo de entrenamiento. Sin embargo, en esta situación, cualquier algoritmo elegido proporciona una tasa de selección muy baja.

3.4.1. Comparación con otros métodos de selección

La meta principal de este trabajo fue el desarrollo de un selector basado en reglas difusas en una práctica y novedosa aplicación. Sin embargo, no existen métodos aplicados a este problema que sirvan como referencia. Por esa razón, únicamente usamos métodos clásicos para la evaluación de la tasa correcta de selección.

Se tomaron en cuenta algunos métodos clásicos para clasificación a efectos de comparación. El algoritmo *K*-Nearest-Neighbor (*K*-NN) es un método no paramétrico utilizado para la clasificación y la regresión. La entrada consiste en los *K* ejemplos de entrenamiento más cercanos en el espacio de características. *K*-means es otro método básico para el análisis de agrupaciones en la minería de datos. Este método tiene como objetivo particionar *n* observaciones en *k* agrupaciones en las que cada observación pertenece a la agrupación con la media más cercana, siendo un prototipo de la agrupación. Un clasificador discriminante lineal (LDC) es un método paramétrico basado en el valor de una combinación lineal de las características. Un clasificador discriminante cuadrático (QDC) se utiliza para mediciones separadas de dos o más clases de objetos o eventos por una superficie cuadrática. Éste es una versión más general del método LDC. Más detalles acerca de estos métodos, se pueden encontrar en [Kuncheva, 2000; Duda et al., 2012].

La Tabla 3.9 muestra los resultados obtenidos por cada método de selección. Los resultados muestran que nuestro enfoque proporciona la mejor selección. El algoritmo *K*-NN presenta una ligera mejora en el desempeño si se usa solo un vecino en lugar de tres, con 62.4 % y 60.5 %, respectivamente. Nuestro selector usa 15 reglas (5 por algoritmo), por lo que el algoritmo *k*-means también considera *k* = 15. Este método muestra un mejor desempeño con el 64.0 %. Los métodos QDC y LDC alcanzan un desempeño más

3.4 Resultados del sistema

Tabla 3.9: Comparación con otros métodos de selección.

Enfoque	Tasa de selección (%)
K-NN (K=3)	60.5
K-NN (K=1)	62.4
<i>k</i> -means	64.0
QDC	66.9
Sistema difuso basado en reglas (Mamdani)	69.2
LDC	72.3
Sistema difuso basado en reglas (Larsen)	77.5

alto, con 66.9 % y 72.3 %, respectivamente. A pesar que los últimos métodos proporcionan una mejor tasa de selección que K-NN y *k*-means, nuestro enfoque, usando la inferencia de Larsen, es el mejor método logrando un 77.5 %. De acuerdo a esta tasa de selección, podemos decir que nuestra propuesta es razonablemente mejor que el resto de los clasificadores. La Figura 3.4 muestra 5 ejemplos de imágenes seleccionadas correctamente.

La clasificación o selección de imágenes usando algoritmos de constancia de color es una tarea difícil y desafiante. Por ejemplo, en un trabajo análogo extrayendo contenido visual de alto nivel [Bianco et al., 2010] la tasa de éxito fue aproximadamente del 40 % usando 87 características de la imagen. El propósito de nuestro estudio fue ligeramente diferente, orientado a cuestiones de mejora de la imagen en lugar de la estimación del iluminante. Sin embargo, se logró un significativo 77.5 % de tasa de selección, con un intervalo de confianza del 95 % de [74.1, 81.9]. A pesar de que utilizamos aproximadamente el veinte por ciento de las imágenes en el proceso de aprendizaje, en lugar del sesenta o setenta por ciento que se usa normalmente para entrenar. Otra ventaja de nuestro enfoque consiste en el uso de sólo doce características con un sentido intuitivo. Creemos que estas características son muy importantes para obtener la tasa de selección aceptable.

Si se revisa la Tabla 3.5, podemos hacer algunas conclusiones acerca de la contribución de las características en las reglas. Todas las características se consideran en las reglas para el algoritmo WP. Aunque el Número de colores (x_1) ó la Probabilidad de especularidad (x_{12}) solamente son tomadas en cuenta en la primera regla. Las reglas para el algoritmo GW no necesitan incluir la Fuerza del EC (x_5), Probabilidad de color dominante (x_7), Contraste (x_9) y Probabilidad de especularidad (x_{12}). Sin embargo, el Radio del EC (x_3) y la Entropía (x_8) son las más significativas para este algoritmo. Finalmente, en el caso del algoritmo GE1, el Número de colores (x_1) es aparentemente importante para las reglas.

En resumen, las características con mayor ocurrencia en las reglas son el Valor promedio del espectro de potencia (x_2), la Distancia al EC (x_4), la Entropía (x_8) y el Promedio de iluminación (x_{11}). Podemos concluir que estas características son muy importantes

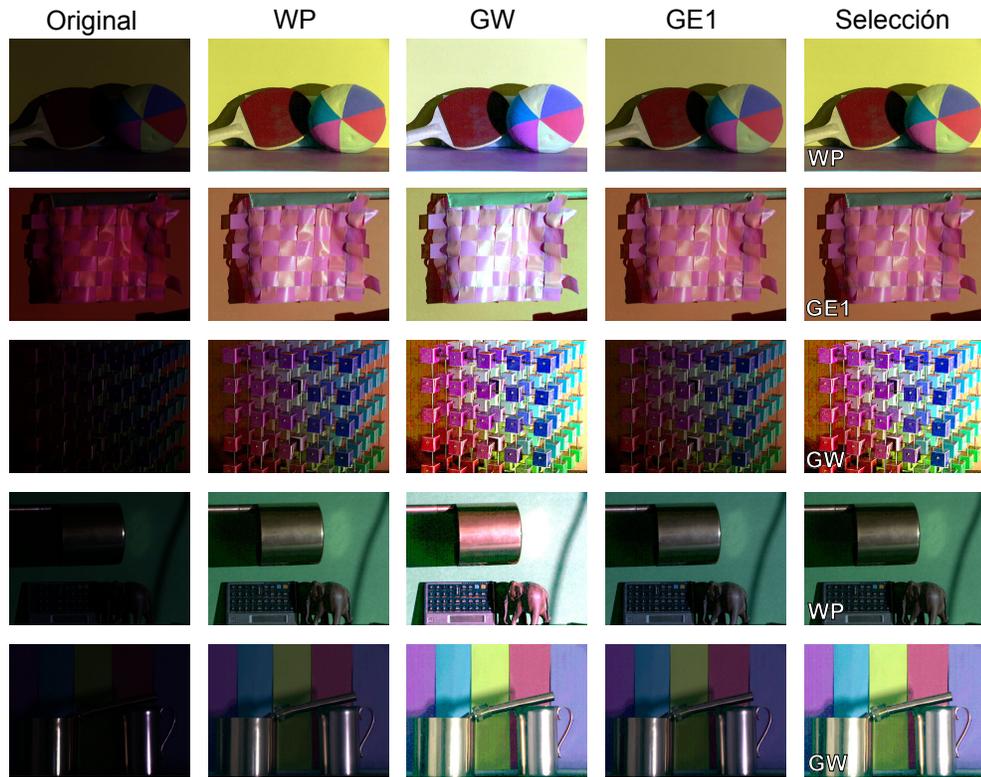


Figura 3.4: Cinco ejemplo de los 429 usados en los experimentos. Éstos fueron seleccionados correctamente. De arriba a abajo: imágenes “ball_solux-4100”, “weave_syl-50MR16Q”, “clr_cubes3-solux-4700”, “elephant_solux-4100+3202” y “fcooper_syl-wwf”. De izquierda a derecha: Imagen original, resutantes del WP, resultantes del GW, resultantes del GE1 y la imagen seleccionada por el sistema.

para el desempeño de nuestro sistema. Es importante señalar que los resultados de este trabajo ya han sido publicados [Cepeda-Negrete y Sanchez-Yanez, 2015].

Resumen

A lo largo de este capítulo fue introducido un sistema experto que selecciona algoritmos de constancia de color. Esta tarea está orientada a la selección entre tres algoritmos que mejoren el contenido cromático de imágenes oscuras: White-Patch, Gray-World, y Gray-Edge. Estos algoritmos han sido ampliamente usados en tareas de constancia de color gracias a sus buenos resultados. Aunque también han mostrado que son muy útiles en el mejoramiento de imágenes, especialmente en imágenes bajo condiciones de iluminación bajas.

El diseño de un sistema automático no es una tarea trivial cuando muchas caracte-

3.4 Resultados del sistema

rísticas están involucradas en la selección. Es por esto que se desarrolló un sistema basado en reglas difusas, con la intención de modelar la información a través de reglas simples. Una ventaja de este enfoque es que solamente se usan doce características. Consideramos que estas características son muy importantes para la tasa de selección alcanzada. Los experimentos se realizaron de manera separada usando inferencias de Mamdani y Larsen. Específicamente, el modelo de la inferencia de Larsen proporcionó una tasa de selección sobresaliente para esta tarea.

La meta principal de este trabajo ha sido el desarrollo de un sistema selector basado en reglas difusas para una tarea novedosa y práctica. El enfoque propuesto resulta en una tarea muy adecuada para resolver dos problemas al mismo tiempo: la constancia de color y el mejoramiento de la imagen. Especialmente para imágenes con bajas condiciones de iluminación. El marco de trabajo es fácil de replicar para posibles tareas posteriores.

Transferencia de color perceptual para mejorar imágenes de noche

Existe un problema con respecto al uso de imágenes oscuras en aplicaciones de transferencia de color. La información contenida en una imagen RGB es el producto de la intensidad de luz registrada por el sensor. Sin embargo, la medición de la intensidad de luz no corresponde a la percepción del color en humanos. Los colores y niveles de intensidad percibidos por el ser humano no son lineales, y no corresponden a la intensidad real de la luz que el sensor alcanza. Este problema es más fuerte en imágenes oscuras (o en zonas oscuras de la imagen).

Aunque estudios previos han abordado el mejoramiento de imágenes de noche, varios problemas aún se siguen presentando. En este trabajo, extendemos la aplicabilidad de una técnica de transferencia de color muy conocida, ahora para transformar una imagen oscura a otra con una apariencia de iluminación natural. Nuestro enfoque transfiere los colores de una imagen deseada en una imagen oscura, obtenida a partir de un sensor RGB común. En nuestro enfoque, para obtener un resultado que muestra colores similares a los de la imagen de destino, sólo se requieren estadísticas a partir de una sola imagen. En un espacio de color perceptual, la representación del color de la imagen se ajusta para emular la forma en que el ojo humano percibe el color. Los espacios de color perceptuales se utilizan para conseguir correcciones precisas de color, porque la transformación de una imagen RGB a tal espacio de color elimina los errores producidos por la percepción no lineal inducida en las zonas oscuras. El rendimiento de la metodología de transferencia de color se compara en cuatro espacios de color perceptuales. Además, comparamos nuestros resultados con los generados por métodos de referencia. Debido a que no existe una base de datos de referencia para este fin, proporcionamos un conjunto de datos que consiste en imágenes oscuras en RGB, que se obtiene utilizando una cámara de uso común. Las pruebas realizadas apoyan nuestra hipótesis: los resultados de la transferencia de color son mejores cuando se utilizan espacios de color perceptuales.

4.1 Marco de trabajo

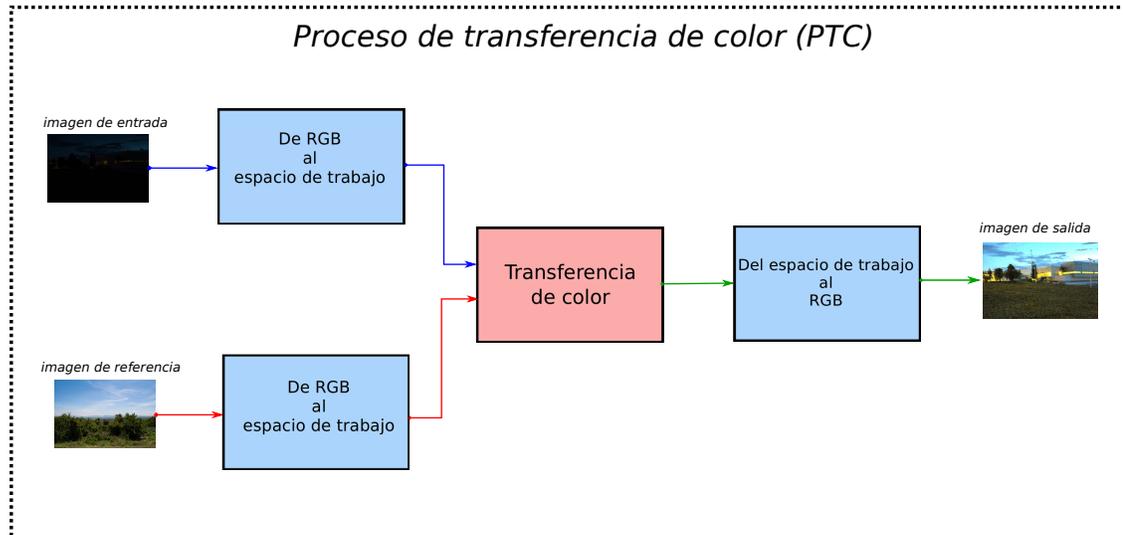


Figura 4.1: Proceso de la transferencia de color para una imagen oscura.

Entre los espacios de color, RLAB muestra los mejores resultados. RLAB fue diseñado específicamente para emular la percepción del color que tiene un humano bajo condiciones de luz no comunes, incluyendo escenas oscuras [Fairchild, 1996].

4.1. Marco de trabajo

En esta sección se presenta la técnica de transferencia de color (CT) usada en este trabajo. Las imágenes usadas para pruebas se obtienen de un sensor RGB común. La técnica de transferencia de color es la propuesta por Reinhard et al. [2001]. Aunque hay muchos enfoques de transferencia de color, este método fue elegido por su simplicidad y rapidez. La Figura 4.1 muestra el proceso de la transferencia de color usando esta técnica. Los detalles con respecto de la metodología son descritos más tarde.

Se comparan las metodologías usando el proceso de la transferencia de color en espacios de color perceptuales y en RGB. Esta comparación se realiza para determinar qué metodología es la más adecuada para ser usada en imágenes bajo condiciones de baja iluminación. La Figura 4.2 representa la metodología seguida para la comparación de los diferentes enfoques. Los métodos que usamos para comparar son: el algoritmo White-Patch (WP), el algoritmo Gray-World (GW), la Igualación de Histograma (HE), y la transferencia de color (CT) en diferentes espacios de color (RGB, $l\alpha\beta$, CIELUV, CIELAB y RLAB).

Es importante mencionar que, para una imagen dada, las componentes de color se procesan separadamente. Si se usa otro espacio de color, la imagen RGB necesita ser transformada a ese espacio de color antes de realizar la transferencia de color. Después, la imagen se transforma de vuelta a RGB para visualizar el resultado.

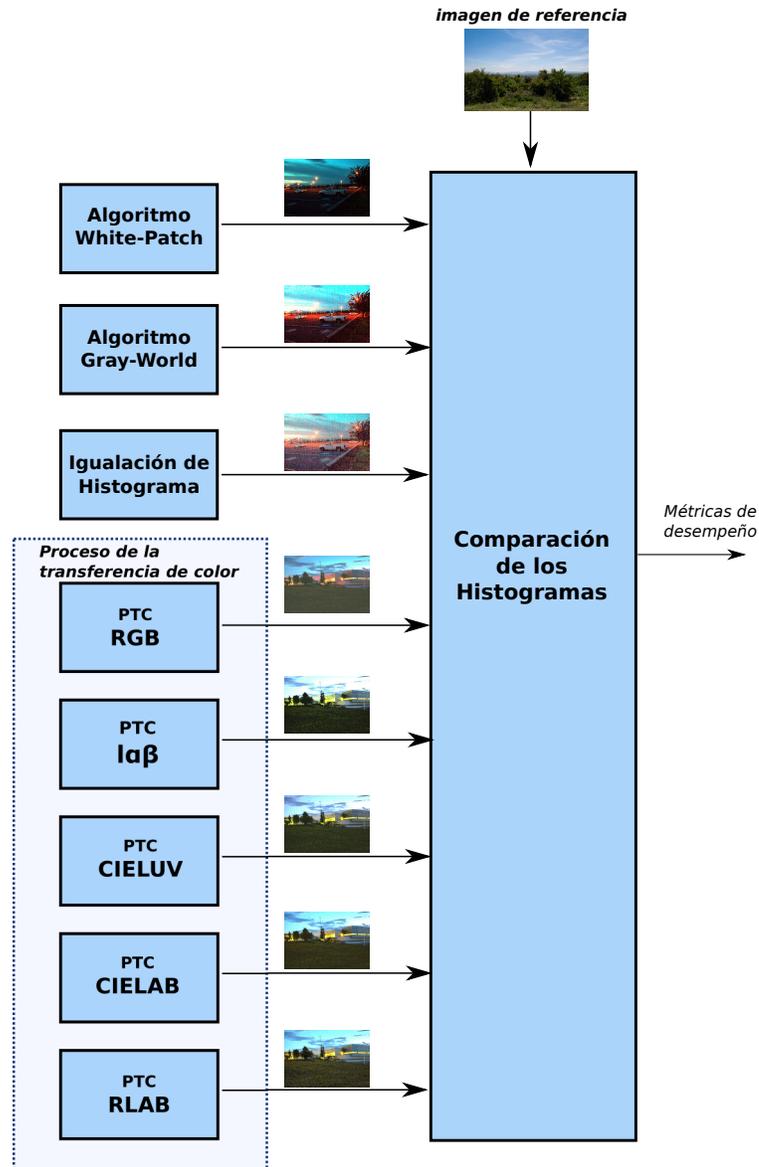


Figura 4.2: Diagrama a bloques de la metodología seguida para la comparación de enfoques y transferencia de color en distintos espacios.

4.1.1. Espacios de color perceptuales usados en la transferencia de color

El método propuesto por Reinhard et al. [2001] se aplicó originalmente en el espacio de color $l\alpha\beta$. Varios trabajos que han adoptado este enfoque [Toet, 2003; Yin y Coopers-tock, 2004; Xiang et al., 2009]. También, Reinhard y Pouli [2011] realizaron una comparación de espacios de color, concluyendo que es recomendado el uso del espacio CIELAB para la transferencia de color en imágenes naturales. Este proceso puede ser mejorado usando otros espacios de color perceptuales.

Se realizan pruebas usando cinco espacios de color. Las imágenes son obtenidas de

Tabla 4.1: Las componentes correspondientes al índice i de acuerdo a cada espacio de color.

Espacio de color	índice i
RGB	$i \in \{R, G, B\}$
$l\alpha\beta$	$i \in \{l, \alpha, \beta\}$
CIELUV	$i \in \{l, u, v\}$
CIELAB	$i \in \{l, a, b\}$
RLAB	$i \in \{l, a, b\}$

escenas bajo condiciones de escasa iluminación. Los espacios de color considerados en este estudio son: RGB, $l\alpha\beta$, CIELUV, CIELAB y RLAB. Otros espacios de color también fueron considerados: HSI, YIQ, CIECAM y YCbCr. Sin embargo, los resultados en este estudio mostraron que esos espacios no son adecuados en nuestra aplicación, y por esa razón no se reportaron sus números en el documento. La Tabla 4.1 muestra las componentes de la imagen con su correspondiente índice i para cada espacio de color usado. El segundo espacio considerado es el $l\alpha\beta$, inspirado de los estudios [Reinhard et al., 2001; Toet y Franken, 2003].

En 1976, debido a la proliferación de ecuaciones para calcular las diferencias del color, el CIE recomendó el uso de dos espacios de color perceptuales para promover una mejor ecuación de color [Fairchild, 2005]: El espacio de color CIE 1976 (L, u^*, v^*), comúnmente conocido como CIELUV, y el espacio CIE 1976 (L, a^*, b^*), mejor conocido como CIELAB. Para esos espacios, la distancia Euclidiana entre dos puntos en el espacio es proporcionalmente uniforme a las diferencias perceptuales de los colores correspondientes en esos puntos.

Como se mencionó antes, el espacio de color RLAB es una mejora del CIELAB. El RLAB fue diseñado con el fin de arreglar los problemas mostrados por el CIELAB, que se presentaban en condiciones de iluminación inusuales [Fairchild, 1996]. Este espacio mantiene las propiedades perceptuales bajo condiciones de luz normales (luz natural) y también bajo condiciones extremas. Imágenes oscuras son un ejemplo de tales casos extremos. Para mayor información, favor de consultar el trabajo de Fairchild [1996].

Todas las ecuaciones de transformación entre los espacios de color RGB y el de trabajo pueden ser revisadas en el Apéndice A.

4.1.2. Valoración de la transferencia de color

Cuando diferentes algoritmos se aplican a una imagen, se necesita una medida objetiva para comparar las imágenes resultantes. En este estudio usamos métrica de distancia para calcular la similitud entre la imagen objetivo y la resultantes. La medida de comparación usa los histograma de las imágenes. Estos histogramas se comparan calcu-

4.2 Resultados de la comparación entre enfoques

lando la distancia entre ellos. Aquí probamos cuatro diferentes distancias de histogramas: Euclidiana (d_{L_2}), Bhattacharyya (d_B), chi-square (d_{chi-s}) y la de intersección (d_{\cap}).

$$d_{L_2}(h_o, h_t) = \sqrt{\sum_j (h_o(j) - h_t(j))^2}, \quad (4.1)$$

$$d_B(h_o, h_t) = \sqrt{1 - \frac{1}{\sqrt{\mu_{h_o}\mu_{h_t}N^2}} \sum_j \sqrt{h_o(j) \cdot h_t(j)}}, \quad (4.2)$$

$$d_{chi-s}(h_o, h_t) = \sum_j \frac{(h_o(j) - h_t(j))^2}{h_o(j)}, \quad (4.3)$$

$$d_{\cap}(h_o, h_t) = \sum_j \min(h_o(j), h_t(j)), \quad (4.4)$$

donde h_o y h_t son los histogramas de color normalizados de la imagen de salida y de la referencia, respectivamente. Para las distancias Euclidiana, Bhattacharyya y chi-square (d_{L_2} , d_B , d_{chi-s}), un valor pequeño corresponde a una mayor similitud, es decir, una mejor transferencia de color. A diferencia de la distancia de intersección (d_{\cap}), donde a mayor sea la distancia, mejor es la transferencia de color.

4.2. Resultados de la comparación entre enfoques

Las pruebas fueron realizadas usando diferentes enfoques para el mejoramiento de una imagen oscura dada como entrada a nuestro sistema. Por un lado y con propósitos comparativos, se usaron tres métodos para mejorar la entrada sin la necesidad de ninguna imagen de referencia. Estos métodos son el algoritmo White-Patch (WP), el algoritmo Gray-World (GW) y la Igualación de Histograma (HE). Por otro lado, se obtienen resultados usando la transferencia de color sobre la misma entrada a partir una imagen objetivo. Este último proceso se efectúa en cada uno de los espacios de color perceptuales en discusión: $l\alpha\beta$, CIELUV, CIELAB y RLAB. Adicionalmente se incluye el espacio RGB con propósitos comparativos. Ejemplos de los resultados obtenidos por este procedimientos se pueden apreciar en la Figura 4.3.

Según nuestro conocimiento, no hay bases de datos para las pruebas particulares que hacemos sobre imágenes oscuras. Debido a esto, creamos una colección de imágenes oscuras tomadas por una cámara simple y común (celular, digital, reflex, webcam, etc.). La base de datos de imágenes usada consiste de una colección de 100 imágenes oscuras. Un subconjunto de 60 imágenes fueron tomadas en la noche. Las 40 imágenes restantes se tomaron en el día y después se oscurecieron artificialmente. Adicionalmente, el conjunto de imágenes contiene 20 imágenes objetivo, tomadas en condiciones de iluminación natural. Muchas de ellas corresponden a las escenas tomadas en la noche.

Se realizaron una serie de pruebas midiendo la distancia entre dos histogramas. Un

4.2 Resultados de la comparación entre enfoques

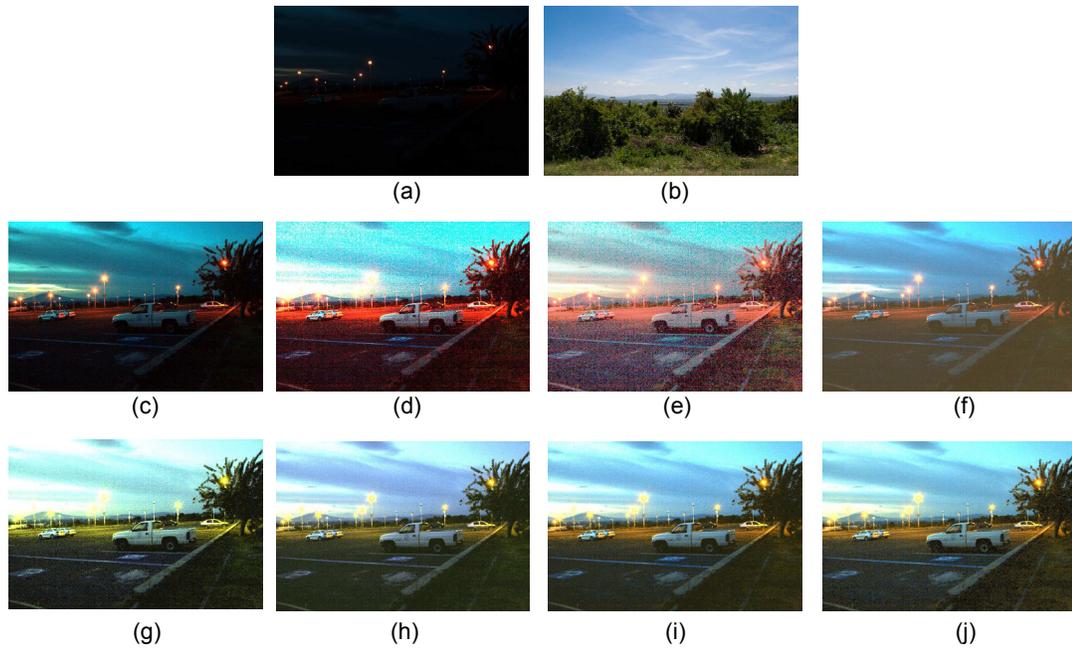


Figura 4.3: Una imagen oscura y sus resultantes obtenidas usando diferentes métodos. (a) Entrada no. 92 (paisaje); (b) imagen de referencia no. 19; resultantes usando (c) WP, (d) GW, y (e) HE; resultantes usando (b) como objetivo y (f) CT en RGB, (g) CT en $l\alpha\beta$, (h) CT en CIELUV, (i) CT en CIELAB, y (j) CT en RLAB.

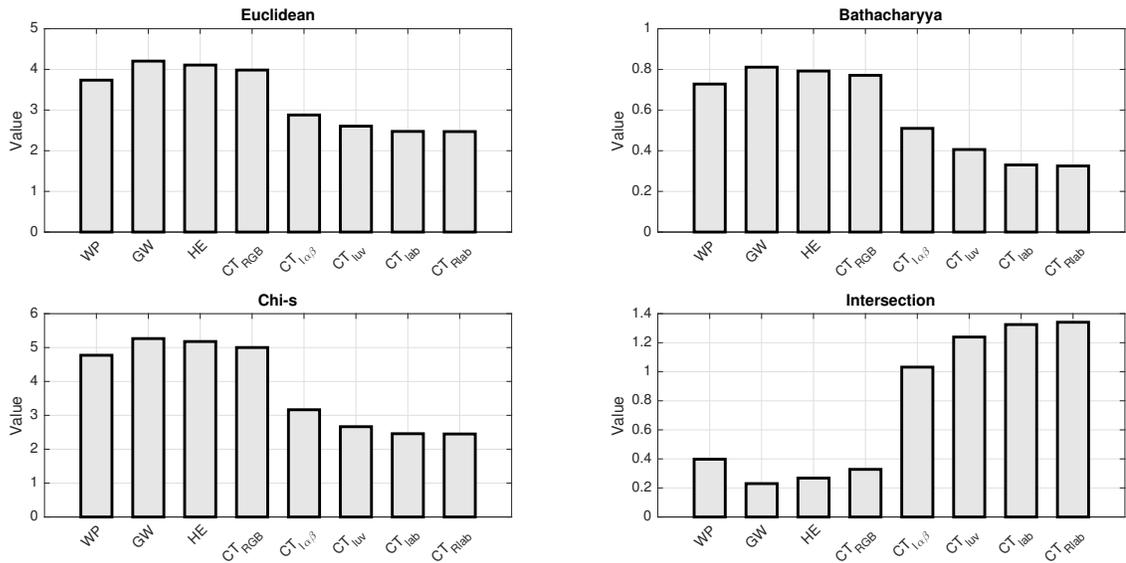


Figura 4.4: Resultados para la Imagen no. 92 (paisaje) mostrado en la Figura 4.3. Las medidas usando una métrica específica para los resultantes de diferentes métodos. Detalles de esos resultados son reportados en la Tabla 4.2.

4.2 Resultados de la comparación entre enfoques

Tabla 4.2: Medidas de distancia entre la imagen objetivo (Figure 4.3b) y los resultantes después de procesar la entrada (Figure 4.3a) usando los diferentes enfoques (Figure 4.3c-j).

Método	Euclidiana	Bhathacharyya	Chi-s	Intersección
WP	3.736	0.728	4.773	0.398
GW	4.204	0.811	5.265	0.230
HE	4.107	0.792	5.178	0.268
CT _{RGB}	3.984	0.771	5.001	0.328
CT _{$l\alpha\beta$}	2.879	0.510	3.165	1.032
CT _{CIELUV}	2.604	0.406	2.665	1.239
CT _{CIELAB}	2.475	0.330	2.458	1.325
CT _{RLAB}	2.471	0.325	2.450	1.341

histograma correspondiente a la imagen objetivo y otros correspondientes a las resultantes. La Tabla 4.2 muestra los resultados cuantitativos, obtenidos de la imagen muestra que podemos observar en la Figura 4.3. Cada celda en la tabla muestra el valor de la comparación entre la resultante y la imagen objetivo dadas en la Figura 4.3b. Cada valor se obtiene usando un método de mejoramiento y una métrica de distancia específica. Esta tabla muestra que los mejores valores se obtienen usando el espacio RLAB. Adicionalmente, se incluye una gráfica con fines de visualización en la Figura 4.4. Esta figura muestra que la transferencia de color en RGB es peor que los resultados obtenidos usando alguno de los espacios perceptuales.

El procedimiento antes mencionado para cada par de entrada-objetivo se reproduce usando diferentes imágenes objetivo. La Figura 4.5 muestra un ejemplo de los resultantes obtenidos por una única imagen de entrada después de aplicar seis imágenes objetivo. La figura muestra algunas imágenes objetivo altamente no similares a la escena de entrada, generando colores falsos en los resultantes. Esta característica podría ser útil en aplicaciones particulares (e.g. generación de efectos artísticos), sin embargo de momento esto no es deseable en este trabajo.

La transferencia de color se aplicó a las 100 imágenes oscuras, usando para cada imagen las 20 objetivo. Las medidas de distancia fueron calculadas para todos los pares de resultante-referencia. Un total de 2000 resultantes se obtienen para cada método de transferencia de color y para cada espacio de color (RGB, $l\alpha\beta$, CIELUV, CIELAB o RLAB). Adicionalmente, las resultantes de los métodos de referencia se comparan con las 20 imágenes objetivo, obteniendo 2000 medidas para cada método de referencia (WP, GW y HE). Posteriormente, el valor promedio de las 2000 medidas se calcula para cada enfoque bajo evaluación. En la Tabla 4.3, las celdas muestran el valor promedio de cada enfoque y para cada medida de distancia.

4.2 Resultados de la comparación entre enfoques

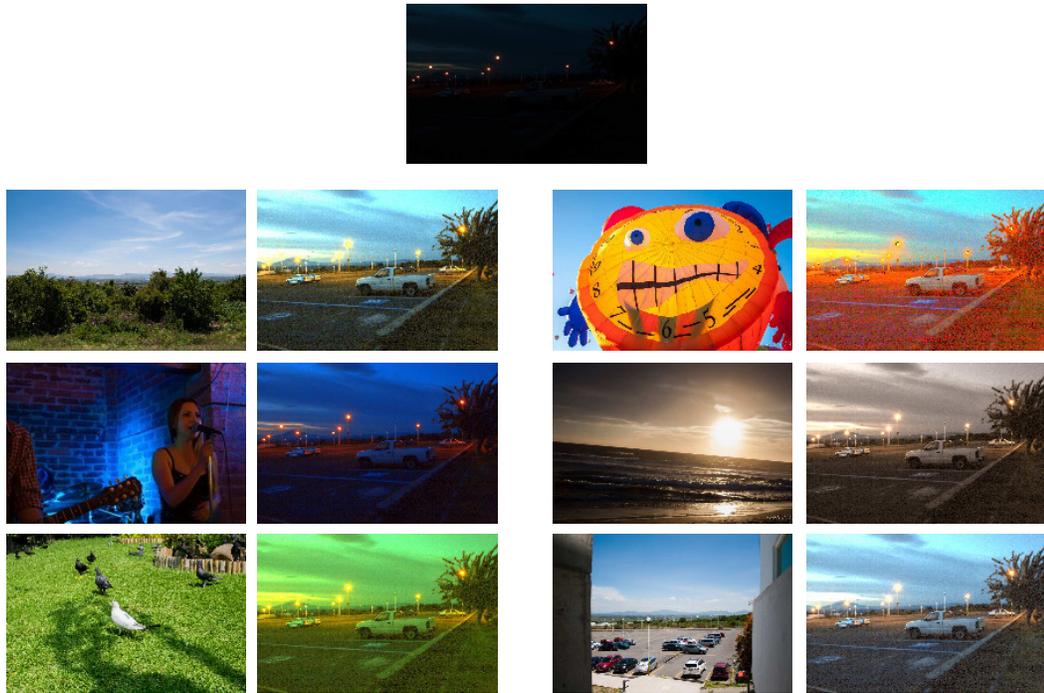


Figura 4.5: Ejemplos de imágenes objetivo usadas en la transferencia de color para la Imagen no. 92, y sus resultados correspondientes.

Hemos encontrado que aplicar la metodología de transferencia de color en un espacio de color perceptual es mejor que aplicarla en el espacio RGB. De todos los métodos,

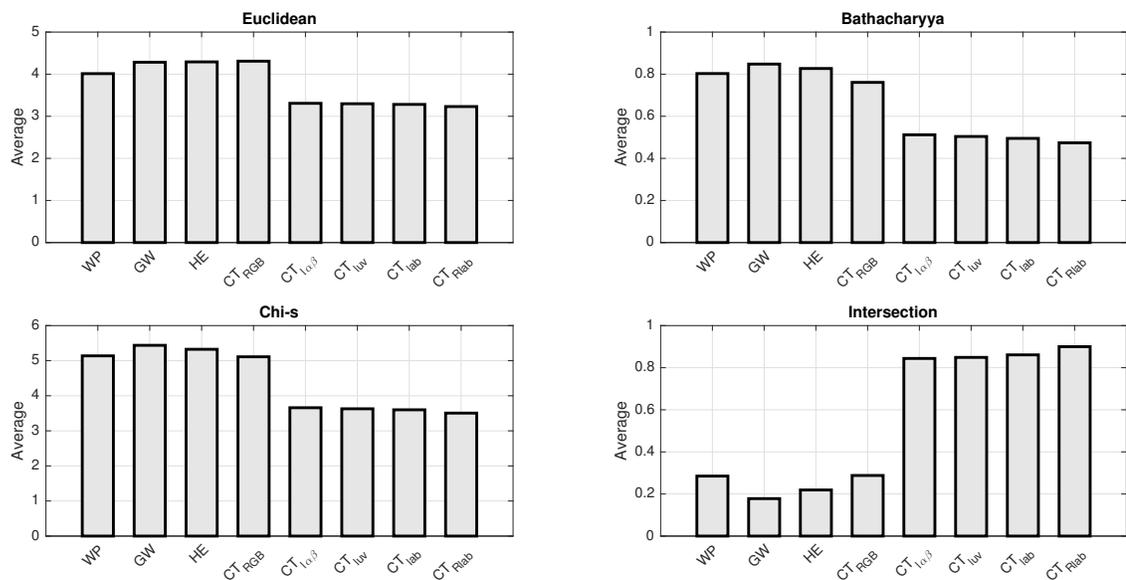


Figura 4.6: Valores promedios de las 2000 resultantes en cada metodología usando las cuatro medidas. Detalles de los resultados son reportados en la Tabla 4.3.

4.2 Resultados de la comparación entre enfoques

Tabla 4.3: Valores promedio de las 2000 medidas de cada método bajo análisis. Se dan los datos para las cuatro métricas de distancia entre los histogramas

Método	Euclidiana	Bhathacharyya	Chi-s	Intersección
WP	4.013	0.803	5.138	0.285
GW	4.284	0.848	5.438	0.178
HE	4.292	0.827	5.325	0.219
CT _{RGB}	4.309	0.761	5.111	0.288
CT _{l$\alpha$$\beta$}	3.308	0.512	3.659	0.844
CT _{CIELUV}	3.298	0.504	3.629	0.849
CT _{CIELAB}	3.282	0.495	3.600	0.861
CT _{RLAB}	3.231	0.474	3.506	0.900

la transferencia de color usando el espacio RLAB obtiene consistentemente los mejores resultados en todas las métricas consideradas. Adicionalmente a estas pruebas, realizamos un estudio de significancia estadística con la prueba z entre los resultados obtenidos por RLAB y CIELAB (que obtiene los segundos mejores resultados). Encontrando que la diferencia entre los resultados es significativa usando un $\alpha = 1$ como nivel de confianza. Podemos concluir que la transferencia de color en el espacio de color RLAB es la mejor opción para el mejoramiento de imágenes oscuras.

Resumen

En este estudio, hemos discutido un enfoque de transferencia de color para transformar las imágenes oscuras en imágenes con una apariencia de luz del día. Este enfoque utiliza una imagen de entrada en RGB, capturada con una cámara común. Nuestra principal aportación es el uso de la transferencia de color en espacios de color perceptuales para aclarar las imágenes oscuras, aprovechando la propiedad de las metodologías de transferencia de color para disminuir la producción de colores falsos. También hemos construido un conjunto de imágenes oscuras como un base de referencia. Los resultados muestran que, en general, la transferencia de color en los espacios perceptuales da mejores resultados que la transferencia en RGB. Además, la transferencia de color aplicada en el RLAB alcanza los mejores resultados. La metodología de transferencia de color en RLAB se puede aplicar para mejorar el reconocimiento y la interpretación de las imágenes de noche, en tareas como la videovigilancia.

Iluminando imágenes de noche para mejorar la segmentación por color

La segmentación por color ha sido un problema central en la visión por computadora y el reconocimiento de patrones por varias décadas. Esta tarea es importante como un paso previo en muchas aplicaciones como es el reconocimiento de objetos, seguimiento, entendimiento de escenas, recuperación de imágenes, entre otras. La segmentación se refiere al proceso de particionar una imagen en múltiples regiones que suponemos corresponden a objetos significativos en la escena [Unnikrishnan et al., 2007]. La partición consiste de asignar una etiqueta de trabajo a cada píxel de una imagen, de tal manera que los píxeles semejantes tengan la misma etiqueta y estén espacialmente conectados [Shapiro y Stockman, 2001]. De acuerdo con Ilea y Whelan [2011], la tarea de segmentación de imágenes se ha desarrollado sobre todo mediante las propiedades de color y textura en un gran número de métodos, incluyendo el crecimiento de regiones, detección de bordes y técnicas de agrupamiento, entre otros. De hecho, la investigación actual se está inclinando hacia la integración de características y métodos.

Un considerable número de enfoques han abordado el problema de la segmentación de imágenes [Fu y Mui, 1981; Pal y Pal, 1993; Cheng et al., 2001; Lucchese y Mitra, 2001; Lizarraga-Morales et al., 2014b,a; Yi y Eramian, 2016]. De acuerdo a Vantaram y Saber [2012], los métodos pueden ser separados en tres grupos principales: (I) guiados espacialmente, (II) ciegos espacialmente, y (III) diversos (misceláneos), donde las primeras dos categorías comprenden la mayoría de los métodos. Por un lado, la idea principal de los métodos guiados es que píxeles vecinos tengan características en común. La meta de estos métodos es aglomerar píxeles adyacentes con estrategias como, por ejemplo, *split-and-merge* [Wu, 1993; Ojala y Pietikäinen, 1999] y crecimiento de regiones [Fan et al., 2001; Wan y Higgins, 2003]. Su principal inconveniente es que, incluso cuando las regiones resultantes de la segmentación son espacialmente bien conectados y compactadas, no hay garantía de que los segmentos sean homogéneos en un espacio de características

específico. Por otro lado, tenemos los algoritmos espacialmente ciegos. Estos métodos suponen que las características en la superficie de un objeto son invariables y, por lo tanto, el objeto se puede representar como un conjunto de puntos en un espacio de características dado. Debido a su simplicidad y bajo costo computacional, este tipo de métodos se ha adoptado ampliamente en el desarrollo de algoritmos de segmentación. Ejemplos de estos enfoques incluyen la agrupación [Comaniciu y Meer, 2002; Kanungo et al., 2002; Mignotte, 2008; Chen et al., 2005] y aproximaciones basadas en histogramas [Kurugollu et al., 2001; Cheng et al., 2002].

El uso de información por color y textura, colectivamente, tiene fuertes lazos con la percepción humana, y en varios escenarios el uso de color o textura, independientes, no es suficiente para describir el contenido de una imagen. La segmentación de imágenes naturales ejemplifica este problema, principalmente porque ellas no presentan homogeneidades en color y textura [Ilea y Whelan, 2011; Mobahi et al., 2011]. Además, las escenas usualmente son complejas, con cierto grado de aleatoriedad e irregularidad. De igual manera, la influencia de la textura y el color puede variar considerablemente de imagen a imagen cuando se incorporan distorsiones debido a la iluminación, cambios de escala, y otras fuentes.

En general, la mayoría de los métodos han sido exitosos en un amplio rango de condiciones. Ya que han podido segmentar rostros humanos, muebles, animales, arboles, edificios, etc., tanto en escenarios de interiores como de exteriores. Sin embargo, una de las condiciones más retadoras es la situación de baja iluminación, como en imágenes de noche.

En este capítulo, proponemos utilizar los métodos de mejoramiento anteriormente revisados, para iluminar imágenes nocturnas antes de aplicarles un proceso de segmentación. La idea consiste en mostrar que se produce una mejor segmentación sobre una imagen mejorada, que sobre aquella que sea oscura o de noche. Para esto, se utiliza el conjunto de datos de Berkeley (BSDS300) Martin et al. [2001] nuevamente. Sin embargo, en esta ocasión, a toda la base de datos se le aplicará una etapa de oscurecimiento artificial para emular un efecto de noche. Los resultados se evalúan utilizando el índice NPR propuesto en [Unnikrishnan et al., 2005, 2007]. Esta medida mide la calidad de la imagen segmentada. Las segmentaciones se obtienen de un algoritmo dado, y se comparan con una o más referencias hechas por personas. En este estudio, se comparan los resultados obtenidos por cuatro métodos de segmentación. Cada método se aplica sobre la imagen resultante por cada método de mejoramiento, sobre la imagen oscura y sobre la imagen original (sin ser oscurecida). Con los resultados obtenidos intentamos mostrar que la aplicación de métodos de segmentación al aplicarse después de un pre-procesamiento de mejora, se obtiene una calidad de segmentación similar a aquella realizada sobre las imágenes originales.

5.1. Marco de trabajo

En este trabajo se usan cuatro métodos de mejoramiento de imágenes: White-Patch (WP), Gray-World (GW), Igualación de Histograma (HE) y Transferencia de Color (CT). Los algoritmos de constancia de color han demostrado ser exitosos en el mejoramiento de imágenes oscuras, como se revisó en el Capítulo 3. Por otra parte, la transferencia de color aplicada en el espacio RLAB también proporcionó buenos resultados en el Capítulo 4, y es por eso que adoptamos esta técnica como método de mejora. Es importante señalar, que para la realización de la transferencia se usa una paleta de colores estándar usada para calibración y aplicada en el espacio RLAB. La Figura 5.1 muestra el diagrama de la metodología empleada en este estudio.

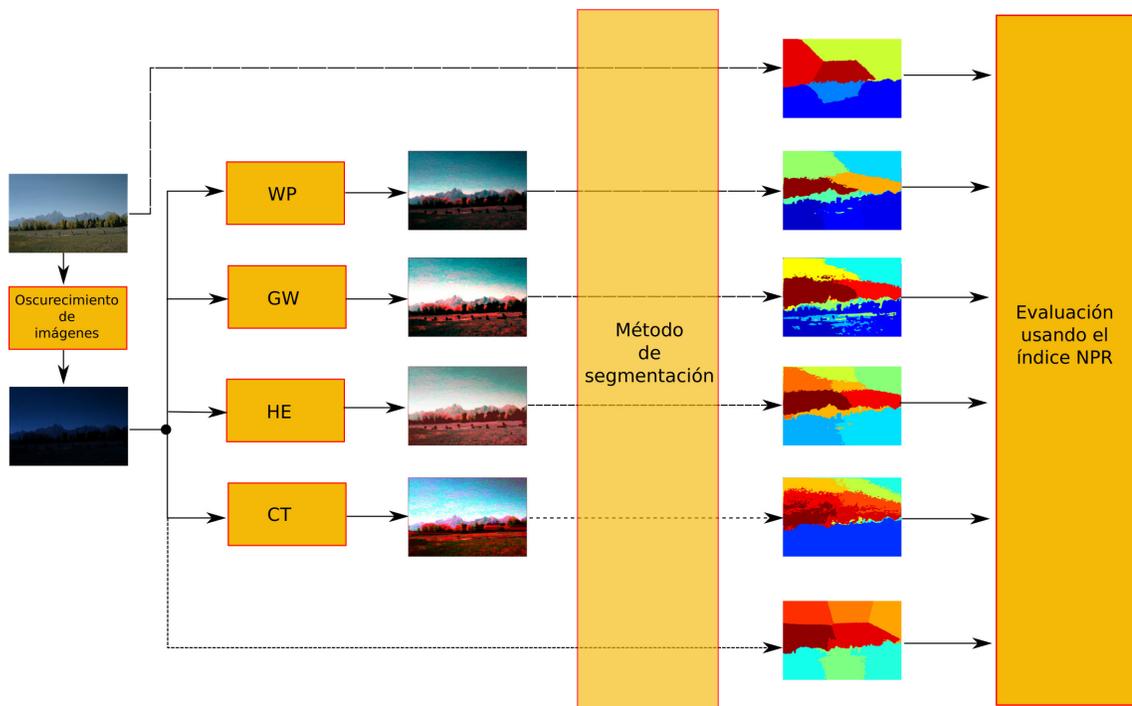


Figura 5.1: Diagrama a bloques de la metodología seguida para la evaluación de la segmentación por un método sobre diferentes enfoques.

Primeramente, una imagen natural, tomada de una base de datos para segmentación, debe convertirse a una escena nocturna. Más tarde, se le aplicarán los cuatro métodos de mejoramiento, obteniendo así cuatro imágenes resultantes. A cada uno de éstas, se les aplicará un método de segmentación y después se mide la segmentación realizada (con respecto a la verdad de referencia). El objetivo del estudio es comparar las segmentaciones en cada imagen resultante y concluir que efectivamente es mejor aplicar un pre-procesamiento de mejora a las imágenes oscuras antes de realizar una segmentación. Además, de que la calidad de las segmentaciones son similares a aquellas realizadas

sobre las imágenes originales.

5.1.1. Creación de base de datos nocturna para segmentación

Según nuestro conocimiento, no hay bases de datos para segmentación de imágenes nocturnas u oscuras. Debido a esto, en este estudio usamos una metodología para transformar imágenes naturales en escenas nocturna. La base de datos que oscurecemos es la de Berkeley (BSDS300) [Martin et al., 2001].

El marco de trabajo propuesto por Thompson et al. [2002] emula la pérdida de detalle y efectos de ruido asociados por la visión nocturna. Antes que nada, debemos asumir que las entradas son imágenes en RGB. Si nuestra imagen de entrada tiene un amplio rango dinámico, primero debe de transferirse un tono de noche usando alguna técnica estándar, como por ejemplo la aquella propuesta por Reinhard et al. [2001]. Así, la imagen RGB es mapeada a una escena con luz de noche. Esta imagen debe tener una apariencia con tendencia azulada, ya que debe emular la forma en que percibimos las escenas nocturnas los humanos. Nuestros bastones de los ojos son más sensibles a los tonos azules, que a los rojos y verdes [Thompson et al., 2002]. Posteriormente, se lleva a cabo una etapa de filtrado para simular la pérdida de agudeza visual. Una forma es usando una difusión controlada, la cual tiene el efecto de difuminar la geometría de los bordes sin emborronar a través de ellos. Finalmente, en este estudio aplicamos ruido Gaussiano a las imágenes ya que una imagen nocturna real presenta una gran cantidad de ruido, dependiendo del sensor. La Figura 5.2 muestra tres ejemplos del nuevo conjunto de imágenes, ahora oscurecidas y sus respectivas imágenes originales con luz natural.

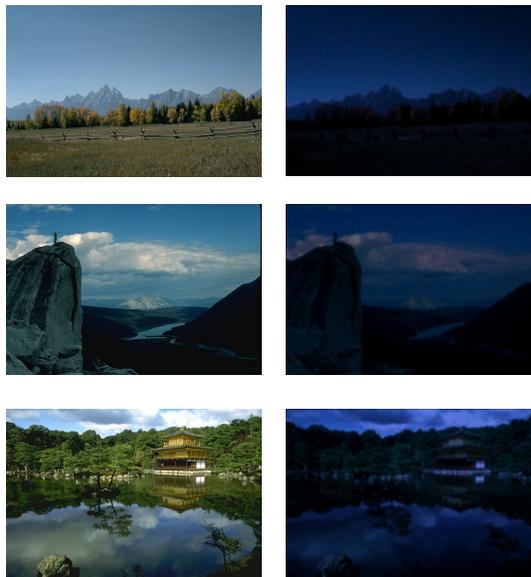


Figura 5.2: Tres ejemplos de la nueva base de datos BSDS300 de noche y sus respectivas imágenes originales con luz natural.

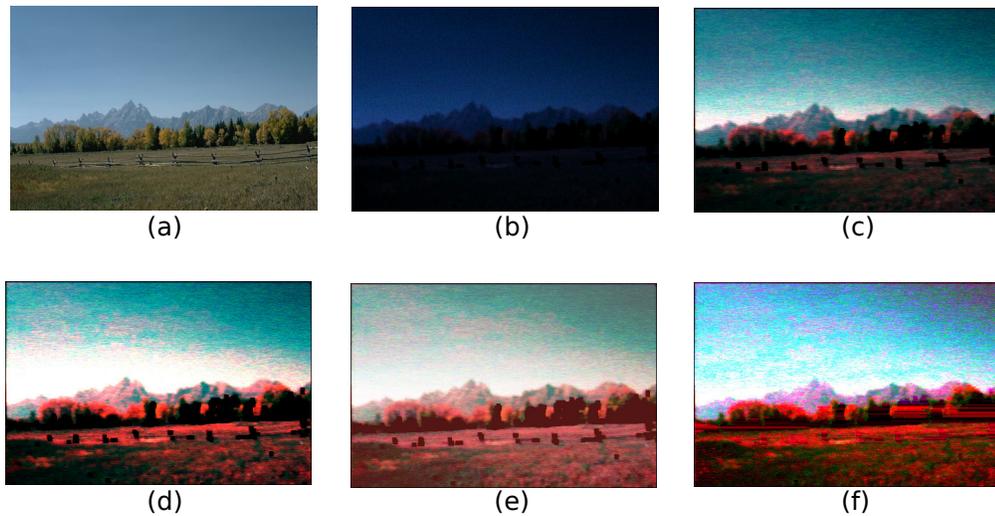


Figura 5.3: Un ejemplo mostrando los resultados por los métodos de mejoramiento. (a) Imagen original, (b) imagen de noche, (c) resultante por WP, (d) resultante por GW, (e) resultante por HE y (f) resultante por CT.

5.1.2. Métodos de mejoramiento en imágenes nocturnas

Los métodos contemplados han sido descritos en los Capítulos 1 y 2 . La Figura 5.3 muestra un ejemplo de la mejora proporcionada por los cuatro métodos, en una imagen oscura. Se incluye la imagen original con propósitos comparativos.

5.2. Pruebas y Resultados

Para evaluar, consideramos solamente una métrica. Esta métrica ha sido ampliamente usada para la evaluación de los enfoques de segmentación; el índice *Normalize Probabilistic Rand (NPR)* [Unnikrishnan et al., 2007]. El índice NPR, tiene cuatro propiedades deseables para la evaluación. La primera es que el método permite comparar segmentaciones que tengan diferente número de regiones. La segunda es que si se presentan casos raros en donde las malas segmentaciones obtengan buenos resultados, este índice detecte esto. En tercer lugar, que es capaz de utilizar múltiples segmentaciones de referencia, para obtener un resultado más objetivo, y es capaz de acoplar las regiones que se encuentran en las referencias. Por último, el valor obtenido del índice NPR permite la comparación entre segmentaciones de diferentes imágenes o entre segmentaciones de diferentes algoritmos.

Nosotros usamos el código proporcionado por Yang et al. [2008] para calcular el índice PR [Unnikrishnan et al., 2005]. Y así, con este valor podemos calcular el índice

5.2 Pruebas y Resultados

NPR usando la Ecuación (5.1):

$$NPR = \frac{PR - E[PR]}{\max[PR] - E[PR]}, \quad (5.1)$$

donde $\max[PR] = 1$, y $E[PR] = 0.6064$, de acuerdo a Vantaram y Saber [2012], para el conjunto de 300 imágenes en la base de datos BSDS. Importante señalar que a mayor el valor del índice NPR, mejor la calidad de la segmentación.

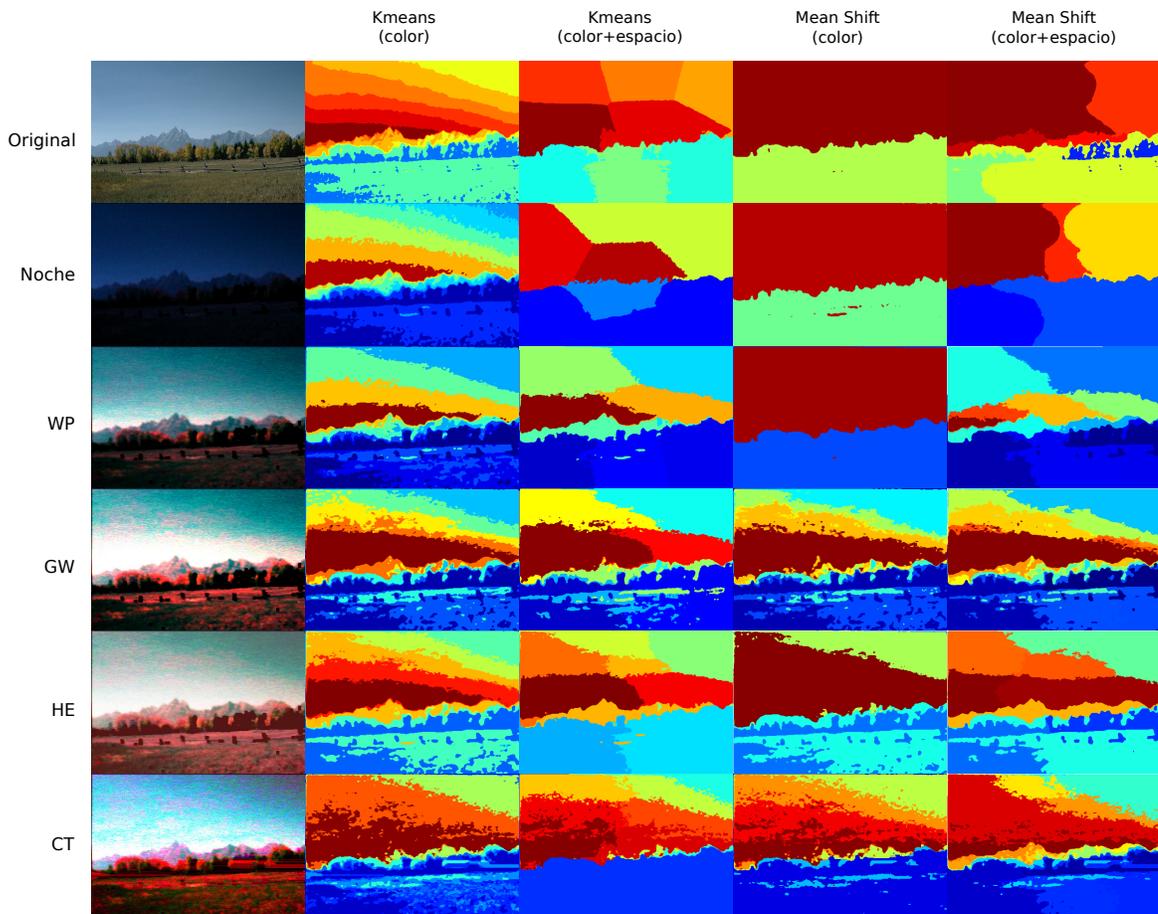


Figura 5.4: Un catálogo mostrando las imágenes señaladas en la Figura 5.3 y sus respectivas segmentaciones por cuatro métodos diferentes.

Los algoritmos de agrupamiento operan adecuadamente en diversos rangos de datos. Esto es, permiten segmentar imágenes oscuras aún si un observador humano sólo percibe una zona “negra” en la escena. En la Figura 5.4 se muestra un catálogo de imágenes mostrando las imágenes señaladas en la Figura 5.3 y sus respectivas segmentaciones por cuatro métodos diferentes. Además, en la Tabla 5.1 se presentan los valores respectivos de la medición por el índice NPR, para cada imagen segmentada mostrada en la Figura 5.4. Para este caso en particular, los valores indican que la mejor segmentación la proporciona el método Mean Shift usando únicamente color, sobre la imagen original.

5.2 Pruebas y Resultados

Tabla 5.1: Valores respectivos de la medición por el índice NPR para las imágenes mostradas en la Figura 5.4.

Imagen	Kmeans (color)	Kmeans (color+espacio)	Mean Shift (color)	Mean Shift (color+ espacio)
Original	0.1875	0.2081	0.5499	0.4723
Noche	0.2057	0.3289	0.5282	0.2915
WP	0.2638	0.2761	0.5167	0.3025
GW	0.2411	0.2617	0.2383	0.2738
HE	0.2081	0.1925	0.3778	0.2673
CT	0.3162	0.1089	0.2147	0.3265

5.2.1. Pruebas estadísticas

Para todas las pruebas realizadas, los resultados de la segmentación obtenidos mediante el índice NPR se promediaron. Es decir, cada método de pre-procesamiento proporciona un conjunto de 300 resultantes. Y a cada uno de esos conjuntos se les aplica un método de segmentación. Al final, tendremos un conjunto de 300 imágenes segmentadas, correspondiente a un algoritmo de mejora. La evaluación se hace para cada conjunto promediando el índice NPR obtenido en las 300 resultantes del mismo. Cada promedio es comparado con cada uno de los promedios de los seis conjuntos de imágenes mejoradas por los métodos utilizados, con el fin de determinar sus diferencias en la calidad. Sin embargo, una simple comparación del promedio no es suficiente para proporcionar una conclusión a partir de estas comparaciones. En los casos en que los resultados son muy cercanos, la diferencia puede no ser significativa. Para identificar tales casos, pruebas de significancia estadística han de llevarse a cabo para todas las comparaciones realizadas. En este estudio, utilizamos una prueba-z de dos muestras, donde se asumen diferentes variaciones. Nuestra hipótesis nula es que los resultados promedio de dos conjuntos de pruebas son los mismos. Por otro lado, nuestra hipótesis alternativa es que los resultados promedio de dos conjuntos de prueba son en realidad diferentes.

Tabla 5.2: Resultados del índice NPR usando el método de segmentación Kmeans por color.

Conjunto de imágenes	i	μ_i	σ_i	pruebas-z	
Original	1	0.2246	0.3367	$\mu_2 = \mu_1$	
Noche	2	0.1922	0.3291	$\mu_1 = \mu_2$	
WP	3	0.2099	0.3374	$\mu_2 = \mu_3$	$\mu_1 = \mu_3$
GW	4	0.2171	0.3437	$\mu_2 = \mu_4$	$\mu_1 = \mu_4$
HE	5	0.2054	0.3530	$\mu_2 = \mu_5$	$\mu_1 = \mu_5$
CT	6	0.2091	0.3423	$\mu_2 = \mu_6$	$\mu_1 = \mu_6$

5.2 Pruebas y Resultados

Las Tablas (5.2-5.5) muestran la siguiente información. Para cada conjunto de imágenes segmentadas, es decir, tanto el original, el de noche, y los de las resultantes por los métodos de mejora, se presentan su valores promedio y sus desviaciones estandar. Además, el estudio de la significancia estadística de cada uno con respecto el de noche (μ_2, μ_i), y el de cada uno con respecto al original (μ_1, μ_i). De acuerdo a los resultados obtenidos con el método Kmeans usando solo color, no hay diferencia alguna entre un enfoque u otro. Ya que la validación estadística indica que la diferencia entre los resultados no es significativa.

Tabla 5.3: Resultados del índice NPR usando el método de segmentación Kmeans por color e información espacial.

Conjunto de imágenes	i	μ_i	σ_i	pruebas-z
Original	1	0.2847	0.3509	$\mu_2 = \mu_1$
Noche	2	0.2517	0.3315	$\mu_1 = \mu_2$
WP	3	0.2762	0.3579	$\mu_2 = \mu_3$ $\mu_1 = \mu_3$
GW	4	0.2715	0.3674	$\mu_2 = \mu_4$ $\mu_1 = \mu_4$
HE	5	0.2756	0.3721	$\mu_2 = \mu_5$ $\mu_1 = \mu_5$
CT	6	0.2758	0.3679	$\mu_2 = \mu_6$ $\mu_1 = \mu_6$

Con base en los resultados obtenidos con el método Kmeans usando color e información espacial, tampoco hay diferencia alguna entre un enfoque u otro. La validación estadística indica nuevamente que la diferencia entre los resultados no es significativa.

Tabla 5.4: Resultados del índice NPR usando el método de segmentación Mean shift por color.

Conjunto de imágenes	i	μ_i	σ_i	pruebas-z
Original	1	-0.1529	0.5170	$\mu_2 \neq \mu_1$
Noche	2	-0.6026	0.5359	$\mu_1 \neq \mu_2$
WP	3	-0.1565	0.5162	$\mu_2 \neq \mu_3$ $\mu_1 = \mu_3$
GW	4	0.0995	0.3293	$\mu_2 \neq \mu_4$ $\mu_1 \neq \mu_4$
HE	5	0.1173	0.2805	$\mu_2 \neq \mu_5$ $\mu_1 \neq \mu_5$
CT	6	0.1020	0.2795	$\mu_2 \neq \mu_6$ $\mu_1 \neq \mu_6$

Según los resultados obtenidos con el método Mean Shift usando sólo color, todas las hipótesis se rechazan. Es decir, la validación estadística indica que la diferencia entre los resultados es significativa.

Finalmente, los resultados obtenidos en el método Mean Shift usando color e información espacial, las hipótesis se rechazan entre la comparación con las segmentaciones dadas por las originales, las de WP y las de HE. Es decir, la validación estadística indica que la diferencia entre esos resultados es significativa. Lo cual nos dice, que sólo el WP

Tabla 5.5: Resultados del índice NPR usando el método de segmentación Mean Shift por color e información espacial.

Conjunto de imágenes	i	μ_i	σ_i	pruebas-z	
Original	1	0.3490	0.3249	$\mu_2 \neq \mu_1$	
Noche	2	0.2538	0.3065		$\mu_1 \neq \mu_2$
WP	3	0.2967	0.3357	$\mu_2 \neq \mu_3$	$\mu_1 = \mu_3$
GW	4	0.2793	0.3545	$\mu_2 = \mu_4$	$\mu_1 \neq \mu_4$
HE	5	0.3056	0.3742	$\mu_2 \neq \mu_5$	$\mu_1 = \mu_5$
CT	6	0.2506	0.3407	$\mu_2 = \mu_6$	$\mu_1 \neq \mu_6$

y el HE funcionan para mejorar la segmentación sobre imágenes de noche usando este enfoque de segmentación.

Los resultados obtenidos y la validación realizada nos permiten llegar a las siguientes conclusiones. Los algoritmos de mejoramiento de imágenes no ayudan si la segmentación se hace con Kmeans en cualquiera de sus dos variantes. En cambio con el método Mean Shift, se aprecia que cuando se realiza la segmentación en imágenes mejoradas y con luz natural, se obtienen mejores resultados al aplicar uno de los algoritmos de constancia de color y la validación estadística lo ratifica.

Adicionalmente, se hicieron pruebas estadísticas entre los resultados por la segmentación realizada en las imágenes originales y entre las segmentaciones sobre las imágenes resultantes por los algoritmos de mejora. Las pruebas indicaron que la diferencia de los resultados no es significativa, ya que las medias son consideradas iguales. Por lo cual, al aplicar un pre-procesamiento a una imagen nocturna, específicamente con el WP, es posible obtener una segmentación equivalente a la que se obtendría a partir de la escena diurna.

Resumen

En este trabajo, se propuso utilizar métodos de mejoramiento, para iluminar imágenes nocturnas y aplicarles un proceso de segmentación posterior. La idea consistió en mostrar que se produce una mejor segmentación sobre una imagen mejorada, que sobre aquella que sea oscura o de noche. Aunque los métodos de segmentación pueden funcionar bien, incluso en condiciones de baja iluminación, el aplicar un pre-procesamiento a una imagen nocturna es posible obtener una segmentación equivalente a la que se obtendría a partir de la escena diurna. Para esto, se utilizó el conjunto de datos de Berkeley (BSDS300). Primeramente, una imagen natural se convirtió a una escena nocturna. Después, se le aplicaron los cuatro métodos de mejoramiento, obteniendo así cuatro resultantes. A éstas, se les aplicó un método de segmentación y después se midió la exactitud con

5.2 Pruebas y Resultados

la que fue realizada. Los resultados indicaron que cuando se aplica el método Kmeans, no hay diferencia entre aplicar un método de mejora previo a la imagen. Lo cual apunta que los métodos de mejora no ayudan a la segmentación al usar Kmeans. Sin embargo, para el método Mean Shift es mejor aplicar previamente un algoritmo de constancia de color, específicamente el WP, para mejorar los resultados de la segmentación. Lo que ahora apunta a que los métodos de mejora sí ayudan a la segmentación pero no podemos definir alguno de manera general, y habría que realizar experimentos con otros métodos de segmentación. También concluimos que para el Mean Shift los resultados nos dicen que al aplicar un pre-procesamiento a una imagen nocturna es posible obtener una segmentación equivalente a la que se obtendría a partir de la escena diurna.

Conclusiones y productos generados

En el Capítulo 1 se establecieron los conceptos y temas principales que abordan el problema de la visión nocturna y el procesamiento de imágenes de noche en esta tesis. La constancia de color, la transferencia de color y los sistemas basados en reglas son herramientas que nos ayudan a la solución de los problemas que surgen en la visión nocturna.

En el Capítulo 2, se propuso una variante de un método utilizando la hipótesis de GW para la constancia del color en los espacios de color CIELAB y CIELUV. Dada una imagen de prueba, se estimó el iluminante y después se comparó con un valor de referencia. Las pruebas experimentales se realizaron comparando los resultados de los métodos propuestos con los obtenidos utilizando el GW estándar en RGB y el algoritmo de GE. Según las pruebas, se concluyó que los resultados del GW en un espacio de color perceptual, son mejores que los obtenidos utilizando el procedimiento estándar en RGB. A pesar de que los resultados utilizando el algoritmo de GE fueron ligeramente mejores que los de nuestro enfoque, para aplicaciones prácticas se puede optar por el enfoque propuesto, ya que es mucho más rápido y no requiere un proceso de sintonía. También, se pudo apreciar que el GW en CIELUV es marginalmente mejor que en CIELAB de acuerdo a la precisión de la fuente luminosa estimada. Por otra parte, el tiempo de procesamiento es considerablemente menor en CIELUV. Por estas razones, el enfoque propuesto, en particular el uso del GW en el espacio CIELUV, puede considerarse un método de pre-procesamiento adecuado para aplicaciones en tiempo real en los campos del reconocimiento de patrones y visión por computadora.

En el Capítulo 3, fue introducido un sistema experto que selecciona algoritmos de constancia de color. Esta tarea está orientada a la selección entre tres algoritmos que mejoran el contenido cromático de imágenes oscuras: el White-Patch, el Gray-World, y el Gray-Edge. Estos algoritmos han sido ampliamente usados en tareas de constancia de color gracias a sus buenos resultados. Aunque también han mostrado que son muy útiles en el mejoramiento de imágenes, especialmente en imágenes bajo condiciones de baja iluminación.

El diseño de un sistema automático de selección no es una tarea trivial cuando

muchas características están involucradas en la selección. Es por esto que se desarrolló un sistema basado en reglas difusas, con la intención de modelar la información a través de reglas simples. Una ventaja del enfoque es que solamente se usan doce características, lo cual permite una mejor interpretación del problema. Y pensamos que estas características son muy importantes para la tasa de selección alcanzada, ya que se probaron otras y los resultados no fueron buenos. Los experimentos se realizaron de manera separada usando inferencias de Mamdani y Larsen. Específicamente, el modelo de la inferencia de Larsen proporcionó una tasa de selección sobresaliente para esta tarea.

La meta principal de este capítulo ha sido el desarrollo de un sistema selector basado en reglas difusas para una tarea novedosa y práctica. El enfoque propuesto es una tarea muy adecuada para resolver dos problemas al mismo tiempo: la constancia de color y el mejoramiento de la imagen. Especialmente para imágenes con bajas condiciones de iluminación. El marco de trabajo es fácil de replicar para posibles tareas posteriores.

En el Capítulo 4, se propuso un enfoque de transferencia de color para transformar las imágenes oscuras en imágenes con una apariencia de luz del día. Este enfoque utiliza una imagen de entrada en RGB, capturada con una cámara común. La principal aportación es el uso de la transferencia de color en espacios de color perceptuales para aclarar las imágenes oscuras, aprovechando la propiedad de las metodologías de transferencia de color para disminuir la producción de colores falsos. También se ha propuesto un conjunto de imágenes oscuras como un base de referencia. Los resultados muestran que, en general, la transferencia de color en los espacios perceptuales da mejores resultados que la transferencia en RGB. Además, la transferencia de color aplicada en el espacio RLAB alcanza los mejores resultados. La metodología de transferencia de color en RLAB se puede aplicar para mejorar el reconocimiento y la interpretación de las imágenes de noche, en tareas como la video-vigilancia.

En el Capítulo 5 se propuso utilizar métodos de mejoramiento para iluminar imágenes nocturnas y aplicarles un proceso de segmentación posterior. La idea consistió en demostrar que se produce una mejor segmentación sobre una imagen iluminada, que sobre aquella que sea oscura o de noche. Aunque los métodos de segmentación pueden funcionar bien, incluso en condiciones de baja iluminación, el aplicar un pre-procesamiento de mejora a una imagen nocturna es posible obtener una segmentación equivalente a la que se obtendría a partir de la escena diurna. Para esto, se utilizó el conjunto de datos de Berkeley (BSDS300). Primeramente, una imagen natural se convirtió a una escena nocturna. Después, se le aplicaron los cuatro métodos de mejoramiento, obteniendo así cuatro resultantes. A éstas, se les aplicó un método de segmentación y después se midió la exactitud con la que fue realizada. Los resultados indicaron que cuando se aplica el método Kmeans, no hay diferencia entre aplicar un método de mejora previo a la imagen. Sin embargo, para el método Mean Shift es mejor aplicar previamente un algoritmo

de constancia de color para mejorar los resultados de la segmentación. También concluimos que para este método los resultados nos dicen que al aplicar un pre-procesamiento a una imagen nocturna es posible obtener una segmentación equivalente a la que se obtendría a partir de la escena diurna. Aunque es necesario realizar experimentos usando otros métodos de segmentación para afirmar totalmente estas conclusiones.

Productos generados

A continuación se enlistan los productos científicos que se lograron durante el desarrollo de esta tesis.

- Cepeda-Negrete J., Sanchez-Yanez R.E., "Gray-World assumption on perceptual color spaces", **Lecture Notes in Computer Science**, 8333, pp. 493-504, (2013).
- Cepeda-Negrete J., Sanchez-Yanez R.E., "Automatic selection of color constancy algorithms for dark image enhancement by fuzzy rule-based reasoning", **Applied Soft Computing**, 28, March, pp. 1-10, (2015).

Comentarios finales

A lo largo del desarrollo de este trabajo de tesis surgieron distintos retos. Comenzando por el mejoramiento de habilidades de redacción para artículos científicos, como para este manuscrito de tesis. Explorando las diferentes estrategias a seguir para el mejoramiento de imágenes oscuras y al mismo tiempo el uso de espacios perceptuales, fueron surgiendo ideas que concluyeron en distintos estudios con resultados publicados.

En el caso de la constancia de color, además de funcionar muy bien para la adaptación cromática y mejora de la imagen, se propuso la estimación del color de la fuente de iluminación en espacios perceptuales. Al final de este estudio, obtuvimos muy buenos resultados con un enfoque perceptual, tanto en CIELAB como en CIELUV, siendo este último más rápido en ejecución.

Aprovechando el uso de la constancia de color como estrategia de mejoramiento de imágenes, se desarrolló un sistema experto para la selección de estos, para una escena dada. Se considera una gran ventaja el uso de pocas características, ya que así podemos darle una mejor interpretación y además intuitiva a nuestro modelo basado en reglas. Es decir, es más fácil la comprensión de una regla al considerar si la escena tiene baja iluminación, un número de colores grande, una iluminación dominante alta, etc., para la elección de una escena. Caso contrario a clasificadores o selectores basados en cientos o miles de características que sólo representan un número y se pierde ese concepto intuitivo al que pertenece.

Una vez explorados los espacios de color y algunos enfoques que mejoran la calidad cromática de las imágenes, especialmente aquellas oscuras, el proyecto se centró a combinar esto para la obtención de una imagen con colores naturales de día a partir de una nocturna. El uso de la transferencia de color fue nuestra mejor opción, sin embargo la idea no tenía demasiada novedad si se usaba tal cual se propuso en el 2001. Es por ello que llevamos a cabo otro trabajo que nos permitió encontrar el mejor espacio de color, en el que se puede aplicar la transferencia para producir un aspecto más similar a la imagen de referencia usada. La conclusión a partir de los resultados es que usar el espacio RLAB, el cual es una mejora del CIELAB, nos produce la mejor apariencia en una imagen diurna. Esto coincide con la teoría propuesta basada en que CIELAB tiene problemas en imágenes con iluminaciones extremas. Lo que llama la atención es que casi no hay trabajos que hagan uso de este espacio, a pesar de tratarse de una mejora del ampliamente usado CIELAB.

En cuanto a la segmentación de las imágenes oscuras, se puede decir que aún falta trabajo por desarrollar para poder afirmar con seguridad que un algoritmo de mejora aplicado previamente nos permite obtener una mejor segmentación. Especialmente cuando diferentes métodos de segmentación proporcionan resultados distintos. Sin embargo, creemos que la aportación de convertir una base de datos conocida a escenas nocturnas abre las puertas para que otros investigadores aborden temas relacionados con la segmentación nocturna.

Me gustaría enfatizar, que las aportaciones principales de este trabajo de tesis son en varios sentidos. (1) Un estudio sobre las técnicas más convenientes en el mejoramiento de imágenes oscuras. (2) La verificación del espacio más adecuado para utilizar con las técnicas de mejoramiento. (3) La construcción de dos bases de datos de imágenes nocturnas para el mejoramiento y segmentación, respectivamente. (4) Finalmente, la propuesta de usar sistemas basados en reglas inteligentes para una tarea específica. En nuestro caso, para la selección de algoritmos para mejorar una escena oscura.

Ecuaciones de transformación entre espacios de color

A.1. LMS

El espacio de color LMS representa la respuesta de los tres tipos de conos del ojo humano, llamado así por su capacidad de respuesta (sensibilidad) a longitudes de onda largas (L), medias (M) y cortas (S). Es común utilizar el espacio de color LMS cuando se realiza la adaptación cromática (la estimación de la aparición de una muestra bajo un iluminante diferente). Las ecuaciones de transformación se dan a continuación.

Transformación directa

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3811 & 0.5783 & 0.0402 \\ 0.1967 & 0.7244 & 0.0782 \\ 0.0241 & 0.1288 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (\text{A.1})$$

Transformación inversa

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 4.4679 & -3.5873 & 0.1193 \\ -1.2186 & 2.3809 & -0.1624 \\ 0.0497 & -0.2439 & 1.2045 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix}. \quad (\text{A.2})$$

A.2. $l\alpha\beta$

El espacio $l\alpha\beta$ es un conversión a partir del espacio LMS. Este es un espacio descorrelacionado, donde el canal l representa la luminancia y las componentes α y β el color

amarillo-azul y rojo-verde, respectivamente. Las ecuaciones de transformación se dan a continuación.

Transformación directa

Primero, los datos se transforman a un espacio de color intermedio, el LMS. $\mathbf{L} = \log L$, $\mathbf{M} = \log M$ y $\mathbf{S} = \log S$ [Ruderman et al., 1998].

$$\begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} & 0 & 0 \\ 0 & 1/\sqrt{6} & 0 \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{L} \\ \mathbf{M} \\ \mathbf{S} \end{bmatrix}, \quad (\text{A.3})$$

Transformación inversa

$$\begin{bmatrix} \mathbf{L} \\ \mathbf{M} \\ \mathbf{S} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{3}/3 & 0 & 0 \\ 0 & \sqrt{6}/6 & 0 \\ 0 & 0 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix}, \quad (\text{A.4})$$

los valores de los píxeles se calculan como $L = \mathbf{L}^{10}$, $M = \mathbf{M}^{10}$ y $S = \mathbf{S}^{10}$.

A.3. CIECAM97s

CIECAM97s es uno de varios modelos de apariencia del color. Fue desarrollado como un estándar internacional que toma las mejores ideas de entre varios espacios. Este modelo representa el color de estímulo usando la iluminación, croma y matiz. Es un modelo descorrelacionado y las componentes de color están representadas por c_1 y c_2 . Las ecuaciones de transformación se dan a continuación.

Transformación directa

La matriz de transformación para la conversión de LMS a CIECAM97s es la recomendada por Pattanaik et al. [1998].

$$\begin{bmatrix} A \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2.00 & 1.00 & 0.05 \\ 1.00 & -1.09 & 0.09 \\ 0.11 & 0.11 & -0.22 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \quad (\text{A.5})$$

Transformación inversa

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.33 & 0.32 & 0.21 \\ 0.33 & -0.63 & -0.18 \\ 0.33 & -0.16 & -4.53 \end{bmatrix} \begin{bmatrix} A \\ c_1 \\ c_2 \end{bmatrix} \quad (\text{A.6})$$

A.4. HSI

El espacio de color intenta producir una representación más intuitiva del color a partir del espacio RGB. La componente H representa el matiz (tono). La componente S la saturación, que corresponde a la intensidad colorida del tono. Por último, la I corresponde a la intensidad luminosa. Este espacios también es descorrelacionado. Las ecuaciones de transformación se dan a continuación.

Transformación directa

$$\begin{bmatrix} I \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (\text{A.7})$$

$$H = \tan^{-1}(c_2/c_1) \quad (\text{A.8})$$

$$S = \sqrt{c_1^2 + c_2^2} \quad (\text{A.9})$$

Transformación inversa

$$c_1 = S \sin(H) \quad (\text{A.10})$$

$$c_2 = S \cos(H) \quad (\text{A.11})$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 1/\sqrt{6} & 1/\sqrt{2} \\ 1 & 1/\sqrt{6} & -1/2 \\ 1 & -2/\sqrt{6} & 0 \end{bmatrix} \begin{bmatrix} I \\ c_1 \\ c_2 \end{bmatrix} \quad (\text{A.12})$$

A.5. YIQ

YIQ es otro espacio descorrelacionado que tiene una transformación directa a partir del espacio RGB sin pérdida de información. Este modelo fue diseñado para aprovechar una mayor sensibilidad a los cambios de luminancia del sistema visual humano con respecto a cambios en la tonalidad o la saturación. Para YIQ, la componente Y representa la luminancia de un color, mientras que el color se denota por las componentes I y Q [Kumar et al., 2000]. Las ecuaciones de transformación se dan a continuación.

Transformación directa

Transformación recomendada en [Kumar et al., 2000].

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.30 & 0.59 & 0.11 \\ 0.60 & -0.27 & -0.32 \\ 0.21 & -0.52 & 0.31 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (\text{A.13})$$

Transformación inversa

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.00 & 0.96 & 0.62 \\ 1.00 & -0.27 & -0.65 \\ 1.00 & -1.11 & 1.70 \end{bmatrix} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix} \quad (\text{A.14})$$

A.6. YCbCr

Este espacio se utiliza en sistemas de vídeo y fotografía digital. Y' representa a la componente de luminancia mientras que Cb y Cr son las componentes cromáticas. La componente Y' ha sido linealizada mediante la corrección de gamma. Este espacio es otra manera de representar al espacio RGB. Las ecuaciones de transformación se dan a continuación.

Transformación directa

La ecuación usada para la transformación en la recomendada en [ITU-R Rec, 1995].

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.74 & 129.06 & 25.06 \\ -37.95 & -74.50 & 112.44 \\ 112.44 & -94.15 & -18.29 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (\text{A.15})$$

Transformación inversa

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = - \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 1.16 & 0.00 & 25.06 \\ 1.16 & -0.39 & -0.81 \\ 1.16 & 2.01 & 0.00 \end{bmatrix} \begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} \quad (\text{A.16})$$

A.7. XYZ

El espacio de color CIE 1931, es el nombre de uno de los primeros espacios de color definidos matemáticamente. Fue establecido en 1931, por la Comisión Internationale de l'Éclairage (CIE), y sus resultados experimentales fueron incluidos en la especificación CIE RGB. Con él se definieron con precisión los tres colores primarios de la síntesis aditiva de color, a partir de los cuales pueden crearse todos los demás.

En este modelo, Y significa luminosidad, Z es aproximadamente igual al estímulo del azul (conos de onda corta), y X es una mezcla tendiente a la curva de sensibilidad del rojo al verde (conos de onda larga y media). De esta manera, XYZ puede confundirse con las respuestas de los conos en RGB. Sin embargo, en el espacio de color CIE XYZ, los valores de tri-estímulo no equivalen a las respuestas del ojo humano, incluso teniendo en cuenta que X y Z son aproximadamente rojo y azul, cuando realmente deben verse como parámetros derivados de los colores rojo, verde y azul. Las ecuaciones de transformación se dan a continuación.

Transformación directa

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} r \\ g \\ b \end{bmatrix}, \quad (\text{A.17})$$

donde $r, g, b \in [0, 1]$, obtenidos de dividir cada componente R, G, B entre 255,

Transformación inversa

$$\begin{bmatrix} r \\ g \\ b \end{bmatrix} = \begin{bmatrix} 3.2410 & -1.5374 & -0.4986 \\ -0.9692 & 1.8760 & 0.0416 \\ 0.0556 & -0.2040 & 1.0570 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (\text{A.18})$$

A.8. CIELUV

El espacio de color CIE 1976 (L^* , u^* , v^*), comúnmente conocido por como CIELUV, es un espacio de color adoptado por la Comisión Internationale de l'Éclairage en 1976, como una transformación simple del espacio CIE 1931 XYZ, pero que intentaron darle uniformidad perceptual. Se utiliza ampliamente para aplicaciones tales como gráficos por computadora donde hay luces de colores. La componente L representa la luminancia, mientras que u y v son las componentes de color. Las ecuaciones de transformación se dan a continuación.

Transformación directa

$$u' = \frac{4X}{X + 15Y + 3Z} \quad (\text{A.19})$$

$$v' = \frac{9Y}{X + 15Y + 3Z} \quad (\text{A.20})$$

En este trabajo de tesis, se usa el iluminante E ($X_n = 1$, $Y_n = 1$ y $Z_n = 1$) como blanco de referencia.

$$L^* = \begin{cases} (2/\sigma)^3 Y/Y_n & \text{if } Y/Y_n \leq \sigma^3 \\ 116(Y/Y_n)^3 - 16 & \text{otherwise} \end{cases}, \quad (\text{A.21})$$

$$u^* = 13L^*(u' - u'_n), \quad (\text{A.22})$$

$$v^* = 13L^*(v' - v'_n), \quad (\text{A.23})$$

donde $\sigma = 6/29$.

Transformación inversa

$$u' = \frac{u^*}{13L^*} + u'_n \quad (\text{A.24})$$

$$v' = \frac{v^*}{13L^*} + v'_n \quad (\text{A.25})$$

$$Y = \begin{cases} Y_n L^* (\sigma/2)^3 & \text{if } L^* \leq 8 \\ Y_n \left(\frac{L^* + 16}{116} \right)^3 & \text{otherwise} \end{cases}, \quad (\text{A.26})$$

$$X = Y \left(\frac{9u'}{4v'} \right), \quad (\text{A.27})$$

$$Z = Y \left(\frac{12 - 3u' - 20v'}{4v'} \right). \quad (\text{A.28})$$

A.9. CIELAB

CIELAB (estrictamente CIE 1976 L^*, a^*, b^*) Es un espacio derivado del espacio CIE 1931 XYZ. Sin embargo, CIELAB se calcula usando raíces cúbicas. Se recomienda utilizar CIELAB para nuevas aplicaciones. La componente L representa la luminancia, mientras que u y v son las componentes de color. Las ecuaciones de transformación se dan a continuación.

Transformación directa

$$L^* = 116f(Y/Y_n) - 16, \quad (\text{A.29})$$

$$a^* = 500 [f(X/X_n) - f(Y/Y_n)], \quad (\text{A.30})$$

$$b^* = 200 [f(Y/Y_n) - f(Z/Z_n)], \quad (\text{A.31})$$

$$f(t) = \begin{cases} t^{1/3} & \text{if } t > \sigma^3 \\ t/(3\sigma^2) + 16/116 & \text{otherwise} \end{cases}, \quad (\text{A.32})$$

donde t puede ser X/X_n , Y/Y_n or Z/Z_n , y $\sigma = 6/29$.

$$f_Y = (L^* + 16)/166, \quad (\text{A.33})$$

$$f_X = f_Y + (a^*/500), \quad (\text{A.34})$$

$$f_Z = f_Y - (b^*/200). \quad (\text{A.35})$$

Transformación inversa

$$X = \begin{cases} X_n f_X^3 & \text{if } f_X > \sigma \\ f_X - 16/116 & \text{otherwise} \end{cases}, \quad (\text{A.36})$$

$$Y = \begin{cases} Y_n f_Y^3 & \text{if } f_Y > \sigma \\ f_Y - 16/116 & \text{otherwise} \end{cases}, \quad (\text{A.37})$$

$$Z = \begin{cases} Z_n f_Z^3 & \text{if } f_Z > \sigma \\ f_Z - 16/116 & \text{otherwise} \end{cases}. \quad (\text{A.38})$$

A.10. RLAB

El espacio de color RLAB es una mejora del CIELAB. El RLAB fue diseñado con el fin de arreglar los problemas mostrados por el CIELAB, que se presentaban en condi-

ciones de iluminación inusuales [Fairchild, 1996]. Este espacio mantiene las propiedades perceptuales bajo condiciones de luz normales (luz natural) y también bajo condiciones extremas.

Transformación directa

$$\mathbf{RAM} = \begin{bmatrix} 1.0020 & -0.0401 & 0.0084 \\ -0.0042 & 0.9666 & 0.0008 \\ 0.0000 & 0.0000 & 0.9110 \end{bmatrix}, \quad (\text{A.39})$$

$$\begin{bmatrix} X_{ref} \\ Y_{ref} \\ Z_{ref} \end{bmatrix} = \mathbf{RAM} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \quad (\text{A.40})$$

$$L^R = 100(Y_{ref})^\sigma, \quad (\text{A.41})$$

$$a^R = 430[(X_{ref})^\sigma - (Y_{ref})^\sigma], \quad (\text{A.42})$$

$$b^R = 170[(Y_{ref})^\sigma - (Z_{ref})^\sigma]. \quad (\text{A.43})$$

En esta tesis, se usa $\sigma = 1/3.5$. Este valor es sugerido para imágenes bajo condiciones de iluminación muy bajas [Fairchild, 1996].

Transformación inversa

$$Y_{ref} = \left(\frac{L^R}{100} \right)^{1/\sigma}, \quad (\text{A.44})$$

$$X_{ref} = \left[\left(\frac{a^R}{430} \right) + (Y_{ref})^\sigma \right]^{1/\sigma}, \quad (\text{A.45})$$

$$Z_{ref} = \left[(Y_{ref})^\sigma - \left(\frac{b^R}{170} \right) \right]^{1/\sigma}, \quad (\text{A.46})$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = (\mathbf{RAM})^{-1} \begin{bmatrix} X_{ref} \\ Y_{ref} \\ Z_{ref} \end{bmatrix}. \quad (\text{A.47})$$

Bibliografía

- A. Abadpour y S. Kasaei. An efficient pca-based color transfer method. *Journal of Visual Communication and Image Representation*, 18(1):15–34, 2007. 6
- V. Agarwal, B. R. Abidi, A. Koshan, y M. A. Abidi. An overview of color constancy algorithms. *Journal of Pattern Recognition Research*, 1:42–54, 2006. 3, 4, 13
- K. Barnard, L. Martin, B. Funt, y A. Coath. A data set for colour research. *Color research and applications*, 27(3):148–152, 2002. URL <http://kobus.ca/research/data/>. 25
- J. Bernard. Use of a rule-based system for process control. *IEEE Control Systems Magazine*, 8(5):3–13, 1988. 10
- S. Bianco, G. Ciocca, C. Cusano, y R. Schettini. Improving color constancy using indoor-outdoor image classification. *IEEE Transactions on Image Processing*, 17(12):2381–2392, 2008. 3, 20, 21, 24, 28, 29, 30
- S. Bianco, G. Ciocca, C. Cusano, y R. Schettini. Automatic color constancy algorithm selection and combination. *Pattern Recognition*, 43(3):695–705, 2010. 22, 24, 28, 30, 31, 40
- D. H. Brainard y W. T. Freeman. Bayesian color constancy. *Journal of the Optical Society of America A*, 14(7):1393–1411, 1997. 4
- G. Buchsbaum. A spatial processor model for object colour perception. *Journal of The Franklin Institute*, 310:1–26, 1980. 4, 16
- J. Cepeda-Negrete y R. E. Sanchez-Yanez. Experiments on the white patch retinex in RGB and CIELAB color spaces. *Acta Universitaria*, 22(NE-1):21–26, 2012. 18
- J. Cepeda-Negrete y R. E. Sanchez-Yanez. Gray-world assumption on perceptual color spaces. *Lecture Notes in Computer Science (Image and Video Technology)*, 8333:493–504, 2014. 23

- J. Cepeda-Negrete y R. E. Sanchez-Yanez. Automatic selection of color constancy algorithms for dark image enhancement by fuzzy rule-based reasoning. *Applied Soft Computing*, 28:1–10, 2015. 2, 41
- Y.-S. Chen, B.-T. Chen, y W.-H. Hsu. Efficient fuzzy c-means clustering for image data. *Journal of Electronic Imaging*, 14(1):013017–013017, 2005. 53
- H.-D. Cheng, X. Jiang, Y. Sun, y J. Wang. Color image segmentation: advances and prospects. *Pattern recognition*, 34(12):2259–2281, 2001. 52
- H.-D. Cheng, X. Jiang, y J. Wang. Color image segmentation based on homogram thresholding and region merging. *Pattern recognition*, 35(2):373–393, 2002. 53
- F. Ciurea y B. Funt. A large image database for color constancy research. In *Proceedings of the IS&T/SID Eleventh Color Imaging Conference*, pages 160–164, 2003. 20
- D. Cohen-Or, O. Sorkine, R. Gal, T. Leyvand, y Y.-Q. Xu. Color harmonization. *ACM Transactions on Graphics (TOG)*, 25(3):624–630, 2006. 6
- D. Comaniciu y P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619, 2002. 53
- O. Cordon, M. J. del Jesus, y F. Herrera. A proposal on reasoning methods in fuzzy rule-based classification systems. *International Journal of Approximate Reasoning*, 20(1):21–45, 1999. 10
- R. O. Duda, P. E. Hart, y D. G. Stork. *Pattern classification*. John Wiley & Sons, 2012. 39
- M. D’Zmura, G. Iverson, y B. Singer. Probabilistic color constancy. *Geometric Representations of Perceptual Phenomena*, pages 187–202, 1995. 4
- M. Ebner. Evolving color constancy. *Pattern Recognition Letters*, 27(11):1220–1229, 2006. 4
- M. Ebner. *Color Constancy*. John Wiley & Sons, 2007. 15, 18
- M. Ebner. Color constancy based on local space average color. *Machine Vision and Applications*, 20:283–301, 2009. 4
- E. A. Essock, M. J. Sinai, J. S. McCarley, W. K. Krebs, y J. K. DeFord. Perceptual ability with real-world nighttime scenes: Image-intensified, infrared, and fused-color imagery. *Human Factors*, 41:438–452, 1999. 2
- M. M. Faghieh y M. E. Moghaddam. Multi-objective optimization based color constancy. *Applied Soft Computing*, 17(0):52–66, 2014. 24

BIBLIOGRAFÍA

- M. D. Fairchild. Refinement of the RLAB color space. *Color Research and Applications*, 21: 338–346, 1996. 44, 46, 73
- M. D. Fairchild. *Color Appearance Models*. John Wiley & Sons, 2 edition, 2005. 3, 13, 46
- J. Fan, D. K. Yau, A. K. Elmagarmid, y W. G. Aref. Automatic image segmentation by integrating color-edge extraction and seeded region growing. *IEEE Transactions on Image Processing*, 10(10):1454–1466, 2001. 52
- G. D. Finlayson y E. Trezzi. Shades of gray and colour constancy. In *Color and Imaging Conference*, volume 2004, pages 37–41, 2004. 4
- G. D. Finlayson, D. M. S., y B. Funt. Spectral sharpening: sensor transformations for improved color constancy. *Journal of the Optical Society of America A*, 11(5):1553–1563, 1994. 14
- G. D. Finlayson, S. D. Hordley, y P. M. Hubel. Color by correlation: A simple, unifying framework for color constancy. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 23(11):1209–1221, 2001. 4
- G. D. Finlayson, S. D. Hordley, C. Lu, y M. S. Drew. On the removal shadows from images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(1):59–68, 2006a. 15
- G. D. Finlayson, S. D. Hordley, y I. Tastl. Gamut constrained illuminant estimation. *International Journal of Computer Vision*, 67:93–109, 2006b. 4
- D. A. Forsyth. A novel algorithm for color constancy. *International Journal of Computer Vision*, 5(1):5–36, 1990. 4
- K.-S. Fu y J. Mui. A survey on image segmentation. *Pattern recognition*, 13(1):3–16, 1981. 52
- B. Funt y B. C. Lewis. Diagonal versus affine transformations for color correction. *Journal of the Optical Society of America A*, 17(11):2108–2112, 2000. 14
- F. Gasparini y R. Schettini. Color balancing of digital photos using simple image statistics. *Pattern Recognition*, 37(6):1201 – 1217, 2004. 3, 24, 30
- T. Gevers y A. Smeulders. Pictoseek: Combining color and shape invariant features for image retrieval. *IEEE Transactions on Image Processing*, 9(1):102–119, 2000. 3
- A. Gijsenij y T. Gevers. Color constancy using natural image statistics. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007. 3

- A. Gijsenij, T. Gevers, y J. van de Weijer. Generalized gamut mapping using image derivative structures for color constancy. *International Journal of Computer Vision*, 86(2): 127–139, 2010. 4
- A. Gijsenij, T. Gevers, y J. van de Weijer. Computational color constancy: Survey and experiments. *IEEE Transactions on Image Processing*, 20(9):2475–2489, 2011. 3, 13
- V. Goffaux, C. Jacques, A. Mouraux, A. Oliva, P. Schyns, y B. Rossion. Diagnostic colours contribute to the early stages of scene categorization: Behavioural and neurophysiological evidence. *Visual Cognition*, 12(6):878–892, 2005. 1, 2
- J. W. Grzymala-Busse. Lers—a system for learning from examples based on rough sets. In *Intelligent decision support*, pages 3–18. 1992.
- Y. HaCohen, E. Shechtman, D. B. Goldman, y D. Lischinski. Non-rigid dense correspondence with applications for image enhancement. In *ACM Transactions on Graphics (TOG)*, volume 30, pages 70:1–70:9, 2011. 6
- R. M. Haralick, K. Shanmugam, y I. Dinstein. Textural features for image classification. *IEEE Transaction on System Man and Cybernetics*, 3(6):610–621, 1973. 31
- M. A. Hogervorst y A. Toet. Fast natural color mapping for night-time imagery. *Information Fusion*, 11(2):69–77, 2010. v, v, 1, 2, 8, 9
- S. D. Hordley y G. D. Finlayson. A Re-evaluation of colour constancy algorithms. In *Proceedings of 17th International Conference on Pattern Recognition (ICPR)*, pages 76–79, 2004. 20
- D. Ilea y P. Whelan. Image segmentation based on the integration of colour–texture descriptors—a review. *Pattern Recognition*, 44(10):2479–2501, 2011. 52, 53
- H. Ishibuchi, T. Nakashima, y M. Nii. *Classification and Modeling with Linguistic Information Granules*. Springer-Verlag, Secaucus, NJ, USA, 2004. 10
- B. . ITU-R Rec. Encoding parameters of digital television for studios. Technical report, ITU, Geneva, Switerland, 1995. 69
- D. Joshi, R. Datta, E. Fedorovskaya, Q.-T. Luong, J. Wang, J. Li, y J. Luo. Aesthetics and emotions in images. *IEEE Signal Processing Magazine*, 28(5):94–115, 2011. 27
- T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, y A. Y. Wu. An efficient k-means clustering algorithm: Analysis and implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):881–892, 2002. 53

- G. K. Kloss. Colour constancy using von kries transformations colour constancy goes to the lab. *Res. Lett. Inf. Math. Sci.*, 13:19–33, 2009. 18
- A. S. Kumar, B. Kartikeyan, y K. Majumdar. Band sharpening of irs-multispectral imagery by cubic spline wavelets. *International Journal of Remote Sensing*, 21(3):581–594, 2000. 69
- L. I. Kuncheva. *Fuzzy classifier design*. Springer, 2000. 39
- F. Kurugollu, B. Sankur, y A. E. Harmanci. Color image segmentation using histogram multithresholding and fusion. *Image and vision computing*, 19(13):915–928, 2001. 53
- K. Lam. *Metamerism and colour constancy*. PhD thesis, University of Bradford, 1985. 14
- E. H. Land. The retinex theory of color vision. *Scientific American*, 237(6):108–128, 1977. 14
- E. H. Land y J. J. McCann. Lightness and retinex theory. *Journal of the Optical Society of America*, 61(1):1–11, 1971. 4, 14
- A. Levin, D. Lischinski, y Y. Weiss. Colorization using optimization. *ACM Transactions on Graphics*, 23(3):689–694, 2004. 6
- S. Liu, H. Sun, y X. Zhang. Selective color transferring via ellipsoid color mixture map. *Journal of Visual Communication and Image Representation*, 23(1):173–181, 2012. doi: <http://dx.doi.org/10.1016/j.jvcir.2011.09.006>. 6
- R. A. Lizarraga-Morales, R. E. Sanchez-Yanez, V. Ayala-Ramirez, y F. E. Correa-Tome. Integration of color and texture cues in a rough set-based segmentation method. *Journal of Electronic Imaging*, 23(2):023003–023003, 2014a. 52
- R. A. Lizarraga-Morales, R. E. Sanchez-Yanez, V. Ayala-Ramirez, y A. J. Patlan-Rosales. Improving a rough set theory-based segmentation approach using adaptable threshold selection and perceptual color spaces. *Journal of Electronic Imaging*, 23(1):013024–013024, 2014b. 52
- L. Lucchese y S. Mitra. Colour image segmentation: a state-of-the-art survey. In *Proceedings of the National Academy of Sciences, India*, volume 2, pages 207–222, 2001. 52
- L. Lucchese y S. Mitra. A new class of chromatic filters for color image processing. theory and applications. *IEEE Transactions on Image Processing*, 13(4):534–548, 2004. 5
- J. Maines. Surveillance and night vision. *Electronics & Power*, 30(9):679–683, 1984. 1
- E. Mamdani. Advances in the linguistic synthesis of fuzzy controllers. *International Journal of Man-Machine Studies*, 8(6):669–678, 1976. 11, 25

- D. Martin, C. Fowlkes, D. Tal, y J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings of 8th International Conference of Computer Vision*, volume 2, pages 416–423, 2001. 53, 55
- M. Mignotte. Segmentation by fusion of histogram-based-means clusters in different color spaces. *IEEE Transactions on Image Processing*, 17(5):780–787, 2008. 53
- H. Mobahi, S. R. Rao, A. Y. Yang, S. S. Sastry, y Y. Ma. Segmentation of natural images by texture and boundary compression. *International journal of computer vision*, 95(1):86–98, 2011. 53
- N. Moroney, M. Fairchild, R. Hunt, C. Li, M. R. Luo, y T. Newman. The ciecam02 color appearance model. In *IS&T/SID 10th Color Imaging Conference*, pages 23–27, 2002. 14
- M. Nilashi y O. Ibrahim. A model for detecting customer level intentions to purchase in B2C websites using TOPSIS and fuzzy logic rule-based system. *Arabian Journal for Science and Engineering*, 39(3):1–16, 2013. 10
- T. Ojala y M. Pietikäinen. Unsupervised texture segmentation using feature distributions. *Pattern Recognition*, 32(3):477–486, 1999. 52
- N. R. Pal y S. K. Pal. A review on image segmentation techniques. *Pattern recognition*, 26(9):1277–1294, 1993. 52
- N. Papadakis, E. Provenzi, y V. Caselles. A variational model for histogram transfer of color images. *IEEE Transactions on Image Processing*, 20(6):1682–1695, 2011. 6
- S. N. Pattanaik, J. A. Ferwerda, M. D. Fairchild, y D. P. Greenberg. A multiscale model of adaptation and spatial vision for realistic image display. In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '98*, pages 287–298, 1998. 67
- Z. Pawlak. Rough sets. *International Journal of Computer & Information Sciences*, 11(5): 341–356, 1982.
- T. Pouli y E. Reinhard. Progressive color transfer for images of arbitrary dynamic range. *Computers & Graphics*, 35(1):67–80, 2011. 6
- S. Premoze y M. Ashikhmin. Rendering natural waters. In *Computer graphics forum*, volume 20, pages 189–200. Wiley Online Library, 2001. v, 7
- E. Provenzi, C. Gatta, M. Fierro, y A. Rizzi. A spatially variant white-patch and gray-world method for color image enhancement driven by local contrast. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 30(10):1757–1770, 2008. 5

BIBLIOGRAFÍA

- X. Qian, Y. Wang, y B. Wang. Effective contrast enhancement method for color night vision. *Infrared Physics & Technology*, 55(1):130–136, 2012. 1, 2
- X. Qian, L. Han, Y. Wang, y B. Wang. Color contrast enhancement for color night vision based on color mapping. *Infrared Physics & Technology*, 57:36–41, 2013. 1, 2
- Y. Rao, W. Lin, y L. Chen. Image-based fusion for video enhancement of night-time surveillance. *Optical Engineering*, 49(12):120501–120501, 2010. 1
- E. Reinhard y T. Pouli. Colour spaces for colour transfer. In *Computational Color Imaging*, pages 1–15. Springer, 2011. 45
- E. Reinhard, M. Ashikhmin, B. Gooch, y P. Shirley. Color transfer between images. *IEEE Computer graphics and applications*, 21(5):34–41, 2001. 6, 8, 44, 45, 46, 55
- 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 (JOSA) A*, 15(8):2036–2045, 1998. 67
- C. Sauvaguet, J.-N. Vittaut, J. Suarez, V. Boyer, y S. Manuel. Automated colorization of segmented images based on color harmony. *Journal of Multimedia Processing and Technologies (JMPT)*, 1(4):228–244, 2010. 6
- M. Schroder y S. Moser. Automatic color correction based on generic content-based image analysis. In *Proceedings of Color Imaging Conference*, pages 41–45, 2001. 3, 24
- C. Senanayake y D. C. Alexander. Colour transfer by feature based histogram registration. In *British Machine Vision Conference (BMVC)*, pages 1–10, 2007. 6
- L. Shapiro y G. Stockman. *Computer vision*. Pearson, 2001. 52
- G. Shobha, J. Gubbi, K. S. Raghavan, L. K. Kaushik, y M. Palaniswami. A novel fuzzy rule based system for assessment of ground water potability: A case study in South India. *Magnesium (Mg)*, 30:35–41, 2013. 10
- J. Stefanowski. On rough set based approaches to induction of decision rules. *Rough sets in knowledge discovery*, 1(1):500–529, 1998.
- W. B. Thompson, P. Shirley, y J. A. Ferwerda. A spatial post-processing algorithm for images of night scenes. *Journal of Graphics Tools*, 7(1):1–12, 2002. 55
- A. Toet. Natural colour mapping for multiband nightvision imagery. *Information Fusion*, 4(3):155–166, 2003. 2, 45

- A. Toet y E. M. Franken. Perceptual evaluation of different image fusion schemes. *Displays*, 24(1):25–37, 2003. 2, 46
- A. Toet y M. A. Hogervorst. Progress in color night vision. *Optical Engineering*, 51(1):010901, 2012. 2
- A. Toet y J. Walraven. New false color mapping for image fusion. *Optical Engineering*, 35(3):650–658, 1996. 1, 2
- K. Trawinski, O. Cordon, L. Sanchez, y A. Quirin. A genetic fuzzy linguistic combination method for fuzzy rule-based multiclassifiers. *IEEE Transactions on Fuzzy Systems*, 21(5):950–965, 2013. 10
- V. Tsagaris, G. Ghirstoulas, y V. Anastassopoulos. A measure for evaluation of the information content in color images. In *IEEE International Conference on Image Processing*, volume 1, pages 417–20, 2005. 27, 31
- R. Unnikrishnan, C. Pantofaru, y M. Hebert. A measure for objective evaluation of image segmentation algorithms. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, (CVPR 2005)*, pages 34–34, 2005. 53, 56
- R. Unnikrishnan, C. Pantofaru, y M. Hebert. Toward objective evaluation of image segmentation algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6):929–944, 2007. 52, 53, 56
- M. Unser. Sum and difference histograms for texture classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):118–125, 1986. 31
- N. Vadiie. *Fuzzy Logic and control: Software and hardware applications*, chapter 4. Prentice Hall, Englewood, Cliffs, N.J., 1993. 12, 25
- J. van de Weijer, T. Gevers, y A. Gijsenij. Edge-based color constancy. *IEEE Transactions on Image Processing*, 16(9):2207–2214, 2007a. 4, 17
- J. van de Weijer, C. Schmid, y J. Verbeek. Using high-level visual information for color constancy. In *Proceedings of the International Conference on Computer Vision*, pages 1–8, 2007b. 3, 24
- S. R. Vantaram y E. Saber. Survey of contemporary trends in color image segmentation. *Journal of Electronic Imaging*, 21(4):040901–1, 2012. 52, 57
- J. Von Kries. *Sources of Color Vision*, chapter Influence of adaptation on the effects produced by luminous stimuli Vision, pages 109–119. MIT Press, 1970. 14

- S.-Y. Wan y W. E. Higgins. Symmetric region growing. *IEEE Transactions on Image Processing*, 12(9):1007–1015, 2003. 52
- G. West y M. H. Brill. Necessary and sufficient conditions for von kries chromatic adaptation to give color constancy. *Journal of Mathematical Biology*, 15:249–258, 1982. 14
- X. Wu. Adaptive split-and-merge segmentation based on piecewise least-square approximation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(8):808–815, 1993. 52
- Y. Xiang, B. Zou, y H. Li. Selective color transfer with multi-source images. *Pattern Recognition Letters*, 30(7):682–689, 2009. 45
- X. Xiao y L. Ma. Color transfer in correlated color space. In *Proceedings of the 2006 ACM international conference on Virtual reality continuum and its applications*, pages 305–309. ACM, 2006. 6
- A. Yamasaki, H. Takauji, S. Kaneko, T. Kanade, y H. Ohki. Denighting: Enhancement of nighttime images for a surveillance camera. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4, 2008. 1
- A. Yang, J. Wright, Y. Ma, y S. Sastry. Unsupervised segmentation of natural images via lossy data compression. *Computer Vision and Image Understanding*, 110(2):212–225, 2008. 56
- J. Yang, R. Stiefelhagen, U. Meier, y A. Waibel. Visual tracking for multimodal human computer interaction. In *Proceedings of the Conference on Human factors in computing systems*, pages 140–147, 1998. 3
- X. Yi y M. Eramian. Lbp-based segmentation of defocus blur. *IEEE Transactions on Image Processing*, 25(4):1626–1638, 2016. 52
- J. Yin y J. R. Cooperstock. Color correction methods with applications to digital projection environments. 12(1-3):499–506, 2004. 45
- L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965. 10
- L. A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on System Man and Cybernetics*, 3(1):28–44, 1973. 10
- S. Zeki. *A vision of the brain*. John Wiley & Sons, 1993. 3, 13
- Y. Zhang, P. An, Q. Zhang, L. Shen, y Z. Zhang. A no-reference image quality evaluation based on power spectrum. In *3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*, pages 1–4, 2011. 30