



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA  
INGENIERÍA ELÉCTRICA – PROCESAMIENTO DIGITAL DE SEÑALES

INDEXACIÓN DE IMÁGENES DIGITALES BASADA EN CONTENIDO

TESIS  
QUE PARA OPTAR POR EL GRADO DE:  
MAESTRA EN INGENIERÍA

PRESENTA:  
LAURA JOSEFINA REYES RUIZ

TUTOR  
DR. FRANCISCO JAVIER GARCÍA UGALDE  
FACULTAD DE INGENIERÍA

MÉXICO, D. F. DICIEMBRE 2015

**JURADO ASIGNADO:**

Presidente: Dra. Medina Gómez Lucía  
Secretario: Dr. Moctezuma Flores Miguel  
Vocal: Dr. García Ugalde Francisco  
1<sup>er.</sup> Suplente: Dra. Baydyk Mykolaivna Tetyana  
2<sup>d o.</sup> Suplente: Dra. Nakano Miyatake Mariko

Lugar donde se realizó la tesis:

Universidad Nacional Autónoma de México, México, D.F.

**TUTOR DE TESIS:**

Dr. Francisco Javier García Ugalde

---

**FIRMA**

## *Agradecimientos*

*Agradezco infinitamente a Dios por brindarme los medios para realizar este trabajo y por las bendiciones que me otorga diariamente.*

*Mi más sincera gratitud al Dr. Francisco García Ugalde por su orientación, dedicación y apoyo en la realización de esta tesis. De la misma manera quiero agradecer a todos los sinodales que contribuyeron en la revisión del trabajo.*

*Doy gracias a mis padres Laura Elvira Ruiz y José Reyes, a familia, a mis amigos y a todas las personas que me han acompañado en este trayecto, por el apoyo que me han brindado.*

*Esta investigación fue realizada en parte gracias al Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PÁPPII) de la UNAM IN-112513 "Diseño e implementación de algoritmos de marca de agua digital para la autenticación y protección de derechos de autor en imágenes digitales". Por lo cual agradezco a la DGAPA-UNAM la beca recibida.*

*Por último doy gracias al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico otorgado durante mis estudios de maestría.*

# *Índice general*

1. Introducción .....	1
1.1. Objetivos .....	1
2. Estado del arte .....	2
2.1. Indexación .....	2
2.2. Tipos de indexación.....	3
2.3. Algoritmos basados en conceptos .....	5
2.4. Algoritmos basados en el contenido de la imagen .....	5
2.4.1. Técnicas en el dominio de los píxeles .....	5
2.4.1.1. Basadas en las relaciones espaciales.....	5
2.4.1.2. Basados en la textura .....	6
2.4.1.3. Basados en la forma .....	6
2.4.1.4. Basados en histogramas de color.....	7
2.4.2. Técnicas en el dominio comprimido .....	7
2.4.2.1. Transformada discreta de Fourier (DFT) .....	8
2.4.2.2. Transformada de Kaurhunen-Loeve (KLT).....	8
2.4.2.3. Transformada Coseno Discreta (DCT) .....	9
2.4.2.4. Transformada Wavelet Discreta (DWT)/ Sub-bandas .....	9
2.4.2.5. Cuantización Vectorial (VQ) .....	10
2.4.3. Consideraciones de eficiencia de los algoritmos de indexación .....	11
2.5. Sistemas de búsqueda.....	11
2.5.1. QBIC.....	11
2.5.2. Virage .....	12
2.5.3. Photobook.....	12
2.5.4. MARS.....	12
3. Marco Teórico .....	13
3.1. Modelos de color.....	13
3.1.1. Modelo RGB .....	13
3.1.2. Modelos HSI, HSV, HSB y HSL.....	14
3.1.3. Modelo YCbCr.....	17
3.1.4. Modelo CIE XYZ .....	18
3.1.5. Modelo CIE LAB .....	19

3.1.6. Transformaciones entre modelos de color .....	20
3.1.6.1. Conversión de RGB a HSI .....	20
3.1.6.2. Conversión de HSI a RGB .....	20
3.1.6.3. Conversión de RGB a HSV .....	21
3.1.6.4. Conversión de HSV a RGB .....	22
3.1.6.5. Conversión de RGB a HSL .....	22
3.1.6.6. Conversión de HSL a RGB .....	23
3.1.6.7. Conversión de RGB a YCbCr .....	23
3.1.6.8. Conversión de YCbCr a RGB .....	23
3.1.6.9. Conversión de RGB a CIE XYZ .....	24
3.1.6.10. CIE XYZ a RGB .....	24
3.1.6.11. CIE XYZ A CIE LAB .....	24
3.1.6.12. CIE LAB a CIE XYZ .....	25
3.2. Descriptores .....	25
3.2.1. Basados en color .....	25
3.2.1.1. Histograma de Color Convencional (CCH) .....	26
3.2.1.2. Color Dominante (DCD) .....	27
3.2.1.3. Esquema de Color (CLD) .....	28
3.2.1.4. Correlograma de Color (CC) y Auto Correlograma de Color (ACC) .....	29
3.2.1.5. Correlograma Compacto Generalizado (CGC) .....	31
3.2.1.6. Auto Correlograma de Colores Dominantes (DCCD) .....	32
3.3. Medidas de similitud .....	35
3.3.1. Norma $L_p$ .....	35
3.3.1.1. Norma $L_1$ .....	36
3.3.1.2. Norma $L_2$ .....	36
3.3.1.3. Norma $L_\infty$ .....	36
3.3.2. Intersección de histograma .....	36
3.3.3. Distancia de Color Dominante .....	37
3.3.4. Distancia de Esquema de Color .....	38
3.3.5. Distancia de Correlograma .....	38
3.4. Métricas de evaluación .....	39
3.4.1. Característica Operativa del Receptor (ROC) .....	39
3.4.2. Precisión Promedio de Recuperación (ARP) .....	41

3.4.3. Tasa promedio de Recuperación (ARR).....	41
3.4.4. Rango Promedio de Recuperación Normalizado Modificado (ANMRR) .....	42
4. Propuesta .....	44
4.1. Algoritmo.....	44
4.2. Distancia de Auto correlograma de color dominante (ADCDd) .....	45
4.3. Imágenes obtenidas de una búsqueda .....	46
5. Experimentación y Resultados .....	49
5.1. Bases de datos utilizadas.....	49
5.1.1. DB01 .....	49
5.1.2. DB02 .....	50
5.1.3. DB03 .....	51
5.1.4. DB04 .....	52
5.2. Configuración del experimento.....	53
5.3. Resultados del algoritmo propuesto.....	57
5.4. Evaluación .....	60
5.4.1. Mediante ROC .....	60
5.4.1.1. Para la base DB01.....	61
5.4.1.2. Para la base DB02.....	62
5.4.1.3. Para la base DB03.....	63
5.4.1.4. Para la base DB04.....	64
5.4.2. Evaluación con la métrica ARP .....	65
5.4.2.1. Para la base DB01.....	65
5.4.2.2. Para la base DB02.....	65
5.4.2.3. Para la base DB03.....	66
5.4.2.4. Para la base DB04.....	66
5.4.1. Evaluación con la métrica ARR.....	67
5.4.1.1. Para la base DB01.....	67
5.4.1.2. Para la base DB02.....	67
5.4.1.3. Para la base DB03.....	68
5.4.1.4. Para la base DB04.....	68
5.4.2. Evaluación con la métrica ANMRR .....	68
5.4.3. Evaluación Conjunta.....	69
6. Conclusiones.....	71

7. Referencias.....	72
ANEXO I. Bases Utilizadas.....	76
I-1. DB01.....	76
I-2. DB02.....	76
I-3. DB03.....	77
I-4. DB04.....	77

# Índice de Figuras

Figura 2-1 Indexación automática de imágenes. ....	4
Figura 2-2 Varios métodos de indexación de imágenes basados en contenidos [2]. ....	4
Figura 2-3 Representación de la transformada wavelet de una imagen [8]. ....	10
Figura 3-1 Cubo RGB. ....	14
Figura 3-2 Matiz (H).....	15
Figura 3-3 Saturación (S) .....	15
Figura 3-4 Modelo de color HSI.....	16
Figura 3-5 Modelo de color HSV .....	16
Figura 3-6 Modelo de color HSL.....	17
Figura 3-7 Modelo de color YCbCr .....	18
Figura 3-8 Cromaticidad de CIE .....	18
Figura 3-9 Espacio de color CIE LAB .....	19
Figura 3-10 Ejemplos de colores en el espacio LAB .....	19
Figura 3-11 Imagen RGB con su histograma .....	26
Figura 3-12 Imagen 50.jpg con sus colores N dominantes.....	27
Figura 3-13 Proceso de extracción del CLD de la imagen 696.jpg .....	29
Figura 3-14 Proceso de extracción del CC de una imagen de 8x8 píxeles. ....	30
Figura 3-15 Imágenes con sus Auto Correlogramas.....	31
Figura 3-16 Proceso de extracción del CGC .....	32
Figura 3-17 Proceso de extracción del DCD mediante el histograma de color el espacio HSV cuantizado con 72 colores.....	34
Figura 3-18 Extracción de DCCD a partir de la matriz de colores dominantes representados mediante índices. ....	35
Figura 3-19 Tipos de Curvas ROC .....	39
Figura 4-1 Imagen original 15.jpg y sus imágenes de Colores dominantes obtenidas mediante tres diferentes representaciones en espacios de color. ....	44
Figura 4-2 Imagen de búsqueda 134001.jpg (izquierda), e imágenes recuperadas empleando el descriptor de colores dominantes $CDC_{RGB}$ con la distancia de colores dominantes $CDCd$ . ....	46
Figura 4-3 Imágenes recuperadas con DCCD y L1 normalizada. ....	47
Figura 4-4 Imágenes recuperadas con el Correlograma de colores dominantes (auto correlograma en el espacio HSV) DCCD, con el auto correlograma a una distancia de 1 píxel, y la distancia de auto correlograma de colores dominantes $ADCDD$ . ....	47
Figura 4-5 Imágenes recuperadas $ADCDD_{RGB}$ , con el auto correlograma a una distancia de 1 píxel, y la distancia de auto correlograma de colores dominantes $ADCDD$ . ....	47
Figura 4-6 Imágenes recuperadas con el Auto correlograma compacto generalizado de colores dominantes $ACGDCDD$ , a una distancia de 1 píxel, y la distancia de auto correlograma de colores dominantes $ADCDD$ . ....	48
Figura 4-7 Imágenes recuperadas DCCD, a una distancia de 1 píxel, y la distancia de auto correlograma de colores dominantes $ADCDD$ . ....	48

Figura 5-1 Categorías DB01 con una imagen muestra .....	50
Figura 5-2 Categorías DB02 con una imagen muestra .....	50
Figura 5-3 Categorías DB03 con una imagen muestra .....	51
Figura 5-4 Categorías DB03 con una imagen muestra .....	53
Figura 5-5 Curvas ROC del Color dominate (DCD) con DCDD .....	58
Figura 5-6 Curvas ROC del Auto correlograma de colores dominantes (ADCD) con CCd, L1N, L1, L2, L $\infty$ y ADCDD.....	58
Figura 5-7 Curvas ROC del Auto correlograma compacto generalizado de colores dominantes (ACGDCDD) con CCd, L1N, L1, L2, L $\infty$ y ADCDD.....	59
Figura 5-8 Curvas ROC del Correlograma de colores dominantes (Autot Correlograma-DCCD)) con CCd, L1N, L1, L2 L $\infty$ y ADCDD. ....	59

# *Índice de Tablas*

Tabla 3-1 Graduaciones de saturación.....	16
Tabla 3-2 Pesos usados para el CLD .....	38
Tabla 3-3 Matriz de confusión y algunas métricas de evaluación derivadas.....	40
Tabla 5-1 Nomenclatura de experimentación. ....	54
Tabla 5-2 Parámetros de los diferentes descriptores. ....	55
Tabla 5-3 Nomenclatura de resultados con el Histograma.....	55
Tabla 5-4 Nomenclatura de resultados con el Auto correlograma compacto generalizado de colores dominantes.....	55
Tabla 5-5 Nomenclatura de resultados con el Correlograma compacto generalizado de colores dominantes. ....	56
Tabla 5-6 Nomenclatura de resultados con el Auto correlograma de colores dominantes (en el espacio RGB) .....	56
Tabla 5-7 Nomenclatura de resultados con el Correlograma de colores dominantes (en el espacio RGB).....	56
Tabla 5-8 Nomenclatura de resultados con el Esquema de color, el color dominante, el correlograma de color dominante (en el espacio HSV) y el Auto correlograma de color dominante (en el espacio HSV).....	56
Tabla 5-9 Nomenclatura de resultados con el Correlograma compacto generalizado. ....	56
Tabla 5-10 Nomenclatura de resultados con el Auto correlograma de color.....	57
Tabla 5-11 Nomenclatura de resultados con el Correlograma de color. ....	57
Tabla 5-12 Mejores resultados de los descriptores aplicados a la primera base de datos.....	65
Tabla 5-13 Mejores resultados de los descriptores aplicados a la segunda base de datos.....	65
Tabla 5-14 Mejores resultados de los descriptores aplicados a la tercera base de datos.....	66
Tabla 5-15 Mejores resultados de los descriptores aplicados a la cuarta base de datos. ....	66
Tabla 5-16 Mejores resultados de los descriptores aplicados a la primera base de datos.....	67
Tabla 5-17 Mejores resultados de los descriptores aplicados a la segunda base de datos.....	67
Tabla 5-18 Mejores resultados de los descriptores aplicados a la tercera base de datos.....	68
Tabla 5-19 Mejores resultados de los descriptores aplicados a la cuarta base de datos. ....	68
Tabla 5-20 Mejores resultados de los descriptores aplicados a la tercera base de datos.....	69
Tabla 5-21 Evaluación conjunta para DB01.....	69
Tabla 5-22 Evaluación conjunta para DB02.....	70
Tabla 5-23 Evaluación conjunta para DB03.....	70
Tabla 5-24 Evaluación conjunta para DB04.....	70
Tabla I-1 Imágenes que conforman la DB01 .....	76
Tabla I-2 Imágenes que conforman la DB02 .....	76
Tabla I-3 Imágenes que conforman la DB02 .....	77
Tabla I-4 Imágenes que conforman la DB02 .....	77

# *1. Introducción*

La presente propuesta de trabajo de tesis de maestría, parte de la necesidad de poder encontrar rápidamente y en un contexto dado, imágenes digitalizadas de forma automatizada. Lo que representa una tarea muy importante, dada la cantidad de información digital que existe actualmente, y del uso intensivo que se hace de la misma.

Automatizar las tareas es una tendencia que implica menos costo y tiempo en la realización de algunas tareas como la indexación.

Es posible plantear una solución algorítmica que permita resolver la tarea de la indexación de imágenes digitales para búsqueda rápida de información en un contexto dado. La solución propuesta deberá ser robusta respecto a algunos parámetros, como puede ser la intensidad luminosa y el color de las imágenes digitales registradas en la base de datos.

Se estudió el "estado del arte" de las propuestas algorítmicas que se han utilizado para llevar a cabo la indexación de las imágenes digitales en bases de datos, hacer una selección de uno de ellos, realizar pruebas para comparar su desempeño con trabajos publicados en la literatura internacional del área.

## *1.1. Objetivos*

- Desarrollar un descriptor para la indexación de imágenes digitales basado en características de color.
- Llevar a cabo pruebas con diferentes bases de datos de tamaño y contenido diferente.
- Evaluar el descriptor comparándolo contra los descriptores existentes.

## *2. Estado del arte*

El incremento en la demanda de datos de imágenes y videos digitales en diversas aplicaciones, simultáneamente con el mejoramiento significativo de las tecnologías de procesamiento, subsistemas de redes y la disponibilidad de grandes sistemas de almacenamiento de la información [1], ha generado la necesidad de desarrollar métodos más eficaces para la consulta y búsqueda de la información visual. Al área del conocimiento para realizar estas búsquedas se le conoce con el término de indexación, o también búsqueda de imágenes, o video, en base al contenido, comúnmente llamada Content Based Image Retrieval (CBIR), o Content Based Video Retrieval (CBVR).

### *2.1. Indexación*

La búsqueda automática de imágenes, en bases de datos de gran tamaño a partir de consultas sobre el contenido de estos documentos se puede lograr mediante la indexación de las bases de datos, ordenando una serie de datos, o informaciones, de acuerdo a un criterio común a todos ellos para facilitar su consulta y análisis. Al resultado de este criterio se le conoce como descriptor.

El objetivo de la indexación de imágenes es desarrollar técnicas que proporcionen habilidades para almacenar y recuperar imágenes basadas en sus contenidos [2]. Dicho objetivo ha sido alcanzado en mayor medida mediante diferentes algoritmos de indexación presentados en la literatura [3, 4, 5, 6, 7, 8].

Para identificar los archivos, muchas de las aplicaciones de búsqueda ampliamente difundidas en la actualidad, utilizan encabezados, o metadatos, vinculados a los archivos visuales (imágenes y videos), de esa manera se puede indexar automáticamente a partir de los encabezados que los definen (Ej. Google). Este tipo de algoritmos de indexación, basado en metadatos, requiere de un ordenamiento previo y puede estar desvinculado del concepto perceptivo de las imágenes, ya que se utilizan características que no están basadas en el contenido de la imagen y pueden producir resultados indeseables.

La clave para recuperar satisfactoriamente información visual se encuentra en la disponibilidad de una representación adecuada del contenido visual de las imágenes [9]. Los algoritmos de indexación basados en el contenido utilizan las propiedades visuales de las imágenes que se buscan. Los que actualmente existen solamente funcionan

correctamente con imágenes prediseñadas, o con texturas, pero dan resultados limitados con cualquier otro tipo de imágenes.

## ***2.2. Tipos de indexación***

La indexación manual, o explícita, fue la que se empleó inicialmente y es el método más sencillo pero, también el que requiere mayor inversión de tiempo para su ejecución. Utiliza una semántica de alto nivel, que puede consistir en palabras que señalan algún objeto, o alguna acción.

En este tipo se vinculan las palabras con los archivos y las búsquedas posteriores se realizan con base en dichas palabras, generalmente los resultados de la búsqueda son las imágenes cuyas palabras vinculadas sean las más parecidas a las palabras introducidas para realizar la búsqueda. Una de sus desventajas es que los criterios utilizados para llevarla a cabo, podrían variar con la aplicación, ya que lo que puede ser relevante para una aplicación, puede no serlo para otra.

La indexación automática, descriptiva, o visual, se introdujo más recientemente y es un método que permite una búsqueda más rápida. Utiliza una semántica de bajo nivel, que puede estar constituida, ya sea por el significado de los objetos y de las escenas, o por aspectos visuales, así como por contenido geométrico.

En este tipo de indexación, se genera una base de datos de índices, o descriptores, a partir de la base de datos de las imágenes y cada vez que se realiza una búsqueda, se encuentra el descriptor de la imagen de búsqueda de la misma manera que se hizo con las imágenes de la base de datos. Después se utiliza alguna medida de similitud entre el descriptor de la nueva imagen y los descriptores de las imágenes de la base de datos, el resultado son las imágenes cuyos descriptores sean los más cercanos al descriptor de la nueva imagen (ver Figura 2-1).

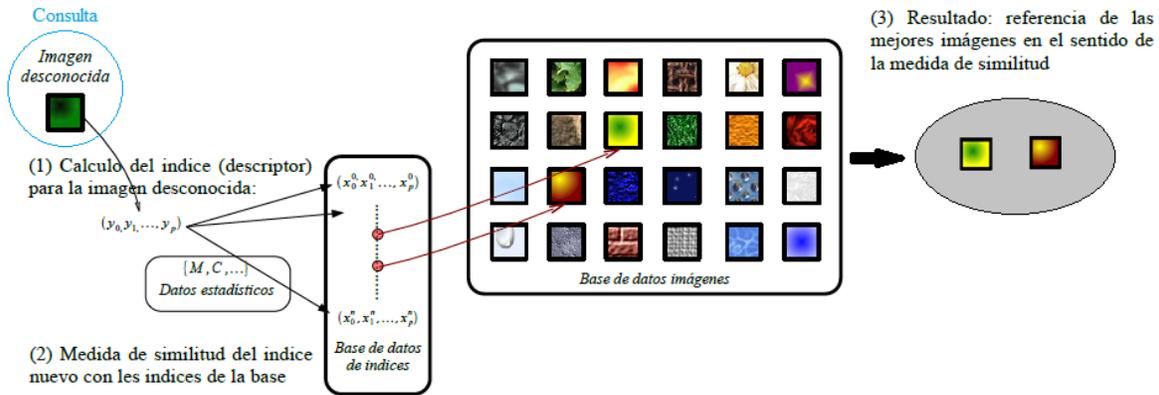


Figura 2-1 Indexación automática de imágenes<sup>1</sup>.

Existen dos grandes grupos de técnicas de indexación automática basándose en el contenido: técnicas en el dominio de los píxeles y técnicas en el dominio comprimido. En la Figura 2-2 se muestran los métodos de cada uno de dichos grupos.

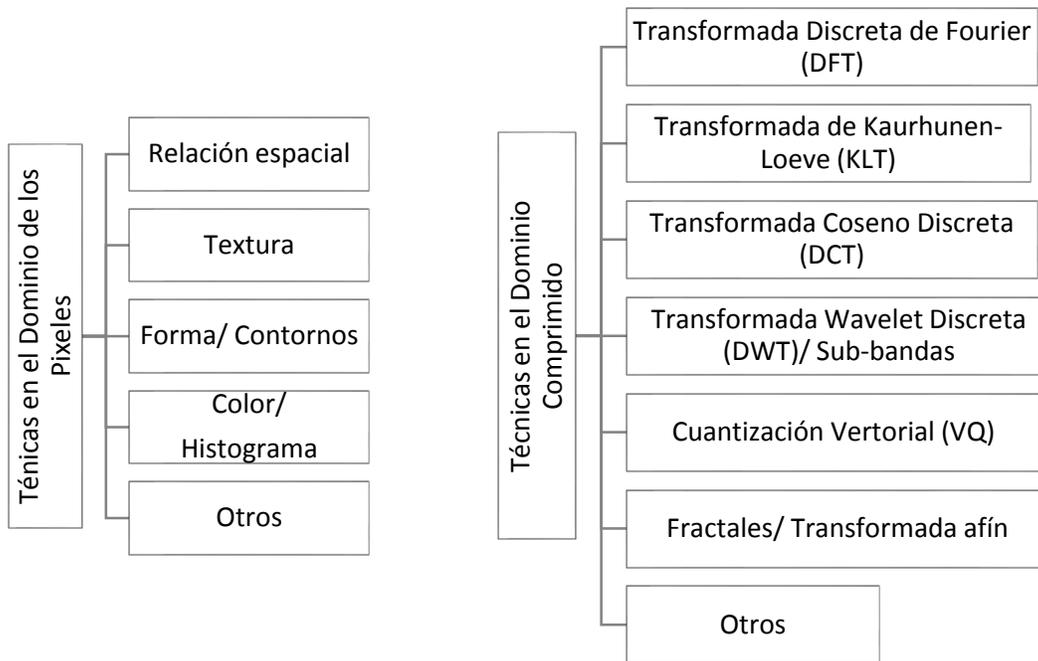


Figura 2-2 Varios métodos de indexación de imágenes basados en contenidos [2].

<sup>1</sup> Imagen tomada de las Notas del curso "Indexación", A. Manzanera, Bogotá, 2008  
[http://www.google.com.mx/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0CCgQFjAA&url=http%3A%2F%2Fwww.ensta-paristech.fr%2F~manzaner%2FDownload%2FTutorials%2FCurso\\_Indexacion\\_Bogota08.pdf&ei=sl6CU4DtOIMqAbC5oLQDg&usg=AFQjCNGTpC7rOBTzjaHbBSQMRzsL5QvpCg](http://www.google.com.mx/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0CCgQFjAA&url=http%3A%2F%2Fwww.ensta-paristech.fr%2F~manzaner%2FDownload%2FTutorials%2FCurso_Indexacion_Bogota08.pdf&ei=sl6CU4DtOIMqAbC5oLQDg&usg=AFQjCNGTpC7rOBTzjaHbBSQMRzsL5QvpCg)

## ***2.3. Algoritmos basados en conceptos***

A este tipo de algoritmos de indexación también se les conoce como basados en descripción, o basados en texto, se hace la indexación a partir del texto que rodea a la imagen y aunque es ampliamente utilizado actualmente por su facilidad de realización y utilización, se considera poco apropiado para el concepto perceptivo de las imágenes, por lo que no es el objetivo de este trabajo.

## ***2.4. Algoritmos basados en el contenido de la imagen***

Indexar basándose en el contenido de la imagen, puede lograrse en varios niveles de abstracción, comenzando con los índices como el nombre del programa y el nombre del tema, hasta aspectos de niveles mucho más bajos como el lugar de edición [10]. La búsqueda, empleando estos algoritmos, puede ser de tipo conceptual, simbólica, de parecido, o semántica. En la de tipo conceptual se emplea una palabra, o una frase, en la simbólica es un diagrama, o boceto, lo que se utiliza, en la de parecido lo que procede es el uso de otra imagen, y por último en la semántica se adjuntan descriptores visuales.

### ***2.4.1. Técnicas en el dominio de los píxeles***

La indexación de datos visuales en el dominio de los píxeles está basada en características como el color, la textura, la forma, los histogramas, momentos, entre otros. Este tipo de técnicas se pueden basar en: las relaciones espaciales, la textura, la forma, e histogramas de color. A continuación se describe cada una de ellas.

#### ***2.4.1.1. Basadas en las relaciones espaciales***

En este tipo de técnicas, los objetos y las relaciones espaciales entre los objetos en una imagen son usados para representar el contenido de la imagen. Primero los objetos de una imagen son segmentados y reconocidos, después la imagen es convertida en una imagen simbólica que es codificada usando códigos de cadenas de dos

dimensiones (2D). Los códigos de cadenas 2D representan las relaciones entre los objetos y la imagen, y son expresados usando un conjunto de operadores, entonces la búsqueda de imágenes se convierte en un problema de comparación de secuencia 2D.

Una desventaja de este tipo de técnicas es que se pierde la información de color y textura de la imagen ya que se reducen a imágenes simbólicas.

#### ***2.4.1.2. Basados en la textura***

Una imagen puede ser considerada como un mosaico de diferentes regiones de textura y las características de la imagen asociadas a estas regiones, pueden ser usadas para la búsqueda y recuperación. La textura es un concepto que se refiere, en la mayoría de los casos, a la repetición de elementos básicos generalmente conocidos como texeles. Los texeles pueden tener naturaleza periódica, cuasi periódica, o más generalmente aleatoria. El modelado y clasificación de la textura se divide en tres grupos principales; estructural, estadística, y espectral.

En el grupo de las estructurales se han utilizado ante todo análisis de la textura usando características y métodos basados en el modelo. Los métodos basados en características, analizan la textura usando una descripción informativa, que se presta más fácilmente a otras tareas, típicamente usando bancos de filtros lineales como sistemas de entrada.

En el grupo de los estadísticos se encuentran los Campos Aleatorios de Markov (MRFs), utilizan modelos manejables para los patrones de textura, y se formula un análisis de textura para la estimación de parámetros [11].

En el grupo de los espectrales se encuentra el Análisis de Componentes Dominantes (DCA).

La principal desventaja en este tipo de técnicas es la dificultad de definir el concepto de textura, lo cual dificulta la elección de los texeles.

#### ***2.4.1.3. Basados en la forma***

Por el modelo de percepción visual, la forma es un criterio importante para la comparación de imágenes basada en su perfil y estructura física. En las aplicaciones de búsqueda de imágenes, las características de forma pueden ser clasificadas en globales y locales.

Las características globales son las que se derivan de las propiedades de la forma entera, tales como redondez, circularidad, momentos centrales, y excentricidad.

Las características locales son aquellas que se derivan del procesamiento parcial de la forma, incluyendo tamaño y orientación de segmentos consecutivos del borde, puntos de curvatura, esquinas, y ángulo de giro.

La búsqueda basada en la similitud de la forma es un problema difícil por la falta de una definición matemática exacta de este parámetro.

Al igual que las técnicas basadas en relaciones espaciales, se pierde la información de color y textura de la imagen.

#### ***2.4.1.4. Basados en histogramas de color***

El color es una de las características más importantes de una imagen. Típicamente el color de una imagen es representado usando el histograma de color de la imagen. El histograma de una imagen con colores en el rango  $[0, L - 1]$  es una función discreta  $p(i) = n/L$ , donde  $i$  es el color de un pixel,  $n$  es el número de pixeles de ese color en la imagen, y el índice  $i = 0, 1, 2, \dots, L - 1$ . Una vez normalizado el histograma, para que el área bajo la curva sea unitaria,  $p(i) = n/L$  da un estimado de la probabilidad de ocurrencia del color  $i$ .

En la búsqueda de imágenes usando histogramas de color, el histograma de la imagen de búsqueda se compara con los histogramas de las imágenes de la base de datos. El proceso de comparación se hace utilizando una medida de similitud. Las métricas de similitud más comunes son la intersección del histograma y la distancia ponderada entre los histogramas de color. La complejidad del proceso de comparación puede ser reducida si se cuantiza el espacio de color, usando los datos dominantes del histograma, es decir, considerando un histograma de dimensión menor mediante la representación del histograma de color a diferentes resoluciones, y representado la información mediante subimágenes [10].

#### ***2.4.2. Técnicas en el dominio comprimido***

Estas técnicas surgieron por la necesidad de usar técnicas de compresión en los enormes volúmenes de datos visuales. Es una tendencia actual que los datos visuales contenidos en las bases de datos multimedia estén archivados en formato comprimido.

Una forma eficiente de indexar este tipo de archivos es hacerlo en el dominio comprimido, para lo cual se requiere un conocimiento muy a fondo de la estructura de los datos codificados.

#### ***2.4.2.1. Transformada discreta de Fourier (DFT)***

La transformada de Fourier es muy importante en el análisis y procesamiento de imágenes y señales. DFT emplea funciones de base exponencial compleja y provee un buen dominio de representación de la información espectral, que puede ser utilizado para la indexación [12].

Stone y Li han propuesto y evaluado un algoritmo de búsqueda de imágenes en el dominio de Fourier. El algoritmo tiene dos umbrales que permiten al usuario modificar independientemente el ajuste de la coincidencia. Un umbral controla la coincidencia de intensidad, mientras que el otro controla la coincidencia de textura. Los umbrales son valores de correlación que pueden ser calculados con bajo costo computacional cuando los coeficientes de Fourier son en su mayoría cero [2].

#### ***2.4.2.2. Transformada de Karhunen–Loeve (KLT)***

La transformada de Karhunen-Loeve, que está estrechamente relacionada con análisis de componentes principales, está basada en las propiedades estadísticas de una imagen. En este caso, las funciones base son eigenvectores de la matriz de auto correlación de la imagen, siendo esta su principal desventaja, dado que requiere un costo computacional alto. La KLT provee compactación máxima de energía y es una transformada estadística óptima. Como las funciones base de la KLT son adaptativas a la imagen, se puede obtener un buen desempeño en la indexación mediante la proyección de las imágenes en el espacio K-L, esta proyección extrae las características más expresivas (MEFs) de una imagen, y la indexación se logra comparando los coeficientes KLT [12].

### ***2.4.2.3. Transformada Coseno Discreta (DCT)***

La DCT, se puede considerar una versión real de la DFT mencionada en el inciso 1.4.2.1, emplea funciones de base real cosenoidal y tiene una eficiencia de compactación cercana a la óptima de la transformada KL para la mayoría de las imágenes naturales, sin el inconveniente del costo computacional, dado que las funciones de base son únicas. Como resultado, muchos de los estándares de compresión de imágenes y video internacionales, como JPEG, MPEG 1 Y 2, H.261/H.263, emplean la DCT. En JPEG por ejemplo, la compresión es realizada en tres pasos principales:

- Cálculo de la DCT,
- Cuantización de los coeficientes de la DCT
- Codificación de longitud variable.

Primero la imagen original se divide en bloques disjuntos de 8x8 pixeles. La DCT 2D de cada bloque es entonces calculada, y cuantizada usando una tabla visualmente adaptada, propuesta por JPEG [13].

### ***2.4.2.4. Transformada Wavelet Discreta (DWT)/ Sub-bandas***

En este método la imagen se pasa por un banco de filtros paso-bajas y paso-altas en ciclos recursivos, y de acuerdo a las propiedades de Fourier y Nyquist, las salidas de los filtros son decimadas para mantener la proporción del espectro. En la DWT, la salida del filtro pasa-bajas es recursivamente filtrada. Como la mayoría de la energía en el dominio de la frecuencia está representada por los coeficientes de la banda pasa-bajas, se logra una alta relación de compresión descartando los coeficientes de alta frecuencia. Con el objetivo de reducir los efectos de aliasing producidos por la decimación, la codificación por sub-bandas es implementada generalmente usando filtros de simetría en cuadratura (QMFs).

Todos los datos de la imagen son pasados a través de los filtros, no hay subdivisión de bloques como en JPEG. La descomposición en sub-bandas tiene muchas ventajas en la codificación como son: característica de multirresolución; alta descorrelación y eficiencia de compactación de energía; mejor adaptación a las señales no estacionarias; reducción de los artefactos de bloques, y ruido tipo "mosquito"; así como una mejor adaptación a las características del sistema visual humano [14].

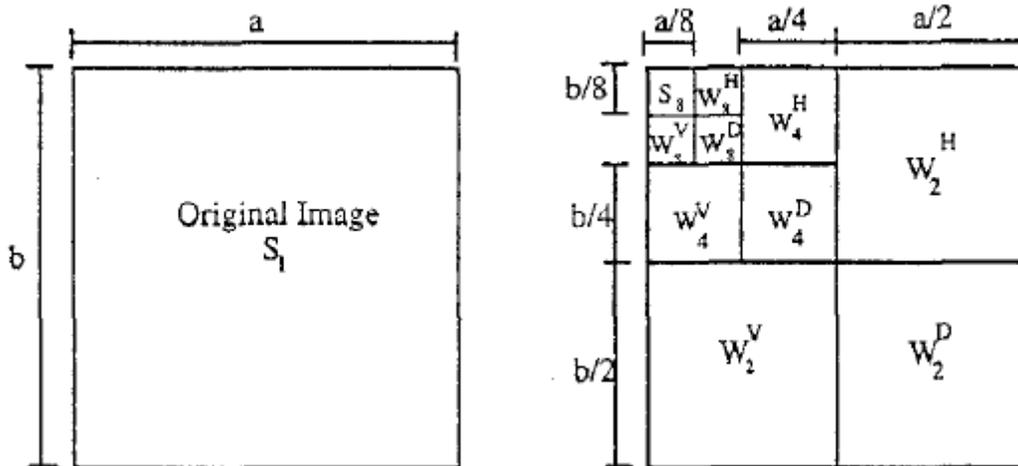


Figura 2-3 Representación de la transformada wavelet de una imagen [8].

#### 2.4.2.5. Cuantización Vectorial (VQ)

En la compresión de imágenes por Cuantización Vectorial (VQ), a un bloque de píxeles de la imagen se le identifica como vector de imagen y se le representa con un número de índice. Para descomprimir la imagen, a cada índice se le asigna de nuevo su correspondiente vector de imagen.

Después de que una imagen ha sido comprimida cada bloque de píxeles está representado por un número índice. Para una imagen dada, se puede calcular el número de ocurrencia de cada índice para obtener un histograma de índices  $H(d_1, d_2, \dots, d_i, \dots, d_n)$ , donde  $d_i$  es el número de veces que el índice identificador  $i$  aparece en la imagen, y  $n$  es el número total de índices.

En la búsqueda de imágenes con VQ, se calcula un histograma de índices para la imagen de búsqueda, al cual se le denominará  $H(b_1, b_2, \dots, b_i, \dots, b_n)$ . Para encontrar las coincidencias de la imagen buscada se calcula la distancia entre la imagen de búsqueda  $B$  y cada una de las imágenes  $D$  de la base de datos, utilizando una medida de distancia, por ejemplo el error acumulado:

$$d(B, D) = \sum_l^n (b_l - d_l)$$

Las imágenes pueden ordenarse ascendentemente según la distancia calculada. Entre más grande sea la distancia significa que más diferentes son [15].

### ***2.4.3. Consideraciones de eficiencia de los algoritmos de indexación***

Entre los elementos utilizados para comparar la eficiencia de los diversos algoritmos de indexación se pueden utilizar: la capacidad de distinguir imágenes diferentes, el número máximo de imágenes que una consulta podría localizar, la cantidad de cálculos requeridos para calcular y comparar las características [2, 16], la robustez respecto a cambios sufridos por la imagen, etc. El factor que decidirá qué tan eficiente es el algoritmo empleado, dependerá de los objetivos particulares del tipo de búsqueda.

## ***2.5. Sistemas de búsqueda***

Con el auge de internet y de las grandes bases de datos, en la última década del siglo pasado, la búsqueda de imágenes basada en contenido se ha convertido en un área muy activa de investigación. Se han construido muchos sistemas de búsqueda, tanto comerciales como académicos [16]. En seguida se presentan algunos ejemplos de los sistemas existentes.

### ***2.5.1. QBIC***

QBIC, de las siglas en inglés de "*query by image content*" (consulta por el contenido de la imagen), fue el primer sistema de búsqueda de imágenes basado en contenido. Su marco de trabajo y técnicas han tenido profundos efectos en los sistemas posteriores de búsqueda de imágenes.

Este sistema puede hacer búsquedas basadas en imágenes de ejemplo, bocetos y dibujos hechos por el usuario, y patrones seleccionados de color y textura, entre otros. Las características de color usadas en QBIC son: la más usual es el espacio (R,G,B); también (L,a,b) y coordenadas MTM (transformada matemática de Mussell), así como un histograma de color de k-elementos. Su característica de textura es una mejora de la representación de textura de Tamura [17]. También su característica de forma se basa en la constitución de la forma del área, la circularidad, la excentricidad, la orientación del eje mayor, y de un conjunto de momentos algebraicos invariantes [18].

### ***2.5.2. Virage***

Este motor de búsqueda que fue desarrollado en Virage Inc. Similar a QBIC, puede soportar búsquedas basadas en color, composición (disposición del color), textura y estructura (información del contorno). Sin embargo Virage va un paso adelante que QBIC, porque también soporta combinaciones arbitrarias de las cuatro consultas atómicas anteriores. Los usuarios pueden ajustar los pesos asociados a las características atómicas de acuerdo con la aplicación respectiva.

### ***2.5.3. Photobook***

Es un conjunto de herramientas interactivas para navegación y búsqueda de información desarrollada en el MIT Media Lab. Photobook consiste en tres sub-libros que incluyen: forma, textura, y características de superficie, una vez que se generan los sub-libros, los usuarios pueden buscar las imágenes basándose en las características correspondientes de los mismos.

### ***2.5.4. MARS***

MARS, de las siglas en inglés de "*multimedia analysis and retrieval system*" (sistema de búsqueda y análisis multimedia) fue desarrollado en la universidad de Illinois. MARS es diferente de otros sistemas, tanto en el ámbito de investigación, como en las técnicas usadas. Es un desarrollo que implicó un esfuerzo interdisciplinario de múltiples comunidades de investigación: visión computacional, sistemas de manejo de bases de datos (DBMS), y búsqueda de información (IR). Las características de investigación de MARS son la integración de DBMS e IR, lo que da como resultado la coincidencia exacta con una búsqueda clasificada, integración de indexación y búsqueda (lo que explica cómo el algoritmo de búsqueda puede tomar ventaja de subrayar la estructura de indexación), y la integración del humano con la computadora [16].

## ***3. Marco Teórico***

La Recuperación de Imágenes Basada en Contenido, CBIR por sus siglas en inglés; también es conocida como Búsqueda por Contenido de imagen, QBIC (siglas en inglés); y como Recuperación de Información Visual Basada en Contenido, CBVIR (siglas en inglés); es una disciplina que se ha incrementado en los últimos años debido a que facilita la búsqueda de imágenes en grandes bases de datos.

Debido a que uno de los objetivos del presente trabajo es desarrollar un algoritmo rápido y que contribuya a la investigación en el área de la Recuperación de Imágenes Basada en Contenido CBIR, en esta sección se describen los conceptos en los que se fundamenta el presente trabajo de investigación.

### ***3.1. Modelos de color***

Como se ha visto en los párrafos anteriores, los modelos de color en el procesamiento digital de imágenes son una herramienta importante, ya que permiten analizar y aprovechar toda la información presente dentro de la imagen [19].

El objetivo de un Modelo de Color, también conocido como Espacio de Color, o Sistema de Color, es facilitar la descripción de la percepción humana del color. En general un modelo de color es un espacio matemático tridimensional.

#### ***3.1.1. Modelo RGB***

En este modelo, cada color aparece en sus componentes espectrales primarios: Rojo (*Red*), Verde (*Green*) y Azul (*Blue*) [12]. La gran mayoría de las imágenes digitales se encuentran representadas con este modelo, ya que es el que utilizan los monitores electrónicos modernos.

El modelo RGB está basado en un sistema de coordenadas cartesianas, el cual se ilustra en la Figura 3-1. Cada canal, representado por un eje del cubo respectivamente, es de 8 bits y por lo tanto puede tomar  $2^8=256$  tonos en el rango de  $[0, 255]$ , si se usan valores normalizados el rango es de  $[0, 1]$ . Al considerarse los tres ejes, el cubo posee  $(2^8)^3=2^{24}=16,777,216$  colores.

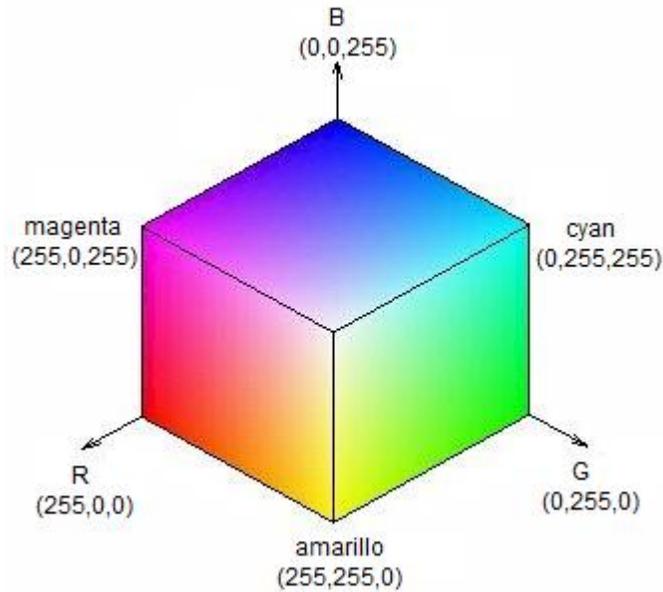


Figura 3-1 Cubo RGB.

En este modelo, la triada (0,0,0) corresponde al color negro, en tanto que la triada (255,255,255) representa al blanco y la diagonal entre estos dos puntos, cuando el valor de  $R$  es igual a  $G$  y  $B$ , representa todos los tonos de grises, y la combinación de las diferentes coordenadas dentro del cubo, representan los distintos colores.

### 3.1.2. Modelos HSI, HSV, HSB y HSL

Estos modelos son una deformación no lineal del espacio RGB, en ellos cada color se representa por medio de atributos percibidos por la vista humana ( $H, S, I, V, B$  y  $L$ ), estos atributos se describen en detalle más adelante, y dos de los cuales ( $H$  y  $S$ ) se encuentran presentes en todos estos modelos.

Atributos percibidos por la vista humana

$H$	$S$	$I$	Intesidad ( <i>Intesity</i> )
Matiz ( <i>Hue</i> )	Saturación ( <i>Saturation</i> )	$V$	Valor ( <i>Value</i> )
		$B$	Brillo ( <i>Brightness</i> )
		$L$	Luminosidad ( <i>Luminosity</i> )

El primer atributo compartido es el Matiz  $H$  que se refiere a un color puro, en la Figura 3-2 se puede observar que está propiedad toma valores en un rango de  $[0^\circ, 360^\circ]$ , donde el  $0^\circ/360^\circ$  corresponde al color rojo, el  $120^\circ$  al verde, y  $240^\circ$  al color azul.

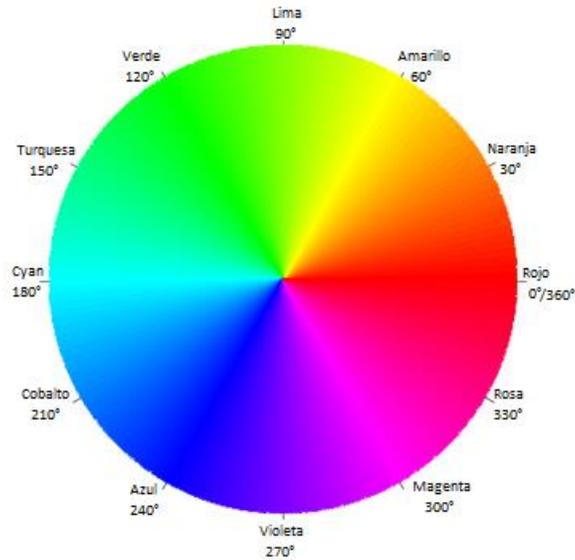


Figura 3-2 Matiz (H).

El segundo atributo compartido es la Saturación  $S$ , que indica el grado de pureza del color, en términos comunes indica que tan decolorado está, los valores se encuentran en el rango de  $[0,1]$  como se aprecia en la Figura 3-3. Aquí se observa que en el 1 el color se mantiene completamente puro, y en el 0 el color está totalmente decolorado.

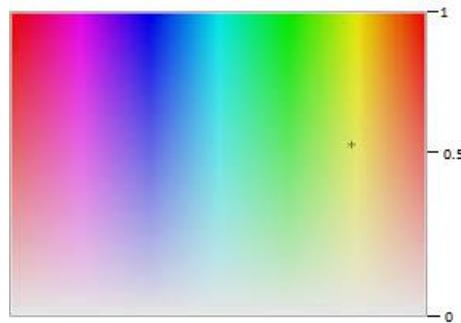


Figura 3-3 Saturación (S)

Se pueden apreciar algunas diferencias entre las Saturaciones de los diferentes modelos de color, y habrá mayor variación dependiendo del valor del tercer componente ( $I, V, B$  o  $L$ ). En la Tabla 3-1 se puede observar la saturación del color rojo, dependiendo de la luminosidad, o del valor.

Tabla 3-1 Graduaciones de saturación

Modelo HSL				
matiz 100% puro	75% de saturación	saturación media	25% de saturación	0% de saturación
Modelo HSV				
matiz 100% puro	75% de saturación	saturación media	25% de saturación	0% de saturación

Los atributos restantes ( $I, V, B, L$ ) son muy similares entre sí, cabe aclarar que el modelo HSV es igual al modelo HSB. Todos estos atributos pueden ir del negro al blanco.

La Intensidad puede entenderse como la presencia, o ausencia de luz, entre mayor es la intensidad de la iluminación más cerca está del blanco, y viceversa.

El Valor, o Brillo, puede entenderse como que tan claro, u oscuro, pueden volverse los colores.

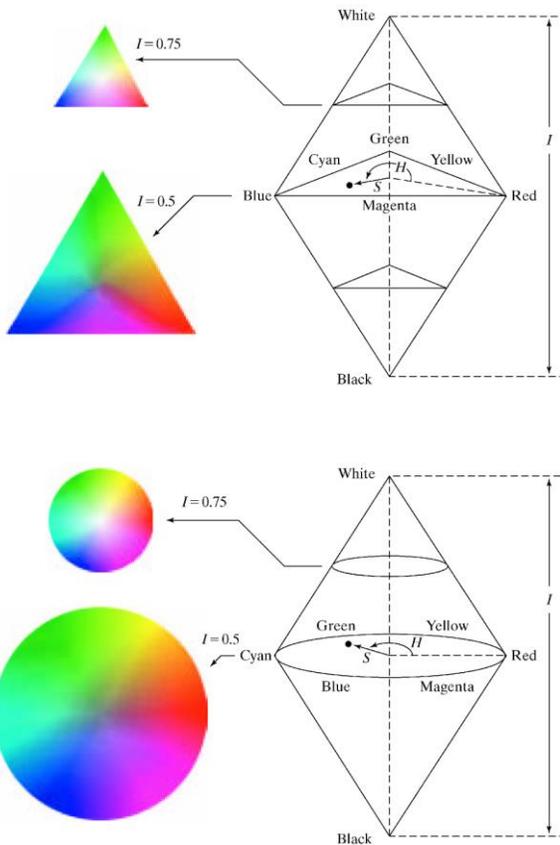


Figura 3-4 Modelo de color HSI<sup>2</sup>

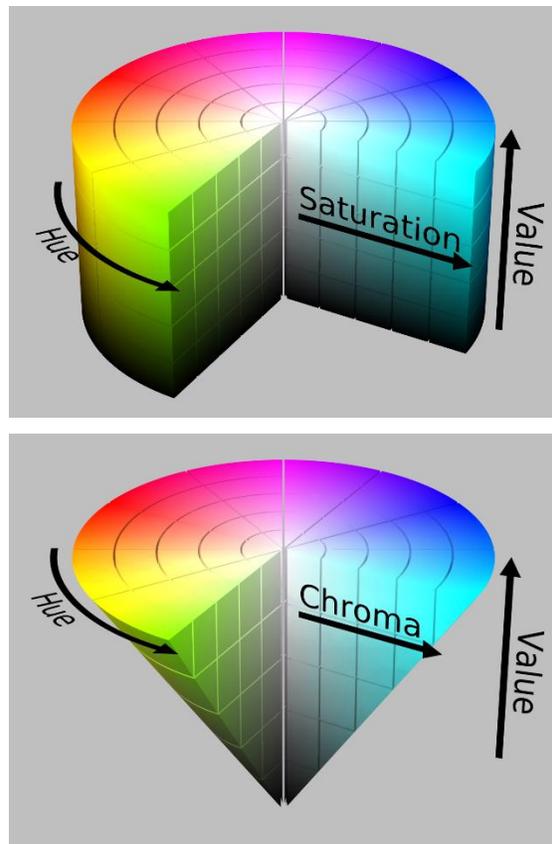


Figura 3-5 Modelo de color HSV<sup>3</sup>

<sup>2</sup>[https://www.google.com.mx/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&cad=rja&uact=8&ved=0CAcQjRxqFQoTCLrAiPKt8sYCFYs\\_Pgodr7UD6A&url=https%3A%2F%2Fmiac.unibas.ch%2F5IP%2F02-Fundamentals.html&ei=S3KxVfrujlv\\_-AGv647ADg&bvm=bv.98476267,d.cGU&psig=AFQjCNGa7ca79\\_10tw4FtaSF3wTVPIfOpw&ust=1437778820920590](https://www.google.com.mx/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&cad=rja&uact=8&ved=0CAcQjRxqFQoTCLrAiPKt8sYCFYs_Pgodr7UD6A&url=https%3A%2F%2Fmiac.unibas.ch%2F5IP%2F02-Fundamentals.html&ei=S3KxVfrujlv_-AGv647ADg&bvm=bv.98476267,d.cGU&psig=AFQjCNGa7ca79_10tw4FtaSF3wTVPIfOpw&ust=1437778820920590)

<sup>3</sup>[https://en.wikipedia.org/wiki/HSL\\_and\\_HSV](https://en.wikipedia.org/wiki/HSL_and_HSV)

La Luminosidad también puede entenderse como la presencia, o ausencia, de luz, cuando está muy oscuro, lo único que se aprecia es el color negro; cuando aumenta la iluminación se pueden observar los colores, pero si hay demasiada luminosidad, es deslumbrante, y solamente se aprecia el color blanco.

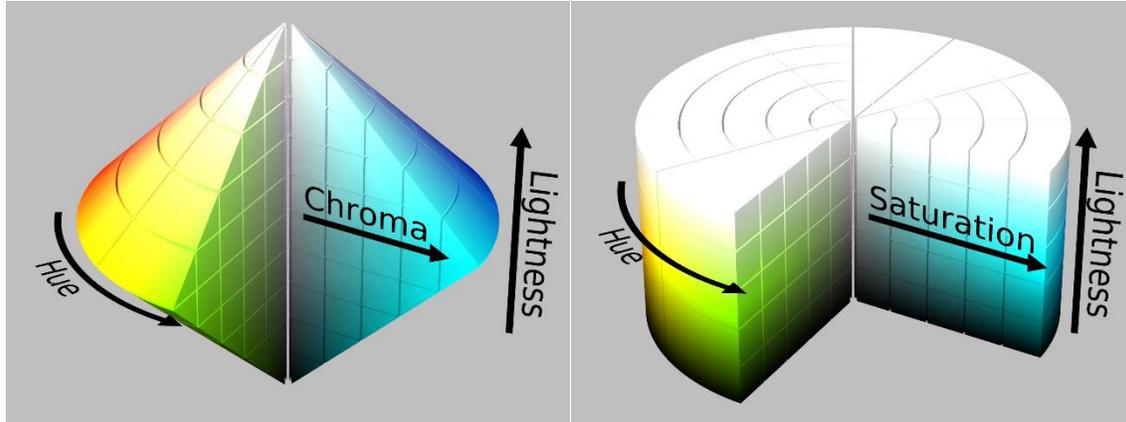


Figura 3-6 Modelo de color HSL<sup>4</sup>

### 3.1.3. Modelo YCbCr

En este modelo la  $Y$  simboliza la luminancia, puede tomar valores en el rango  $[0, 255]$ ; el término  $Cb$  representa la crominancia en el color azul, y  $Cr$  también alude a la crominancia pero del color rojo, ambos pueden tomar valores en el rango  $[-128, 127]$ , si se utiliza el signo. Si se utiliza la versión sin signo el rango es de  $[0, 255]$ .

En la Figura 3-7 se muestra este modelo de color como un cubo, donde la luminancia va del blanco al negro, al igual que en algunos de los modelos explicados anteriormente. También se muestra un plano donde se puede observar los colores a lo largo de los ejes  $Cr$  y  $Cb$ .

Este modelo es utilizado en la televisión digital y en la televisión de Alta Definición (HDTV), aunque en estas aplicaciones el rango de  $Y$  se limita al rango  $[16, 235]$ , y el de  $Cb$  y  $Cr$  al rango  $[16, 240]$ .

Cuando los valores del modelo YCbCr se encuentran en el rango de  $[0, 255]$ , estos pueden normalizarse al rango de  $[0, 1]$ .

---

<sup>4</sup>Idem

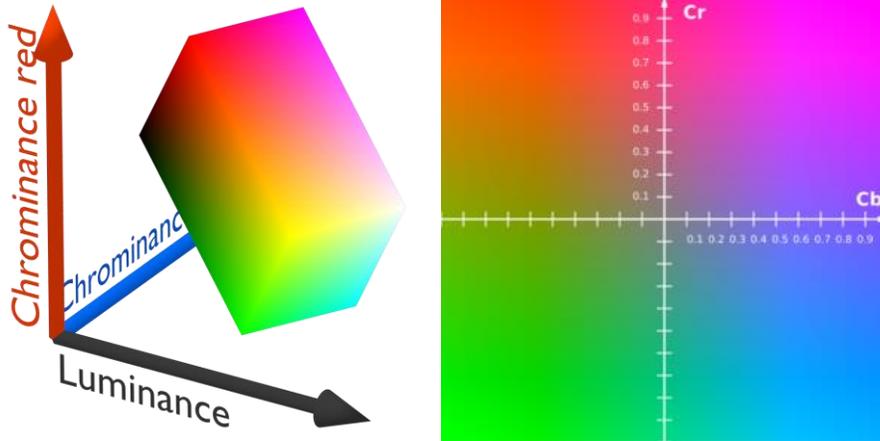


Figura 3-7 Modelo de color YCbCr<sup>5</sup>

### 3.1.4. Modelo CIE XYZ

Este modelo fue establecido en 1931 por la *Comision Internationale de l'Éclairage* (CIE). Este modelo también es tridimensional, la  $X$  tiende a una curva de la sensibilidad del ojo humano del rojo al verde, colores detectados por los conos  $L$  y  $M$  respectivamente. La  $Y$  se refiere a la luminosidad, y  $Z$  se aproxima al estímulo del azul detectado por los conos  $S$ . Se utiliza como referencia para definir el rango de colores que percibe el ojo humano.

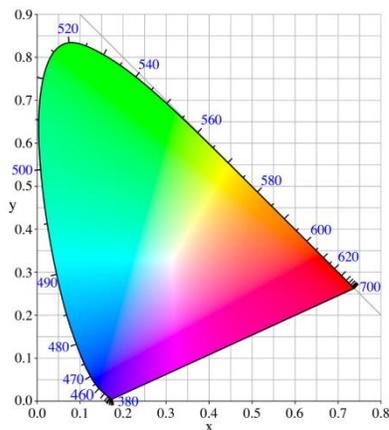


Figura 3-8 Cromaticidad de CIE<sup>6</sup>

En la Figura 3-8 se muestra la cromaticidad del modelo *CIE XYZ*, el contorno curvado negro se refiere a la zona monocromática y los números marcados en ella son las longitudes de onda en nanómetros.

La luminosidad, se refiere a que tan alumbrado, o iluminado, está un color. Por ejemplo el color blanco está muy iluminado, mientras que los diferentes tonos de gris podrían considerarse como el mismo color blanco gradualmente oscurecido, o con menor luminosidad.

<sup>5</sup>[https://upload.wikimedia.org/wikipedia/en/f/f4/YCbCrColorSpace\\_Perspective.png](https://upload.wikimedia.org/wikipedia/en/f/f4/YCbCrColorSpace_Perspective.png)  
[https://upload.wikimedia.org/wikipedia/commons/thumb/3/34/YCbCr-CbCr\\_Scaled\\_Y50.png/300px-YCbCr-CbCr\\_Scaled\\_Y50.png](https://upload.wikimedia.org/wikipedia/commons/thumb/3/34/YCbCr-CbCr_Scaled_Y50.png/300px-YCbCr-CbCr_Scaled_Y50.png)

<sup>6</sup> [https://es.wikipedia.org/wiki/Espacio\\_de\\_color\\_CIE\\_1931](https://es.wikipedia.org/wiki/Espacio_de_color_CIE_1931)

### 3.1.5. Modelo CIE LAB

Este modelo es derivado del CIE XYZ y se calcula empleando raíces cúbicas, fue propuesto en 1976 con el objetivo de describir todos los colores que puede percibir el sistema de visión humana, el cual se muestra en la Figura 3-9.

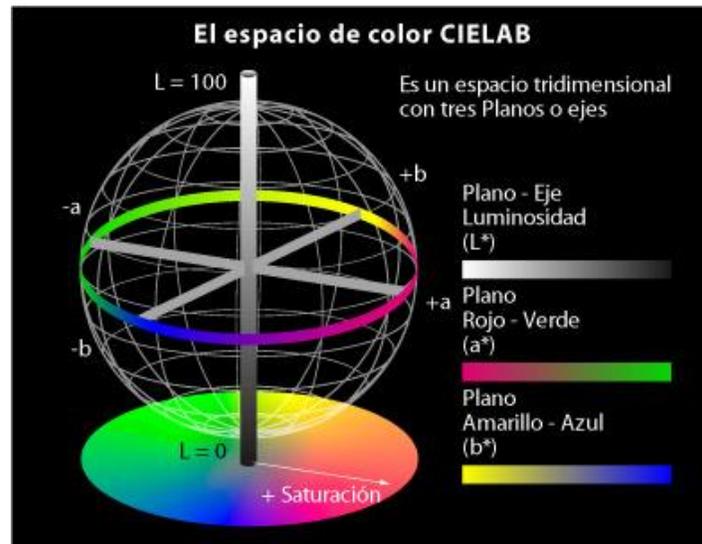


Figura 3-9 Espacio de color CIE LAB<sup>7</sup>

En este modelo la  $L$  se refiere a la luminosidad, que se representa en un rango de  $[0,100]$ , donde 0 es sin iluminación (negro) y 100 totalmente iluminado (blanco); la  $A$  se refiere a la sensibilidad que tiene el ojo entre el rojo y el verde, donde los valores negativos tienden al verde y los positivos al rojo; y  $B$  se refiere a la sensibilidad del ojo entre el azul y el amarillo, donde los valores negativos tienden al azul y los positivos al amarillo.

En la Figura 3-10 se pueden observar secciones de color con diferentes luminosidades.

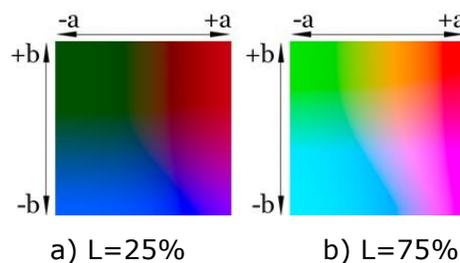


Figura 3-10 Ejemplos de colores en el espacio LAB<sup>8</sup>

<sup>7</sup> [http://www.gusgsm.com/espacio\\_color\\_cie\\_lab](http://www.gusgsm.com/espacio_color_cie_lab)

<sup>8</sup> [https://es.wikipedia.org/wiki/Espacio\\_de\\_color\\_Lab](https://es.wikipedia.org/wiki/Espacio_de_color_Lab)

### ***3.1.6. Transformaciones entre modelos de color***

Algunos métodos de extracción de características de imágenes digitales requieren la transformación del modelo de color original de la imagen a otros modelos convencionales, por lo que en los siguientes párrafos se presentan algunos métodos de conversión.

#### ***3.1.6.1. Conversión de RGB a HSI***

Dados los valores de RGB normalizados al rango [0,1] y el ángulo  $\theta$  medido como se muestra en la Figura 3-2; la transformación del modelo RGB al modelo HSI se obtiene mediante las siguientes ecuaciones:

$$H = \begin{cases} \theta & \text{si } B \leq G \\ 360 - \theta & \text{si } B > G \end{cases}$$

con

$$\theta = \cos^{-1} \left\{ \frac{\frac{1}{2} [(R - G) + (R - B)]}{\sqrt{[(R - G)^2 + (R - B)(G - B)]}} \right\}, \quad (3-1)$$

$$S = 1 - \frac{3}{(R + G + B)} [\min(R, G, B)],$$

$$S = 1 - \frac{[\min(R, G, B)]}{I}, \quad (3-2)$$

$$I = \frac{1}{3}(R + G + B). \quad (3-3)$$

La imagen resultante HSI tendrá  $H$  en el rango de  $[0^\circ, 360^\circ]$ , y  $S$  e  $I$  en el rango de  $[0,1]$ .

#### ***3.1.6.2. Conversión de HSI a RGB***

Dados los valores de  $H$  en el rango de  $[0^\circ, 360^\circ]$ , y los valores de  $S$  e  $I$  en el rango  $[0,1]$ ; para la transformación del modelo HSI al modelo RGB se divide al Matiz ( $H$ ) en tres regiones:

I. RG ( $0^\circ \leq H < 120^\circ$ )

$$B = I(1 - S), \quad (3-4)$$

$$R = I \left[ 1 + \frac{S \cos H}{\cos(60^\circ - H)} \right], \quad (3-5)$$

$$G = 3I - (R + B). \quad (3-6)$$

II. GB ( $120^\circ \leq H < 240^\circ$ )

$$H = H - 120^\circ, \quad (3-7)$$

$$R = I(1 - S), \quad (3-8)$$

$$G = I \left[ 1 + \frac{S \cos H}{\cos(60^\circ - H)} \right], \quad (3-9)$$

$$B = 3I - (R + G). \quad (3-10)$$

III. BR ( $240^\circ \leq H \leq 360^\circ$ )

$$H = H - 240^\circ, \quad (3-11)$$

$$G = I(1 - S), \quad (3-12)$$

$$B = I \left[ 1 + \frac{S \cos H}{\cos(60^\circ - H)} \right], \quad (3-13)$$

$$R = 3I - (G + B). \quad (3-14)$$

La imagen resultante RGB estará normalizada con todos sus valores en el rango de  $[0,1]$ .

### 3.1.6.3. Conversión de RGB a HSV

Una vez normalizados los valores de RGB al rango  $[0,1]$ ; la transformación del modelo RGB al modelo HSV se obtiene mediante las siguientes ecuaciones:

$$C_{max} = \max(R, G, B), \quad (3-15)$$

$$C_{min} = \min(R, G, B), \quad (3-16)$$

$$\Delta = C_{max} - C_{min}, \quad (3-17)$$

$$H = \begin{cases} 0^\circ & \text{si } \Delta = 0 \\ 60^\circ \times \left( \frac{G - B}{\Delta} \text{ (mod } 6) \right) & \text{si } C_{max} = R \\ 60^\circ \times \left( \frac{B - R}{\Delta} + 2 \right) & \text{si } C_{max} = G \\ 60^\circ \times \left( \frac{R - G}{\Delta} + 6 \right) & \text{si } C_{max} = B \end{cases} \quad (3-18)$$

$$S = \begin{cases} 0 & \text{si } \Delta = 0 \\ \frac{\Delta}{C_{max}} & \text{si } \Delta \neq 0 \end{cases} \quad (3-19)$$

$$V = C_{max}. \quad (3-20)$$

La imagen resultante HSV tendrá  $H$  en el rango de  $[0^\circ, 360^\circ]$ ,  $S$  y  $V$  en el rango de  $[0,1]$ .

### 3.1.6.4. Conversión de HSV a RGB

Con los valores de  $H$  en el rango de  $[0^\circ, 360^\circ]$ , y los valores de  $S$  e  $I$  en el rango  $[0,1]$ ; para la transformación del modelo HSV al modelo RGB se emplean las siguientes ecuaciones:

$$C = V \times S, \quad (3-21)$$

$$X = C \times \left(1 - \left| \left( \frac{H}{60^\circ} \right) \bmod 2 - 1 \right| \right), \quad (3-22)$$

$$m = V - C, \quad (3-23)$$

$$(R, G, B) = \begin{cases} (C, X, 0) + m & \text{si } 0^\circ \leq H < 60^\circ \\ (X, C, 0) + m & \text{si } 60^\circ \leq H < 120^\circ \\ (0, C, X) + m & \text{si } 120^\circ \leq H < 180^\circ \\ (0, X, C) + m & \text{si } 180^\circ \leq H < 240^\circ \\ (X, 0, C) + m & \text{si } 240^\circ \leq H < 300^\circ \\ (C, 0, X) + m & \text{si } 300^\circ \leq H < 360^\circ. \end{cases} \quad (3-24)$$

La imagen resultante RGB estará normalizada con todos sus valores en el rango de  $[0,1]$ .

### 3.1.6.5. Conversión de RGB a HSL

Dados los valores de RGB normalizados al rango  $[0,1]$ ; la transformación del modelo RGB al modelo HSL se obtiene mediante las siguientes ecuaciones:

$$C_{max} = \max(R, G, B), \quad (3-15)$$

$$C_{min} = \min(R, G, B), \quad (3-16)$$

$$\Delta = C_{max} - C_{min}, \quad (3-17)$$

$$H = \begin{cases} 0^\circ & \text{si } \Delta = 0 \\ 60^\circ \times \left( \frac{G - B}{\Delta} \bmod 6 \right) & \text{si } C_{max} = R \\ 60^\circ \times \left( \frac{B - R}{\Delta} + 2 \right) & \text{si } C_{max} = G \\ 60^\circ \times \left( \frac{R - G}{\Delta} + 4 \right) & \text{si } C_{max} = B \end{cases} \quad (3-18)$$

$$S = \begin{cases} 0 & \text{si } \Delta = 0 \\ \frac{\Delta}{1 - |2L - 1|} & \text{si } \Delta \neq 0 \end{cases}, \quad (3-25)$$

$$L = \frac{(C_{max} + C_{min})}{2}. \quad (3-26)$$

La imagen resultante HSL tendrá H en el rango de  $[0^\circ, 360^\circ]$ , y S e V en el rango de  $[0, 1]$ .

### 3.1.6.6. Conversión de HSL a RGB

Dados los valores de H en el rango de  $[0^\circ, 360^\circ]$ , y los valores de S e I en el rango  $[0, 1]$ ; para la transformación del modelo HSV al modelo RGB se emplean las siguientes ecuaciones:

$$C = (1 - |2L - 1|) \times S, \quad (3-27)$$

$$X = C \times \left(1 - \left| \left( \frac{H}{60^\circ} \right) \bmod 2 - 1 \right| \right), \quad (3-28)$$

$$m = L - \frac{C}{2}, \quad (3-29)$$

$$(R, G, B) = \begin{cases} (C, X, 0) + m & \text{si } 0^\circ \leq H < 60^\circ \\ (X, C, 0) + m & \text{si } 60^\circ \leq H < 120^\circ \\ (0, C, X) + m & \text{si } 120^\circ \leq H < 180^\circ \\ (0, X, C) + m & \text{si } 180^\circ \leq H < 240^\circ \\ (X, 0, C) + m & \text{si } 240^\circ \leq H < 300^\circ \\ (C, 0, X) + m & \text{si } 300^\circ \leq H < 360^\circ. \end{cases} \quad (3-30)$$

La imagen resultante RGB estará normalizada con todos sus valores en el rango de  $[0, 1]$ .

### 3.1.6.7. Conversión de RGB a YCbCr

Con valores de RGB en el rango de  $[0, 255]$  la transformación del modelo RGB al modelo YCbCr se obtiene mediante las siguientes ecuaciones:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3-31)$$

La imagen resultante YCbCr estará con todos sus valores en el rango de  $[0, 255]$ .

### 3.1.6.8. Conversión de YCbCr a RGB

Dados los valores de YCbCr en el rango de  $[0, 255]$  la transformación del modelo YCbCr al modelo RGB se obtiene mediante las siguientes ecuaciones:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & 0.000 & 1.400 \\ 1.000 & -0.343 & -0.711 \\ 1.000 & 1.765 & 0.000 \end{bmatrix} \begin{bmatrix} Y \\ (Cb - 128) \\ (Cr - 128) \end{bmatrix} \quad (3-32)$$

La imagen resultante RGB estará con todos sus valores en el rango de [0, 255].

### 3.1.6.9. Conversión de RGB a CIE XYZ

Cuando los valores de RGB están normalizados al rango [0, 1]; la transformación del modelo RGB al modelo CIE XYZ se obtiene mediante las siguientes ecuaciones [20]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3-33)$$

### 3.1.6.10. CIE XYZ a RGB

Dados los valores de XYZ normalizados al rango [0, 1]; la transformación del modelo CIE XYZ al modelo RGB se obtiene mediante las siguientes ecuaciones:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.240479 & -1.537150 & -0.498535 \\ -0.969256 & 1.875992 & 0.041556 \\ 0.055648 & -0.204043 & 1.057311 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (3-34)$$

### 3.1.6.11. CIE XYZ A CIE LAB

Dados los valores de XYZ, la transformación del modelo CIE XYZ al modelo CIE LAB está dada por las siguientes ecuaciones:

$$L = \begin{cases} 116 \times \sqrt[3]{\left(\frac{Y}{Y_n}\right)} - 16 & \text{para } \left(\frac{Y}{Y_n}\right) > 0.008856 \\ 903.3 \times \left(\frac{Y}{Y_n}\right) & \text{cualquier otro caso} \end{cases} \quad (3-35)$$

$$A = \begin{cases} 500 \times \sqrt[3]{\left(\frac{X}{X_n}\right)} - \sqrt[3]{\left(\frac{Y}{Y_n}\right)} & \text{para } \left(\frac{Y}{Y_n}\right) > 0.008856 \\ 500 \times \left[ \left( 7.787 \times \left(\frac{X}{X_n}\right) + \frac{16}{116} \right) - \left( 7.787 \times \left(\frac{Y}{Y_n}\right) + \frac{16}{116} \right) \right] & \text{cualquier otro caso} \end{cases} \quad (3-36)$$

$$B = \begin{cases} 200 \times \left[ \sqrt[3]{\left(\frac{Y}{Y_n}\right)} - \sqrt[3]{\left(\frac{Z}{Z_n}\right)} \right] & \text{para } \left(\frac{Y}{Y_n}\right) > 0.008856 \\ 200 \times \left[ \left(7.787 \times \left(\frac{Y}{Y_n}\right) + \frac{16}{116}\right) - \left(7.787 \times \left(\frac{Z}{Z_n}\right) + \frac{16}{116}\right) \right] & \text{cualquier otro caso} \end{cases} \quad (3-37)$$

Donde  $X_n$ ,  $Y_n$  y  $Z_n$  conforman la triada de valores del blanco de referencia.

### 3.1.6.12. CIE LAB a CIE XYZ

Dados los valores de LAB, la transformación del modelo CIE LAB al modelo CIE XYZ está dada por las siguientes ecuaciones, para  $\left(\frac{Y}{Y_n}\right) > 0.008856$ :

$$X = X_n \times \left( \frac{L + 16}{116} + \frac{A}{500} \right)^3 \quad (3-38)$$

$$Y = Y_n \times \left( \frac{L + 16}{116} \right)^3 \quad (3-39)$$

$$Z = Z_n \times \left( \frac{L + 16}{116} - \frac{b}{500} \right)^3 \quad (3-40)$$

## 3.2. Descriptores

En el caso de la CBIR un descriptor es por lo general un vector, un conjunto de números, que representa el contenido de una imagen; dicho descriptor permite comparar imágenes velozmente, y localizar las que son similares entre sí. Existen diversas formas de obtener el descriptor de imagen basadas en color, forma y textura. En seguida se describen algunos descriptores basados en color y forma.

### 3.2.1. Basados en color

Como se ha mencionado anteriormente, el color es una de las características más importantes de una imagen, por lo cual es uno de los rasgos más utilizados en el área de la búsqueda por contenido CBIR. En los siguientes párrafos se describen descriptores basados en este elemento.

### 3.2.1.1. Histograma de Color Convencional (CCH)

Este descriptor representa la distribución de colores en una imagen. En el caso de las imágenes digitales lo que indica es el número de píxeles que corresponden a cada color, o en su caso a cada uno de los rangos seleccionados. El histograma es una función discreta:

$$h(c_k) = n_k, \quad (3-41)$$

donde  $k = 0, 1, \dots, L - 1$ ,  $L$  es el número de colores determinados para calcular el histograma,  $c_k$  es la triada correspondiente al  $k$ -ésimo color, y  $n_k$  es el número de píxeles en la imagen que son de color  $c_k$ .

Es una práctica muy frecuente normalizar el histograma dividiéndolo entre el total de píxeles de la imagen, dando la probabilidad de ocurrencia del color  $k$ :

$$p(c_k) = \frac{n_k}{n}, \quad (3-42)$$

donde  $n$  es el número total de píxeles de la imagen,  $\sum_{k=0}^{L-1} h(c_k) = n$  y  $\sum_{k=0}^{L-1} p(c_k) = 1$ .

Como ejemplo de lo anterior en la Figura 3-11 inciso *a*) se muestra una imagen RGB de 10x10 píxeles, es decir  $n = 100$ , con  $L = 5$  colores, y en el inciso *b*) se muestra el histograma de color de dicha imagen. El CCH de este ejemplo sería  $h = [36, 20, 28, 12, 4]$ .

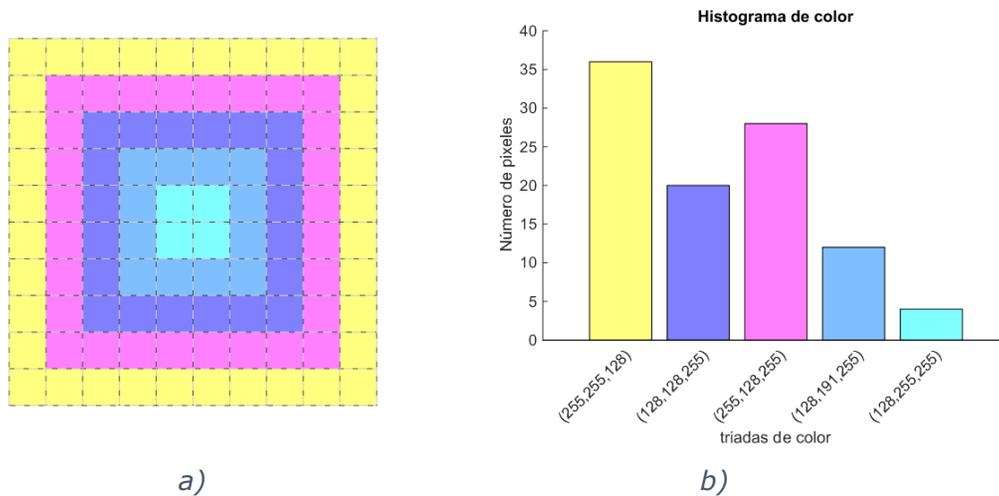


Figura 3-11 Imagen RGB con su histograma

El histograma de color es invariante a la traslación, a la rotación alrededor de un eje perpendicular a la imagen. Y cambia sólo lentamente con respecto a la rotación en los otros ejes, a la oclusión, y al cambio de la distancia al objeto [21].

### 3.2.1.2. Color Dominante (DCD)

El Descriptor de Color Dominante (*Dominant Color Descriptor*) proporciona la distribución de los colores sobresalientes de la imagen [22]. El DCD es un descriptor de color compacto, ya que en él solamente se recopilan los colores dominantes (máximo  $N$  colores) en lugar de un histograma de color. El DCD almacena los colores dominantes, sus porcentajes, y varios parámetros opcionales [23].

$$F_I = \{(c_i, p_i), i = 1, \dots, N\}, \quad (3-43)$$

donde  $F_I$  es el Descriptor del Color Dominante de la imagen  $I$ ,  $c_i$  es la triada del  $i$ -ésimo color dominante,  $p_i$  es el porcentaje del  $i$ -ésimo color dominante, y  $N$  es el número total de colores dominantes donde  $\sum_{i=1}^N p_i = 1$ .

Para extraer los colores dominantes de una imagen, se debe predeterminedir un algoritmo de cuantización de color, uno de los más usados es el Algoritmo de Lloyd Generalizado (GLA) [24, 25], también conocido como k-medias. Sin embargo existen otros métodos tales como el Algoritmo de Bloque Lineal (LBA) [26], y mediante el histograma de color en el espacio HSV cuantizado con 72 colores [27].

Como ejemplo de lo anterior en la

Figura 3-12 se muestra una imagen de 256x384 pixeles, y la misma imagen representada solamente con  $N=4$  colores dominantes, extraídos con LBA.



Figura 3-12 Imagen 50.jpg<sup>9</sup> con sus colores  $N$  dominantes.

Las triadas de los colores dominantes de la imagen anterior son  $c = \{(79 \ 40 \ 34), (174 \ 66 \ 59), (217 \ 162 \ 100), (215 \ 191 \ 176)\}$  y los porcentajes de cada triada de color son  $p = \{0.4059 \ 0.3589 \ 0.1028 \ 0.1324\}$ .

<sup>9</sup> De la base de Datos de Wang (image.orig)

Por lo que el descriptor de color dominante, de la imagen de color dominante LBA sería:

$$F = \left\{ \begin{array}{l} ((79 \ 40 \ 34), 0.4059), \\ ((174 \ 66 \ 59), 0.3589), \\ ((217 \ 162 \ 100), 0.1028), \\ ((215 \ 191 \ 176), 0.1324) \end{array} \right\}$$

La primera vez que se propuso el DCD se trabajó en el modelo de color CIE LAB, sin embargo de acuerdo a numerosos experimentos, la selección del modelo de color no es un parámetro crítico para la extracción del DCD [28].

### ***3.2.1.3. Esquema de Color (CLD)***

El Descriptor de Esquema de Color (*Color Layout Descriptor*) especifica la distribución espacial de los colores [29]. La extracción de este descriptor se inicia con una imagen en el modelo RGB, y se obtiene mediante los siguientes pasos [30]:

- I. Se divide la imagen en 64 regiones, dividiendo entre 8 el largo y el ancho de la imagen. En la Figura 3-13 se utiliza una imagen de 384x256 pixeles.
- II. Para cada región se selecciona un color dominante, o color representativo, este único color se obtiene promediando el color de cada pixel de la región en cuestión. De esta manera se obtiene una imagen pequeña de 8x8 pixeles.
- III. La imagen reducida se transforma al modelo de color YCbCr, dando valores enteros en el rango de [0, 255].
- IV. Se aplica la transformada de coseno discreta (DCT [31]) a cada canal de la imagen de 8x8.
- V. Se extraen 6 coeficientes de la matriz DCT(Y), 3 de la DCT(Cb) y 3 de la DCT(Cr), haciendo una lectura en cada matriz de acuerdo a un patrón de zigzag.

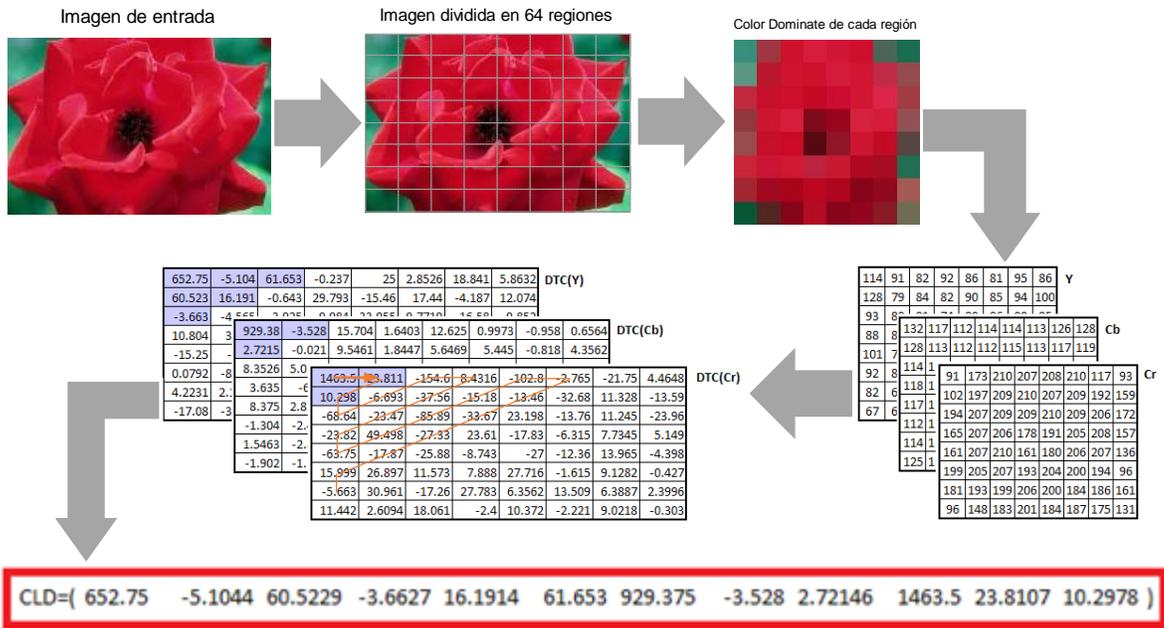


Figura 3-13 Proceso de extracción del CLD de la imagen 696.jpg<sup>10</sup>

### 3.2.1.4. Correlograma de Color (CC) y Auto Correlograma de Color (ACC)

El correlograma de color de una imagen es una tabla indexada por pares de colores, donde la k-ésima entrada (i,j) especifica la probabilidad de encontrar un pixel de color 'j' a una distancia 'k', de un pixel de color 'i' [10, 31]. En otras palabras representa el cambio de la correlación espacial entre las parejas de colores con respecto a la distancia. Se expresa mediante la siguiente ecuación:

$$\gamma_{c_i, c_j}^{(k)}(I) \triangleq \Pr_{p_1 \in I_{c_i}, p_2 \in I_{c_j}} [p_2 \in I_{c_j} | |p_1 - p_2| = k], \quad (3-44)$$

donde  $(i, j) \in \{1, 2, \dots, m\}$  son los colores cuantizados,  $k \in \{1, 2, \dots, d\}$  son las distancias, las cuales pueden no ser consecutivas,  $p_1$  y  $p_2$  son los pixeles de la imagen  $I$ , e  $I_{c_j}$  y  $I_{c_i}$  son el color  $c_j$  y el color  $c_i$ , correspondientemente, de la imagen  $I$ .

Se puede calcular el correlograma con la matriz de coocurrencia  $\Gamma_{c_i, c_j}^{(k)}(I)$  mediante la siguiente ecuación:

$$\gamma_{c_i, c_j}^{(k)}(I) = \frac{\Gamma_{c_i, c_j}^{(k)}(I)}{h_{c_i}^{(I)} \times 8k}, \quad (3-45)$$

<sup>10</sup> De la base de Datos de Wang (image.org)

donde

$$\Gamma_{c_i c_j}^{(k)}(I) \triangleq \left\{ \left\{ p_1 \in I_{c_i}, p_2 \in I_{c_j} \mid |p_1 - p_2| = k \right\} \right\} \quad (3-46)$$

y  $h_{c_i}(I)$  es el número de pixeles que son del color  $c_i$  de la imagen  $I$ .

El correlograma se puede obtener de una imagen en escala de grises, o de una imagen a color. La extracción del descriptor para una imagen a color  $I$  de  $N \times M = n$  pixeles se obtiene de la siguiente manera:

- I. De ser necesario se cuantiza la imagen a  $m$  colores  $c_1, c_2, \dots, c_m$ , de tal forma que si  $p \in I, p \in c$ , a cada pixel de la imagen 'I' le corresponde un color 'c'. En el caso de la imagen de la Figura 3-14 este paso no es necesario, ya que es bicolor.
- II. Se convierten las triadas de color a índices para obtener una sola matriz en lugar de las tres que tiene una imagen a color.
- III. Se calcula el histograma y las matrices de coocurrencia [33], a  $d$  distancias  $k_1, k_2, \dots, k_d$ , en las cuatro direcciones ( $0^\circ, 45^\circ, 90^\circ$  y  $135^\circ$ ), y en ambos sentidos (positivo y negativo). En el ejemplo de la Figura 3-14 los valores de las distancias son consecutivos y el número de distancias  $d = 8$ .
- IV. Se emplea la fórmula ( 3-45 ) para calcular los correlogramas correspondientes.
- V. Se acomodan los elementos de los correlogramas en un vector para obtener un descriptor que facilite su posterior comparación con los descriptores de otras imágenes.

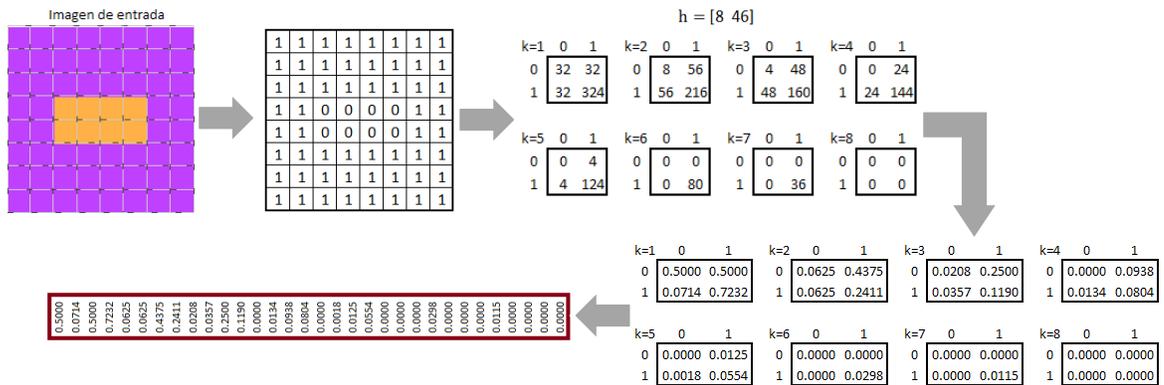


Figura 3-14 Proceso de extracción del CC de una imagen de 8x8 pixeles.

El auto correlograma ACC de la imagen  $I$  captura la correlación espacial de las parejas de colores idénticos, es decir  $i = j$ , se representa con la siguiente ecuación:

$$\alpha_c^{(k)}(I) \triangleq \gamma_{c,c}^{(k)} \quad (3-47)$$

Para obtenerlo se toman los valores de las diagonales de los correlogramas de una imagen, obtenidos en el punto IV del proceso de extracción del correlograma, por lo que para varias distancias se obtiene un vector por cada color de la imagen. En la Figura 3-15 se muestran dos imágenes cuyos histogramas son iguales, por lo que para distinguirlas es conveniente el uso del ACC. En la gráfica de la Figura 3-15; las líneas son del mismo color de los píxeles en la imagen, adicionalmente los vértices de la imagen A están marcados con asteriscos azules, y los de la imagen B están indicados con estrellas negras.

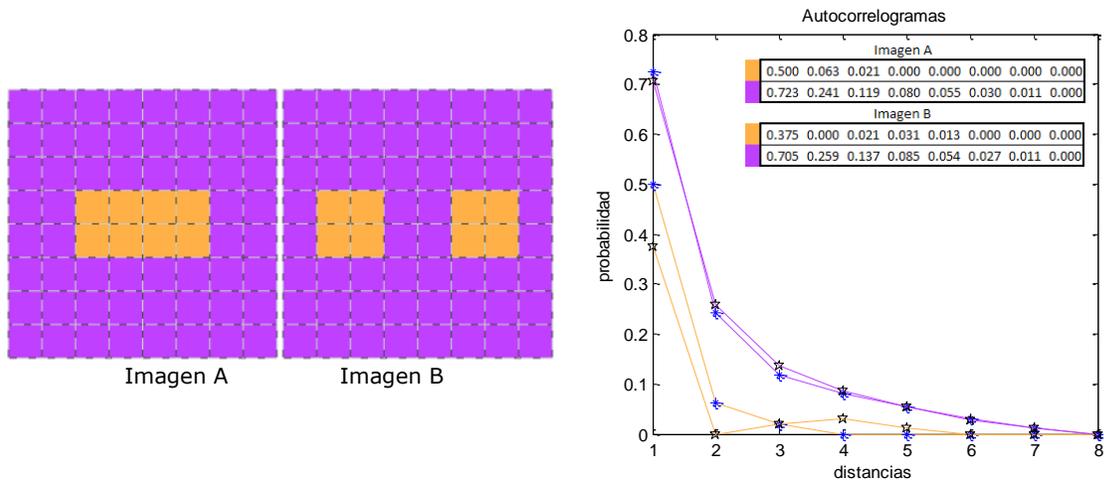


Figura 3-15 Imágenes con sus Auto Correlogramas

### 3.2.1.5. Correlograma Compacto Generalizado (CGC)

El descriptor del Correlograma Compacto Generalizado (*Compact Generalized Correlogram*) se basa conceptualmente en la definición de correlograma. Para obtener el CGC se sigue un procedimiento similar al del ACC, el cual se ilustra en la Figura 3-16, y se explica a continuación:

- I. De ser necesario se cuantiza la imagen a  $m$  colores, para disminuir el número de colores de la imagen, exactamente como en el correlograma del ACC.
- II. Se convierten las triadas de color a índices, igual que en el correlograma del ACC.
- III. Se calcula una matriz de coocurrencia única en una sola dirección ( $0^\circ$ ), en ambos sentidos (positivo y negativo), y a una distancia única que se selecciona entre el 10% y 50% de la dimensión más pequeña de la imagen; por ejemplo,

si la dimensión más pequeña es de 20 píxeles, la distancia podría ser  $k$  de entre 2 y 10 píxeles.

IV. Se calcula el correlograma (CC) dividiendo la matriz de coocurrencia entre el total de parejas de la matriz de coocurrencia.

$$\gamma_{c_i,c_j}^{(k)}(I) = \frac{\Gamma_{c_i,c_j}^{(k)}(I)}{\sum_{i,j} \Gamma_{c_i,c_j}^{(k)}(I)} \quad (3-48)$$

V. Se extrae la matriz triangular superior de CC y se le suman los valores de la matriz triangular inferior sin la diagonal, con el objetivo de no duplicar los valores de la diagonal; de esta manera se obtiene el CGC.

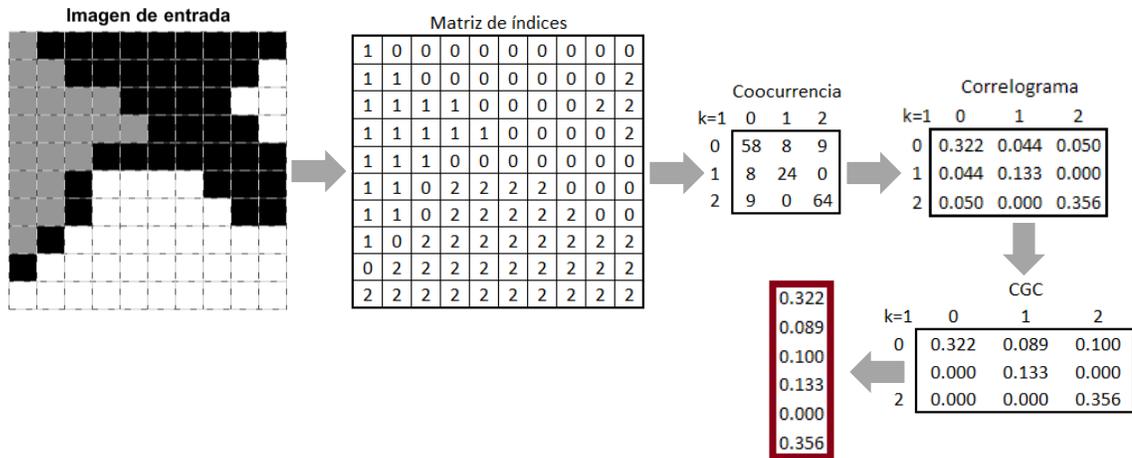


Figura 3-16 Proceso de extracción del CGC

### 3.2.1.6. Auto Correlograma de Colores Dominantes (DCCD)

Este descriptor (*Dominat Color Correlogram Descriptor*) se obtiene al combinar el DCD con el ACC, primero se obtiene una imagen con  $N = 8$  colores dominantes, y después se calcula el correlograma de dichos colores dominantes. Para obtener este descriptor se siguen los siguientes pasos:

I. Se obtiene el DCD de la imagen de entrada mediante el histograma de color en el espacio HSV cuantizado con 72 colores [27], en adelante referido como DCD-HSV.

El procedimiento de extracción del DCD-HSV se muestra en la

Figura 3-17, donde se puede apreciar que todos los píxeles de la imagen que no corresponden a un color dominante son igualados a -1, por lo cual se ven de color negro cuando se representan en el modelo RGB.

II. Se extrae el ACC con una sola distancia  $k=1$  de la matriz de índices de colores dominantes de la imagen. Para esta operación no se consideran los elementos iguales a -1 en la matriz de índices de colores dominantes.

Para obtenerlo se considera la matriz de índices obtenida en el punto anterior, se calcula la matriz de coocurrencia, se elimina la fila y columna correspondientes a los elementos iguales a -1, se calcula el correlograma, y finalmente se obtiene el auto correlograma de los colores dominantes (DCCD). Este proceso se ilustra en la Figura 3-18.



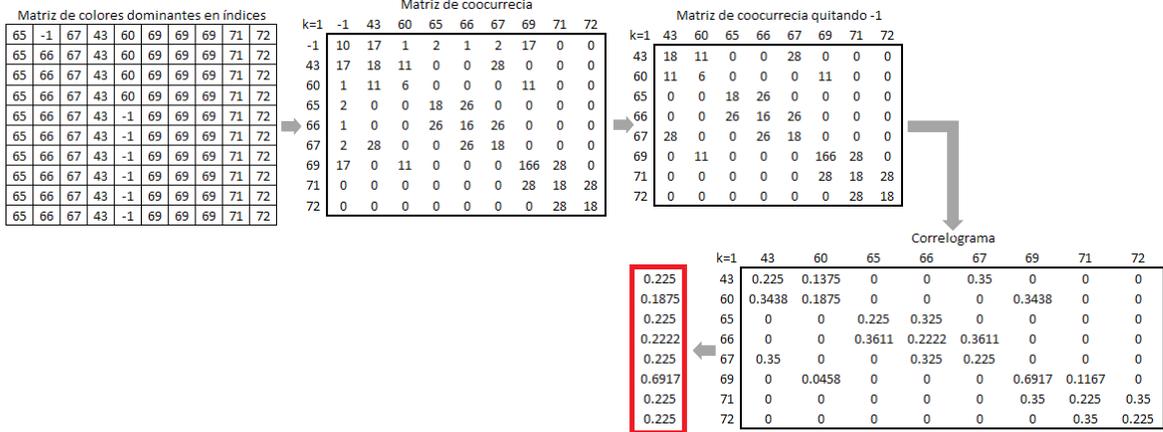


Figura 3-18 Extracción de DCCD a partir de la matriz de colores dominantes representados mediante índices.

### 3.3. Medidas de similitud

Una vez que se han obtenido los descriptores correspondientes, la extracción o recuperación de imágenes se hace mediante las medidas de similitud. Existen diversas medidas de similitud que se explicarán a continuación.

#### 3.3.1. Norma $L_p$

La norma  $L_p$ , también conocida como distancia Minkowski se representa mediante la siguiente ecuación:

$$d_p(Q, T) = \left( \sum_{i=0}^{N-1} (Q_i - T_i)^p \right)^{\frac{1}{p}}, \quad (3-49)$$

donde  $Q$  y  $T$  son los vectores descriptores de la imagen de búsqueda y la imagen de la base de datos respectivamente,  $Q = \{Q_0, Q_1, Q_2, \dots, Q_{N-1}\}$ ,  $T = \{T_0, T_1, T_2, \dots, T_{N-1}\}$  y  $N$  es el número de elementos del descriptor, o el tamaño del vector descriptor.

En esta medida la similitud es mayor, entre menor sea la diferencia entre los vectores  $Q$  y  $T$ .

### 3.3.1.1. Norma L1

La norma L1, también llamada distancia Manhattan, *cityblock*, o *taxicab*, es un caso particular de la Norma Lp, donde p=1.

$$d_1(Q, T) = \left( \sum_{i=0}^{N-1} (Q_i - T_i)^1 \right)^{\frac{1}{1}} = \sum_{i=0}^{N-1} |Q_i - T_i|. \quad (3-50)$$

### 3.3.1.2. Norma L2

La norma L2, también conocida como distancia Euclidiana, es un caso particular de la Norma Lp, donde p=2.

$$d_2(Q, T) = \left( \sum_{i=0}^{N-1} (Q_i - T_i)^2 \right)^{\frac{1}{2}}. \quad (3-51)$$

### 3.3.1.3. Norma L $\infty$

La norma L $\infty$ , también conocida como distancia de Chebyshev, es un caso particular de la Norma Lp, donde p tiende a  $\infty$ .

$$d_\infty(Q, T) = \left( \sum_{i=0}^{N-1} (Q_i - T_i)^\infty \right)^{\frac{1}{\infty}} = \max_{0 \leq i \leq N-1} \{|Q_i - T_i|\}. \quad (3-52)$$

## 3.3.2. Intersección de histograma

La intersección de histograma es capaz de manejar coincidencias parciales cuando el tamaño del objeto (con descriptor Q) es menor que el tamaño de la imagen (con descriptor T). La definición original de la intersección de histograma está dada por [34]:

$$d_{hi}(Q, T) = 1 - \frac{\sum_{i=0}^{N-1} \min(Q_i, T_i)}{|Q|}. \quad (3-53)$$

También se puede utilizar una versión simplificada [35]:

$$d_{hi}(Q, T) = \sum_{i=1}^N \min(Q_i, T_i). \quad (3-54)$$

Por otro lado, se ha extendido a una métrica de distancia como [34]:

$$d_{hi}(Q, T) = 1 - \frac{\sum_{i=0}^{N-1} \min(Q_i, T_i)}{\min(|Q|, |T|)}. \quad (3-55)$$

Y una manera alternativa de calcularlo es [35]:

$$d_{hi}(Q, T) = \frac{1}{2} \sum_{i=1}^N (Q_i + T_i - |Q_i - T_i|). \quad (3-56)$$

En esta medida el valor máximo de la similitud está dado por  $d_{hi}(Q, T) = 1$ . Entre más cercano a cero sea el valor de  $d_p(Q, T)$ , mayor será la diferencia entre los vectores  $Q$  y  $T$ .

### 3.3.3. Distancia de Color Dominante

Esta medida de similitud se calcula entre los Descriptores de Color Dominante, de dos imágenes  $q$  y  $t$ ,  $F_q = \{(c_i, p_i), i = 1, \dots, N_1\}$  y  $F_t = \{(c_j, p_j), i = 1, \dots, N_2\}$  como [26]:

$$d_{DCD}(F_1, F_2) = 1 - SIM(F_1, F_2), \quad (3-57)$$

$$SIM(F_1, F_2) = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} a_{i,j} S_{i,j}, \quad (3-58)$$

$$a_{i,j} = \begin{cases} 1 - \frac{d_{i,j}}{d_{max}} & \text{si } d_{i,j} \leq T_d, \\ 0 & \text{si } d_{i,j} > T_d \end{cases}, \quad (3-59)$$

$$S_{i,j} = [1 - |p_q(i) - p_t(j)|] \times \min(p_q(i), p_t(j)), \quad (3-60)$$

donde  $p_q(i)$  es el porcentaje del  $i$ -ésimo color dominante del descriptor  $F_q$ ,  $p_t(j)$  es el porcentaje del  $j$ -ésimo color dominante del descriptor  $F_t$ , el término  $\min(p_q(i), p_t(j))$  representa la similitud entre los porcentajes de los colores dominantes, el término  $[1 - |p_q(i) - p_t(j)|]$  mide la diferencia del porcentaje de dos colores,  $d_{i,j}$  es la distancia euclidiana entre dos triadas de color, que representan cada color dominante, y  $T_d$  es el umbral de la distancia máxima usada para juzgar si dos colores dominantes son similares, o no.

$$d_{max} = \alpha T_d, \quad (3-61)$$

en [26]  $T_d = 25$  y el parámetro  $\alpha = 2$ .

En esta medida las imágenes son iguales cuando  $d_{DCD}(F_1, F_2) = 0$ , entre mayor sea el valor de  $d_{DCD}(F_1, F_2)$  mayor será la diferencia entre los descriptores  $F_1$  y  $F_2$ .

### 3.3.4. Distancia de Esquema de Color

La distancia entre descriptores  $I$  e  $I'$  de esquema de color (CLD) se calcula de la siguiente manera [29]:

$$d_{CLD}(I, I') = \sqrt{\sum_{i \in (Y)} w1_i (Y_i - Y_i')^2} + \sqrt{\sum_{i \in (Cb)} w2_i (Cb_i - Cb_i')^2} + \sqrt{\sum_{i \in (Cr)} w3_i (Cr_i - Cr_i')^2}, \quad (3-62)$$

Donde  $Y_i$  es el  $i$ -ésimo coeficiente de  $Y$ ,  $Cb_i$  es el  $i$ -ésimo coeficiente de  $Cb$ ,  $Cr_i$  es el  $i$ -ésimo coeficiente de  $Cr$ , y  $w1_i, w2_i, w3_i$  son los valores del peso para cada  $i$ -ésimo coeficiente respectivamente.

Los pesos según el estándar se muestran en la Tabla 3-2.

Tabla 3-2 Pesos usados para el CLD

		Coeficientes de cada matriz $i$					
		1	2	3	4	5	6
Pesos	$w1_i$	2	2	2	1	1	1
	$w2_i$	2	1	1			
	$w3_i$	4	2	2			

En esta medida las imágenes son iguales cuando  $d_{CLD}(I, I') = 0$ , entre mayor sea el valor de  $d_{CLD}(I, I')$ , mayor será la diferencia entre los descriptores  $I$  e  $I'$ .

### 3.3.5. Distancia de Correlograma

Mide la diferencia espacial entre dos descriptores,  $Q$  e  $I$ , mediante el correlograma de color [32]:

$$d_{Corr}(Q, I) = \sum_{i \in c}^m \sum_{j \in c}^m \sum_k^d \left\{ \begin{array}{ll} 0 & \text{si } \gamma_{c_i, c_j}^k(Q) = \gamma_{c_i, c_j}^k(I) = 0 \\ \left| \frac{\gamma_{c_i, c_j}^k(Q) - \gamma_{c_i, c_j}^k(I)}{\gamma_{c_i, c_j}^k(Q) + \gamma_{c_i, c_j}^k(I)} \right| & \text{cualquier otro caso} \end{array} \right\}, \quad (3-63)$$

donde  $m$  es el número total de colores en la imagen,  $d$  es la distancias entre los pixeles, y  $\gamma_{c_i, c_j}^k$  es la probabilidad de que dado un pixel de color  $c_i$  exista un pixel de color  $c_j$  a una distancia  $k$ .

En esta medida las imágenes son iguales cuando  $d_{Corr}(Q, I) = 0$ , entre más cercano sea el valor de  $d_{Corr}(Q, I)$  a 1, mayor será la diferencia entre los descriptores  $Q$  e  $I$ .

### 3.4. Métricas de evaluación

Las métricas más ampliamente utilizadas para evaluar el desempeño de los algoritmos de recuperación de imágenes se describen a continuación.

#### 3.4.1. Característica Operativa del Receptor (ROC)

El análisis ROC por sus siglas en inglés, *Receiver Operating Characteristic*, es una herramienta gráfica útil para evaluar la exactitud de un modelo estadístico. Para obtener las curvas ROC en la Figura 3-19 se muestra la gráfica de la {1-especificidad} (FPR) Razón de Falsos Positivos, contra la sensibilidad (TPR) Razón de Verdaderos Positivos.

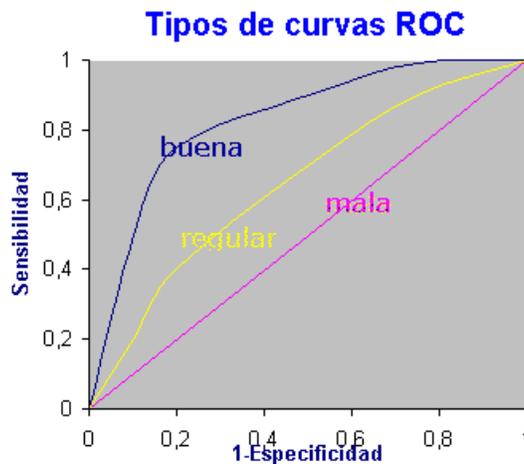


Figura 3-19 Tipos de Curvas ROC<sup>11</sup>

En la

Tabla 3-3 se muestra la matriz de confusión de 2x2 (celdas grises claro) para un evento de clasificación binaria como es la recuperación de imágenes (reconocido, no reconocido), conformada por cuatro eventos: Verdadero positivo (VP), Falso positivo (FP), Falso negativo (FN) y Verdadero negativo (VN). A partir de estos eventos se pueden obtener diferentes métricas de evaluación como se puede observar en la misma tabla.

<sup>11</sup> Tomado de [http://www.hrc.es/bioest/roc\\_1.html](http://www.hrc.es/bioest/roc_1.html)

Tabla 3-3 Matriz de confusión y algunas métricas de evaluación derivadas.

		Condición			
		positiva (p)	negativa (n)	Predominio $\frac{VP + FN}{Total}$	
Recuperado	Total				
	positivo (p')	Verdadero Positivo (VP)	Falso Positivo (FP) Error tipo I	Valor Predictivo Positivo (PPV) o Precisión $\frac{VP}{VP + FP}$	Razón de falsos descubrimientos (FDR) $\frac{FP}{VP + FP}$
	Negativo (n')	Falso Negativo (FN) Error tipo II	Verdadero Negativo (VN)	Razón de Falsas Omisiones (FOR) $\frac{FN}{FN + VN}$	Valor Predictivo Negativo (NPV) $\frac{VN}{FN + VN}$
	Exactitud (ACC) $\frac{VP + VN}{Total}$	Razón de Verdaderos Positivos (TPR), recall o sensibilidad $\frac{VP}{VP + FN}$	Razón de Falsos Positivos (FPR) o <u>fall-out</u> $\frac{FP}{FP + VN}$	Razón de Similitud Positiva (LR+) $\frac{TPR}{FPR}$	Razón de probabilidades de Diagnóstico (DOR) $\frac{LR +}{LR -}$
	Razón de Falsos Negativos (FNR) o Razón de pérdida $\frac{FN}{VP + FN}$	Razón de Verdaderos Negativos (TNR) o Especificidad (SPC) $\frac{VN}{FP + VN} = 1 - FPR$	Razón de similitud Negativa (LR-) $\frac{FNR}{TNR}$		

Para la métrica ROC [36] se utiliza la Razón de Verdaderos Positivos (TPR), sensibilidad, o *recall*:

$$TPR = \frac{VP}{VP + FN} \quad (3-64)$$

La Razón de Falsos Positivos (FPR), o *fall-out*, que es igual al complemento de la Especificidad, o Razón de Verdaderos Negativos (TNR)

$$especificidad = \frac{VN}{FP + VN} = 1 - FPR, \quad (3-65)$$

$$FPR = \frac{FP}{FP + VN} = 1 - especificidad. \quad (3-66)$$

Para graficar una curva ROC es necesario considerar diferentes cantidades de p' (recuperado como positivo). Es decir, por ejemplo, en una consulta con una imagen

determinada, para una base de datos de 200 imágenes, podemos considerar las primeras 10, 20, 50, 100, y 200 imágenes obteniendo así 5 matrices de confusión, para calcular 5 diferentes valores de TPR y de FPR. Los cuales podemos graficar para obtener la curva ROC de dicha consulta.

### ***3.4.2. Precisión Promedio de Recuperación (ARP)***

La métrica ARP (*Average Retrieval Precision*), como lo indica su nombre es el porcentaje promedio y nos da información de qué tan preciso es un descriptor.

La ecuación utilizada para calcularla es:

$$ARP = \frac{1}{N_B} \sum_{b=1}^{N_B} PR(b), \quad (3-67)$$

donde  $N_B$  es el número de búsquedas, o consultas realizadas, que puede ser igual al número total de imágenes en la base de datos; y  $PR(b)$  es la precisión de recuperación (*retrieval precision*) de la consulta  $b$ , y se define como:

$$PR(b) = \frac{VP(b)}{N_R(\alpha, b)}, \quad (3-68)$$

donde  $VP(b)$  son los verdaderos positivos, que se refiere al número de imágenes recuperadas que son relevantes en la consulta  $b$ ; y  $N_R(\alpha, b)$  es el número de imágenes que se considera como recuperadas en la consulta  $b$ , y se establece como:

$$N_R(\alpha, b) = \alpha \times N_G(b), \quad (3-69)$$

donde  $N_G(b)$  es el número de imágenes de la categoría de la búsqueda  $b$  en la base de datos, usualmente llamadas imágenes de *ground truth*, y  $\alpha$  indica la fracción del número de las imágenes de *ground truth* que se considera como extraídas. Los valores más ampliamente usados son:

$$\alpha = [1 \ 0.5 \ 0.25].$$

La métrica ARP puede tomar valores en el rango  $[0 \ 1]$ , cuando  $ARP = 1$  significa que el descriptor evaluado realizó una extracción perfecta.

### ***3.4.3. Tasa promedio de Recuperación (ARR)***

La métrica ARR (*Average Retrieval Rate*) se calcula con la siguiente ecuación:

$$ARR(b) = \frac{1}{N_B} \sum_{b=1}^{N_B} RR(b) , \quad ( 3-70 )$$

donde  $N_B$  es el número de búsquedas, o consultas, realizadas y puede ser igual al número total de imágenes en la base de datos;  $RR(b)$  es la tasa de recuperación (*retrieval rate*) de la consulta  $b$ , y se define como:

$$RR(b) = \frac{VP(b)}{N_G(b)} , \quad ( 3-71 )$$

donde  $VP(b)$  son los verdaderos positivos, se refiere al número de imágenes recuperadas que son relevantes en la consulta  $b$ , dentro de las primeras  $\alpha \times N_G(b)$  imágenes; y  $N_G(b)$  es el número de imágenes de la categoría de la búsqueda  $b$  en la base de datos, usualmente llamadas imágenes de *ground truth*.

La métrica ARR puede tomar valores en el rango [0 1], el desempeño del descriptor evaluado es mejor cuando  $ARR = 1$ .

### 3.4.4. Rango Promedio de Recuperación Normalizado Modificado (ANMRR)

La métrica ANMRR [22] (*Average Normalized Modified Retrieval Rank*), su rango de valores es de [0, 1], entre menor sea su valor, mejor es el desempeño del descriptor evaluado, y se calcula con la siguiente ecuación:

$$ANMRR(b) = \frac{1}{N_B} \sum_{b=1}^{N_B} NMRR(b) , \quad ( 3-72 )$$

donde  $N_B$  es el número de búsquedas o consultas realizadas, que puede ser igual al número total de imágenes en la base de datos; y  $NMRR(b)$  = Rango de Recuperación Normalizado Modificado (*normalized modified retrieval rank*) de la consulta  $b$ , y se define como:

$$NMRR(b) = \frac{MRR(b)}{K + 0.5 - 0.5N_G(b)} , \quad ( 3-73 )$$

donde  $N_G(b)$  es el número de imágenes de la categoría de la búsqueda  $b$  en la base de datos, usualmente llamadas imágenes de *ground truth*;  $MRR(b)$  es el Rango Modificado de Recuperación (*normalized modified retrieval rank*) de la consulta  $b$ , y se define como:

$$MRR(b) = AVR(b) - 0.5[1 + N_G(b)] , \quad ( 3-74 )$$

donde  $AVR(b)$  es el Rango Promedio (*average rank*) de la consulta  $b$ , y se define como:

$$AVR(b) = \frac{1}{N_G(b)} \sum_{k=1}^{N_G(b)} R(k) , \quad ( 3-75 )$$

donde  $R(k)$  es el Rango calculado y se define como:

$$R(k) = \begin{cases} \text{Rango}(k) & \text{si } \text{Rango}(k) \leq K(b) \\ 1 + K & \text{si } \text{Rango}(k) > K(b) \end{cases} \quad ( 3-76 )$$

donde  $\text{Rango}(k)$  es el número correspondiente al lugar que ocupa la  $k$ -ésima imagen *ground truth* en la búsqueda  $b$ ; y  $K(b)$  = Número seleccionado para especificar los rangos que son relevantes en la búsqueda  $b$ ,  $K \geq N_B$ , para  $N_B > 20$  se puede considerar  $K = 2 \times N_G$  si todas las  $N_G(b)$  son iguales. En MPEG-7  $K = \min(4 \times N_G(b), 2 \times N_G(\text{total}))$ , donde  $N_G(\text{total}) = \max(N_G(b))$ , de todas las categorías de la base de datos.

## 4. Propuesta

En esta sección se describirá el algoritmo propuesto y una medida de similitud para el mismo.

### 4.1. Algoritmo

Se estudia y define un descriptor de color dominante en el espacio RGB, sin utilizar un método recursivo. También se explican las combinaciones con algunos métodos para obtener el correlograma de color.

El procedimiento para extraer el color dominante RGB consiste en los siguientes pasos:

- I. Se precuantiza la imagen en 125 colores, dividiendo cada eje en 5 regiones. Dando un total de 125 combinaciones donde se calcula el histograma obteniendo el promedio de los pixeles en cada región.
- II. Se eligen los  $N=8$  colores con mayor porcentaje.
- III. Se obtiene la imagen de colores dominantes. Haciendo el siguiente agrupamiento: cada pixel de color no dominante se asigna al más cercano de los 'N' colores con mayor porcentaje.

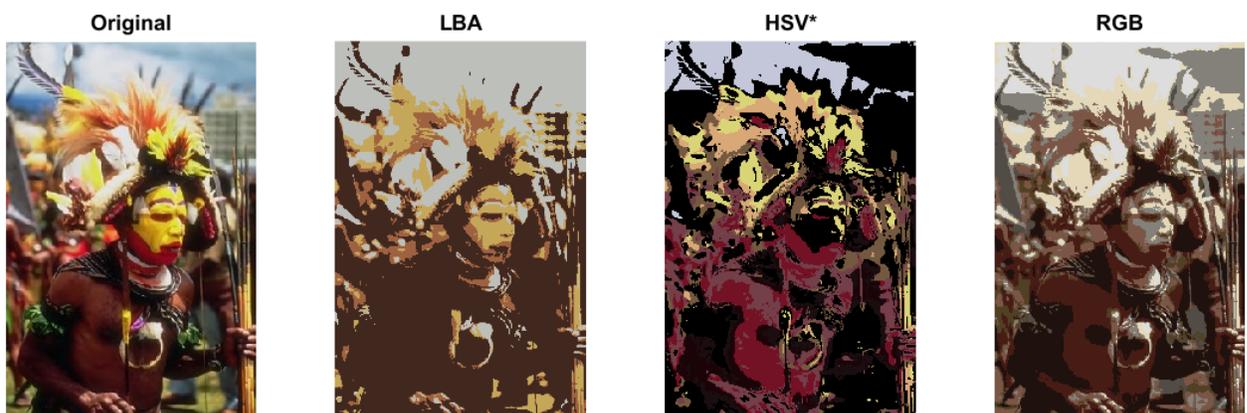


Figura 4-1 Imagen original 15.jpg<sup>12</sup> y sus imágenes de Colores dominantes obtenidas mediante tres diferentes representaciones en espacios de color.

<sup>12</sup> De la base de Datos de Wang (image.org)

En la Figura 4-1 se pueden observar los resultados del algoritmo para obtener los colores dominantes en tres representaciones de espacios de color diferentes. Las representaciones LBA [26] y el RGB trabajan con la imagen en el espacio RGB; en cambio la representación HSV [27] hace una conversión de la imagen, por lo que la imagen mostrada se regresó al espacio RGB para poder visualizarla, denotándola HSV\*.

IV. Una vez que obtenemos la imagen de colores dominantes utilizamos el Correlograma de color, el Auto correlograma (3.2.1.4) y el Correlograma Compacto Generalizado (3.2.1.5), para así obtener el Correlograma de colores dominantes (CDCD), el Auto correlograma de colores dominantes (ADCD), Correlograma Compacto Generalizado de colores dominantes (CGCDCD) y el Auto Correlograma Compacto Generalizado de Colores Dominantes (ACGCDCD). En los cálculos, estos descriptores se obtienen utilizando las distancias de uno y dos pixeles.

En el presente trabajo los correlogramas y los auto correlogramas, CDCD, CGCDCD, ADCD Y ACGCDCD se obtuvieron a las distancias de 1 pixel y 2 pixeles a partir de las imágenes de colores dominantes en el espacio RGB. El DCCD se obtuvo a la distancia de 1 pixel a partir de las imágenes de colores dominantes en el espacio HSV.

## ***4.2. Distancia de Auto correlograma de color dominante (ADCDD)***

Para llevar a cabo el análisis del desempeño se emplean tanto las triadas de los colores dominantes como los valores de los auto correlogramas. Para ello se propone emplear la medida de similitud del descriptor de color dominante, pero utilizando los valores del auto correlograma en lugar de los porcentajes. Los cálculos se hacen únicamente entre los Auto Correlogramas de Colores Dominantes, de dos imágenes  $q$  y  $t$ :  $F_q = \{(c_i, acc_i), i = 1, \dots, N_1\}$  y  $F_t = \{(c_j, acc_j), i = 1, \dots, N_2\}$  como:

$$d_{ADCD}(F_1, F_2) = \max(N_1, N_2) - SIM(F_1, F_2), \quad (4-1)$$

$$SIM(F_1, F_2) = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} a_{i,j} S_{i,j}, \quad (4-2)$$

$$a_{i,j} = \begin{cases} 1 - \frac{d_{i,j}}{d_{max}} & \text{si } d_{i,j} \leq T_d, \\ 0 & \text{si } d_{i,j} > T_d \end{cases}, \quad (4-3)$$

$$S_{i,j} = \begin{cases} 1 - |acc_q(i) - acc_t(j)| & \text{si } d_{i,j} \leq T_d, \\ 0 & \text{si } d_{i,j} > T_d. \end{cases} \quad (4-4)$$

donde  $acc_q(i)$  es el valor del auto correlograma del  $i$ -ésimo color dominante del descriptor  $F_q$ ;  $acc_t(j)$  es el valor del auto correlograma del  $j$ -ésimo color dominante del descriptor  $F_t$ ;  $d_{i,j}$  es la distancia euclidiana entre dos triadas de color, que representan cada color dominante;  $T_d$  es el umbral de la distancia máxima usada para determinar si dos colores dominantes son similares, o no. En nuestra experimentación se emplea  $T_d = 25$  y el parámetro  $\alpha = 2$ .

$$d_{max} = \alpha T_d. \quad (4-5)$$

De acuerdo a la definición de esta medida las imágenes son iguales cuando  $d_{ADCD}(F_1, F_2) = 0$ , entre mayor sea el valor de  $d_{ADCD}(F_1, F_2)$  mayor será la diferencia entre los descriptores  $F_1$  y  $F_2$ .

### 4.3. Imágenes obtenidas de una búsqueda

Como resultado de nuestra experimentación a continuación se muestran la imagen de búsqueda y las imágenes recuperadas con el Descriptor de Color Dominante RGB y utilizando la distancia de Color Dominante.

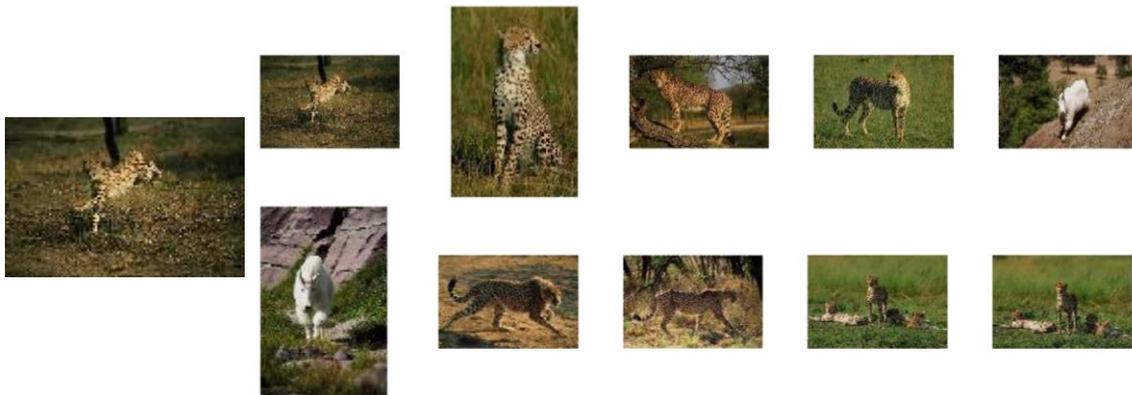


Figura 4-2 Imagen de búsqueda 134001.jpg<sup>13</sup> (izquierda), e imágenes recuperadas empleando el descriptor de colores dominantes  $CDC_{RGB}$  con la distancia de colores dominantes  $CDC_d$ .

<sup>13</sup> De la base de datos CorelDB

En las siguientes figuras se muestran las imágenes obtenidas con la misma imagen de búsqueda, pero con el descriptor de auto correlograma de colores dominantes DCCD, utilizando la Distancia de L1 normalizada L1N y la Distancia de Auto correlograma de Colores Dominates ADCDd; usando los descriptores de Auto correlograma de Colores Dominantes ADCD y de auto correlograma compacto generalizado de colores dominantes AGCGDCD con la distancia ADCDd.



Figura 4-3 Imágenes recuperadas con DCCD y L1 normalizada.

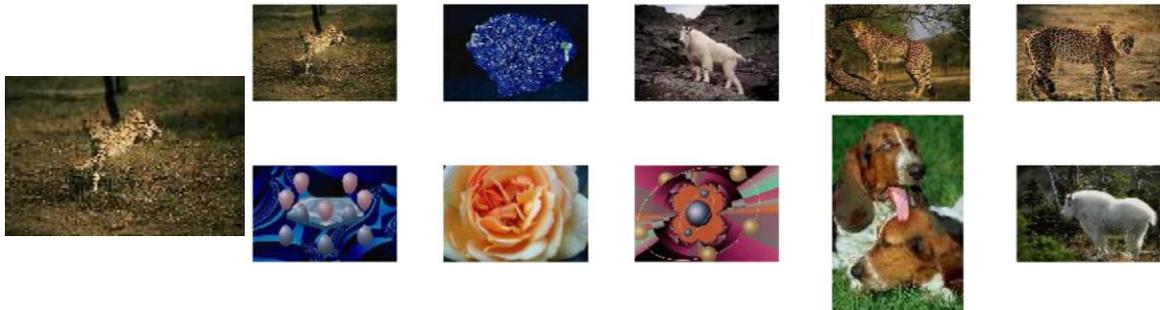


Figura 4-4 Imágenes recuperadas con el Correlograma de colores dominantes (auto correlograma en el espacio HSV) DCCD, con el auto correlograma a una distancia de 1 pixel, y la distancia de auto correlograma de colores dominantes ADCDd.

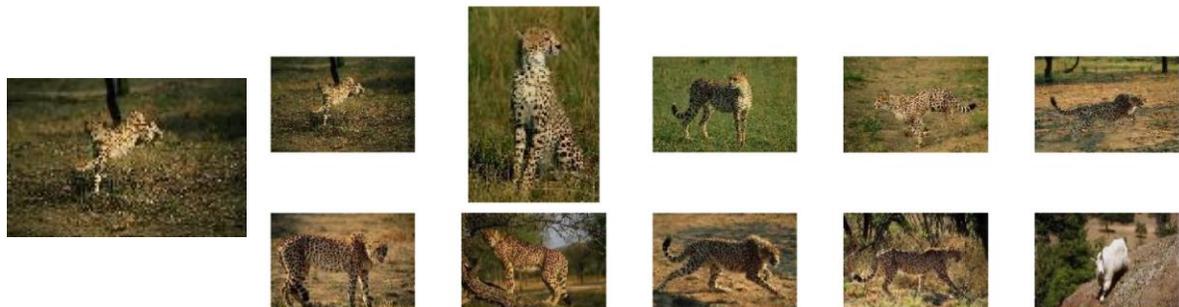
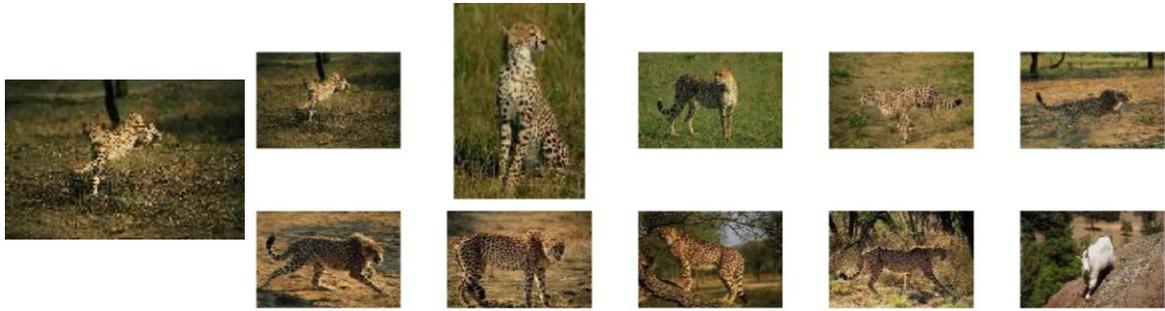
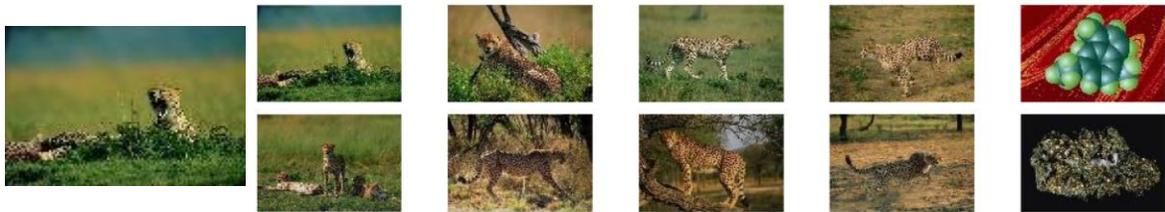


Figura 4-5 Imágenes recuperadas  $ADCD_{RGB}$ , con el auto correlograma a una distancia de 1 pixel, y la distancia de auto correlograma de colores dominantes ADCDd.



*Figura 4-6 Imágenes recuperadas con el Auto correlograma compacto generalizado de colores dominantes ACGCDD, a una distancia de 1 pixel, y la distancia de auto correlograma de colores dominantes ADCDd.*

Los descriptores pueden ser mejores para algunas imágenes y tener fallas con otras como se puede observar en la siguiente figura, donde se muestran las imágenes obtenidas con el decriptor DCCD, utilizando la Distancia de Auto correlograma de Colores Dominantes ADCDd, al igual que en la Figura 4-4, y se obtienen mejores resultados simplemente cambiando la imagen de búsqueda por la 134009.jpg.



*Figura 4-7 Imágenes recuperadas DCCD, a una distancia de 1 pixel, y la distancia de auto correlograma de colores dominantes ADCDd.*

## 5. Experimentación y Resultados

En este capítulo se describen las bases de datos de imágenes utilizadas y la configuración de las pruebas llevadas a cabo con las mismas, a partir de las cuales se obtuvieron los resultados mostrados más adelante.

### 5.1. Bases de datos utilizadas

Las imágenes utilizadas en el presente trabajo se tomaron de la base de datos libre COREL DB, que es ampliamente utilizada [37, 38, 39, 40, 41] y de la base de datos de Wang [42]. Con la primera se formaron subgrupos de imágenes obteniendo las bases denominadas DB01, DB03 y DB04; a la segunda se le utilizó en su totalidad, denominada DB02 en la experimentación.

Las imágenes de COREL DB tienen dimensiones de 120x80 o de 80x120 píxeles, mientras que para las de la base de datos de Wang las dimensiones son de 384x256, o de 256x384 píxeles. A continuación se da una breve descripción de las bases utilizadas: DB01, DB02, DB03 y DB04 (para más información ver ANEXO I).

#### 5.1.1. DB01

Tiene 10 categorías, elegidas aleatoriamente de COREL DB, con 20 imágenes cada una, dando un total de 200 imágenes. La Figura 5-1 ilustra las categorías, junto con una imagen perteneciente a dicha categoría.



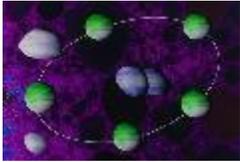
				
247000.jpg	280016.jpg	617012.jpg	533007.jpg	168002.jpg
perros	personas ejercitándose	dibujos órbitas	dibujos moléculas	cabras

Figura 5-1 Categorías DB01 con una imagen muestra

### 5.1.2. DB02

La totalidad de la base de Wang tiene 10 categorías, con 100 imágenes cada una, dando un total de 1000 imágenes. La Figura 5-2 muestra las categorías, junto con una imagen perteneciente a dicha categoría.

				
61.jpg	149.jpg	204.jpg	368.jpg	451.jpg
tribal	playa	arquitectura romana	autobuses	dinosaurios
				
555.jpg	680.jpg	765.jpg	867.jpg	974.jpg
elefantes	rosas	caballos	montañas	Platillos

Figura 5-2 Categorías DB02 con una imagen muestra

### 5.1.3. DB03

Tiene 20 categorías, elegidas aleatoriamente de la base COREL DB, con 50 imágenes cada una, dando un total de 1000 imágenes. La Figura 5-3 ilustra las categorías, junto con una imagen perteneciente a dicha categoría.

 283011.jpg	 114017.jpg	 59042.jpg	 343026.jpg	 40034.jpg
arte espacial	montañas (nevadas)	puertas	brillo metálico	fuegos artificiales
 409006.jpg	 225004.jpg	 113031.jpg	 473012.jpg	 107004.jpg
cuarzos	esquiando	caballos	fractales	elefantes
 461014.jpg	 131016.jpg	 168031.jpg	 186007.jpg	 110019.jpg
banderas del mundo	campestre con flores	cabras	cromado colorido	lobos
 581038.jpg	 75016.jpg	 353028.jpg	 314016.jpg	 52033.jpg
autobuses	lechuzas	bonsáis	delfines	mariposas

Figura 5-3 Categorías DB03 con una imagen muestra

### 5.1.4. DB04

Consta de 40 categorías, elegidas aleatoriamente de la base COREL DB, con 100 imágenes cada una, dando un total de 4000 imágenes. La Figura 5-4 muestra las categorías, junto con una imagen perteneciente a dicha categoría.

 482050.jpg	 364025.jpg	 345039.jpg	 318023.jpg	 533063.jpg
bosques	interiores de las casa	atardeceres	hojas	dibujos moléculas
 108003.jpg	 329004.jpg	 175060.jpg	 59099.jpg	 350091.jpg
tigres	globos	reptiles	puertas	cielo gris
 283055.jpg	 377023.jpg	 435026.jpg	 114017.jpg	 312060.jpg
arte espacial	cascadas	muebles antiguos	montañas (nevadas)	olas del mar
 468027.jpg	 131061.jpg	 109016.jpg	 40045.jpg	 351076.jpg
locomotoras de vapor	campestre con flores	zorros	fuegos artificiales	trenes

 198041.jpg	 29057.jpg	 112032.jpg	 326014.jpg	 184012.jpg
mujeres	carros	hipopótamos y rinocerontes	lince	Paisaje árticos
 287035.jpg	 569011.jpg	 218003.jpg	 275016.jpg	 581093.jpg
noche y luces	comida banquetes	estalactitas	cocteles	autobuses
 407013.jpg	 52028.jpg	 208002.jpg	 353053.jpg	 618080.jpg
desértico	mariposas	hongos	bonsáis	huevos decorados
 425029.jpg	 225030.jpg	 234052.jpg	 141009.jpg	 655088.jpg
barajas	esquiando	adornos festivos	peces	señuelos

Figura 5-4 Categorías DB03 con una imagen muestra

## 5.2. Configuración del experimento

En esta sección se explica brevemente los descriptores que se obtuvieron a partir de cada una de las cuatro bases de imágenes y la notación que se utiliza para referirse a cada uno de ellos.

En la Tabla 5-1 se lista la nomenclatura utilizada para referirse a los elementos más generales utilizados en la experimentación. Las columnas D indican el modo abreviado de referirse a los elementos de las columnas inmediatamente a su izquierda.

*Tabla 5-1 Nomenclatura de experimentación.*

Descriptor	D (modo abreviado)	Imagen cuantizada a $n$ colores: $x^3 = n$	D (modo abreviado)	Medida de similitud	D (modo abreviado)
Histograma	H	$2^3 = 8$	c2	Intersección de Histograma distancia ( 3-54 )	IH1
Esquema de color	CLD	$3^3 = 27$	c3	Intersección de Histograma distancia ( 3-56 )	IH2
Color dominante	DCD	$4^3 = 64$	c4	Norma $L_1$ normalizada	L1N
Correlograma de color	CC	$5^3 = 125$	c5	Norma $L_p-L_1$	L1
Auto correlograma de color	ACC			Norma $L_p-L_2$	L2
Correlograma compacto generalizado	CGC			Norma $L_p-L_\infty$	Linf
Correlograma de colores dominantes (Auto correlograma en el espacio HSV)	DCCD			Distancia de Esquema de color	CLDd
Correlograma de colores dominantes (Correlograma en el espacio HSV)	DCC			Distancia de correlograma	CCd
Auto correlograma de colores dominantes (En el espacio RGB)	ADCD			Distancia de color dominante	DCDd
Correlograma de colores dominantes (En el espacio RGB)	CDCD			Distancia de Auto correlograma de Colores Dominantes	ADCDD
Correlograma compacto generalizado de colores dominantes	CGCD				
Auto correlograma compacto generalizado de colores dominantes	ACGCD				

Algunos de los descriptores se calcularon utilizando diferentes parámetros de entrada. El histograma, por ejemplo, se obtuvo con un diferente número de rangos. En la Tabla 5-2 se listan dichos parámetros.

Tabla 5-2 Parámetros de los diferentes descriptores.

Número de rangos del histograma	Designación	Tipo de DCD	Designación
$5^3 = 125$	rgb5	Algoritmo de Bloque Lineal	LBA
$10^3 = 1000$	rgb10	En HSV cuantizado a 72 colores	HSV
$45^3 = 91125$	rgb45	En HSV cuantizado a 72 colores, y finalmente reconvertida a RGB	HSV*
$256 \times 3 = 768$	256	En RGB cuantizado a 125 colores	RGB
Distancia entre los pixeles para el CGC	Designación	Distancias entre los pixeles para el CC y el ACC	Designación
A 1 pixel	d1	A [1 2 3 4 5 6 7 8] pixeles	D1
A 8 pixeles	d2	A [25 40 76] pixeles	D2a
A 25 pixeles	d3	A [25 40 76 128] pixeles	D2b
A 40 pixeles	d4		
A 76 pixeles	d5		
A 128 pixeles	d6		

En el caso de las distancias entre pixeles para los descriptores CC y ACC, D1 se utiliza para todas las bases de datos; las distancias D2a para las bases DB01, DB03 y DB04; y las distancias D2b y d6 (para el CGC) son solamente para la base DB02.

Con las debidas combinaciones de los elementos de la Tabla 5-1 y la Tabla 5-2 se extrajeron los descriptores listados de las siguientes tablas: de la Tabla 5-3 a la Tabla 5-11, los cuales se comparan utilizando varias medidas de similitud. Se realizaron búsquedas con todas las imágenes de cada base de datos. Como resultado obtenemos las imágenes extraídas ordenadas de mayor a menor parecido respecto a cada búsqueda. En las siguientes tablas se listan las nomenclaturas de los resultados de cada tipo de búsqueda.

Tabla 5-3 Nomenclatura de resultados con el Histograma.

Hrgb5-IH1	Hrgb5-IH2	Hrgb5-L1N	Hrgb5-L1	Hrgb5-L2	Hrgb5-Linf
Hrgb10-IH1	Hrgb10-IH2	Hrgb10-L1N	Hrgb10-L1	Hrgb10-L2	Hrgb10-Linf
Hrgb45-IH1	Hrgb45-IH2	Hrgb45-L1N	Hrgb45-L1	Hrgb45-L2	Hrgb45-Linf
H256-IH1	H256-IH2	H256-L1N	H256-L1	H256-L2	H256-Linf

Tabla 5-4 Nomenclatura de resultados con el Auto correlograma compacto generalizado de colores dominantes.

ACGCDCD <sub>LBA</sub> d1-CCd	ACGCDCD <sub>LBA</sub> d1-L1	ACGCDCD <sub>LBA</sub> d1-Linf
ACGCDCD <sub>RGB</sub> d1-CCd	ACGCDCD <sub>RGB</sub> d1-L1	ACGCDCD <sub>RGB</sub> d1-Linf
ACGCDCD <sub>LBA</sub> d2-CCd	ACGCDCD <sub>LBA</sub> d2-L1	ACGCDCD <sub>LBA</sub> d2-Linf
ACGCDCD <sub>RGB</sub> d2-CCd	ACGCDCD <sub>RGB</sub> d2-L1	ACGCDCD <sub>RGB</sub> d2-Linf
ACGCDCD <sub>LBA</sub> d1-L1N	ACGCDCD <sub>LBA</sub> d1-L2	ACGCDCD <sub>LBA</sub> d1-ADCDD
ACGCDCD <sub>RGB</sub> d1-L1N	ACGCDCD <sub>RGB</sub> d1-L2	ACGCDCD <sub>RGB</sub> d1-ADCDD
ACGCDCD <sub>LBA</sub> d2-L1N	ACGCDCD <sub>LBA</sub> d2-L2	ACGCDCD <sub>LBA</sub> d2-ADCDD
ACGCDCD <sub>RGB</sub> d2-L1N	ACGCDCD <sub>RGB</sub> d2-L2	ACGCDCD <sub>RGB</sub> d2-ADCDD

Tabla 5-5 Nomenclatura de resultados con el Correlograma compacto generalizado de colores dominantes.

CGCDCD <sub>LBA</sub> d1-CCd	CGCDCD <sub>LBA</sub> d1-L1	CGCDCD <sub>LBA</sub> d1-Linf
CGCDCD <sub>RGB</sub> d1-CCd	CGCDCD <sub>RGB</sub> d1-L1	CGCDCD <sub>RGB</sub> d1-Linf
CGCDCD <sub>LBA</sub> d2-CCd	CGCDCD <sub>LBA</sub> d2-L1	CGCDCD <sub>LBA</sub> d2-Linf
CGCDCD <sub>RGB</sub> d2-CCd	CGCDCD <sub>RGB</sub> d2-L1	CGCDCD <sub>RGB</sub> d2-Linf
CGCDCD <sub>LBA</sub> d1-L1N	CGCDCD <sub>LBA</sub> d1-L2	CGCDCD <sub>LBA</sub> d1-ADCd
CGCDCD <sub>RGB</sub> d1-L1N	CGCDCD <sub>RGB</sub> d1-L2	CGCDCD <sub>RGB</sub> d1-ADCd
CGCDCD <sub>LBA</sub> d2-L1N	CGCDCD <sub>LBA</sub> d2-L2	CGCDCD <sub>LBA</sub> d2-ADCd
CGCDCD <sub>RGB</sub> d2-L1N	CGCDCD <sub>RGB</sub> d2-L2	CGCDCD <sub>RGB</sub> d2-ADCd

Tabla 5-6 Nomenclatura de resultados con el Auto correlograma de colores dominantes (en el espacio RGB)

ACDCD <sub>LBA</sub> d1-CCd	ACDCD <sub>LBA</sub> d1-L1	ACDCD <sub>LBA</sub> d1-Linf
ACDCD <sub>RGB</sub> d1-CCd	ACDCD <sub>RGB</sub> d1-L1	ACDCD <sub>RGB</sub> d1-Linf
ACDCD <sub>LBA</sub> d2-CCd	ACDCD <sub>LBA</sub> d2-L1	ACDCD <sub>LBA</sub> d2-Linf
ACDCD <sub>RGB</sub> d2-CCd	ACDCD <sub>RGB</sub> d2-L1	ACDCD <sub>RGB</sub> d2-Linf
ACDCD <sub>LBA</sub> d1-L1N	ACDCD <sub>LBA</sub> d1-L2	ACDCD <sub>LBA</sub> d1-ADCd
ACDCD <sub>RGB</sub> d1-L1N	ACDCD <sub>RGB</sub> d1-L2	ACDCD <sub>RGB</sub> d1-ADCd
ACDCD <sub>LBA</sub> d2-L1N	ACDCD <sub>LBA</sub> d2-L2	ACDCD <sub>LBA</sub> d2-ADCd
ACDCD <sub>RGB</sub> d2-L1N	ACDCD <sub>RGB</sub> d2-L2	ACDCD <sub>RGB</sub> d2-ADCd

Tabla 5-7 Nomenclatura de resultados con el Correlograma de colores dominantes (en el espacio RGB)

CDCD <sub>LBA</sub> d1-CCd	CDCD <sub>RGB</sub> d2-L1N	CDCD <sub>LBA</sub> d2-L2
CDCD <sub>RGB</sub> d1-CCd	CDCD <sub>LBA</sub> d1-L1	CDCD <sub>RGB</sub> d2-L2
CDCD <sub>LBA</sub> d2-CCd	CDCD <sub>RGB</sub> d1-L1	CDCD <sub>LBA</sub> d1-Linf
CDCD <sub>RGB</sub> d2-CCd	CDCD <sub>LBA</sub> d2-L1	CDCD <sub>RGB</sub> d1-Linf
CDCD <sub>LBA</sub> d1-L1N	CDCD <sub>RGB</sub> d2-L1	CDCD <sub>LBA</sub> d2-Linf
CDCD <sub>RGB</sub> d1-L1N	CDCD <sub>LBA</sub> d1-L2	CDCD <sub>RGB</sub> d2-Linf
CDCD <sub>LBA</sub> d2-L1N	CDCD <sub>RGB</sub> d1-L2	

Tabla 5-8 Nomenclatura de resultados con el Esquema de color, el color dominante, el correlograma de color dominante (en el espacio HSV) y el Auto correlograma de color dominante (en el espacio HSV).

CLD	DCD	DCC	DCCD
CLD-CLDd	DCDLBA-DCDd	DCC-CCd	DCCD-CCd
CLD-L1N	DCDHSV-DCDd	DCC-L1N	DCCD-L1N
CLD-L1	DCDRGB-DCDd	DCC-L1	DCCD-L1
CLD-L2		DCCD-L2	DCCD-L2
CLD-Linf		DCCD-Linf	DCCD-Linf
		DCCD-ADCd	DCCD-ADCd

Tabla 5-9 Nomenclatura de resultados con el Correlograma compacto generalizado.

CGC-CGCd-d1c2	CGC-CGCd-d3c2	CGC-CGCd-d5c2
CGC-CGCd-d1c3	CGC-CGCd-d3c3	CGC-CGCd-d5c3
CGC-CGCd-d1c4	CGC-CGCd-d3c4	CGC-CGCd-d5c4
CGC-CGCd-d1c5	CGC-CGCd-d3c5	CGC-CGCd-d5c5
CGC-CGCd-d2c2	CGC-CGCd-d4c2	CGC-CGCd-d6c2
CGC-CGCd-d2c3	CGC-CGCd-d4c3	CGC-CGCd-d6c3
CGC-CGCd-d2c4	CGC-CGCd-d4c4	CGC-CGCd-d6c4
CGC-CGCd-d2c5	CGC-CGCd-d4c5	CGC-CGCd-d6c5

*Tabla 5-10 Nomenclatura de resultados con el Auto correlograma de color.*

ACCc2D1-CCd	ACCc2D1-L1N	ACCc2D1-L1	ACCc2D1-L2	ACCc2D1-Linf
ACCc2D2a-CCd	ACCc2D2a-L1N	ACCc2D2a-L1	ACCc2D2a-L2	ACCc2D2a-Linf
ACCc2D2b-CCd	ACCc2D2b-L1N	ACCc2D2b-L1	ACCc2D2b-L2	ACCc2D2b-Linf
ACCc3D1-CCd	ACCc3D1-L1N	ACCc3D1-L1	ACCc3D1-L2	ACCc3D1-Linf
ACCc3D2a-CCd	ACCc3D2a-L1N	ACCc3D2a-L1	ACCc3D2a-L2	ACCc3D2a-Linf
ACCc3D2b-CCd	ACCc3D2b-L1N	ACCc3D2b-L1	ACCc3D2b-L2	ACCc3D2b-Linf
ACCc4D1-CCd	ACCc4D1-L1N	ACCc4D1-L1	ACCc4D1-L2	ACCc4D1-Linf
ACCc4D2a-CCd	ACCc4D2a-L1N	ACCc4D2a-L1	ACCc4D2a-L2	ACCc4D2a-Linf
ACCc4D2b-CCd	ACCc4D2b-L1N	ACCc4D2b-L1	ACCc4D2b-L2	ACCc4D2b-Linf
ACCc5D1-CCd	ACCc5D1-L1N	ACCc5D1-L1	ACCc5D1-L2	ACCc5D1-Linf
ACCc5D2a-CCd	ACCc5D2a-L1N	ACCc5D2a-L1	ACCc5D2a-L2	ACCc5D2a-Linf
ACCc5D2b-CCd	ACCc5D2b-L1N	ACCc5D2b-L1	ACCc5D2b-L2	ACCc5D2b-Linf

*Tabla 5-11 Nomenclatura de resultados con el Correlograma de color.*

CCc2D1-CCd	CCc2D1-L1N	CCc2D1-L1	CCc2D1-L2	CCc2D1-Linf
CCc2D2a-CCd	CCc2D2a-L1N	CCc2D2a-L1	CCc2D2a-L2	CCc2D2a-Linf
CCc2D2b-CCd	CCc2D2b-L1N	CCc2D2b-L1	CCc2D2b-L2	CCc2D2b-Linf
CCc3D1-CCd	CCc3D1-L1N	CCc3D1-L1	CCc3D1-L2	CCc3D1-Linf
CCc3D2a-CCd	CCc3D2a-L1N	CCc3D2a-L1	CCc3D2a-L2	CCc3D2a-Linf
CCc3D2b-CCd	CCc3D2b-L1N	CCc3D2b-L1	CCc3D2b-L2	CCc3D2b-Linf
CCc4D1-CCd	CCc4D1-L1N	CCc4D1-L1	CCc4D1-L2	CCc4D1-Linf
CCc4D2a-CCd	CCc4D2a-L1N	CCc4D2a-L1	CCc4D2a-L2	CCc4D2a-Linf
CCc4D2b-CCd	CCc4D2b-L1N	CCc4D2b-L1	CCc4D2b-L2	CCc4D2b-Linf
CCc5D1-CCd	CCc5D1-L1N	CCc5D1-L1	CCc5D1-L2	CCc5D1-Linf
CCc5D2a-CCd	CCc5D2a-L1N	CCc5D2a-L1	CCc5D2a-L2	CCc5D2a-Linf
CCc5D2b-CCd	CCc5D2b-L1N	CCc5D2b-L1	CCc5D2b-L2	CCc5D2b-Linf

A manera de ejemplo, de las tablas anteriores se puede decir que:

- Hrgb10-IH1 corresponde a las imágenes recuperadas mediante el descriptor del Histograma con 1000 rangos, teniendo como entrada las imágenes originales, y empleando la intersección de histograma descrita en el punto 3.3.2.
- CCc5D1-CCd corresponde a las imágenes recuperadas mediante el descriptor del Correlograma de color a las distancias [1 2 3 4 5 6 7 8] entre pixeles, teniendo como entrada las imágenes cuantizadas con 125 colores, y empleando la distancia de Correlograma.

### ***5.3. Resultados del algoritmo propuesto***

En esta sección se muestran los mejores resultados del Color Dominante RGB  $DCD_{RGB}$  comparado con los diferentes descriptores de color dominante mediante la distancia de colores dominantes. También se muestra el Auto correlograma de color dominante (en el espacio RGB) ADCD, el Auto correlograma del correlograma compacto

generalizado ACGDCD y el DCCD utilizando de la distancia propuesta, y comparándola contra varias de las distancias existentes.

Los resultados técnicos completos de todos los descriptores con las diferentes distancias se encuentran en [43].

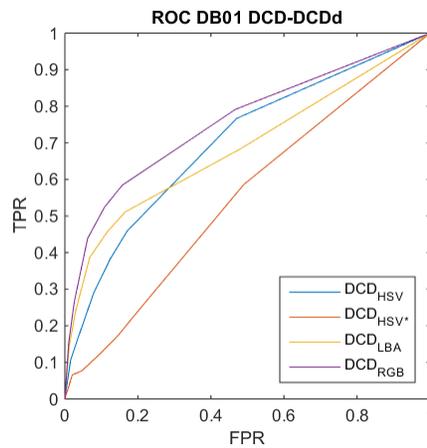


Figura 5-5 Curvas ROC del Color dominante (DCD) con DCDd

Como se puede observar en la el Figura 5-5, el descriptor  $DCD_{RGB}$  resulto ser el mejor de los descriptores de colores dominantes, para la base de datos DB01.

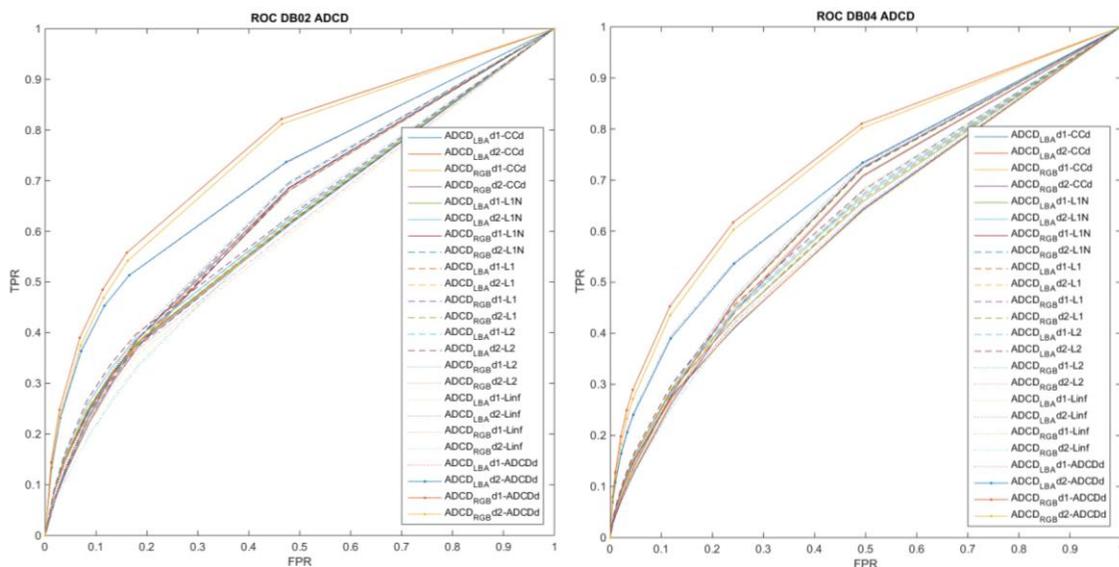


Figura 5-6 Curvas ROC del Auto correlograma de colores dominantes (ADCD) con CCd, L1N, L1, L2,  $L\infty$  y ADCDd.

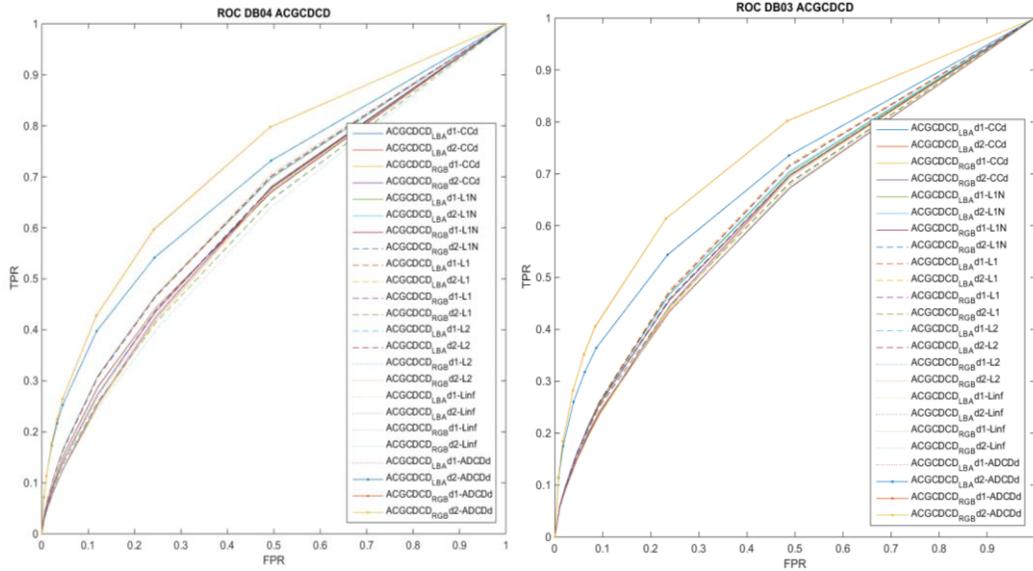


Figura 5-7 Curvas ROC del Auto correlograma compacto generalizado de colores dominantes (ACGDCD) con CCd, L1N, L1, L2, L $\infty$  y ADCDd.

En las curvas de la Figura 5-6 a la Figura 5-8, se puede observar que la distancia propuesta (ADCDd) mejora el rendimiento de los descriptores de Auto correlogramas de colores dominantes, sobre todo aquellos que utilizaron como base el CDC<sub>RGB</sub>.

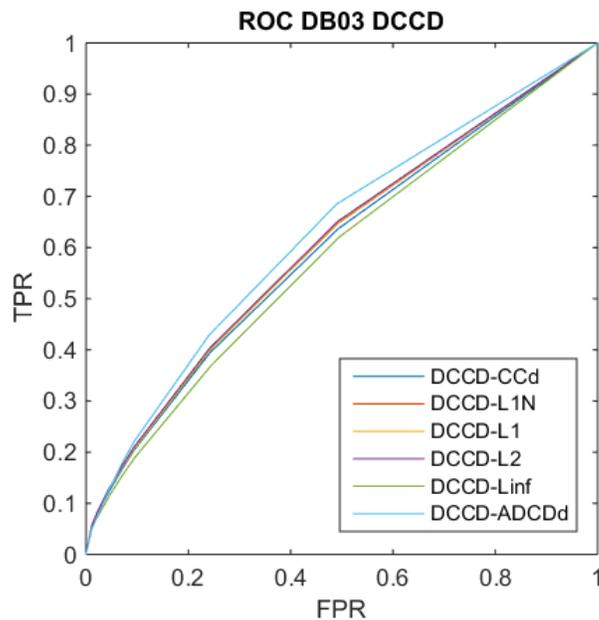


Figura 5-8 Curvas ROC del Correlograma de colores dominantes (Aut Correlograma-DCCD) con CCd, L1N, L1, L2 L $\infty$  y ADCDd.

## ***5.4. Evaluación***

Empleando las cuatro bases de datos descritas en el punto anterior se obtuvieron los descriptores definidos en el punto 5.2.

Cada descriptor se compara contra todos los descriptores de la base de datos a la que pertenece mediante algunas de las medidas de similitud descritas en la sección 3.3. Finalmente se utilizan las métricas de evaluación para determinar cuál descriptor tiene el mejor desempeño.

### ***5.4.1. Mediante ROC***

Para obtener las curvas ROC de los descriptores, se consideraron las primeras 25, 50, 100, 150, 200 imágenes para DB01; las primeras 25, 50, 100, 150, 200, 500, 1000 imágenes para DB02 y DB03; y las primeras 25, 50, 100, 150, 200, 500, 1000, 2000 y 4000 imágenes para DB04; y de esa manera determinar las imágenes como reconocidas.

Considerando los datos del párrafo anterior se calcularon las tasas de verdaderos positivos y falsos positivos TPR y FPR respectivamente, de cada búsqueda, y se promediaron para graficar una curva ROC por cada descriptor obtenido, como se muestra en las figuras siguientes.

### 5.4.1.1. Para la base DB01

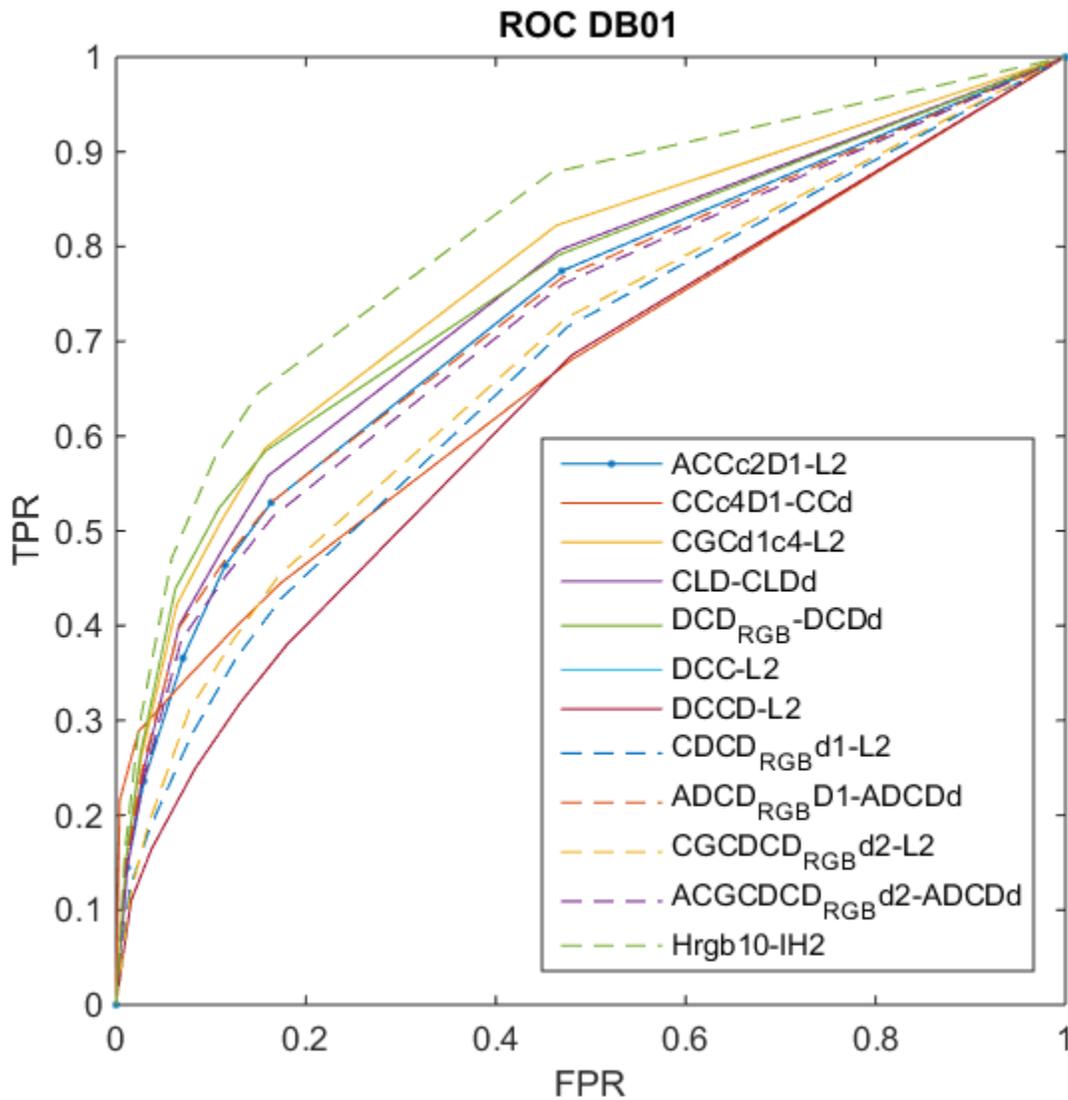


Figura 5-5 Indexación automática de imágenes

Se puede observar que el mejor descriptor es el Histograma con 1000 rangos utilizando la distancia de la Intersección del histograma ( 3-56 ) Hrgb10-IH2, que tiene un comportamiento muy similar al del Correlograma de colores dominantes (auto correlograma-DCCD) utilizando la distancia de Auto correlograma de colores dominantes DCCD-ADCDD.

### 5.4.1.2. Para la base DB02

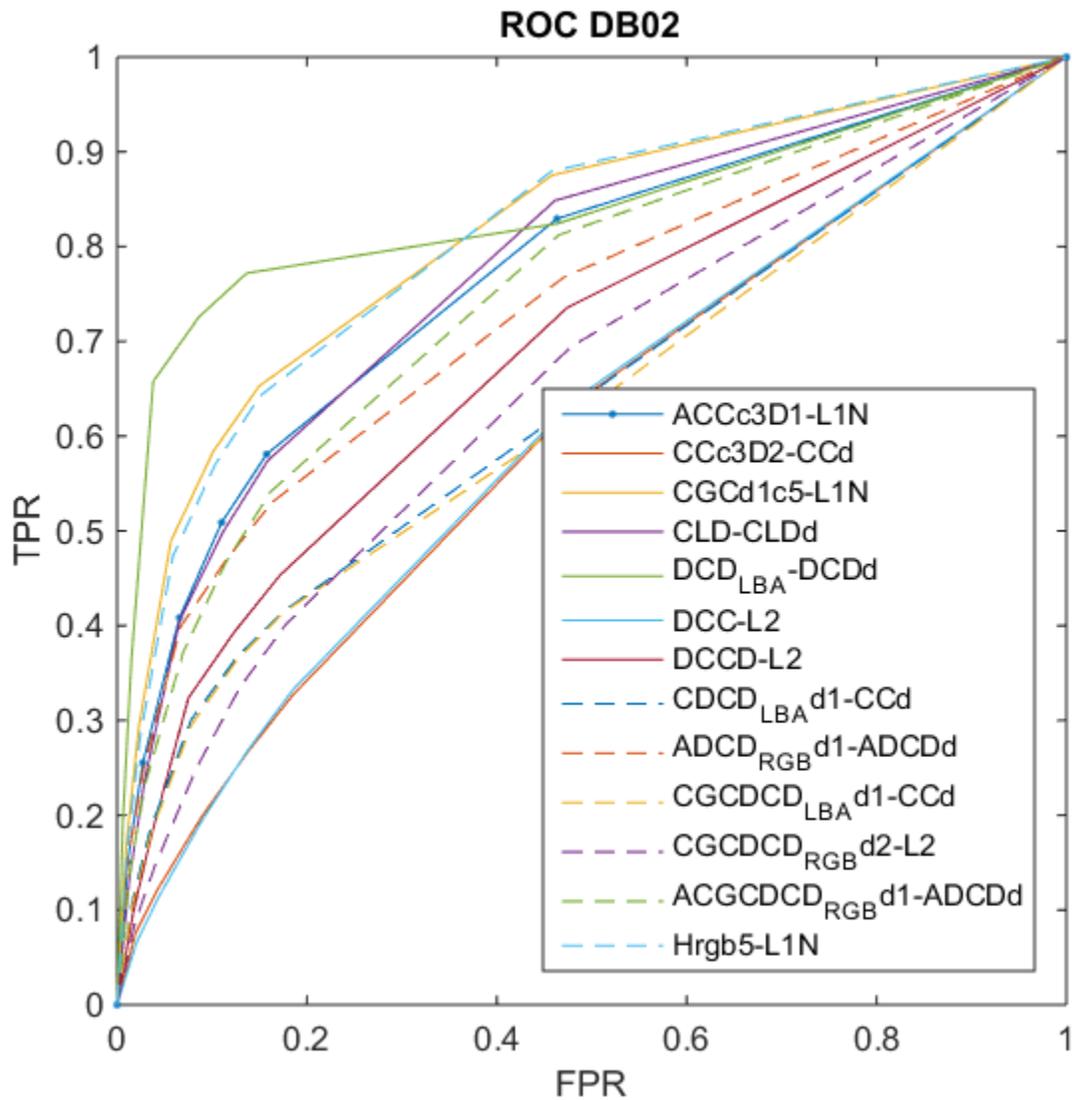


Figura 5-6 Indexación automática de imágenes

Para esta base de datos se podría considerar que uno de los mejores descriptores es el de Colores Dominantes usando la distancia de Colores Dominantes  $DCD_{LBA}-DCDd$ , sin embargo se observa un efecto de saturación abrupto. En cambio, el Correlograma Compacto Generalizado con la distancia L1 normalizada  $CGCd1c5-L1N$  y el Histograma con 125 rangos usando la distancia L1 normalizada  $Hrgb5-L1N$ , tienen buenos desempeños y mantienen un crecimiento más uniforme, sin cambios tan grandes.

5.4.1.3. Para la base DB03

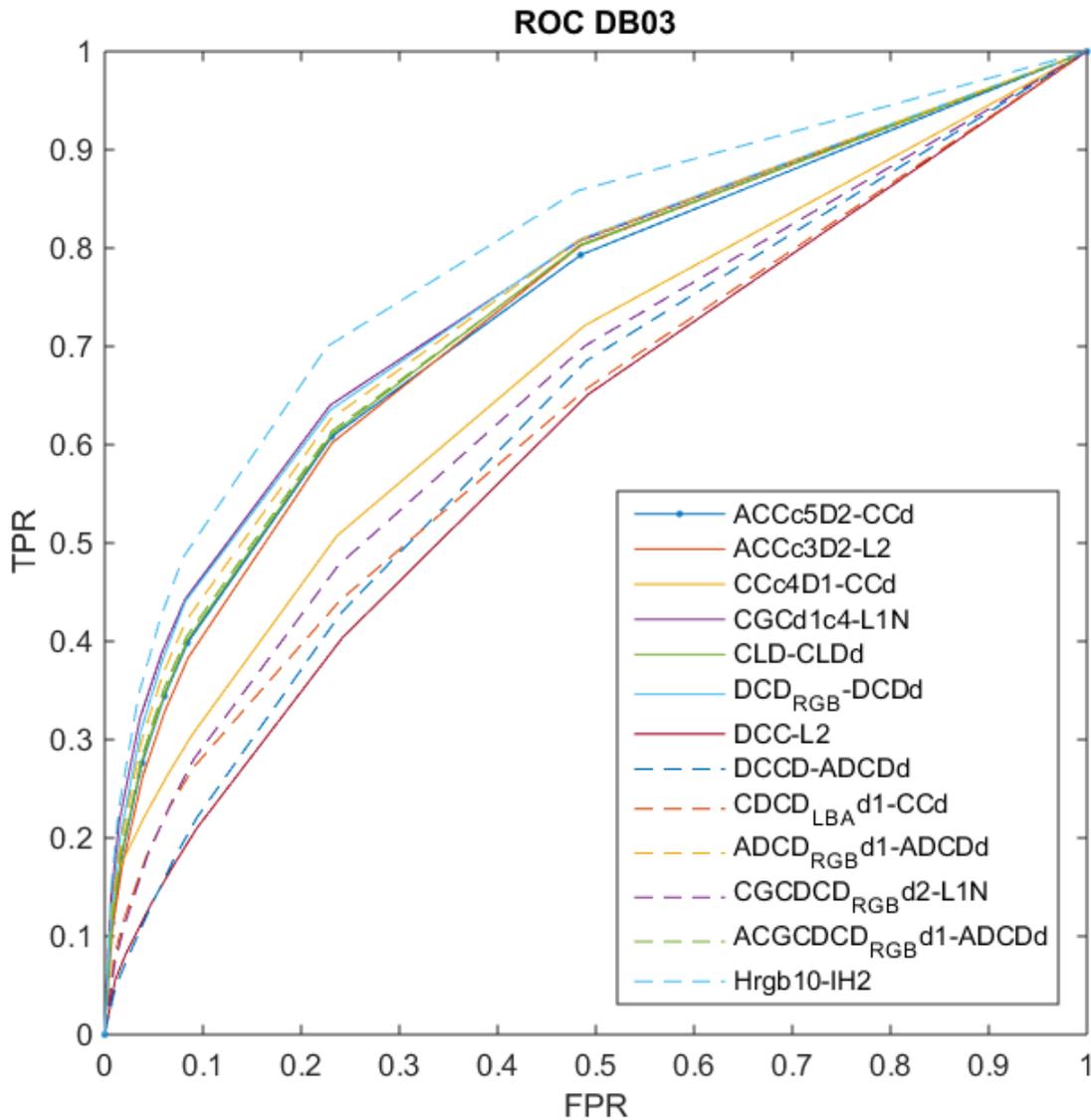


Figura 5-7 Indexación automática de imágenes

En ese caso el descriptor con mejor desempeño es el Histograma con 1000 rangos utilizando la distancia de la Intersección del histograma ( 3-56 ) Hrgb10-IH2.

#### 5.4.1.4. Para la base DB04

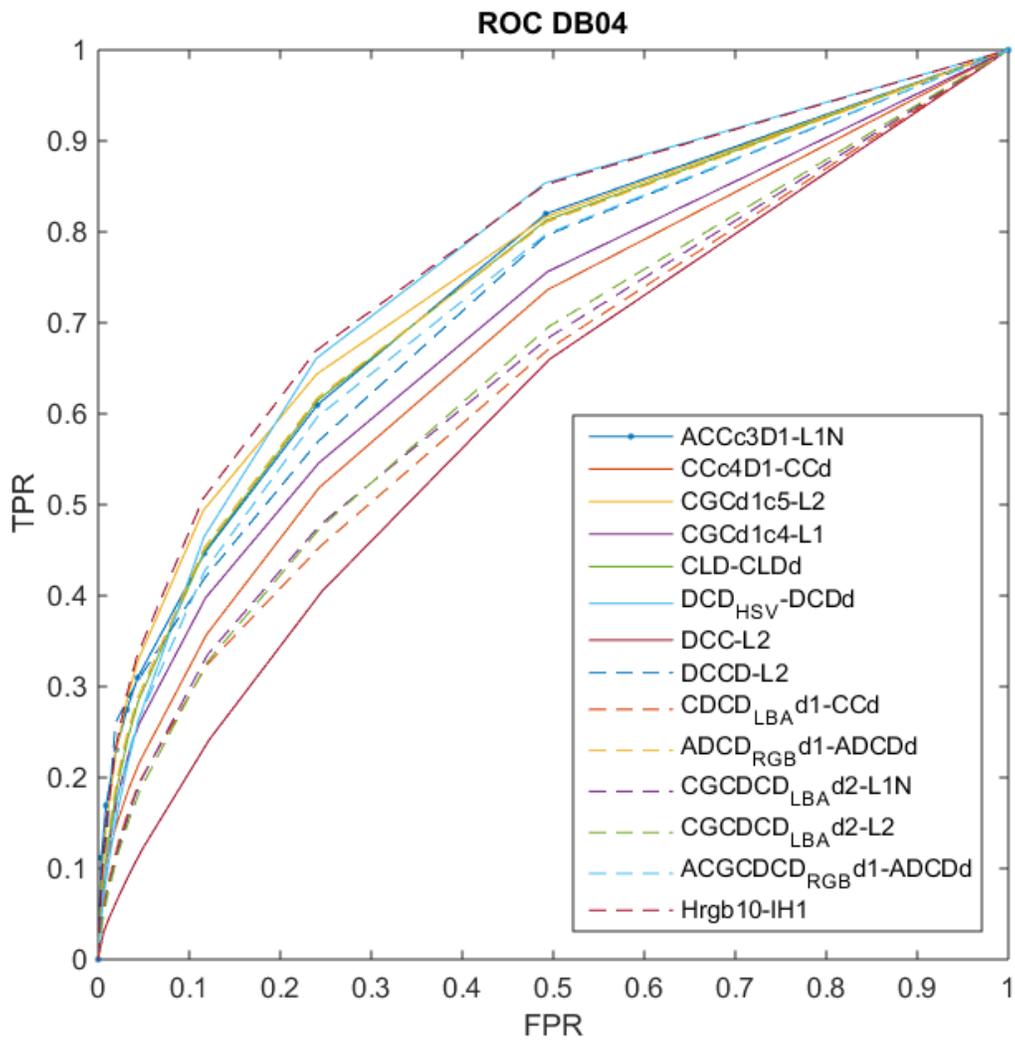


Figura 5-8 Indexación automática de imágenes

Para esta base, es difícil seleccionar el descriptor con el mejor desempeño, sin embargo el que consideramos como mejor es el Histograma con 1000 rangos utilizando la distancia de Intersección de Histograma con ecuación ( 3-54 ) Hrgb10-IH1.

## 5.4.2. Evaluación con la métrica ARP

Como se mencionó en el punto 3.4.2, para la métrica ARP entre más cercano es su valor a 1, mejor es el desempeño del descriptor. En la experimentación se emplea los parámetros:  $\alpha_1 = 0.25$ ,  $\alpha_2 = 0.5$  y  $\alpha_3 = 1$ .

En las tablas que se enumeran a continuación se resaltan los resultados de los descriptores con mejor desempeño, y resaltados en negritas se presentan los descriptores del algoritmo propuesto.

### 5.4.2.1. Para la base DB01

Tabla 5-12 Mejores resultados de los descriptores aplicados a la primera base de datos.

Designación	$\alpha_1$	$\alpha_2$	$\alpha_3$
ACc3D1-L2	0.6570	0.5575	0.4240
CCc5D1-CCd	<b>0.8750</b>	<b>0.5985</b>	0.3455
CGCd1c5-L2	0.6500	0.5520	0.4370
CLD-L2	0.5760	0.4890	0.4008
DCD <sub>RGB</sub> -DCDd	0.6170	0.5325	0.4387
DCC-L1N	0.4610	0.3340	0.2355
DCCD-L2	0.4370	0.3285	0.2495
CDCD <sub>LBA</sub> d1-CCd	0.4930	0.3815	0.2892
ADCD <sub>RGB</sub> d1-ADCd	0.6030	0.4990	0.3987
CGCD <sub>RGB</sub> d2-L1N	0.4360	0.3885	0.3113
ACGCD <sub>RGB</sub> d2-ADCd	0.5860	0.4750	0.3842
Hrgb45-IH1	0.6810	0.5815	0.4665
Hrgb10-IH1	0.6590	0.5795	<b>0.4832</b>

Cabe mencionar que Hrgb10-IH1, Hrgb10-IH2 y Hrgb10-L2 obtuvieron exactamente los mismos resultados. Lo mismo ocurre con Hrgb45-IH1 y Hrgb45-IH2.

### 5.4.2.2. Para la base DB02

Tabla 5-13 Mejores resultados de los descriptores aplicados a la segunda base de datos.

Designación	$\alpha_1$	$\alpha_2$	$\alpha_3$
ACc5D1-L1N	0.6176	0.5240	0.4074
CCc5D1-CCd	0.4050	0.2849	0.2071
CGCd2c5-L1N	0.6652	0.5891	0.4905
CLD-CLDd	0.5319	0.4755	0.4058
DCD <sub>LBA</sub> -DCDd	<b>0.7800</b>	<b>0.7282</b>	<b>0.6576</b>
DCC-CCd	0.2941	0.2362	0.1888
DCCD-L2	0.3893	0.3552	0.3237
CDCD <sub>LBA</sub> d2-CCd	0.4317	0.3708	0.3007
ADCD <sub>RGB</sub> d1-ADCd	0.5773	0.4958	0.3900
CGCD <sub>LBA</sub> d2-CCd	0.4232	0.3631	0.2960
ACGCD <sub>LBA</sub> d1-ADCd	0.5495	0.4746	0.3710
Hrgb45-L2	0.6780	0.6022	0.4817

### 5.4.2.3. Para la base DB03

Tabla 5-14 Mejores resultados de los descriptores aplicados a la tercera base de datos.

Designación	$\alpha_1$	$\alpha_2$	$\alpha_3$
ACCc5D1-L1	0.4917	0.3962	0.2883
CCc5D1-CCd	0.5476	0.3486	0.2222
CGCd1c5-L2	0.5368	0.4419	0.3333
CLD-CLDd	0.4391	0.3646	0.2789
DCD <sub>RGB</sub> -DCDd	0.4852	0.4011	0.3074
DCC-L1	0.2265	0.1708	0.1309
DCCD-L2	0.2144	0.1658	0.1302
CDCL <sub>LBA</sub> d1-CCd	0.3088	0.2369	0.1828
ADCD <sub>RGB</sub> d1-ADCd	0.4752	0.3886	0.2965
CGCDCL <sub>LBA</sub> d1-CCd	0.2888	0.2256	0.1735
ACGCDCL <sub>RGB</sub> d1-ADCd	0.4463	0.3689	0.2825
Hrgb10-IH1	<b>0.5542</b>	<b>0.4578</b>	<b>0.3454</b>

Cabe mencionar que Hrgb10-IH1, Hrgb10-IH2 y Hrgb10-L2 obtuvieron exactamente los mismos resultados.

### 5.4.2.4. Para la base DB04

Tabla 5-15 Mejores resultados de los descriptores aplicados a la cuarta base de datos.

Designación	$\alpha_1$	$\alpha_2$	$\alpha_3$
ACCc5D1-L1N	<b>0.4613</b>	<b>0.3504</b>	0.2341
CCc4D1-CCd	0.3482	0.2265	0.1532
CGCd1c4-L1N	0.3718	0.3008	0.2317
CLD-CLDd	0.2951	0.2404	0.1896
DCD <sub>RGB</sub> -DCDd	0.3183	0.2597	0.2037
DCC-CCd	0.1331	0.0950	0.0707
DCCD-L2	0.2985	0.2754	<b>0.2602</b>
CDCL <sub>LBA</sub> d1-CCd	0.2125	0.1666	0.1297
ADCD <sub>RGB</sub> d1-ADCd	0.3178	0.2552	0.1980
CGCDCL <sub>LBA</sub> d1-CCd	0.2049	0.1621	0.1282
ACGCDCL <sub>RGB</sub> d1-ADCd	0.2835	0.2272	0.1774
Hrgb10-L1N	0.3735	0.3053	0.2332

## 5.4.1. Evaluación con la métrica ARR

### 5.4.1.1. Para la base DB01

Tabla 5-16 Mejores resultados de los descriptores aplicados a la primera base de datos.

Designación	$\alpha_1$	$\alpha_2$	$\alpha_3$
ACCc3D1-L1N	0.4220	0.5033	0.5562
CCc4D1-CCd	0.3430	0.3953	0.4445
CGCd1c5-L2	0.4370	0.5247	0.5912
CLD-L2	0.4008	0.4823	0.5582
DCD <sub>RGB</sub> -DCDd	0.4387	0.5240	0.5843
DCC-L2	0.2495	0.3193	0.3800
DCCD-L2	0.2495	0.3193	0.3800
CDCL <sub>BAD</sub> 1-CCd	0.2892	0.3695	0.4300
ADCD <sub>RGB</sub> d1-ADCDd1	0.3987	0.4690	0.5303
CGCD <sub>RGB</sub> d2-L1N	0.3113	0.3852	0.4532
CGCD <sub>RGB</sub> d1-L2	0.3078	0.3880	0.4518
CGCD <sub>RGB</sub> d2-L2	0.3098	0.3858	0.4533
ACGCD <sub>RGB</sub> d1-ADCDd	0.3842	0.4543	0.5051
Hrgb10-L2	<b>0.4832</b>	<b>0.5748</b>	<b>0.6405</b>

El histograma Hrgb10-IH1 y Hrgb10-IH2 son casi iguales a Hrgb10-L2. El descriptor con mejor desempeño es el Correlograma de colores dominantes (Auto correlograma-DCCD), utilizando la distancia de Auto correlograma de colores dominantes DCCD-ADCDd.

### 5.4.1.2. Para la base DB02

Tabla 5-17 Mejores resultados de los descriptores aplicados a la segunda base de datos.

Designación	$\alpha_1$	$\alpha_2$	$\alpha_3$
ACCc3D1-L1N	0.4084	0.5088	0.5811
CCc3D2b-CCd	0.1985	0.2660	0.3276
CCc5D1-CCd	0.2071	0.2615	0.3110
CCc5D2b-CCd	0.2004	0.2661	0.3269
CGCd1c5-L1N	0.4895	0.5846	0.6528
CLD-CLDd	0.4058	0.4984	0.5739
DCD <sub>LBA</sub> -DCDd	<b>0.6576</b>	0.4256	<b>0.7719</b>
DCC-L2	0.1946	0.2674	0.3326
DCCD-L2	0.3237	0.3927	0.4535
CDCL <sub>BAD</sub> 2-CCd	0.3007	0.3666	0.4156
ADCD <sub>RGB</sub> d1-ADCDd	0.3900	0.4851	0.5579
CGCD <sub>LBA</sub> d1-CCd	0.2953	0.3638	0.4148
ACGCD <sub>RGB</sub> d1-ADCDd	0.3723	0.4671	0.5416
Hrgb10-L1N	0.4822	<b>0.5806</b>	0.6493

El descriptor con mejor desempeño es el Descriptor de Colores dominantes utilizando la distancia de Colores Dominantes DCD<sub>LBA</sub>-DCDd.

### 5.4.1.3. Para la base DB03

Tabla 5-18 Mejores resultados de los descriptores aplicados a la tercera base de datos.

Designación	$\alpha_1$	$\alpha_2$	$\alpha_3$
ACCc3D1-L1N	0.2869	0.3517	0.4032
CCc5D1-CCd	0.2222	0.2676	0.3091
CGCd1c5-L2	0.3333	0.4078	0.4667
CLD-CLDd	0.2789	0.3463	0.4013
DCD <sub>RGB</sub> -DCDd	0.3074	0.3813	0.4406
DCC-L2	0.1302	0.1719	0.2104
DCCD-ADCd	0.1279	0.1769	0.2204
CDCL <sub>LBAd</sub> 1-CCd	0.1828	0.2326	0.2723
ADCD <sub>RGBd</sub> 1-ADCd	0.2965	0.3676	0.4214
CGCD <sub>RGBd</sub> 2-L1N	0.1809	0.2336	0.2799
ACGCD <sub>RGBd</sub> 1-ADCd	0.2825	0.3519	0.4058
Hrgb10-IH2	<b>0.3454</b>	<b>0.4235</b>	<b>0.4852</b>

Cabe mencionar que Hrgb10-IH1 y Hrgb10-IH2 obtuvieron resultados muy similares.

### 5.4.1.4. Para la base DB04

Tabla 5-19 Mejores resultados de los descriptores aplicados a la cuarta base de datos.

Designación	$\alpha_1$	$\alpha_2$	$\alpha_3$
ACCc5D1-L1	0.2370	0.2775	0.3087
CCc3D2a-CCd	0.1269	0.1438	0.1584
CGCd1c4-L1N	0.2317	0.2874	0.3295
CLD-CLDd	0.1896	0.2418	0.2839
DCD <sub>RGB</sub> -DCDd	0.2037	0.2551	0.2961
DCC-L2	0.0721	0.0971	0.1208
DCCD-L2	<b>0.2602</b>	0.2837	0.3063
CDCL <sub>LBAd</sub> 1-CCd	0.1297	0.1666	0.1974
ADCD <sub>RGBd</sub> 1-ADCd	0.1980	0.2492	0.2892
CGCD <sub>LBAd</sub> 1-CCd	0.1282	0.1654	0.1966
ACGCD <sub>RGBd</sub> 1-ADCd	0.1774	0.2250	0.2643
Hrgb10-IH1	0.2347	<b>0.2900</b>	<b>0.3339</b>

El histograma Hrgb10-IH1 y Hrgb10-IH2 son casi iguales a Hrgb10-L2. El descriptor con mayor desempeño es el Histograma con 1000 rangos utilizando la Intersección de histograma Hrgb10-IH1.

## 5.4.2. Evaluación con la métrica ANMRR

Como se mencionó en el 3.4.4, para esta métrica entre más cercano es su valor a 0, mejor es el desempeño del descriptor.

Tabla 5-20 Mejores resultados de los descriptores aplicados a la tercera base de datos.

Designación	DB01	DB02	DB03	DB04
ACCc3D1-L2/ ACCc3D1-L2	0.4945	0.5056	0.6488	0.7154
ACCc3D1-L1N-7 ACCc3D1-L1N	0.4948	0.7159	0.6475	<b>0.5032</b>
ACCc5D1-L1/ ACCc5D1-L1	0.5479	0.5189	0.6495	0.7110
CCc5D1-CCd/ CCc5D1-CCd	0.5672	0.7354	0.7139	0.8065
CGCd1c4-L1N	0.5037	0.4387	0.6080	0.7155
CGCd1c5-L1N	0.5008	0.4269	0.6102	0.7195
CGCd1c5-L2	0.4801	0.4379	0.5922	0.7186
CLD-CLDd	0.5240	0.5174	0.6564	0.7639
CLD-L2	0.5224	0.5479	0.6715	0.7753
DCD <sub>LBA</sub> -DCDd	0.5441	<b>0.2830</b>	0.6798	0.7718
DCD <sub>RGB</sub> -DCDd	0.4831	0.4973	0.6222	0.7491
DCC-L1	0.7301	0.6305	0.8294	0.7216
DCC-L2	0.6831	0.6231	0.8322	0.7198
DCCD-L2	0.6831	0.6231	0.8322	0.7198
CDCD <sub>LBA</sub> d1-CCd	0.6343	0.6401	0.7708	0.8368
CDCD <sub>RGB</sub> d1-CCd	0.6720	0.7453	0.8303	0.8959
ADCDCD <sub>RGB</sub> d1-ADCDCd	0.5289	0.5245	0.6354	0.7547
CGCDCD <sub>LBA</sub> d1-CCd	0.6519	0.6447	0.7811	0.8386
CGCDCD <sub>RGB</sub> d1-L1N	0.6168	0.6895	0.7739	0.8671
CGCDCD <sub>RGB</sub> d2-L1N	0.6170	0.6803	0.7732	0.8662
ACGCDCD <sub>RGB</sub> d1-ADCDCd	0.5458	0.5444	0.6521	0.7791
Hrgb10-IH1	<b>0.4354</b>	0.4380	<b>0.5773</b>	0.7125
Hrgb10-L1N	0.4692	0.4311	0.5979	0.7136

El histograma Hrgb10-IH1, Hrgb10-IH2 y 10-L2, obtuvieron evaluaciones iguales con esta métrica.

### 5.4.3. Evaluación Conjunta

A continuación se muestran los descriptores con mejor desempeño según cada una de la métricas, ROC, ARP, ARR y ANMRR; y se resaltan aquellos que tienen los mejores resultados.

Tabla 5-21 Evaluación conjunta para DB01.

Designación	ROC	ARP			ARR			ANMRR
		$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 2$	
CCc5D1-CCd		<b>0.8750</b>	<b>0.5985</b>	0.3455	0.3538	0.4187	0.4717	0.5672
Hrgb10-IH1		0.6590	0.5795	<b>0.4832</b>	<b>0.4832</b>	<b>0.5748</b>	0.6403	<b>0.4354</b>
Hrgb10-IH2	<b>X</b>	0.6590	0.5795	<b>0.4832</b>	<b>0.4832</b>	<b>0.5748</b>	0.6403	<b>0.4354</b>
Hrgb10-L2		0.6590	0.5795	<b>0.4832</b>	<b>0.4832</b>	<b>0.5748</b>	<b>0.6405</b>	<b>0.4354</b>

Para base DB01 los histogramas Hrgb10-IH1, Hrgb10-IH2, y Hrgb10-L2, tienen resultados casi iguales, por lo que el histograma de 1000 rangos se perfila como el mejor descriptor, en este caso particular con la distancia IH2.

Tabla 5-22 Evaluación conjunta para DB02.

Designación	ROC	ARP			ARR			ANMRR
		$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 2$	
DCD <sub>LBA</sub> -DCDd	1	<b>0.7800</b>	<b>0.7282</b>	<b>0.6576</b>	<b>0.6576</b>	0.4256	<b>0.7719</b>	<b>0.2830</b>
CGCd1c5-L1N	2	0.6626	0.5869	0.4895	0.4895	<b>0.5846</b>	0.6528	0.4269
Hrgb10-L1N		0.6530	0.5645	0.4545	0.4822	0.5806	0.6493	0.4311

Para la base DB02 el mejor descriptor resulta ser el de color dominante en el espacio HSV con la distancia de color dominante, DCD<sub>LBA</sub>-DCDd.

Tabla 5-23 Evaluación conjunta para DB03.

Designación	ROC	ARP			ARR			ANMRR
		$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 2$	
Hrgb10-IH1		<b>0.5542</b>	<b>0.4578</b>	<b>0.3454</b>	<b>0.3454</b>	0.4234	<b>0.4852</b>	<b>0.5773</b>
Hrgb10-IH2	X	<b>0.5542</b>	<b>0.4578</b>	<b>0.3454</b>	<b>0.3454</b>	<b>0.4235</b>	<b>0.4852</b>	<b>0.5773</b>

El histograma de color es claramente superior para la base DB03. Nuevamente el histograma con 1000 rangos se presenta como el mejor en este caso con la distancia de intersección de histograma IH2.

Tabla 5-24 Evaluación conjunta para DB04.

Designación	ROC	ARP			ARR			ANMRR
		$\alpha = 0.25$	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 1$	$\alpha = 1.5$	$\alpha = 2$	
ACCC5D1-L1N		<b>0.4613</b>	<b>0.3504</b>	0.2341	0.2341	0.2726	0.3015	0.7143
DCCD-L2		0.2985	0.2754	<b>0.2602</b>	<b>0.2602</b>	0.2837	0.3063	0.7198
Hrgb10-IH1	1	0.3715	0.3037	0.2347	0.2347	<b>0.2900</b>	<b>0.3339</b>	0.7125
ACCC3D1-L1N		0.4465	0.3383	0.2306	0.2306	0.2744	0.3099	<b>0.5032</b>
DCDhsv-DCDd	2	0.0999	0.0706	0.0513	0.1634	0.2174	0.2637	0.7923

Para la base DB04, es un poco más complejo elegir el mejor descriptor ya que tienen desempeños muy similares; en este caso gracias a la métrica ROC se elige al histograma con 1000 rangos (Hrgb10-IH1) como el más sobresaliente.

## 6. Conclusiones

Aunque parece controvertido respecto a otros estudios publicados del área, de acuerdo a los datos que se obtuvieron en el presente trabajo de tesis, el histograma de color mostró ser un método muy confiable, lo cual nos permite concluir que es el descriptor con mejor desempeño.

Una de las aportaciones del presente trabajo de tesis es, la modificación de la Distancia de Color Dominante DCDd, lo que permitió obtener la Distancia de Auto correlograma de Colores dominantes ADCDd; lo anterior resultó en un mejor desempeño de los descriptores de Auto Correlograma de Colores Dominantes, tanto en el espacio RGB como HSV: ADCD, DCCD, y ACGCDC.

Una segunda aportación de este trabajo es, la generación de la base de resultados técnicos completos, que se pueden consultar en [43], ya que después de una búsqueda exhaustiva no se encontró una información parecida ya publicada, que permita reproducir los diferentes algoritmos. En la literatura revisada cada investigador utiliza diferentes parámetros y muchos no indican cuáles son éstos, lo cual dificulta su posterior comparación de manera confiable. Esto se refleja en los resultados tan dispares obtenidos en este trabajo respecto a los de otros investigadores, por lo que sería conveniente que en trabajos futuros los parámetros con los que se trabaje, tanto para los descriptores como para las métricas, fuesen más más detallados.

Una propuesta para trabajos futuros es llevar a cabo una experimentación más completa y con parámetros que pudieran estandarizarse, además de llevar a cabo una depuración de los programas desarrollados para efectos de optimización y una validación más exhaustiva. Asimismo sería conveniente realizar el mismo estudio para descriptores de textura, forma y puntos característicos que en este trabajo no se consideraron por falta de tiempo.

## 7. Referencias

- [1] S. Atani, R. Kasturi y R. Jain, «A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video,» *Pattern Recognition*, vol. 35, nº 4, pp. 945-965, 2002.
- [2] M. K. Mandal, F. Idris y S. Panchanathan, «A critical evaluation of image and video indexing techniques in the compress domain,» *Image and Vision Computing*, vol. 17, nº 7, pp. 513-529, 1999.
- [3] C. YU, High-Dimensional Indexing. Transformational Approches to High Dimensional Range and Similarity Searches, Verlag, Berlin, Heidelberg, Alemania: Springer, 2002.
- [4] W. H. Adams, G. Iyengar, C.-Y. Lin, M. R. Naphade, C. Neti, H. J. Nock y R. Smith, «Semantic indexing of multimedia content using visual, audio and text cues,» *EURASIP Journal on Applied Signal Processing*, vol. 2003, nº Enero, pp. 170-185, 2003.
- [5] R. Brunelli, O. Mitch y C. M. Modena, «A survey on the Automatic Indexing of Video Data,» *Journal of Visual Communication and image Representation*, vol. 10, nº 2, pp. 78-112, 1999.
- [6] H. Y. Lee, H. K. Lee y Y. H. Ha, «Spatial Color Descriptor for Image Retrieval and Video Segmentation,» *IEEE Transactions on Multimedia*, vol. 5, nº 3, pp. 358-367, Septiembre 2003.
- [7] H. Lejsek, F. H. Ásmundsson, B. Y. Jónsson y L. Amsaleg, «NV-tree: An efficient disk-based index for approximate search in very large high-dimensional collections,» Universidad de Reikiavik, Reikiavik, Islandia, 2007.
- [8] M. K. Mandal, T. Aboulnasr y S. Panchanathan, «Image indexing using moments and wavelets,» *Transactions on Consumer Electronics*, vol. 42, nº 3, pp. 557-565, Agosto 1996.
- [9] V. Sánchez Luna, *Indexado de Imágenes aplicando la transformada polinomial*, México: UNAM, 2008.
- [10] J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu y R. Zabih, «Image Indexing using Color Correlograms,» de *Conference on Computer Society Conference on Computer Vision and Pattern Recognition*, Ft. Collins, CO, USA, 1999.
- [11] I. Kokkinos, G. Evangelopoulos y P. Maragos, «Texture Analysis and Segmentation using modulation features, generative models, and Weighted curve evolution,» *IEEE Transactions on pattern Analysis and machine intelligence*, vol. 31, nº 1, pp. 142-157, Enero 2009.
- [12] R. C. González y R. E. Woods, *Digital image Processing*, New Yersey, USA: Prentice Hall, 2002.
- [13] J. W. O'Brien, *The JPEG Image Compression Algorithm*, Boulder, CO, USA: Universidad de Colorado, 2005.

- [14] O. Rioul y M. Vetterli, «Wavelets and Signal Processing,» *Signal Processing Magazine, IEEE*, vol. 8, nº 4, pp. 14-38, 1991.
- [15] G. Lu y S. Teng, «A novel Image Retrieval Technique based on Vector Quantization,» de *Computational Intelligence for Modelling, Control & Automation*, Viena, Austria, 1999.
- [16] Y. Rui y T. S. Huang, «Image Retrieval: Current Techniques, Promising directions, and Open issues,» *Journal of Visual Communication and Image Representation*, vol. 10, nº 1, pp. 39-62, 1999.
- [17] H. Tamura, S. Mori y T. Yamawaki, «Textural Features Corresponding to Visual Perception,» *IEEE Transactions on system, man, and cybernetics*, vol. 8, nº 6, pp. 460-473, Junio 1978.
- [18] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafine, D. Lee, D. Steele y P. Yanker, «Query by image and video content: The QBIC system,» *IEEE Computer*, vol. 28, nº 9, pp. 23-32, 1995.
- [19] J. A. Vega Uribe y M. A. Reyes Figueroa, «Transformaciones lineales y no lineales para espacios de color en procesamiento digital de imágenes,» *Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería*, vol. 22, nº 3, pp. 223-240, 2006.
- [20] «Color Coverision Algorithm,» Rochester Institute of Technology, 2000 Febrero 20. [En línea]. Available: [http://www.cs.rit.edu/~ncs/color/t\\_convert.html](http://www.cs.rit.edu/~ncs/color/t_convert.html). [Último acceso: 26 07 2015].
- [21] M. J. Swain y D. H. Ballard, «Color Indexing,» *International journal of Computer*, vol. 7, nº 1, pp. 11-32, 1991.
- [22] B. S. Manjunath, J.-R. Ohm, V. V. Vasudevan y A. Yamada, «Color and Texture Descriptors,» *Transactions on circuits and systems for Video Technology*, vol. 11, nº 6, pp. 703-715, Junio 2001.
- [23] K.-M. Wong, L.-M. Po y K.-W. Cheung, «Dominant Color Structure Descriptor for Image Retrieval,» de *International Conference on image Processing*, San Antonio, TX, USA, 2007.
- [24] Y. Linde, A. Buzo y R. M. Gray, «An Algorithm for Vector Quantizer Design,» *IEEE Transactions on communications*, vol. 28, nº 1, pp. 84-95, Enero 1980.
- [25] M. S. Jayanthi, K. Swapnil Marothi, T. M. Ishaq Mohammed Abbas y D. A. Shanmugam, «Performance Analysis of Vector Quantizer using Modified Generalized Lloyd Algorithm,» *International journal of Imaging Science and Engineering*, vol. 1, nº 1, pp. 11-15, Enero 2007.
- [26] N.-C. Yang, C.-M. Kuo, W.-H. Chang y T.-H. Li, «A fast MPEG-7 dominant color extraction with new similarity measure for image retrieval,» *Journal of Visual Communication & Image Representation*, vol. 19, nº 2, pp. 92-105, 2008.
- [27] A. N. Fierro-Radilla, M. Nakano-Miyatake, K. Perez-Daniel, H. Perez-Menea, F. Garcia-Ugalde y M. Cedillo-Hernandez, «Image Retrieval System Based on Combination of Color, Texture

and Shape Features,» de *International Conferences on Advances in Multimedia*, Barcelona, España, 2015.

- [28] M. B. Rao, B. P. Rao y A. Govardhan, «Content based image retrieval using dominant color, texture and shape,» *International Journal of Engineering Science and Technology*, vol. 3, nº 4, pp. 2887-2896, 2011.
- [29] E. Kasutani y A. Yamada, «The MPEG-7 Color Layout Descriptor: A compact image feature description for High-Speed Image/Video segment retrieval,» de *International Conference on image Processing*, Salónica (Thessaloniki ), Grecia, 2001.
- [30] R. Balasubramani y V. Kannan, «Efficient use of MPEG-7 Color Layout and Edge Histogram Descriptors in CBIR Systems,» *Global Journal of Computer Science and Technology*, vol. 9, nº 5, pp. 157-163, 2009.
- [31] D. Marshall, «The Discrete Cosine Transform (DCT),» The JNT Association, 10 Abril 2001. [En línea]. Available: <http://www.cs.cf.ac.uk/Dave/Multimedia/node231.html>. [Último acceso: 2015 julio 30].
- [32] A. Talib, M. Mahmuddin, H. Husni y L. E. George, «Efficient, Compact, and Dominant Color Correlogram Descriptors for Content-based Image Retrieval,» de *International Conferences on Advances in Multimedia*, Venecia, Italia, 2013.
- [33] B. Sebastian V, A. Unnikrishnan y K. Balakrishnan, «Grey level co-occurrence matrices: Generalisation and some new features,» *International Journal of Computer Science, Engineering and Information Technology*, vol. 2, nº 2, pp. 151-157, 2012.
- [34] D. Zhang y G. Lu, «Evaluation of Similarity Measurement for Image Retrieval,» de *International Conference on Neural Networks and Signal Processing*, Nanjing, China, 2003.
- [35] J. C. Caicedo, «Histogram Intersection,» 26 Enero 2012. [En línea]. Available: <http://jccaicedo.blogspot.mx/2012/01/histogram-intersection.html>. [Último acceso: 2015 Julio 20].
- [36] M. Molina, «Ciencia sin seso...locura doble,» 27 Noviembre 2012. [En línea]. Available: <http://www.cienciasinseeso.com/tag/curva-roc/>. [Último acceso: 2015 Agosto 6].
- [37] W. Bian y D. Tao, «Biased Discriminant Euclidean Embedding for Content based Image Retrieval,» *IEEE Transactions on image Processing*, vol. 19, nº 2, pp. 545-554, Febrero 2010.
- [38] D. Tao, X. Li y S. J. Maybank, «Negative Samples Analysis in Relevance Feedback,» *IEEE Transactions on Knowledge and Data Engineering*, vol. 19, nº 4, pp. 568-580, Abril 2007.
- [39] D. Tao, X. Tang, X. Li y X. Wu, «Asymmetric Bagging and Random Subspace for Support Vector Machines based Relevance Feedback in image Retrieval.,» *Transactions on Pattern Analysis and machine Intelligence*, vol. 28, nº 7, pp. 1088-1099, 2006.

- [40] D. Tao, X. Tang, X. Li y Y. Rui, «Direct Kernel Biased Discriminant Analysis A New Content based Image Retrieval Relevance Feedback Algorithm,» *Transactions on Multimedia*, vol. 8, nº 4, pp. 716-727, Agosto 2006.
- [41] J. Li, N. Allinson, D. Tao y X. Li, «Multitraining Support Vector Machine for Image Retrieval,» *Transactions on Image Processing*, vol. 15, nº 11, pp. 3597-3601, Noviembre 2006.
- [42] S. Agarwal, A. K. Verma y N. Dixit, «Content Based Image Retrieval using Color Edge detection and Discrete Wavelet Transform,» de *International Conference on issues and Challenges in Intelligent Computing Techniques*, Ghaziabad, India, 2014.
- [43] L. J. Reyes Ruiz, «Resultados técnicos de los diferentes descriptores combinados con las diferentes distancias,» 9 octubre 2015. [En línea]. Available: <http://scccerezo.wix.com/listadotecnico>. [Último acceso: 9 octubre 2015].

# ANEXO I. Bases Utilizadas

La base de imágenes COREL DB (CorelDB) se puede consultar en:

<https://sites.google.com/site/dctresearch/Home/content-based-image-retrieval>

La base de imágenes de Wang (Image.Orig) se puede consultar en:

<https://sites.google.com/site/dctresearch/Home/content-based-image-retrieval>

## I-1. DB01

De Corel DB, se tomaron 10 categorías elegidas aleatoriamente con 20 imágenes cada una, dando un total de 200 imágenes.

*Tabla I-1 Imágenes que conforman la DB01*

	Carpeta	Categoría	Imágenes
1	wl_leopard	chitas	134000.jpg-134019.jpg
2	obj_train	trenes	351000.jpg-351019.jpg
3	obj_aviation	aviones	360000.jpg-360019.jpg
4	obj_mineral	cuarzos	409000.jpg-409019.jpg
5	pl_flower	flores	84000.jpg-84019.jpg
6	pet_dog	perros	247000.jpg-247019.jpg
7	fitness	personas ejercitándose	280000.jpg-280019.jpg
8	obj_orbits	dibujos órbitas	617000.jpg-617019.jpg
9	wl_goat	cabras	168000.jpg-168019.jpg
10	obj_moleculr	dibujos moléculas	533000.jpg-533019.jpg

Se consideraron la primeras 20 imágenes de cada categoría seleccionada.

## I-2. DB02

De la base de imágenes de Wang, se tomaron todas las imágenes, divididas en 10 categorías con 100 imágenes cada una, dando un total de 1000 imágenes.

*Tabla I-2 Imágenes que conforman la DB02*

	Categoría	Imágenes
1	tribal	0.jpg-99.jpg
2	playa	100.jpg-199.jpg
3	arquitectura romana	200.jpg-299.jpg
4	autobuses	300.jpg-399.jpg
5	dinosaurios	400.jpg-499.jpg
6	elefantes	500.jpg-599.jpg
7	rosas	600.jpg-699.jpg
8	caballos	700.jpg-799.jpg
9	montañas	800.jpg-899.jpg
10	platillos	900.jpg-999.jpg

Se utilizaron todas las imágenes.

## I-3. DB03

De Corel DB, 20 categorías elegidas aleatoriamente con 50 imágenes cada una, dando un total de 1000 imágenes.

*Tabla I-3 Imágenes que conforman la DB02*

	Carpeta	Categoría	Imágenes
1	sp_ski	esquiando	225000.jpg-225049.jpg
2	sc_mountain	montañas (nevadas)	114000.jpg-114049.jpg
3	obj_door	puertas	59000.jpg-59049.jpg
4	texture_5	brillo metálico	343000.jpg-343049.jpg
5	sc_firewrk	fuegos artificiales	40000.jpg-40049.jpg
6	obj_mineral	cuarzos	409000.jpg-409049.jpg
7	art_cybr	arte espacial	283000.jpg-283049.jpg
8	wl_horse	caballos	113000.jpg-113049.jpg
9	wl_elephant	elefantes	107000.jpg-107049.jpg
10	texture_3	fractales	473000.jpg-473049.jpg
11	obj_flags	banderas del mundo	461000.jpg-461049.jpg
12	sc_rural	campestre con flores	131000.jpg-131049.jpg
13	wl_porp	delfines	314000.jpg-314049.jpg
14	texture_1	cromado colorido	186000.jpg-186049.jpg
15	wl_wolf	lobos	110000.jpg-110049.jpg
16	obj_bus	autobuses	581000.jpg-581049.jpg
17	wl_owls	lechuzas	75000.jpg-75049.jpg
18	obj_bonsai	bonsáis	353000.jpg-353049.jpg
19	wl_goat	cabras	168000.jpg-168049.jpg
20	wl_butterfly	mariposas	52000.jpg-52049.jpg

Se consideraron la primeras 50 imágenes de cada categoría seleccionada.

## I-4. DB04

De Corel DB, 40 categorías elegidas aleatoriamente con 100 imágenes cada una, dando un total de 4000 imágenes.

*Tabla I-4 Imágenes que conforman la DB02*

	Carpeta	Categoría	Imágenes
1	sc_forests	bosques	482000.jpg-482099.jpg
2	sc_indoor	interiores de las casas	364000.jpg-364099.jpg
3	sc_sunset	atardeceres	191010.jpg, 191028.jpg, 191034.jpg, 345002.jpg, 345014.jpg-345016.jpg, 345018.jpg, 345023.jpg, 345024.jpg, 345028.jpg, 345030.jpg, 345031.jpg, 345033.jpg, 345036.jpg, 345039.jpg-345044.jpg, 345048.jpg, 345053.jpg, 345055.jpg-345058.jpg, 345060.jpg, 345063.jpg, 345064.jpg, 345066.jpg, 345073.jpg, 345074.jpg, 345077.jpg, 345079.jpg, 345081.jpg, 345090.jpg, 345092.jpg-345094.jpg, 345097.jpg, 460000.jpg, 460001.jpg, 460003.jpg-460006.jpg, 460013.jpg, 460014.jpg, 460017.jpg, 460022.jpg, 460024.jpg, 460034.jpg-460036.jpg, 460039.jpg, 460043.jpg, 460059.jpg, 460061.jpg, 460065.jpg, 460066.jpg, 460070.jpg, 460074.jpg, 460082.jpg-460085.jpg, 460091.jpg, 460093.jpg, 460096.jpg, 460097.jpg, 460099.jpg, 604002.jpg, 604004.jpg, 604009.jpg-604011.jpg, 604020.jpg, 604063.jpg, 604068.jpg, 604079.jpg, 604080.jpg, 664066.jpg-664068.jpg, 790003.jpg, 790005.jpg, 790009.jpg, 790015.jpg, 790021.jpg,

			790028.jpg, 790031.jpg, 790036.jpg, 790037.jpg, 790039.jpg, 790040.jpg, 790047.jpg, 790081.jpg, 790082.jpg, 790084.jpg
4	obj_moleculr	dibujos moléculas	533000.jpg-533099.jpg
5	pl_foliage	hojas	318000.jpg-318099.jpg
6	wl_tiger	tigres	108000.jpg-108099.jpg
7	obj_balloon	globos	329000.jpg-329006.jpg, 329008.jpg, 329009.jpg, 329012.jpg, 329013.jpg, 329017.jpg, 329019.jpg, 329021.jpg, 329022.jpg, 329024.jpg-329029.jpg, 329031.jpg, 329032.jpg, 329035.jpg-329040.jpg, 329042.jpg-329045.jpg, 329048.jpg-329051.jpg, 329051.jpg, 329053.jpg-329099.jpg, 528001.jpg-528008.jpg, 528010.jpg-528017.jpg
8	wl_lizard	reptiles	175000.jpg-175099.jpg
9	obj_door	puertas	59000.jpg-59099.jpg
10	texture_4	cielo gris	350000.jpg-350099.jpg
11	art_cybr	arte espacial	283000.jpg-283099.jpg
12	sc_waterfal	cascadas	167014.jpg, 167028.jpg, 167029.jpg, 167031.jpg-167033.jpg, 167035.jpg, 167038.jpg, 167039.jpg, 266005.jpg, 266006.jpg, 266008.jpg-266014.jpg, 377000.jpg-377081.jpg
13	art_antiques	muebles antiguos	435000.jpg-435099.jpg
14	sc_mountain	montañas (nevadas)	114000.jpg-114056.jpg, 114058.jpg-114063.jpg, 114067.jpg-114073.jpg, 114075.jpg, 114076.jpg, 114078.jpg-114081.jpg, 114083.jpg-114086.jpg, 114090.jpg-114095.jpg, 114097.jpg-114099.jpg, 167001.jpg, 167006.jpg, 167013.jpg, 167058.jpg, 167064.jpg, 167076.jpg, 167077.jpg, 167079.jpg-167082.jpg
15	sc_waves	olas del mar	312000.jpg-312099.jpg
16	obj_train	trenes	351000.jpg-351099.jpg
17	sc_rural	campestre con flores	131000.jpg-131099.jpg
18	wl_fox	zorros	109000.jpg-109099.jpg
19	sc_firewrk	fuegos artificiales	40000.jpg-40099.jpg
20	obj_steameng	locomotoras de vapor	468000.jpg-468099.jpg
21	woman	mujeres	181006.jpg, 181009.jpg, 181019.jpg, 181026.jpg, 181029.jpg, 181032.jpg, 181033.jpg, 181035.jpg, 181039.jpg-181043.jpg, 181051.jpg, 181052.jpg, 181055.jpg, 181063.jpg, 181066.jpg, 181070.jpg, 181082.jpg, 181090.jpg, 181095.jpg, 181096.jpg, 198000.jpg-198077.jpg
22	obj_car	carros	29000.jpg-29003.jpg, 29005.jpg-29015.jpg, 29017.jpg, 29019.jpg-29024.jpg, 29026.jpg, 29027.jpg, 29029.jpg-29045.jpg, 29047.jpg-29099.jpg, 273001.jpg, 273001.jpg, 273011.jpg, 273020.jpg-273022.jpg, 273026.jpg
23	wl_roho	hipopotamos y rinocerontes	112000.jpg-112099.jpg
24	wl_cat	linces	326000.jpg-326099.jpg
25	sc_iceburg	paisajes árticos	184000.jpg-184099.jpg
26	sc_night	noche y luces	287000.jpg-287099.jpg
27	eat_feasts	comida banquetes	569000.jpg-569099.jpg
28	sc_	estalactitas	218000.jpg-218099.jpg
29	eat_drinks	cocteles	275000.jpg-275099.jpg
30	obj_bus	autobuses	581000.jpg-581099.jpg
31	sc_rockform	desértico	407000.jpg-407099.jpg
32	wl_buttrfly	mariposas	52000.jpg-52099.jpg
33	pl_mashroom	hongos	208000.jpg-208099.jpg
34	obj_eastregg	huevos decorados	618000.jpg-618099.jpg
35	obj_bonsai	bonsáis	353000.jpg-353099.jpg
36	obj_cards	barajas	425000.jpg-425099.jpg
37	obj_decoys	señuelos	655000.jpg-655099.jpg

38	obj_234000	adornos festivos	234000.jpg-234099.jpg
39	wl_fish	peces	141000.jpg-141009.jpg, 141020.jpg-141022.jpg, 141024.jpg, 141025.jpg, 141027.jpg-141048.jpg, 141051.jpg, 141054.jpg-141057.jpg, 141060.jpg, 141067.jpg-141071.jpg, 171073.jpg, 141077.jpg, 141079.jpg-141082.jpg, 141085.jpg, 141086.jpg, 141090.jpg-141099.jpg, 164000.jpg, 164007.jpg, 164015.jpg, 164018.jpg, 164019.jpg, 164028.jpg, 164030.jpg, 164034.jpg, 164036.jpg, 164042.jpg, 164076.jpg, 164094.jpg, 427004.jpg-427017.jpg, 427019.jpg-427026.jpg
40	sp_ski	esquiando	225000.jpg-225099.jpg

Se consideraron la primeras 100 imágenes de cada categoría seleccionada.