

CAPÍTULO 2

PROCESAMIENTO DIGITAL DE IMÁGENES

En este capítulo se presentan de manera breve, una explicación de la visión, las imágenes digitales y como son capturadas por medios electrónicos, el campo encargado del procesado digital de imágenes (DIP) y las técnicas de DIP utilizadas en este trabajo. Así mismo se presentan los fundamentos de la morfología matemática y las técnicas más comunes utilizadas con operadores morfológicos. Por último se detallan dos técnicas consideradas para la comparación del método propuesto, la transformación Top-Hat y el filtro gaussiano de desvanecimiento.

2.1 Introducción

Basándose en que una señal de video es una secuencia de imágenes estáticas se determina que es posible realizar un análisis de una señal de video al procesar una cierta cantidad de imágenes estáticas en un intervalo de tiempo.

Las imágenes pueden ser digitalmente analizadas ya sea por técnicas de procesamiento digital de imágenes (DIP) o con teoría de conjuntos, denominada morfología matemática. Ambas herramientas son un campo muy amplio de estudio y se presentan simplemente los fundamentos de ambas teorías, así como las operaciones que fueron utilizadas en el presente trabajo, sin embargo las capacidades de ambos métodos son muy extensas para ser cubiertas en esta tesis.

Se determina que para comparar el método propuesto se debe realizar un análisis similar al propuesto pero con dos técnicas diferentes de análisis de imágenes. La primera de ellas es una técnica DIP, un filtro gaussiano pasa-bajas o también conocido como filtro gaussiano de “suavizado” y la otra técnica que se estableció es la

transformada Top-Hat, una transformación morfológica que resalta las secciones más intensas en una imagen binaria. El uso de tanto técnicas DIP como transformaciones morfológicas determinará las ventajas y desventajas de cada método.

2.2 Visión, imágenes y video

El ojo humano es sensible únicamente a la radiación contenida en una pequeña banda de frecuencias del espectro electromagnético (desde 380 nm hasta 780 nm) y los colores son representados por una longitud de onda dentro de este rango. Para la percepción de los colores, el ojo humano presenta dos tipos de sensores que son denominados las células foto-receptoras. Estas células se dividen en conos y bastones, los primeros de ellos son sensibles en diferentes proporciones a los colores, rojo 65%, verde 33% y azul 2%, mientras que los bastones tienen una respuesta espectral menos selectiva pero más sensible, como es el caso de la visión nocturna. De esta forma el ojo humano percibe los estímulos de los 3 colores por acción de los conos que actúan como un banco de filtros de los estímulos que llegan al ojo humano y la sensación de color es la respuesta aditiva de estas 3 señales [ALBA06].

Una imagen es la representación visual de un objeto y está definida como una función bidimensional $f(x, y)$, donde “ x ” y “ y ” son coordenadas espaciales y la amplitud f en un par de coordenadas (x, y) se llama la intensidad o nivel de gris de la imagen en ese punto. Cuando los valores de “ x ” y “ y ”, así como la amplitud f son cantidades finitas y discretas se habla de una imagen digital (2.1).

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & \dots & f(1,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1,N-1) \end{bmatrix} \quad (2.1)$$

En donde el lado izquierdo de la ecuación denota una imagen digital por definición, mientras que en el lado derecho de la ecuación encontramos que cada uno de los elementos de la imagen y estos se definen como un “elemento de la imagen” o un pixel. La notación de coordenadas que se utiliza en la mayoría de los libros [GONZ02], y la aceptada para este trabajo, es la mostrada en la ecuación (2.1), en donde la imagen contiene M filas y N columnas determinando el origen como el punto $f(0,0)$.

El proceso de digitalización de una imagen consiste en la adquisición que normalmente es realizada por cámaras digitales basadas en sensores CCD o CMOS. Una vez que se adquiere la imagen ocurre el proceso de digitalización en el cual se muestrea (discretización espacial) y se cuantifica (discretización en amplitud) la señal analógica. Este muestreo y cuantificación limitan la resolución y calidad de la imagen respectivamente.

El ojo humano retiene una imagen aproximadamente 1/16 de segundo cuando capta una imagen, de tal manera que si varias imágenes secuenciales son presentadas al ojo humano en un intervalo de tiempo adecuado (entre 15 y 50 imágenes por segundo evitan el parpadeo y la discontinuidad en secuencias de video), el cerebro las asocia como una sola imagen en movimiento. Dado que una señal de video corresponde a una secuencia de imágenes estáticas, es posible realizar el procesamiento de una señal de video mediante el procesamiento de una cierta cantidad de imágenes de la señal por segundo.

2.3 Procesado digital de imágenes

El campo que se encarga del procesado de las imágenes digitales por medios de una computadora digital es el procesado digital de imágenes. La mayor parte de las técnicas

DIP actúan tratando a la imagen como una señal de 2 dimensiones (2D) y después aplicando técnicas estándar de procesamiento de señales de una dimensión (1D). Dentro de las operaciones más comunes de procesamiento encontramos transformaciones geométricas (reducción, rotación, alargamiento), correcciones de color (ajustes de brillo y contraste), alineación de imágenes, segmentación, interpolación, reconocimiento de patrones o características en una imagen, etc.

El desarrollo de tecnología en los campos de adquisición de imágenes y sistemas de cómputo más eficientes, han logrado que las aplicaciones del DIP se extiendan a muy diversas áreas como lo son aplicaciones médicas, monitoreo remoto, reconocimiento de objetos, visión por computadora, manipulación de fotografías, aplicaciones militares, etc.

2.4 Interpolación bicúbica

La interpolación es el proceso de estimar valores intermedios en un evento continuo a partir de muestras discretas. Este método es ampliamente utilizado en DIP para aumentar o reducir imágenes y también para corregir distorsiones geométricas.

El algoritmo de interpolación bicúbica propuesto por [KEY81] es uno de los más utilizados aún hoy en día, debido a su bajo costo computacional y los resultados que se obtienen en comparación con la interpolación bilineal y la interpolación del vecino más próximo [THE00]. Este método estima un pixel analizando la vecindad de 4×4 alrededor de él mediante un algoritmo de polinomios cúbicos definidos en sub-intervalos, tal como se muestra en (2.2) en donde $u(s)$ es la función núcleo o “kernel” de la interpolación bicúbica.

$$u(s) = \begin{cases} +\frac{3}{2}|s|^3 - \frac{5}{2}|s|^2 + 1 & 0 < |s| < 1 \\ -\frac{1}{2}|s|^3 + \frac{5}{2}|s|^2 - 4|s| + 2 & 1 < |s| < 2 \\ 0 & 2 < |s| \end{cases} \quad (2.2)$$

2.5 Tipos de imágenes

2.5.1 Imágenes RGB

Una imagen RGB se define como un arreglo de $3 \times M \times N$ píxeles donde cada píxel corresponde a los componentes rojo, verde y azul de una imagen de color (R-red, G-green, B-blue). El propósito principal del modelo RGB es el sensado, representación y muestra de imágenes en aparatos electrónicos tal como televisores, computadoras, celulares, etc.

El modelo RGB puede verse como una pila de 3 imágenes en escala de intensidades que al ser mostrados por un monitor de color (el cual presenta 3 entradas de color, rojo, verde y azul) representan una imagen de color tal como la percibe un ser humano. Los colores, rojo, verde y azul son conocidos como los colores primarios, y la combinación de estos en diferentes intensidades produce los colores del espectro humano visible.

2.5.2 Imágenes en escala de intensidades

Una escala de intensidades es también conocida como escala de grises o escala monocromática, y para una imagen digital se trata de una sola matriz de $M \times N$ valores en donde cada píxel es una sola muestra que contiene la información de la intensidad de la imagen.

Las 3 características con que se puede definir un color son el tono o matiz (color actual), el valor (luminosidad u oscurecimiento) y la saturación (pureza del color). Por lo tanto la conversión de una imagen de color a una imagen en escala de grises no se realiza de una forma única, sin embargo en su acercamiento más común [POS90], se trata de retener la información de la luminosidad y descartar los valores de tono y saturación. Asumiendo que los colores, rojo, verde y azul son señales de luminosidad, la aproximación de una imagen en escala de grises a partir de una imagen de color está dada por (2.3) en donde 0 es el valor de menos intensidad, refiriéndose al color negro y 1 es el valor de mayor luminosidad o el color blanco.

$$GRAY = (0.30 \cdot R) + (0.59 \cdot G) + (0.11 \cdot B) \quad (2.3)$$

2.5.3 Imágenes indexadas

Esta es una forma práctica, pero limitada, de representar imágenes en color. Una imagen indexada almacena una imagen como dos matrices. La primera de ellas tiene el mismo tamaño que la imagen y un número para cada pixel. La segunda matriz se denomina mapa de color y su tamaño corresponde al número de colores que se desea que tenga la nueva imagen. Esta técnica es utilizada para ahorrar memoria de la computadora, comprimir imágenes y lograr una mejoría en el tiempo de muestra “display”. La información de la imagen pasa de estar guardada en las $3 \times M \times N$ matrices que comúnmente se utilizan en el formato RGB a solamente $1 \times M \times N$ matrices de la información de la imagen y una matriz de $m \times 3$, en donde m es el número de colores del mapa de color que se desea utilizar [GONZ02].

Las ventajas de utilizar esta técnica no sólo se limitan el ahorro de memoria, compresión de imágenes y mejora de tiempos de transmisión o muestra de imágenes en

color, si no que puede dársele un enfoque a ciertas aplicaciones como el caso de esta tesis, en donde se requiere eliminar los detalles de una imagen, que se logra reduciendo el número de tonos del mapa de color para agrupar todos los tonos que estén dentro de un rango hacía un solo color definido. De esta forma aunque normalmente el limitado número de tonos de un mapa de color es una desventaja en el caso de este trabajo se utiliza como una herramienta en el algoritmo de detección de columnas de humo.

2.6 Operaciones morfológicas

En esta sección se expone una descripción básica de las operaciones morfológicas más comunes para comprender la etapa de cierre de imágenes utilizada en el método propuesto. También se explica la Transformada Top-Hat, que fue usada para comparar los resultados de la simulación generada con el método propuesto.

2.6.1 Fundamentos de morfología matemática

La descripción básica de la morfología matemática se basa en la teoría de conjuntos y en la topología. Cada conjunto representa la forma de los objetos en una imagen binaria. Dada esta equivalencia entre conjuntos e imágenes, cualquier transformación sobre un conjunto define una nueva imagen utilizando la correspondencia entre los operadores lógicos básicos y operaciones básicas sobre conjuntos [BEU90].

El objetivo de las transformaciones morfológicas es la extracción de estructuras geométricas en los conjuntos sobre los que se opera, mediante la utilización de otro conjunto conocido como elemento estructurante (EE) (ver Figura 2.1).

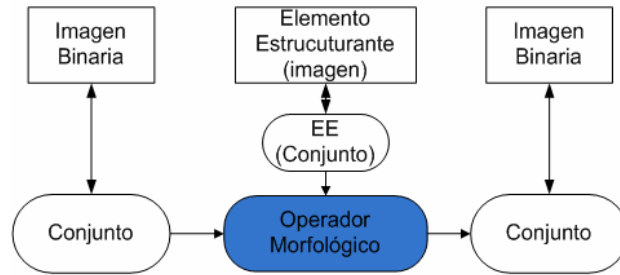


Figura 2.1 Fundamento del procesado morfológico

El tamaño y la forma de este elemento se escogen de acuerdo a la morfología del conjunto sobre el que va a interactuar y de acuerdo a la extracción de formas que se desean obtener, existen diversas formas y tamaños del EE (ver Figura 2.2). Desplazando el EE sobre la imagen, el operador analiza típicamente su posición en relación al primer plano y fondo de la misma.

Actualmente, el ámbito y alcance de los procesos morfológicos es tan amplio como el propio procesamiento de imágenes. Se pueden encontrar aplicaciones tales como la segmentación, restauración, detección de bordes, aumento de contraste, análisis de texturas, compresión, etc. Las aplicaciones básicas de la morfología matemática son operaciones de erosión, dilatación, apertura y cierre.

2.6.2 Operaciones básicas de morfología matemática

La transformación de erosión es el resultado de comprobar si el elemento estructurante Y está totalmente incluido dentro del conjunto X . Cuando esto no ocurre el resultado de

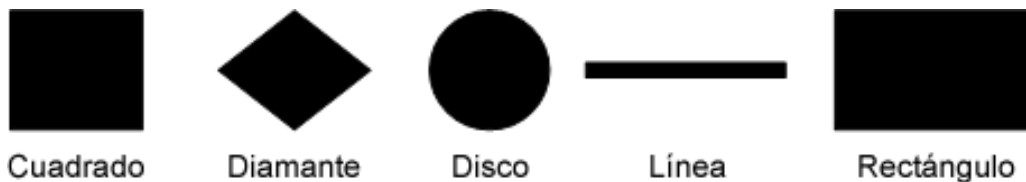


Figura 2.2 Formas típicas de los elementos estructurantes

la erosión es el conjunto vacío. La erosión de un conjunto X por un EE Y , se define como el conjunto de puntos o elementos x , pertenecientes a X , de forma que cuando el EE Y se traslada a ese punto, el elemento queda excluido en X . La ecuación (2.4) muestra la definición de erosión.

$$\varepsilon_Y(X) = \{x \mid Y_x \subseteq X\} \quad (2.4)$$

La dilatación es la operación contraria a la erosión. La dilatación es el conjunto de puntos origen del EE Y , tales que el EE contiene algún elemento del conjunto X , cuando el elemento se desplaza por el espacio que contiene a ambos conjuntos (2.5):

$$\delta_Y(X) = \{x \mid Y_x \cap X \neq \emptyset\} \quad (2.5)$$

La apertura de una señal f por un EE Y se denota por $\gamma_Y(f)$ y se define como la erosión de f por Y , seguida de la dilatación por el mismo EE (2.6):

$$\gamma_Y(f) = \delta_Y(\varepsilon_Y(f)) \quad (2.6)$$

El cierre de una señal f por un EE Y se denota por $\varphi_Y(f)$, se define como la dilatación de f por Y , seguida por la erosión por el mismo EE (2.7):

$$\varphi_Y(f) = \varepsilon_Y(\delta_Y(f)) \quad (2.7)$$

El efecto de las operaciones de dilatación, erosión, apertura y cierre pueden ser observados en la Figura 2.3.

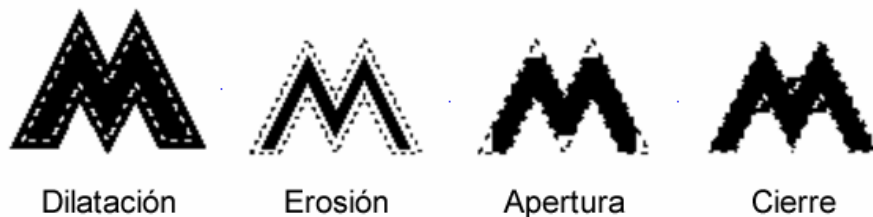


Figura 2.3 Efectos de las operaciones más comunes de morfología matemática

2.6.3 Transformada Top-Hat

La transformación de Top-Hat consiste en descubrir aquellas estructuras de la imagen que han sido eliminadas en el proceso de apertura o cierre. Con la elección de un EE de forma, tamaño y orientación adecuado, es posible filtrar la imagen y eliminar determinados elementos en la imagen original. Esta transformación puede ser vista como el residuo entre la identidad y la apertura morfológica. En la Figura 2.4 se ilustra un ejemplo de la transformada Top-Hat para una imagen 1-D [ORT02]. La Figura 2.4.a presenta la señal original a tratar en gris, el EE Y en la parte superior y la señal generada por el proceso de apertura morfológica, mientras que la Figura 2.4.b presenta la señal generada por la Top-Hat, la señal no eliminada por la apertura.

2.7 Filtro gaussiano de suavizado

Se llama filtro gaussiano de suavizado al operador 2-D utilizado para desvanecer imágenes, removiendo detalles y ruido. Este filtro es un operador de convolución. La convolución de imágenes permite multiplicar dos imágenes de diferentes tamaños para generar una imagen del mismo tamaño. En DIP la operación de convolución es generalmente utilizada para generar pixeles de salida en una combinación lineal de los pixeles de entrada.

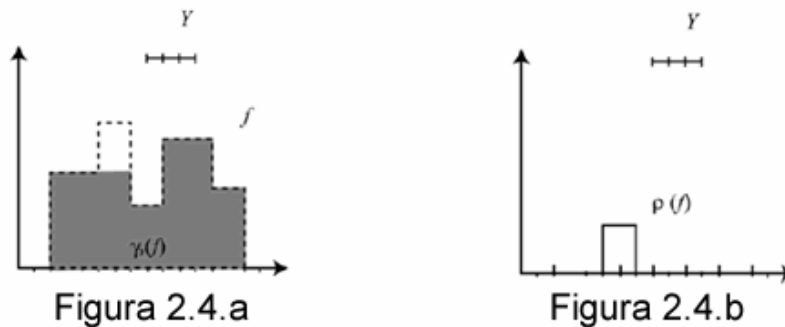


Figura 2.4 Efecto de la transformación Top-Hat

Por lo tanto se tiene una matriz a analizar que debe convolucionarse con una matriz de menor tamaño (generalmente) denominada kernel. Este kernel se genera según la fórmula (2.8) en donde los valores que pueden controlarse son la media (x,y) y la desviación estándar σ . Una distribución con media $(0,0)$ y desviación estándar en 2-D se presenta en la Figura 2.5

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2.8)$$

Como puede observarse la idea de suavizar una imagen 2D es eliminar las altas frecuencias aplicando una convolución de la señal con una función gaussiana con la imagen. Este proceso entrega por tanto una salida de pixeles con un peso ponderado según la vecindad de cada pixel, moviéndose este valor ponderado hacía el valor del pixel central de la ventana definida para la convolución. El grado de suavizado es determinado en gran medida por la desviación estándar de la función gaussiana, sin embargo valores grandes de la desviación requieren convolucionar kernels de mayor tamaño por lo que este filtro puede llegar a ser bastante complejo computacionalmente hablando, pero de igual manera tiene ciertas aplicaciones en DIP.

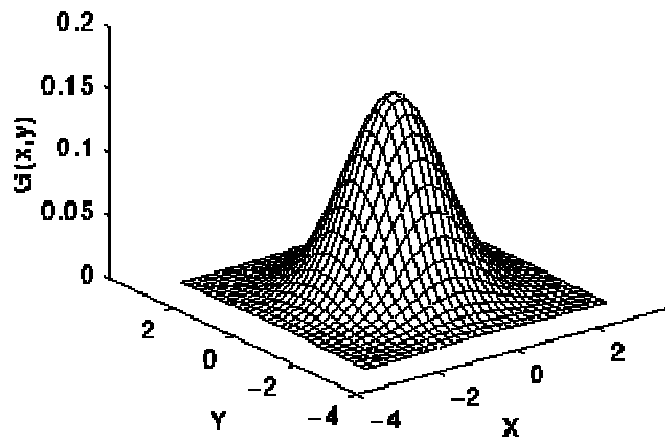


Figura 2.5 Distribución gaussiana en 2-D

2.8 Discusión

El procesamiento digital de imágenes se define como el campo que se encarga de manipular imágenes digitales con el objetivo de mejorar éstas o identificar cierta información relevante, el desarrollo de nuevas tecnologías, tanto en la adquisición de imágenes, como en los sistemas de cómputo para el procesado de dichas imágenes, han hecho de esta rama un campo de estudio interesante y en constante desarrollo.

Existen una gran cantidad de técnicas DIP, sin embargo se selecciona la interpolación bicúbica como el método de mejoramiento de calidad de la imagen tanto por su relación costo-beneficio como por la gran cantidad de documentación de este algoritmo. Así mismo se analizan las ventajas y desventajas de las transformaciones de una imagen de color a una escala de grises o una imagen indexada, resultando ambas apropiadas para el desarrollo de esta tesis.

Las operaciones morfológicas son un campo de estudio que tiene muchas ventajas sobre algunas técnicas DIP, las operaciones básicas son sencillas de entender y aplicar, es por ello que se selecciona las operaciones de cierre de imagen y la transformada Top-Hat como parte de las técnicas utilizadas en el presente trabajo de tesis.