

**ESTUDIO Y ANÁLISIS DE TÉCNICAS PARA PROCESAMIENTO DIGITAL DE
IMÁGENES.**

DIANA VANESSA GÓMEZ TREJOS.

ALEJANDRA GUERRERO GUZMÁN.

DIRECTOR: INGENIERO CARLOS AUGUSTO MENESES ESCOBAR.

UNIVERSIDAD TECNOLÓGICA DE PEREIRA.

INGENIERÍA DE SISTEMAS Y COMPUTACIÓN.

MAYO DEL 2016.

TABLA DE CONTENIDO.

	Página.
AGRADECIMIENTOS.	3
INTRODUCCIÓN.	4
1. GENERALIDADES.	
1.1 DEFINICIÓN DEL PROBLEMA.	5
1.2 JUSTIFICACIÓN DEL TRABAJO DE GRADO.	6
1.3 OBJETIVOS DEL PROYECTO.	
1.3.1 OBJETIVO GENERAL.	7
1.3.2 OBJETIVOS ESPECÍFICOS.	7
2. ESTADO DEL ARTE.	
2.1 MARCO CONCEPTUAL.	8
2.2 MARCO REFERENCIAL.	9-12
2.3 MARCO TEÓRICO.	13-57
3. ANÁLISIS COMPARATIVO ENTRE LAS TÉCNICAS EXPLICADAS	58-62
4. CONCLUSIONES.	63
5. BIBLIOGRAFÍA.	64-66

AGRADECIMIENTOS.

Al equipo director de nuestro proyecto de grado que estuvo conformado por ***Carlos Augusto Meneses***. Ingeniero de Sistemas. Msc. Profesor Asociado. Director principal del proyecto y ***Guillermo Roberto Solarte Martínez***. Ingeniero de Sistemas. Profesor Universidad Tecnológica de Pereira; que fue quien nos guió en la realización del anteproyecto.

INTRODUCCIÓN

El presente trabajo se enfoca en un estudio monográfico acerca de las técnicas del procesamiento de imágenes, sus características, aplicaciones y algunos modelos matemáticos que permiten ampliar y aclarar dichas técnicas. Con esto, se pretende brindar a las personas interesadas en el tema, un material de estudio como referencia bibliográfica o un documento de consulta que pueda ser útil para posteriores investigaciones.

En primer lugar, se brinda una explicación detallada acerca de las técnicas del procesamiento de imágenes, las cuales están agrupadas en 4 principales: Adquisición, realce y mejora, segmentación y extracción de características. Luego, se hace un análisis comparativo de dichas técnicas, que permite destacar los beneficios de utilizar una u otra según sea la aplicación. Por último se elaboran unas conclusiones a partir de todo el estudio realizado y la experiencia del proceso investigativo que se llevó a cabo.

1. GENERALIDADES.

1.1 Definición del problema.

Se ha logrado un desarrollo cuantificable en varios campos a través del procesamiento de imágenes, y uno de sus mayores avances se ha dado en el campo de la medicina; en 1971 se produjo la primera imagen del TAC (Tomografía Computarizada); por medio de la cual haciendo uso de los rayos X se obtienen imágenes del interior del organismo y de esta forma es posible detectar desde un tumor hasta daños en la columna vertebral. [1]

En la arqueología se procesan las imágenes para una observación remota de la superficie de la tierra.

Se han tomado imágenes con satélites y se les ha hecho su respectivo procesamiento, detectando todo tipo de fenómenos naturales, como huracanes, cambios de clima, tornados, entre otros; dándose así el procesamiento de imágenes en el campo climatológico.

La base de datos del consumo global de energía o base de datos de “Luces de noche en el mundo” se ha obtenido a través de procesamiento de imágenes de la banda infrarroja; logrando así una visualización leve de las luces encendidas en países, ciudades, pueblos... [2] También se ha utilizado el procesamiento digital de imágenes en el reconocimiento de la huella digital de una persona, en la autenticidad de los billetes; en estos y en otros campos.

En el momento de desarrollar soluciones que requieran de procesamiento digital de imágenes, es difícil en algunos casos determinar cuáles son las técnicas más adecuadas para el problema en particular, e inclusive a veces se desconocen estas.

Se debe pensar entonces, ¿será posible tener un documento de consulta donde se describan las distintas técnicas del procesamiento digital de imágenes, y sus aspectos relevantes de acuerdo al tipo de solución buscada para problemas reales?.

1.2 Justificación.

En los últimos años se ha logrado un gran avance en el desarrollo de productos que usan técnicas para el procesamiento digital de imágenes y existe una diversidad de áreas donde es aplicado, trayendo beneficios a quienes tienen interés en investigar y/o resolver problemas en el contexto mencionado.

La elaboración de un documento que describa las principales técnicas del procesamiento de imágenes, contribuye a la ampliación del conocimiento y capacitación en el ámbito académico, ya que puede ser usado por los estudiantes, como manual de estudio y documento de consulta en investigaciones de este tipo.

En el ámbito tecnológico, aporta modelos matemáticos que pueden ser aplicados en el desarrollo de nuevas tecnologías o bien, en investigaciones científicas para acotejar dichas técnicas en un tema específico donde se quiera aplicar.

Una descripción detallada de las técnicas para el procesamiento de imágenes digitales, traerá grandes beneficios en el ámbito social, ya que se tendrán avances en la medicina, astronomía, meteorología, y muchos más campos.

Debido al desarrollo tecnológico, se espera tener cada vez nuevos productos que sean aplicados en la industria y a nivel empresarial generando posibles beneficios económicos.

1.3 Objetivos del proyecto.

Objetivo general.

- Elaborar una monografía que describa las principales técnicas existentes para el procesamiento digital de imágenes y algoritmos de análisis de imágenes basados en modelos matemáticos.

Objetivos específicos.

- Realizar un análisis comparativo entre las técnicas de procesamiento de imágenes.
- Conocer las características de las técnicas aplicadas en el procesamiento de imágenes digitales.
- Identificar campos de aplicación para las técnicas del procesamiento digital de imágenes.

2. ESTADO DEL ARTE.

2.1 Marco conceptual.

Imagen digital: La imagen digital es un producto del desarrollo de la informática que tiene como antecesor a la fotografía, (que toma como punto de partida un objeto del mundo real) y a la pintura, (donde la imagen ha sido creada por un artista). [3]

Adquisición: La palabra adquisición proviene del término latino *adquisitio* es la acción de conseguir una determinada cosa, la cosa en sí que se ha adquirido y la persona cuyos servicios o intervención están ampliamente valorados. [4]

Procesamiento: Sometimiento de una cosa a un proceso de elaboración o de transformación. Aplicación de un programa informático a unos datos determinados. [4]

Sistemas Inteligentes: Podemos definir un sistema inteligente como un programa de computación que cuenta con características y comportamientos similares a los de la inteligencia humana o animal. [5]

Almacenamiento: Para almacenar una imagen en disco o memoria se han desarrollado formatos, los cuales se deben usar de acuerdo a la cantidad de color contenida por la imagen. [6]

Histograma de una Imagen: Dada una imagen es posible contar el número de píxeles que corresponden a cada tono en cada canal, a la representación gráfica de esta característica se le llama (como en estadística descriptiva) el Histograma de la imagen. [6]

2.2 Marco referencial.

● *Procesamiento de Imágenes.* [7]

Autor: Diego González Aguilera.

Fecha: Octubre del 2015.

Lugar: Universidad De Salamanca Ingeniería y Arquitectura; Salamanca, España.

Resumen: El procesamiento de imágenes tiene que ver con la adquisición, transmisión, procesamiento y representación de las imágenes. Las técnicas de proceso de imágenes se utilizan para mejorar la apariencia visual de las imágenes para un observador y para preparar convenientemente el contenido fotográfico de cara a la percepción por parte de máquinas. El proceso digital de imágenes se puede dividir en las siguientes áreas:

Adquisición o captura: que se ocupa de los diferentes caminos para la obtención de imágenes; por ejemplo, utilizando cámaras digitales o digitalizando imágenes analógicas (fotografías).

Realce y mejora: son las técnicas que se usan para mejorar la apariencia visual de las imágenes ó para recuperar o restaurar las imágenes degradadas.

Segmentación: Que se ocupa de la división de las imágenes en regiones o áreas significativas.

Extracción de características: Que se ocupa de la detección y localización de entidades geométricas simples y complejas. Desde entidades simples como líneas y puntos hasta geometrías complejas como curvas y cuádricas. El propósito de este tema es presentar los aspectos relevantes del procesamiento de imágenes, centrándose en torno a las herramientas y procesos del Realce y Mejora y la Extracción de características.

Discusión: Elegimos este artículo porque nos ofrece una estructura acerca de las partes más relevantes que conforman el procesamiento digital de imágenes, este nos sirve como guía para ver las características de cada proceso; y así poder integrar las técnicas en su respectiva parte.

● **Técnicas de realce de Imágenes. [8]**

Autor: Desconocido.

Fecha: Marzo del 2014.

Resumen: El objetivo principal de las técnicas que se explicarán es procesar una imagen de forma tal que resulte más adecuada que la imagen original para una aplicación específica.

Se pueden clasificar estas técnicas en dos grupos perfectamente diferenciados:

- Métodos en el dominio del espacio.
- Métodos en el dominio de la frecuencia.

Discusión: Este artículo es una herramienta de gran ayuda para el proceso de realce y mejora, ya que brinda una explicación detallada de algunas técnicas que se integran en esta parte del procesamiento digital de imágenes.

● **Realce y restauración de Imágenes. [9]**

Autor: Pedro Daniel Peguero Núñez.

Fecha: Junio 2010.

Resumen: Una vez que la imagen está digitalizada es posible mejorar la apariencia visual de la misma, o transformarla de tal forma que facilite su análisis computarizado para una aplicación específica. Este es el objetivo de las técnicas de realce. Es importante destacar el término específico, ya que el resultado de estos métodos depende en gran medida de la imagen que se quiere mejorar. As por ejemplo, un algoritmo que obtenga buenos resultados con radiografías, puede no conseguir los mismos resultados en imágenes procedentes del espacio.

Discusión: Realce y restauración de imágenes es un artículo que explica claramente lo que es una imagen digital, la importancia de esta parte del procesamiento digital de imágenes y también contempla algunas técnicas explicadas de una forma resumida, por este motivo elegimos este artículo como apoyo para la realización de nuestro proyecto.

● Segmentación por umbralización. [10]

Autor: Ingenieros en automatización y control industrial.

Fecha: Octubre 2005.

Resumen: La umbralización es una técnica de segmentación ampliamente utilizada en las aplicaciones industriales. Se emplea cuando hay una clara diferencia entre los objetos a extraer respecto del fondo de la escena. Los principios que rigen son la similitud entre los píxeles pertenecientes a un objeto y sus diferencias respecto al resto. Por tanto, la escena debe caracterizarse por un fondo uniforme y por objetos parecidos.

Discusión: La umbralización es muy importante en el procesamiento digital de imágenes, esta abarca algunas de las técnicas de mayor prioridad de la misma; este artículo resulta de gran ayuda en la realización de nuestro proyecto.

● Técnicas clásicas de segmentación de imagen. [11]

Autor: Marcos Martín.

Fecha: 4 de mayo del 2004.

Resumen: El primer paso en cualquier proceso de análisis de imagen es la segmentación. Mediante la segmentación vamos a dividir la imagen en las partes u objetos que la forman. El nivel al que se realiza esta subdivisión depende de la aplicación en particular, es decir, la segmentación terminará cuando se hayan detectado todos los objetos de interés para la aplicación. En general, la segmentación automática es una de las tareas más complicadas dentro del procesado de imagen. La segmentación va a dar lugar en última instancia al éxito o fallo el proceso de análisis. En la mayor parte de los casos, una buena segmentación dará lugar a una solución correcta, por lo que, se debe poner todo el esfuerzo posible en la etapa de segmentación. Los algoritmos de segmentación de imagen generalmente se basan en dos propiedades básicas de los niveles de gris de la imagen: discontinuidad y similitud. Dentro de la primera categoría se intenta dividir la imagen basándonos en los cambios bruscos en el nivel de gris. Las áreas de interés en esta categoría son la detección de puntos, de líneas y de bordes en la imagen. Las áreas dentro de la segunda categoría están basadas en las técnicas de umbrales, crecimiento de regiones, y técnicas de división y fusión.

Discusión: La segmentación permite que la imagen sea dividida para tener un procesamiento más fácil, si esta no es segmentada su tratamiento no podría llevarse a cabo; este artículo explica la importancia de esta parte del procesamiento digital de imágenes y también agrupa una serie de técnicas explicadas detalladamente; por esto resulta de vital importancia tomar en cuenta este documento como una de las guías en la realización de nuestro proyecto.

2.3 Marco teórico.

Técnicas de procesamiento de imágenes.

El proceso digital de imágenes se puede dividir en las siguientes técnicas principales:

- Adquisición.
- Realce y mejora.
- Segmentación.
- Extracción de características.

2.3.1. Adquisición.

Es la obtención de la imagen ya sea por captura a través de cámaras digitales o digitalización por medio de escáner. La adquisición se da mediante 4 procesos:

a) El proceso de formación de imágenes: La formación de imágenes es un proceso mediante el cual una información luminosa en 3D (de la vida real) es proyectada en un plano 2D (imagen digital). Para ellos existen algunos dispositivos como las cámaras digitales que imitan el proceso del ojo humano.

b) Representación de imágenes digitales: Una imagen digital es una matriz o un vector bidimensional de números binarios (0, 1), se representan como pequeños cuadros, en forma de mosaico individual denominados píxeles.

* Imagen binaria: Cada celda de la matriz de una imagen binaria está representada por 1 píxel que equivale 1 bit. Si el bit es 0, el color es negro. Si el bit es 1, el color es blanco.



Ilustración 1. Ejemplo de Imagen en negativo (binaria). Tomada de Imágenes Google.

* Imagen en escala de grises: Cada celda de la matriz de una imagen en escala de grises está representada por 1 píxel que equivale 1 byte, que permite 256 niveles de gris. Desde 0, que es color negro, hasta 255, que es color blanco.



Ilustración 2. Ejemplo imagen en escala de grises. Tomada de imágenes Google.

* Imagen a color: Cada celda de la matriz de una imagen en color está representada por 1 píxel que equivale a 3 bytes. Cada píxel consta de 3 valores (Rojo, Verde, Azul), un byte por color y tiene 16,7 millones de colores posibles. [12]



Ilustración 3. Ejemplo imagen a color. Tomada de imágenes Google.

c) Dispositivos de captura: Existen dos tipos de dispositivos de captura en los que compete a una imagen digital. Cámara fotográfica digital o Escáner. Existen muchos tipos de cámaras, como cámaras de videos, según la resolución y calidad de la fotografía, así como existen muchos tipos de escáner, como los planos, de microfilm, utilizados para diferentes actividades según la necesidad. Pero la definición básica de ellos, es la siguiente:

* Escáner: Es un dispositivo de entrada en el ordenador. Hace una captura de una imagen, documento de texto o fotografía, y lo transfiere en bits de información, los cuales puede entender y manejar un ordenador.

* Cámara fotográfica digital: Es una cámara fotográfica que, en vez de capturar y almacenar fotografías en películas fotográficas como las cámaras fotográficas convencionales, lo hace digitalmente mediante un dispositivo electrónico, o en cinta magnética usando un formato analógico como muchas cámaras de video. [12]

d) Formatos de almacenamiento: Un formato de fichero de imagen es una forma estándar de organizar y almacenar los datos que representan la imagen. Existen diferentes tipos de formatos utilizados en la actualidad según la necesidad de almacenamiento, que se diferencian según los niveles de profundidad admitidos, tipo de compresión y otras características como la posibilidad de definir de transparencias o diferentes imágenes en un mismo archivo (animaciones).[12]

Algunos de los formatos más utilizados y estándares son BMP, TIFF, GIF y JPG.

* BMP (Windows Bitmap): Creado por Microsoft para la permitir una rápida entrada/salida por pantalla. Este formato permite muchos niveles de profundidad: 1 bit por píxel (2 colores), 4 bits (16 colores), 8 bits (escala de grises o paleta), 16 bits (Hi-color) y 24 bits (3 bytes - True-color). A esta clase de archivos puede seleccionarse una compresión RLE (Run Length Encoding) sin pérdida de calidad.[13]

Ventajas:

- El formato BMP no sufre pérdidas de calidad y por tanto resulta adecuado para guardar imágenes que se desean manipular posteriormente.
- Ha sido muy utilizado porque fue desarrollado para aplicaciones Windows.
- La lectura y escritura son muy rápidas.
- Guarda gran cantidad de información de la imagen.

Desventajas:

- El tamaño de las imágenes es excesivamente grande, sobre todo en imágenes fotográficas. Tamaño de imagen (ancho*alto*bits por píxel), por lo que el archivo tiene un tamaño muy grande.
- No adecuado para transmisión por red.

- Poco popular fuera de los entornos de MS Windows.

* Formato TIFF (Tagged Image File Format): Fue creado por Aldus (ahora Adobe) propuesto para trabajos de impresión de alta resolución y calidad. Es un formato muy flexible, basado en tags (etiquetas). Un tag es un bloque de datos (de formato predefinido) que contiene cierto tipo de información sobre la imagen. Existen muchos tipos de tags (y se pueden crear nuevos) y un fichero puede contener muchos tags, uno detrás de otro. Este formato tiene número arbitrario de bits por píxel (hasta enteros o reales de 64 bits), distintos espacios de color, múltiples imágenes por fichero, cualquier tipo de compresión existente, etc. Una imagen se puede almacenar por tiras (cada una 1 tag).[13]

Ventajas:

- Es ideal para archivar archivos originales.
- Es el formato ideal para editar o imprimir una imagen.
- Utiliza cualquier profundidad de color de 1 a 32 bits.

Desventajas:

- Produce archivos muy grandes.
- Demasiado flexible. Difícil crear un programa que soporte todas las opciones y tipos de tags. -Almacenamiento en tiras, inadecuado para ciertos usos.

* Formato GIF (Graphics Interchange Format): Desarrollado por Compuserve en 1987 para la rápida transmisión de imágenes en color por las redes. Está basado en uso de paletas, de hasta 256 colores. Este formato usa el algoritmo de compresión LZW (compresión sin pérdida), pero el uso de paletas implica una pérdida de información de color. Permite definir transparencias. Se puede definir una entrada de la paleta como “transparente”. Los ficheros pueden contener múltiples imágenes, esto permite crear animaciones sencillas y está diseñado específicamente para comprimir imágenes digitales.[13]

Ventajas:

- Admite gamas de menor número de colores y esto permite optimizar el tamaño del archivo que contiene la imagen.

- Adecuado para transmisión en redes.
- Es un formato idóneo para publicar dibujos en la web.
- Muy popular.

Desventajas:

- Poco adecuado para imágenes fotográficas: pérdida de color y tamaños muy grandes.
- Formato poco flexible.
- Problemas de patentes hicieron que apareciera el formato PNG como alternativa al GIF. Hoy día, las patentes existentes sobre GIF han expirado.
- No es recomendable para fotografías de cierta calidad ni originales ya que el color real o verdadero utiliza una paleta de más de 256 colores.

* Formato JPEG (Joint Photographic Experts Group): Es el formato más reciente de los cuatro, más elaborado y orientado al almacenamiento de imágenes fotográficas. Admite imágenes en escala de grises (1 byte por píxel) y RGB (3 bytes por píxel). Incluye un mecanismo avanzado de compresión, que puede ajustarse a distintos ratios de compresión. El fichero puede incluir una versión reducida, para previsualizar la imagen antes de leerla entera. Está libre de patentes.[13]

Ventajas:

- En la mayoría de los casos, consigue un ratio compresión/calidad mucho mejor que los otros formatos.
- Nivel de compresión ajustable. Típicamente entre 1:10 y 1:100.
- Admite una paleta de hasta 16 millones de colores.
- Es ideal para publicar fotografías en la web siempre y cuando se configuren adecuadamente dimensiones y compresión.

Desventajas:

- Debido a su reciente aparición sólo es soportado en navegadores modernos.
- No incluye transparencias ni animaciones.

- Si trabajamos con un JPEG guardando en disco tras cada operación, la imagen se va degradando.
- Si se define un factor de compresión se pierde calidad. Por este motivo no es recomendable para archivar originales.

2.3.2. *Realce y mejora.*

Una vez que la imagen está digitalizada es posible mejorar la apariencia visual y la calidad de la misma, o transformarla de tal forma que facilite su análisis computarizado para una aplicación específica. Esto se logra mediante dos técnicas, realce de imágenes en el dominio espacial y realce en el dominio de la frecuencia. [14]

a) Técnicas de realce en el Dominio espacial.

* *Realce punto a punto:* Son aquellas técnicas cuyas operaciones se efectúan sobre el valor del píxel, sin tener en cuenta los píxeles vecinos. Estas transformaciones se agrupan en dos principales técnicas, transformaciones de intensidad simples y procesado del histograma.

-Transformaciones de intensidad simples: Sirven para mejorar condiciones de bajo contraste, baja luminosidad o demasiada claridad. En estas técnicas de realce, el valor de g en el punto (x, y) depende del valor de f en ese punto. Por tanto, la función T que define la operación sobre el píxel quedaría de la forma de la ecuación:

$$f(x, y) = T(g(x, y))$$

A este operador se le conoce como función de transformación de nivel de gris. Las operaciones más frecuentes son:

i. Imagen en escala de grises: Consiste en obtener una imagen monocroma (Definida en sólo dos colores) a partir de una imagen en color. En función del espacio de color con el que se esté trabajando esta operación se realizará de una forma u otra. En el caso de RGB se puede utilizar la siguiente ecuación para obtener la señal luminancia.

$$Y = 0,299R + 0,587G + 0,114B \quad (1)$$

A continuación un ejemplo de obtención de una imagen a blanco y negro, a partir de una imagen a color, aplicando la ecuación (1):



Ilustración 4. Imagen a color. Tomada de Imágenes Google.



Ilustración 5. Imagen transformada de emplear la ecuación (1) a Ilustración 4.

ii. Negativo de imágenes: Esta sencilla transformación invierte los niveles de blanco a negro con el fin de obtener el negativo de la imagen de entrada. En concreto, si la señal está en escala de grises, y está comprendida dentro del rango $[0, L-1]$, siendo 0 el nivel de negro y $L-1$ el de blanco, bastaría con hacer la transformación lineal que se indica en la siguiente ecuación.

$$s = (L - 1) - r \quad (2)$$

Donde s y r son variables que indican el nivel de gris de $f(x, y)$.

Los negativos de imágenes son útiles para representar imágenes médicas, como las radiografías que se toman a partir de los Rayos X; estas imágenes en negativo se obtienen empleando la Ecuación (2) para su transformación.



Ilustración 6. Ejemplo de imagen médica, radiografía. Tomada de Imágenes de Google.

iii. Aumento del contraste: Las imágenes con poco contraste se deben a muchas causas, como iluminación deficiente o incorrecta selección de apertura del diafragma de la cámara en el momento de la captura, por lo que no es posible visualizar bien el contenido de la imagen [28]. Para solucionar esta situación estas técnicas incrementan el rango dinámico de los niveles de gris de la imagen. En la ilustración 7, se observa una posible función de transformación de esta técnica que queda definida gracias al par de puntos (r_1, s_1) y (r_2, s_2) . Por norma general se supone que $r_1 \geq r_2$ y $s_1 \leq s_2$ dando como resultado una función monótonamente creciente.

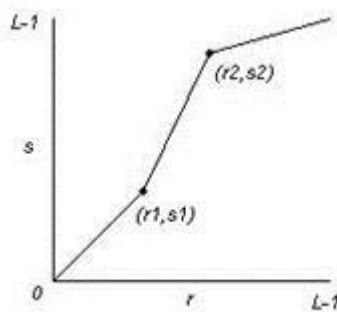


Ilustración 7. Función de transformación de los niveles de gris.

Dentro de las posibles funciones de transformación cabe destacar el caso en el que $r_1 = r_2$, $s_1 = 0$ y $s_2 = L - 1$, (tomando L como el número de niveles de gris empleado) conocida como función umbral dando como resultado una imagen binaria como puede observarse en la ilustración 8.

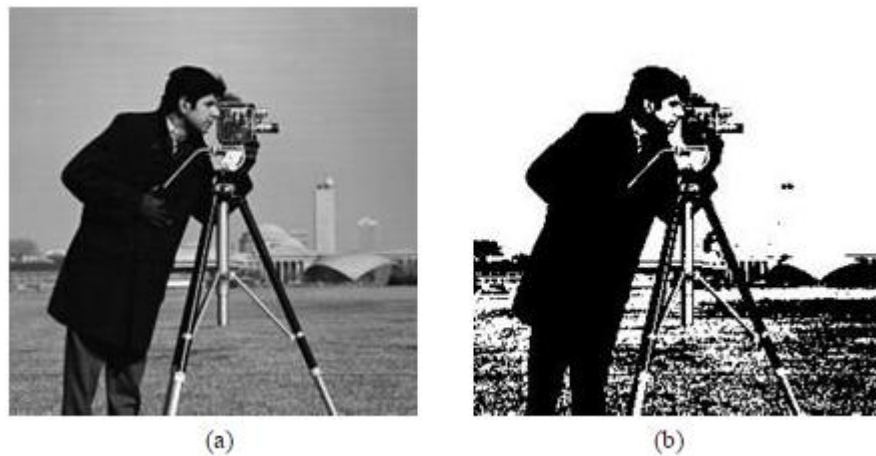
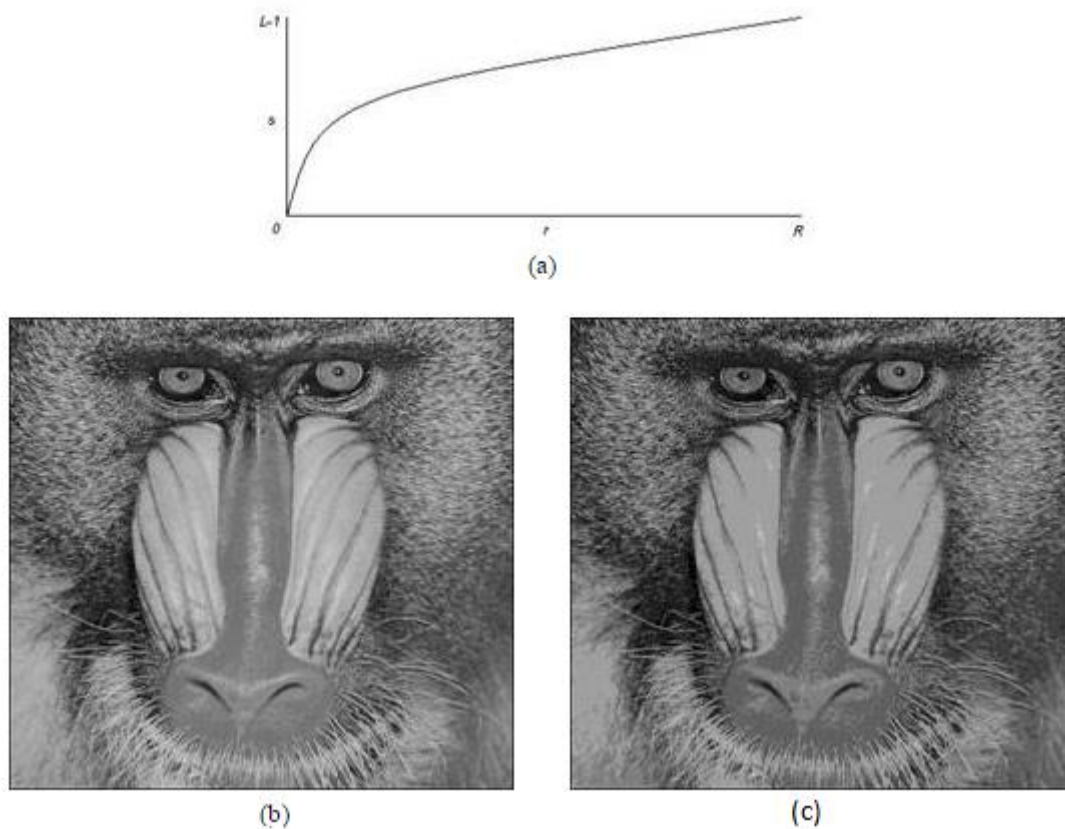


Ilustración 8. Ejemplo de la utilización de la función umbral. (a) Imagen original (b) Imagen binaria correspondiente [28].

iv. Compresión del rango dinámico: A veces el rango dinámico de una imagen procesada excede ampliamente la capacidad del dispositivo de presentación, en cuyo caso sólo las partes más brillantes de la imagen aparecerán en la pantalla. En estos casos será necesario comprimir el rango dinámico de los valores de cada pixel. Para ello se emplea la siguiente transformación:

$$s = c \log(1 + |r|)$$

Donde c es un factor de escala. En la ilustración 9 se puede observar una posible función de transformación para compresión del rango dinámico y los resultados de aplicarla sobre una imagen.



v. Fraccionamiento del nivel de gris: A menudo se desea destacar un rango específico del nivel de gris de una imagen. Existen varias formas para esto, una de ellas es adjudicar un valor alto a todos aquellos niveles de gris del rango de interés y un valor bajo a los restantes

o simplemente intensificar el rango de interés pero preservando el fondo y el resto de las tonalidades de gris.

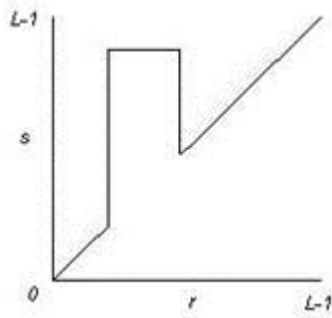


Ilustración 10. Función de transformación de los niveles de gris [28].

- *Procesado de histograma*: Es un proceso por el que se normaliza la imagen de tal forma que el valor de gris más pequeño es llevado a 0 y el máximo valor a 255. El histograma es de gran utilidad en el tratamiento de imágenes ya que con él se puede variar el aspecto de las imágenes. El histograma es una función discreta que representa el número de píxeles en la imagen en función de los niveles de intensidad, g . La probabilidad $P(g)$ de ocurrencia de un determinado nivel g se define como:

$$P(g) = N(g) / M$$

Donde M es el número de píxeles en la imagen y $N(g)$ es el número de píxeles en el n nivel de intensidad g . Como con cualquier distribución de probabilidad, todos los valores de $P(g)$ son menores o iguales que 1 y la suma de todos los valores de $P(g)$ da como resultado 1.

Un histograma con una distribución de los niveles de gris concentrada en una determinada zona presenta un contraste muy bajo, mientras que un histograma con una amplia distribución de los niveles de gris tiene un alto contraste. Un histograma con los niveles de gris concentrados en su parte baja del rango corresponde a una imagen oscura, si están concentrados en su parte alta corresponde a una imagen brillante. El histograma puede ser modificado mediante una serie de funciones que expandan los niveles de gris, los compriman o los desplacen.

En la ilustración 11, (a) y (b) la imagen original tiene un histograma que abarca todo el

rango de valores de niveles de gris (de 0 a 255), la imagen posee un alto contraste. En contrario, el contraste es bajo. En los otros casos, el histograma está concentrado en una determinada zona y las imágenes aparecen poco contrastadas en relación con la primera imagen; esto se debe a que no ocupan todo el rango de niveles de intensidad disponible. En (f) el histograma está situado en la parte izquierda y su imagen (e) aparece oscura. En (h) ocurre lo contrario y su imagen (g) aparece muy brillante (clara). En (c) el histograma ocupa la parte central y su imagen (d) aparece con claridad intermedia.

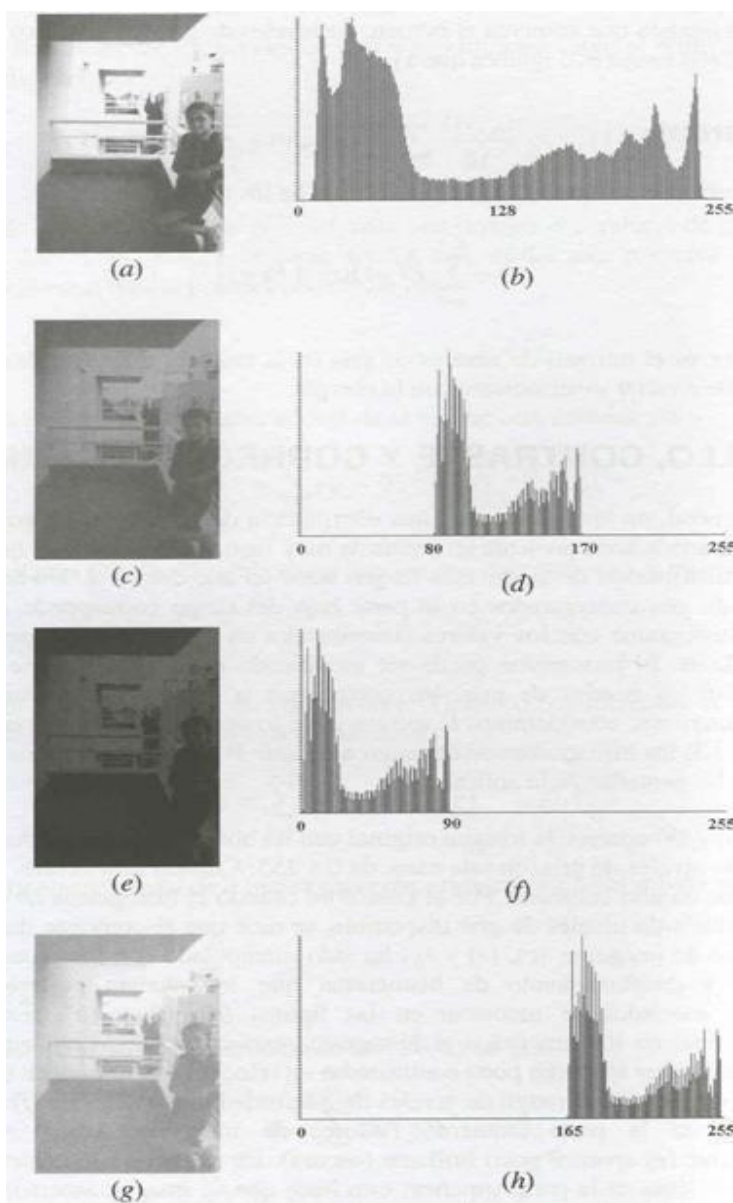


Ilustración 11. Ejemplo de imágenes con su histograma. [8]

La segunda máscara de la ilustración 12 produce un promedio ponderado. Como se puede ver, los píxeles son multiplicados por diferentes coeficientes, dándoles así más importancia (peso) a algunos píxeles que a otros. En el caso particular de la máscara de la ilustración 12, el píxel del centro se multiplica por el valor más alto, mientras que los demás píxeles son

pesados teniendo en cuenta la inversa de su distancia al centro. La razón de darle más importancia al píxel central y menos importancia a los píxeles más alejados al centro, radica en la necesidad de reducir el borronado durante el proceso de suavizado.

En la ilustración 13 pueden verse los efectos del filtro espacial pasa bajos, o suavizado. Teniendo en cuenta la imagen original, figura(a), las figuras (b) a (f) son el resultado de filtrar la imagen original utilizando núcleos cuadrados de $n=3$, 5, 9, 15 y 35, respectivamente. Para la máscara de 3×3 , se nota un leve borronado general de la imagen, pero los detalles que tienen aproximadamente el mismo tamaño que el núcleo han sido afectados considerablemente. Para $n=9$, se ve un borronado mucho mayor, y también que el 20% del círculo negro casi se confunde con el fondo, ilustrando así el efecto de mezclado que el borronado produce sobre aquellos objetos cuyos niveles de gris son parecidos a los de sus vecinos. Otro efecto que se puede notar es la reducción del ruido en los rectángulos. Para $n=15$ y 35, se puede ver que el borronado es excesivo. Este tipo de borronado se utiliza para eliminar los objetos pequeños en una imagen.

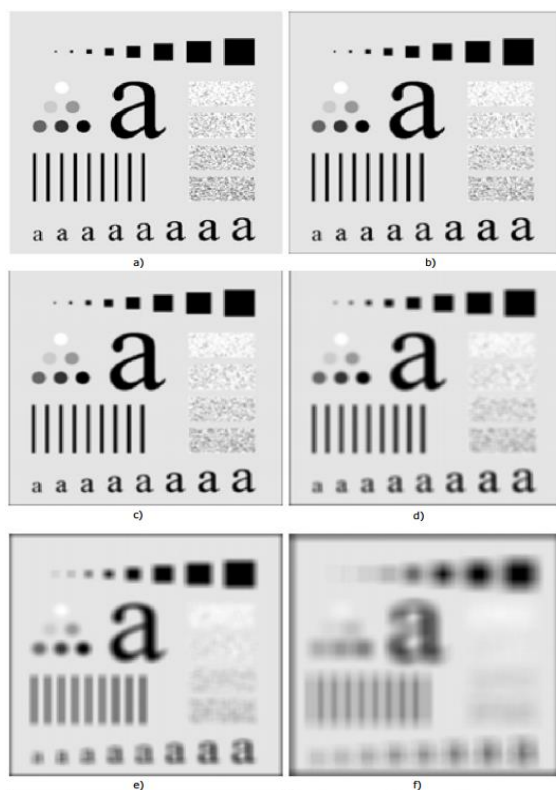


Ilustración 13. (a) Imagen original 500x500 píxeles. La imagen (b) hasta (f) resultados de suavizar con máscaras cuadradas en $n= 3, 5, 9$ y 35 respectivamente. [28]

- *Filtro de mediana:* Los filtros estadísticos son filtros espaciales no lineales cuya respuesta está basada en ordenar los píxeles abarcados por una máscara y luego reemplazar el valor del píxel central con el valor determinado por el resultado del ordenamiento. El más conocido de estos filtros es el filtro de mediana, el cual reemplaza el valor del píxel central por la mediana de los niveles de gris del vecindario de ese píxel (el valor original del píxel es incluido en el cálculo de la mediana). Los filtros de mediana son muy usados debido a que, para ciertos tipos de ruidos aleatorios, proveen una excelente reducción de ruido y un borronado considerablemente menor que los filtros lineales de suavizado del mismo tamaño. Los filtros de mediana son particularmente efectivos cuando el ruido es del tipo impulso (también llamado ruido sal y pimienta) debido a que aparece como puntos negros o blancos superpuestos en la imagen.

La mediana, ξ , de un conjunto de valores es aquella en la que la mitad de los valores en el conjunto son menores o iguales que ξ , y la otra mitad es mayor o igual a ξ . Por ejemplo, si en una imagen se toma un conjunto de píxeles de 3x3 con valores {1, 9, 5, 0, 8, 7, 1, 2, 4} la mediana para este caso será el valor 4, ya que la mitad de este conjunto es menor (o igual) y la otra mitad es mayor (o igual) a éste: {0, 1, 1, 2, 4, 5, 7, 8, 9}. El procedimiento general para realizar el filtro de mediana en cualquier punto consiste en ordenar los valores de dicho píxel y los de su vecindario, determinar la mediana, y asignar éste último valor al píxel en cuestión. Para un vecindario de 3x3, la mediana es el 5to valor más grande; para uno de 5x5 es el 13mo valor más grande, y así. Cuando algunos valores en un vecindario son iguales, éstos se ven agrupados. Por ejemplo, supongamos que un vecindario de 3x3 tiene los valores {10, 20, 20, 20, 15, 20, 20, 25, 100}. Ordenando este conjunto nos queda {10, 15, 20, 20, 20, 20, 20, 25, 100}, de donde se ve que la mediana es 20. Así, la principal función de los filtros de mediana es hacer que los puntos con niveles de gris distintos sean más parecidos a los de su vecindario.

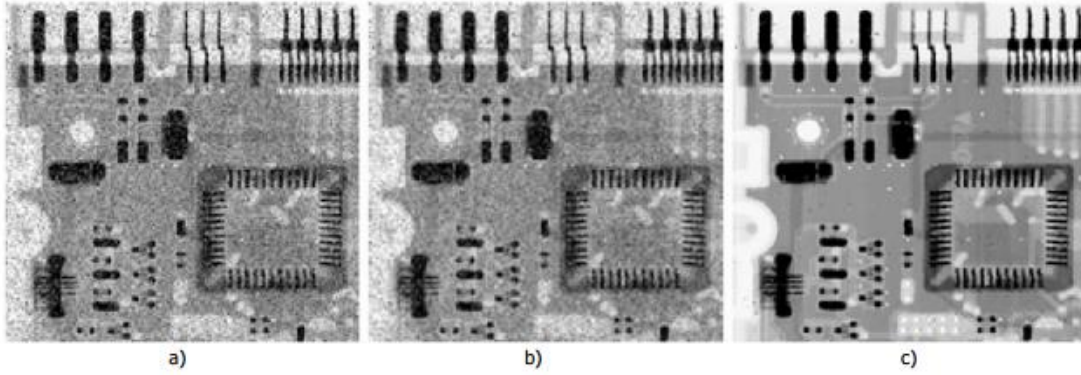


Ilustración 14. (a) Imagen con ruido sal y pimienta. (b) filtrado de (a) con un filtro de suavizado de 3x3. (c) Reducción de ruido con filtro de mediana de 3x3.

b) Técnicas de realce en dominio de frecuencia:

Hasta ahora se ha considerado a la imagen digital como una matriz cuyos píxeles representan un valor proporcional al nivel de brillo de la imagen. Definición dada en la sección del dominio espacial. Sin embargo, la misma imagen puede ser considerada como una función no periódica, y definirse en otro espacio bidimensional cuyos ejes vengan determinados por la amplitud y frecuencia para cada dirección de la imagen. Esto es lo que se conoce como el dominio de la frecuencia.

El fundamento de los métodos del dominio de frecuencias es el uso del teorema de convolución. La hipótesis de partida es que la imagen que queremos mejorar fruto del mecanismo de captación de la imagen ha sufrido una convolución con una determinada función definida por las lentes del aparato de medida, el medio en que se tomó la imagen o cualquier otro efecto y que además está contaminada por ruido aleatorio.

La ecuación que define la imagen que resulta del anterior proceso es la siguiente,

$$g(x, y) = h(x, y) * f(x, y) + n(x, y)$$

Siendo $f(x, y)$ la imagen real, $h(x, y)$ la función de convolución que supondremos que es lineal e invariante respecto de la posición y $n(x, y)$ la función de ruido. Nuestro objetivo es calcular $f(x, y)$ a partir de la anterior ecuación.

Si calculamos la transformada Fourier de esta ecuación obtenemos

$$G(u, v) = H(u, v)F(u, v) + N(u, v)$$

Donde cada letra mayúscula representa a la función TF de la correspondiente función con letra minúscula. A la función $H(u, v)$ en la teoría de sistemas lineales se le denomina **función de transferencia del sistema**, en óptica se llama la función de transferencia óptica. Esta función caracteriza las propiedades del medio en que se transmite la señal.

Para poner de manifiesto de forma aún más clara el papel que juega la función $H(u, v)$ supongamos que la función $f(x, y) = 1$ para un valor dado de (x, y) y cero para el resto. Entonces en ausencia de ruido externo tendríamos que $G(u, v) = H(u, v)$ ya que $F(u, v) = 1$ lo que nos indica que la respuesta del sistema de captación a una imagen con un solo punto distinto de cero nos dará una imagen que corresponde a la perturbación introducida por el sistema de captación. A la función $h(x, y)$ definida por la transformada Fourier de la función de transferencia se le denomina **función impulso respuesta del sistema**. En el caso de sistemas ópticos se le denomina **función de emborronamiento o PFS** (point spread function).

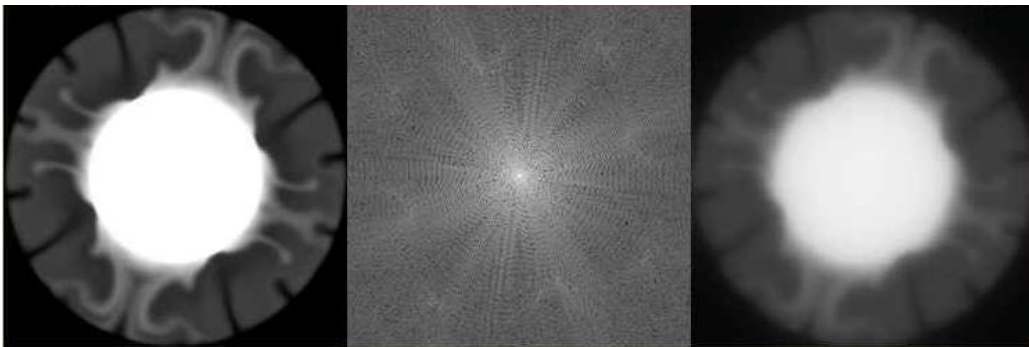


Ilustración 15. Ejemplo de filtrado en el dominio de la frecuencia. Tomada de Imágenes Google.

-Filtro Ideal: Suprime todas las frecuencias más altas que la frecuencia de corte D_o , y no modifica las frecuencias bajas.

$$H(u, v) = \begin{cases} 1 & \text{si } \sqrt{u^2 + v^2} \leq D_o \\ 0 & \text{si } \sqrt{u^2 + v^2} \geq D_o \end{cases}$$

En la práctica D_o , se especifica como una fracción de frecuencia más alta representada en el dominio Fourier.

-Filtro butterworth: Este tipo de filtros en realidad a lo que se dedica es a eliminar ruidos

externos, ayudan a evitar la contaminación de la imagen.

El filtro de Butterworth más básico es el típico filtro pasa bajo de primer orden, el cual puede ser modificado a un filtro pasa alto o añadir en serie otros formando un filtro pasa banda o elimina banda y filtros de mayores órdenes.

2.3.3. Segmentación.

En el proceso de segmentación, la imagen se fracciona con el objetivo de clasificar las partes que la componen, a estas partes las podemos llamar zonas o regiones de la imagen; cada una de ellas está compuesta por diversas características. Y conforme a estas características es que se seleccionan las zonas de interés para las técnicas que se van aplicar.

Los atributos básicos de segmentación de una imagen son: la luminancia en imágenes monocromáticas, los componentes de color en imágenes en color, textura, forma, etc. [15]. Podemos decir que la segmentación automática es una de las fases más complicadas del procesamiento de imágenes, en esta se determina si el análisis de la imagen es exitoso o no, y muy pocas veces se logra una solución confortable.

Dentro de la etapa de segmentación que hace parte del procesamiento digital de imágenes se encuentran muchas técnicas, las cuales podemos agrupar dentro de estos tres grupos:

- * Detección de bordes*

- * Umbralización*

- * Basado en regiones*

Estas serán explicadas cada una a continuación:

a) Detección de bordes: La imagen está compuesta por objetos de diversas características y en la detección de bordes se localiza la intersección de estos objetos para poder determinar sus regiones e información. Los bordes pueden clasificarse por su anchura, ángulo de su pendiente de variación, y las coordenadas de su punto medio. Los diferentes modelos de bordes son: línea, tipo escalón, tipo rampa y tipo tejado. Las técnicas más conocidas para la detección de bordes son: Roberts, Prewitt, Sobel y Canny.

La mayoría de las técnicas de detección de bordes se basa en la primera y la segunda derivada; en la primera derivada se aplica el operador gradiente y en la segunda derivada se aplica el operador laplaciano.

- Derivada de primer orden: Los píxeles que forman los bordes de una imagen en 2D deben ser localizados a través del gradiente. El gradiente representa el conjunto de derivadas direccionales de una función de varias variables con respecto a las diferentes coordenadas de un sistema. Es un operador vectorial dado que como resultado de la operación, resulta un vector que apunta en la dirección de máxima variación de la función dada.

El gradiente se aproxima mediante los valores absolutos para lograr un procesamiento veloz a costa de un error despreciable.

El cálculo del gradiente de una imagen se basa en la obtención de las derivadas parciales dF/dx y dF/dy en cada posición de píxel. Las derivadas de primer orden permiten encontrar lugares en una imagen en donde la intensidad que cambia rápidamente es mayor en magnitud que un umbral especificado. [15]

- Derivada de segundo orden: El laplaciano es un operador diferencial elíptico de segundo orden, denotado como Δ , relacionado con ciertos problemas de minimización de ciertas magnitudes sobre un cierto dominio. Esta es su ecuación:

$$\Delta^2 = \partial^2 / \partial x^2 + \partial^2 / \partial y^2$$

Los píxeles que se encuentran en los bordes de la imagen conciernen a los pasos por cero de la segunda derivada. La laplaciana ofrece un ancho del borde de un píxel y es de carácter isotrópico. [15]

* *Operador Roberts*: Este operador aplica la derivada de primer orden, es decir, el método del gradiente. Está compuesto por dos máscaras o kernels:

$$\begin{array}{cc} \text{Rhorizontal} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} & \text{Rvertical} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \end{array}$$

En el operador horizontal se hace un rastreo de filas y en el operador vertical se hace un rastreo de columnas.

Es una técnica muy sencilla, ya que consiste en calcular el gradiente en la dirección horizontal y/o vertical para cualquier imagen. También se han propuesto kernels o máscaras

para otras direcciones diferentes a la horizontal y vertical [16]. Son las siguientes:

$$\text{Rnorte} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ -1 & -1 & -1 \end{bmatrix}$$

$$\text{Rnoreste} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & -1 & 1 \end{bmatrix}$$

$$\text{Reste} = \begin{bmatrix} -1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & 1 & 1 \end{bmatrix}$$

A continuación se tomará como ejemplo una imagen a la cual se le aplicó los operadores Roberts:

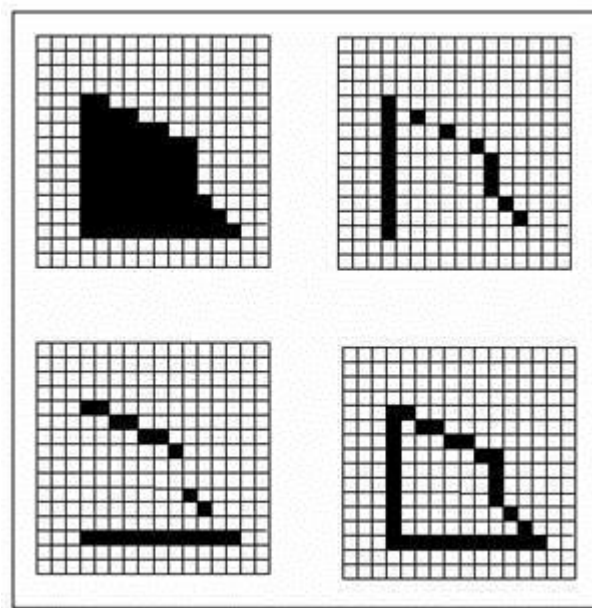


Ilustración 16. Imagen Capturada y sus resultados después de aplicarle el operador Roberts horizontal, vertical y la suma de los dos [16]

En la parte superior izquierda se puede observar la imagen original la cual fue capturada inicialmente, a su lado (superior derecha) se observa el resultado de la imagen después de aplicarle el operador Roberts horizontal, en la parte inferior izquierda se observa el resultado de aplicarle el operador Roberts vertical, y a su derecha (inferior derecha) se puede observar la imagen producto de la suma de los dos operadores.

* *Operador Prewitt:* Este operador al igual que el de Roberts, aplica la derivada de primer orden, es decir, el método del gradiente y está compuesto por dos máscaras o kernels:

Operador de prewitt para gradiente horizontal:

$$\text{PHorizonte} = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\text{Pvertical} = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

El gradiente horizontal permite detectar bordes orientados verticalmente, mientras que para el gradiente vertical se detectan los bordes orientados horizontalmente [17]

El gradiente total de la imagen se puede obtener a través de distintas combinaciones del gradiente horizontal y el gradiente vertical, donde los métodos más utilizados son:

$$\text{Magnitud del gradiente: } G = \sqrt{|G_x|^2 + |G_y|^2} \quad (1)$$

$$\text{Máximo gradiente: } G = \text{Max} \{|G_x|^2 + |G_y|^2\} \quad (2)$$

Después de esto, la imagen del gradiente debe ser binarizada con un umbral adecuado para la detección de los bordes deseados. Nota: La umbralización se explicará más adelante en esta misma monografía.

Pasos para el aplicar el operador de Prewitt:

i. Implementación de filtrado con operador Prewitt: Este paso permite identificar las posibles diferencias entre los bordes resaltados en cada imagen, y también las partes en común que se forman gracias a las componentes diagonales de la imagen. También resalta los cambios en la intensidad sin importar que sean cambios leves, y esto hace posible identificar la presencia de ruido. Este filtrado se hace a través del cálculo del gradiente horizontal y el gradiente vertical (explicados en la parte superior).

ii. Binarización de las imágenes: En esta parte se establece un umbral t entre 0 y 1 de manera que si el umbral es superado en la imagen se tiene un valor 0. En caso contrario, se tiene un valor de 1. De acuerdo a esto se elige el umbral adecuado para percibir los bordes en la imagen.

iii. Gradiente máximo y efecto del umbral: En la aplicación de este operador se obtienen dos imágenes desde el primer paso, por el gradiente vertical y el horizontal; en este paso se consigue el gradiente total de la imagen; es decir, se combinan las dos imágenes mediante

la ecuación (2). Después de obtener el gradiente completo de la imagen se binariza y se busca un umbral para esta, de manera que los bordes pueden ser perceptibles. Después de todo este procedimiento se define el umbral correcto que permite determinar los bordes de la figura únicamente.

* *Operador Sobel*: Este operador se implementa de igual forma que el operador Roberts y Prewitt. El filtro Sobel detecta los bordes horizontales y verticales separadamente sobre una imagen en escala de grises. Las imágenes en color se convierten en RGB en niveles de grises. Como con el filtro Laplace, el resultado es una imagen transparente con líneas negras y algunos restos de color. [18] A continuación se muestra un ejemplo:

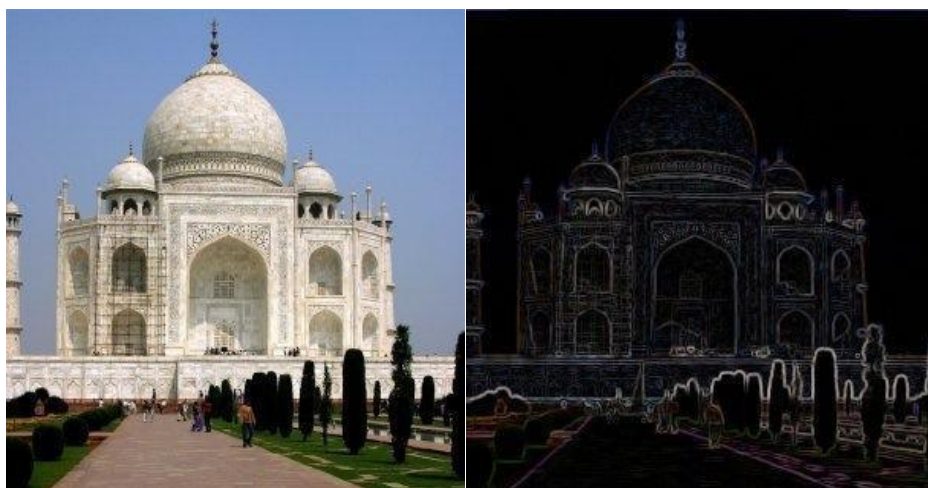


Ilustración 17. Imagen capturada y sus resultados después de aplicarle el operador sobel [18]

* *Operador Canny*: El operador Canny utiliza un filtro basado en la primera derivada de un Kernel G, donde luego permite discriminar los valores en que el gradiente no es máximo local, por lo que se debe a variaciones ruidosas de la intensidad de pixel, a través de este se obtienen mejores resultados que el operador Roberts, Prewitt y Sobel.

Nota: Este se explica detalladamente en la extracción de características en la parte de extracción de líneas.

Dentro de los campos de aplicación que podemos destacar en las técnicas aplicadas para la detección de bordes, está:

Metrología: La detección de bordes es un proceso fundamental en la industria de la

metrología ya que define las fronteras de la característica a ser medida. La precisión en la detección del borde mejora la precisión en los procesos en curso y los procedimientos de control de calidad.

Detección de rostros: Se empezaron a desarrollar algoritmos para detectar rostros y/o caras de personas en una imagen a mediados de la década de los 70s, se basaron en técnicas heurísticas y antropométricas. Pero en ese entonces este tema no era tan relevante y de vital importancia como hoy día, motivo por el cual estas investigaciones se abandonaron. En la década de los 90s fue que el desarrollo de algoritmos que detectaban rostros inició su auge; por los avances tecnológicos que lograron la reproducción de vídeos, videoconferencias, procesadores más rápidos, etc. Gracias al desarrollo de técnicas de detección de bordes en el procesamiento digital de imágenes hoy detectar caras en una imagen puede llevar desde pocas milésimas hasta varios segundos.

Una de las aplicaciones más importantes de las técnicas de detección de bordes, es la del diagnóstico del cáncer de mama, a través del operador Sobel se realiza un procesamiento de las imágenes digitales mamográficas generadas por un detector.

b) Umbralización: Este es una técnica muy importante dentro de la segmentación de imágenes, el umbral se precisa como una función que convierte una imagen con diferentes tonalidades en una imagen en blanco y negro. Se elige un umbral que permita agrupar los píxeles que hacen parte de los objetos de la imagen, diferenciándolos del fondo. El histograma de una imagen es una función discreta que se plantea como el número de píxeles gris sobre el número de píxeles de la imagen. Por lo tanto, la segmentación por histograma consiste en elegir uno o varios umbrales para poder reunir los puntos de la imagen con respecto a sus características y de acuerdo a sus niveles de gris. El método de la Binarización es muy básico pero sirve para diferenciar el objeto, del fondo de la imagen; este consiste en que los píxeles que están marcados con 1 corresponden al objeto de la imagen y los píxeles que están marcados con 0 corresponden al fondo de la imagen. Si el objeto es oscuro con respecto al fondo, el proceso es inverso. Existen varios tipos de umbrales; pero hay 3 de ellos que son los más utilizados; estos son: Umbral global, Umbral local y Umbral dinámico.

El umbral puede depender de $f(x, y)$, de alguna propiedad local del píxel, $p(x,y)$, y hasta de su propia posición: Si el umbral sólo depende de $f(x,y)$ se dice que es un umbral global; en el caso de que además dependa de $p(x,y)$,por ejemplo, el valor medio de los píxeles vecinos, el umbral es denominado local; y si depende también de la posición (x,y) del píxel, se denominará dinámico. [19]



Ilustración 18. Ejemplo de una imagen con umbralización. [19]

Estas son las técnicas y/o métodos de la umbralización:

* Técnica de los dos picos: Hay veces se desea encontrar el umbral más bajo ubicado entre dos picos, en eso se basa esta técnica; descrito en los siguientes dos pasos:

- Encontrar los dos picos altos del histograma.
- Encontrar el punto más bajo entre ellos.

Es sencillo encontrar el primer pico del histograma, pero es menos sencillo cuando se debe encontrar el segundo pico más alto. Veamos el siguiente ejemplo:

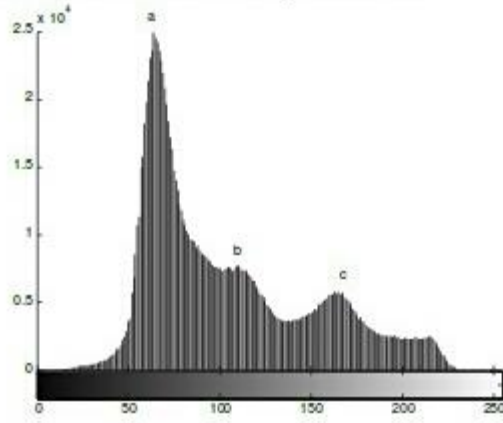


Ilustración 19. Histograma trimodal. [24]

En la ilustración 19 se puede observar que el histograma tiene tres picos, a, b y c. El pico más alto corresponde al punto a, el segundo pico más alto es el punto b, pero el umbral debe ser establecido tomando el punto c [24].

Una solución a este problema es tener en cuenta la distancia al pico máximo; teniendo en cuenta la siguiente ecuación:

$$h(j) = \max_{0 \leq j \leq 255} ((j - k)^2 h(j))$$

Ilustración 20. Ecuación: Segundo pico máximo. [24]

Donde $h(k)$ viene siendo el pico máximo del histograma y $h(j)$ es el segundo pico más alto del histograma; la ecuación 1 incluye una restricción de distancia con el fin de poder calcular $h(j)$ (el segundo pico máximo).

* Algoritmo ISODATA: Es un algoritmo iterativo el cual permite obtener el umbral correcto de manera automática. Los pasos son los siguientes: [25]

- i. Estima un valor umbral inicial T .
- ii. Segmenta la imagen usando el valor del umbral T . $G1$ = píxeles con intensidad $> T$ $G2$ = píxeles con intensidad $\leq T$.
- iii. Se calcula respectivamente la media de la intensidad $m1$ y $m2$ de $G1$ y $G2$.
- iv. Se calcula un nuevo umbral T . $T = \frac{1}{2} (m1 + m2)$.
- v. Se repiten los pasos del ii al iv hasta que cada iteración de los valores de T sea más pequeña que un parámetro predefinido Delta de T .

* Técnica Otsu: Otsu es un método de umbralización que utiliza la distribución de probabilidad, la media y la varianza de las intensidades para definir el umbral óptimo [24].

En este método se utiliza un umbral inicial y se definen dos distribuciones de probabilidad, luego las dos medias de las distribuciones respectivamente. Finalmente se calcula la distribución combinada y la dispersión total. El método de Otsu define, para una umbralización de dos niveles, el umbral óptimo como el valor que maximiza la dispersión total [24] obteniendo como final la siguiente ecuación:

$$T^* = \max_{0 \leq T \leq 255} \left(\sigma_c^2(T) \right).$$

Ilustración 20. Ecuación: umbral óptimo como el valor que maximiza la dispersión total.[24]

El objetivo de este método es calcular el valor umbral de forma que la dispersión dentro de cada clase sea lo más pequeña posible, pero al mismo tiempo la dispersión sea lo más alta posible entre clases diferentes.[25]

Los pasos para llevarlo a cabo son los siguientes: [25]

i. Computar el histograma normalizado de la imagen.

$f(x,y)$ = imagen M x N

$\{0, 1, 2, \dots, L-1\}$ = Niveles de gris de la imagen.

n_i =Número de píxeles con nivel de intensidad i .

$p_i = n_i / MN$ Probabilidad de que ocurra el nivel de intensidad i .

ii. Computar las sumas acumulativas $p_1(k)$, $k= 0, 1, \dots, L-1$.

$$P_1(k) = \sum_{i=0}^k p_i$$

Ilustración 21. Ecuación: Sumas acumulativas. [24]

Si se fija un valor umbral T en el valor de intensidad k ($T= k$), se dividirá el conjunto de píxeles de la imagen en dos clases:

C_1 = píxeles con nivel de intensidad en $[0, k]$

C_2 = píxeles con nivel de intensidad en $[k+1, L-1]$

iii. Computar las medias acumulativas $m(k)$, $k = 0, 1, \dots, L-1$.

$$m(k) = \sum_{i=0}^k i p_i$$

Ilustración 22. Ecuación: Medias acumulativas. [24]

iv. Computar la media global m_G

$$m_G = \sum_{i=0}^{L-1} i p_i$$

Ilustración 23. Ecuación: Media global. [24]

v. Computar la varianza entre clases

$$\sigma_B^2(k) = \frac{[m_G P_1(k) - m(k)]^2}{P_1(k)[1 - P_1(k)]}$$

Ilustración 24. Ecuación: Varianza entre clases. [24]

Este valor mide la dispersión entre los valores de intensidad de las dos clases C_1 y C_2 definidas al tomar como valor umbral el valor de intensidad k .

vi. Maximizar la varianza anterior

$$k^* = \max_{0 \leq k \leq L-1} \sigma_B^2(k)$$

Ilustración 25. Ecuación: Varianza maximizada. [24]

Este paso persigue el objetivo inicial, dispersar lo máximo posible los niveles de intensidad de las dos clases. Si el máximo no es único, definir k^* como la media de todos los máximos posibles.

vii. Obtener la medida de separabilidad

$$\mu^* = \frac{\sigma_B^2(k^*)}{\sigma_G^2}$$

Ilustración 26. Ecuación: Media de separabilidad. [24]

Cociente entre la varianza de clases y la varianza global. Como la varianza global es constante, mientras mayor diversidad de niveles de intensidad haya entre las clases, mayor

será esta medida.

Los anteriores 7 pasos componen el método de Otsu.

La umbralización tiene muchas aplicaciones; en el campo de la medicina permite la detección de células en citología; localización de fracturas, análisis de ecografías, CAT's e imágenes de rayos X en traumatología. En la industria permite la supervisión automática de procesos de control de calidad, vigilancia o detección de fallos.

Esta técnica también permite el análisis de fotografías aéreas o por satélite para su posterior clasificación. [20]

c) Basado en regiones: El objetivo principal de esta técnica de segmentación es dirigirse en la búsqueda de las regiones de interés de la imagen.

La imagen se divide en regiones cada una con características particulares que la diferencia de las otras regiones; esto se da porque los píxeles de una región poseen características similares.

La segmentación basada en regiones es una técnica que verifica las siguientes condiciones:

- La unión de todas las regiones generadas cubren la imagen original a segmentar por completo.
- Las regiones obtenidas son conexas según la distancia seleccionada.
- Las regiones obtenidas son disjuntas.
- Los píxeles pertenecientes a una determinada región cumplen algún criterio de similitud.
- La condición de similitud que se cumple en los píxeles de una región, no se cumple para los píxeles de la unión de dos o más regiones.

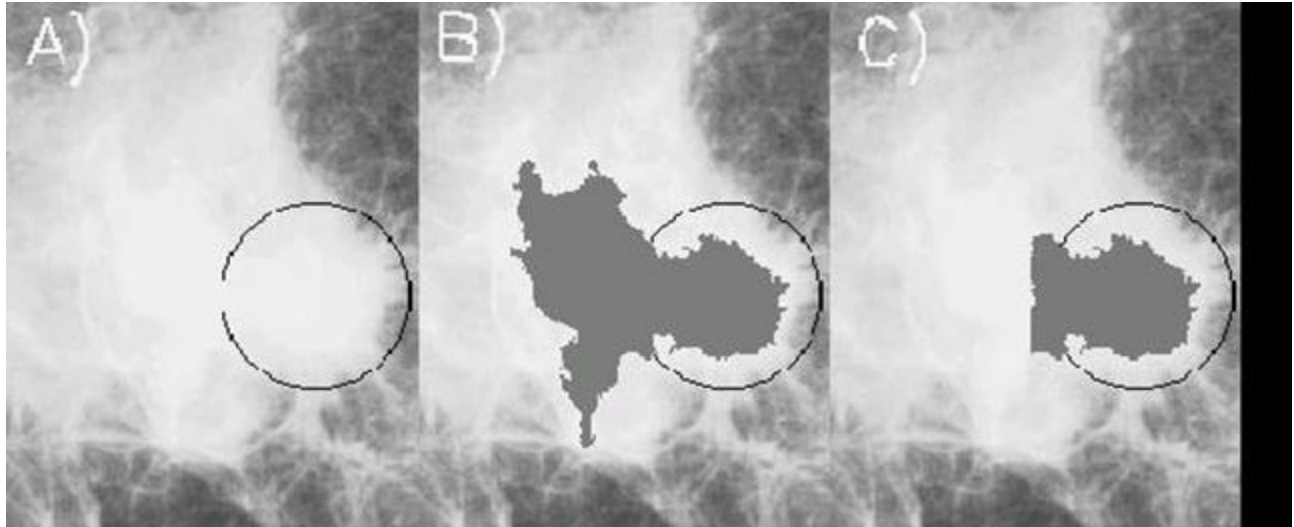


Ilustración 27 de una imagen segmentada por regiones. Tomada de Imágenes de Google.

El algoritmo general que describe paso a paso la segmentación basada en regiones, consta de tres partes las cuales se describirán a continuación: [26]

i. Iniciar la partición: En esta parte inicial se elige la región o regiones que componen la imagen, es muy importante el conocimiento que se tiene sobre las regiones a extraer; por lo que es conveniente la supervisión de un experto que elija los puntos claves para la segmentación.

ii. Definir un criterio de homogeneidad: Aquí se deben seleccionar los criterios que permiten clasificar las regiones seleccionadas en el punto anterior, por partes o clases. Se clasifican los píxeles en grupos por las semejanzas que comparten; Ej: Un píxel pertenece a una región cuando comparte el mismo color, textura y otras características que posea el grupo de píxeles que se encuentran en esa región. Existen dos técnicas para aplicar el criterio de homogeneidad:

- Determinista: [26]

Textura: Varianza respecto a un modelo.

$$C_T = \sum_n \sum_{i,j \in R_n} [x(i,j) - M_n(i,j)]^2$$

Ilustración 28 Ecuación de la varianza respecto a un modelo. [26]

R es la región, (i, j) es la posición y M es el modelo con el que se va a comparar.

Contorno: En este paso determinista se busca minimizar la longitud con el fin de obtener contornos simples. Un contorno simple es aquél que bordea la región objeto de estudio adaptándose a ella mediante la mínima longitud posible. Una vez que se logre describir el contorno de una región posteriormente se comparan las regiones a partir de estos descriptores de contorno.

$$C_c = \text{Longitud de contornos}$$

$$C = aC_t + (1 - a)C_i$$

Ilustración 29. Descriptores de contorno. [26]

- Aleatorio: Estudio del contorno de configuraciones locales. [26]

$$C = C_T \prod_{\text{Transiciones}} e^{(V_c)}$$

Ilustración 30. Descriptores de contorno. [26]

iii. Optimizar la partición mediante modificaciones (división, unión, etc.). En algunos casos después de esta etapa se regresa de forma iterativa al paso 2: Una vez realizados los dos pasos anteriores se procede al último paso en el cual se modifican y actualizan las particiones con el fin de modificar las condiciones de homogeneidad. Los cambios que se realizan en este paso consisten en la división y unión de regiones.

* *Crecimiento de regiones:* consiste en reunir píxeles o regiones con características similares para formar regiones más grandes que las agrupen.

El algoritmo consiste en escoger un conjunto de píxeles, al cual se le llamará píxeles semilla y se le irán agregando píxeles vecinos con características similares a este nuevo grupo de píxeles se le evaluará sus vecinos y se agregaran los que tengan características similares; así se hará hasta que ya no se encuentren píxeles parecidos. El criterio de parada muchas veces representa un problema.

* *División y fusión de regiones:* esta técnica no consiste en la agregación de píxeles como la anterior, por el contrario; consiste en dividir y subdividir la imagen.

Sea R la región correspondiente a la imagen completa y sea P el predicado de similitud elegido. En este caso para segmentar R se subdivide recursivamente cada región en cuadrantes más y más pequeños hasta que se cumpla para cada región R_i que $P(R_i) = \text{CIERTO}$. Es decir si $P(R) = \text{FALSO}$, se divide la imagen en cuatro regiones (cuadrantes). Si para alguno de los cuadrantes P es FALSO, se vuelve a dividir dicho cuadrante en cuatro subcuadrantes y así sucesivamente. [21]

Las técnicas basadas en regiones que hacen parte de la segmentación del procesamiento de imágenes, tienen muchos campos de aplicación entre ellos encontramos la medicina. A través de estas técnicas se puede localizar tumores y otras patologías, realizar medida de volúmenes de tejido, permite realizar cirugías guiada por ordenador, estudiar la estructura anatómica del cuerpo humano, realizar diagnósticos y planificar su respectivo tratamiento.

2.3.4 Extracción de características.

Como su nombre lo indica, esta técnica consiste en extraer las características de la imagen tales como tamaño, colores, área, forma, perímetro; entre otras. También permite conocer las características topográficas como la representación de objetos geográficos y el tamaño de la superficie.

La extracción de características hoy en día constituye un inmenso campo de estudio e investigación en muchas áreas con múltiples aplicaciones. No existe un método o algoritmo general o completo, para esta técnica, pero con la aplicación de varios métodos se logra la extracción de características de la imagen [7].

Local Strokedirection features (LSD) es uno de los muchos algoritmos utilizados en esta técnica que hace parte del procesamiento digital de imágenes; este consiste en procesar cada píxel (uno por uno); los pixeles repetidos en cada una de las cuatro direcciones se agrupan para después normalizarlos.

La normalización consiste en acomodar la forma, el tamaño y la posición de los píxeles con el fin de minimizar la variación entre los grupos de la misma clase. Luego se divide la imagen en secciones de tamaño $(n \times n)$ y se procesan las direcciones de cada área como el promedio de los pixeles en el área.

Una imagen contiene gran cantidad de datos; es por eso importante definir las

características principales para extraerlas y obtener la mayor información posible de la misma; en la elección de las características a extraer se debe tener en cuenta:

- El tiempo y costo de extracción no debe ser desmesurado. El costo no debe representar una complejidad muy grande para el sistema.
- La ubicación debe ser fija o muy precisa y el error cometido en la estimación debe inclinarse a ser muy bajo.
- Deben ser características estables; que puedan permanecer a lo largo de una secuencia.
- Deben contener gran cantidad de datos de los cuales sea posible obtener información de la forma de la imagen.
- La geometría de la imagen que define la forma de la misma, es una de las características más importantes, si estas no están bien definidas afectan en gran medida su calidad. Para el análisis de la geometría de la imagen
- Dentro de la extracción de características, la geometría de la imagen se clasifica entre las más importantes ya que si esta no está bien definida puede afectar en gran medida la calidad de la misma. Para el análisis de la imagen se tiene en cuenta la geometría básica, que son los puntos, líneas y círculos. A continuación se explicará la adecuada extracción para cada uno:

a) Extracción de puntos.

Los sistemas encargados de localizar las características geométricas de la imagen, lo hacen basado en dos métodos:

Métodos que obtienen los puntos como intersección de aristas o como cambio de pendiente sustancial entre dos aristas y por tanto vienen precedidos de una extracción de bordes. [7]

Métodos que trabajan directamente sobre imágenes de gris, esto quiere decir que no requieren una extracción previa de aristas. [7]

Existe una gran clasificación de puntos dentro de una imagen, los más característicos son los puntos de fuga y los puntos de esquina.

* Puntos de fuga: Es el lugar geométrico en el cual las proyecciones de las rectas paralelas a una dirección dada en el espacio, perpendiculares al plano de proyección, convergen. Por cada dirección en el espacio, existe un punto de fuga. Explicaremos algunos de los métodos

más sencillos que se encargan de los puntos de fuga:

- Método de la intersección de rectas: la complejidad computacional de este método es la más sencilla; consiste en delimitar el encuentro de dos rectas perspectivas. La recta es identificada por su distancia al origen y el ángulo que esta forma con los ejes cartesianos.

Este se calcula a través de la ecuación del punto pendiente y la ecuación normal.

Ecuación punto pendiente: $y = ax + b$

Ecuación general: $Ax + By + 1 = 0$

El desarrollo matemático quedaría de la siguiente forma:

$[-a \ 1] \cdot [x] = b$ ó $[A \ B] \cdot [x] = [-1]$ X y Y son las coordenadas del punto de fuga.

$[y] \ [y]$

- Método de la minimización del área del triángulo: Debido a los segmentos extraídos de la imagen y por el punto de fuga, se forman triángulos; este método se basa en el cálculo y minimización de la superficie de estos triángulos. La calidad precisa de este método consiste en localizar correctamente los puntos de aristas para poder detectar de una manera fácil y exacta los puntos de fuga.

* Puntos de esquina: Los puntos de máxima curvatura de una imagen, o los puntos más alejados de la zona factible de la imagen (en donde se localizan las características a extraer); se denominan puntos de esquina.

Uno de los primeros métodos establecidos para la obtención de estos puntos, se desarrolló en 1977; consiste en establecer sumas cuadráticas de las diferencias de los niveles de grises en las cuatro direcciones principales (horizontal, vertical y dos diagonales); a través de un detector y asigna al píxel central el mínimo M de los cuatro valores calculados. [7]

Un años después se desarrolló otro método que consiste en obtener los puntos de las esquinas por medio de un operador que calcula el determinante de una matriz, y luego se le realiza una umbralización (es una técnica de segmentación).

Se realizaron tres métodos para conseguir los puntos de esquina en 1982:

- Utilización del gradiente de la intensidad y del gradiente de la dirección en un punto de la

imagen como una medida de esquinidad.

- Se utilizaron las direcciones de gradiente de puntos vecinos que son perpendiculares a la dirección gradiente del punto; también como una medida de esquinidad.

- En un área de 3x3 se localizan los puntos A y B con un valor de gris similar al punto C. Existe una distancia entre los puntos C y A y los puntos C y B; formándose de esta manera los vectores CA y CB y la diferencia en la dirección entre estos dos vectores es la medida de esquinidad. Después de esto se empieza a eliminar puntos, hasta quedar con un único punto de esquina en el área.

En 1988 se prosigue en el estudio sobre las propiedades del determinante. Pero aun así la ubicación de los puntos no era precisa del todo.

En 1993, se profundiza sobre este método permitiendo saber más sobre la arista y multiresolución; también se desarrolló un estudio analítico del modelo de esquina; en este mismo año se desarrolló un detector de puntos de esquina por Haralick, este selecciona primeramente un cuadro importante dentro de la imagen y luego calcula de forma más exacta la ubicación de los puntos dentro del cuadro elegido.

Los pasos de este detector son: [7]

- i. Selección del tamaño de la máscara por parte del usuario (4 x 4, 8 x 8).
- ii. Cálculo de las derivadas de la imagen usando el operador gradiente para cada fila y columna.
- iii. Especificación de la matriz normal para cada cuadro de acuerdo al tamaño establecido por el usuario para la máscara.
- iv. Cálculo del peso y la circularidad de cada cuadro. Cuando el peso es mayor, mejor definido se encuentra el punto. La circularidad es usada para rechazar puntos localizados en bordes.
- v. Aplicación de la umbralización.
- vi. Supresión no máxima de peso dentro del cuadro seleccionado por el usuario.

b) Extracción de líneas.

Canny + Burns, RANSAC + MMCC y la transformada de Hough son técnicas que permiten extraer de forma automática las líneas de las imágenes. A continuación explicaremos cada una de estas:

* *Canny + Burns*: Técnica que se lleva a cabo a través de un proceso multifase jerárquico que permite extraer los bordes de la imagen por medio del algoritmo de Canny, una vez obtenidos los bordes, se segmentan mediante el algoritmo de Burns.

* *Algoritmo de Canny*: Este algoritmo fue desarrollado en 1986. En imágenes con geometrías regulares resulta ser más útil que en imágenes con otro tipo de geometrías; y esto se debe a tres principios del algoritmo; Precisión, fiabilidad y unicidad. [7]

La precisión permite tener una mejor proximidad de los bordes detectados a los bordes verdaderos. La fiabilidad reduce la posibilidad de detectar falsos bordes causados por el ruido, también la pérdida de bordes reales. Y la unicidad, logra que se detecte un único borde para el contorno.

El proceso multifase llevado a cabo durante este algoritmo, consta de una desviación estándar y dos valores umbral introducidos por el usuario; el resultado obtenido será una imagen binaria con píxeles de color negro y blanco. Los píxeles negros son los que pertenecen a los bordes; y así es como se logra detectar los bordes.

El algoritmo de Canny, también llamado filtro de Canny; consta de 4 fases: [22]

- Fase 1 (obtención del gradiente):

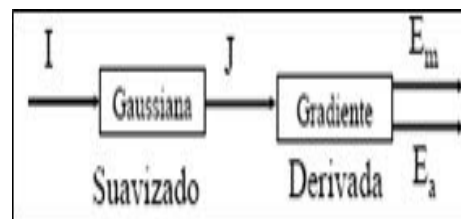


Ilustración 31. Imagen de la perspectiva de cómo se obtiene un gradiente [22]

Antes de obtener el gradiente se debe suavizar la imagen, para suavizar la imagen se aplica un filtro; en este caso aplicaremos el filtro gaussiano; el cual se calcula a través de la siguiente ecuación:

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

Ilustración 32. Ecuación 1 de la obtención de un gradiente [22]

Después de esto se calcula el gradiente para cada pixel que compone la imagen; para esto se

necesita el cálculo de los componentes del gradiente y el cálculo de la magnitud de los bordes. La ecuación que se aplica es la siguiente:

$$e_s(i, j) = \sqrt{J_x^2(i, j) + J_y^2(i, j)}$$

Ilustración 33. Ecuación 2 de la obtención de un gradiente [22]

Por último se estima la orientación de la normal de los bordes:

$$e_o(i, j) = \arctan \frac{J_y}{J_x}$$

Ilustración 34. Ecuación 3 de la obtención de un gradiente [22]

- Fase 2 (Supresión no máxima al resultado del gradiente):_Esta fase tiene como objetivo obtener bordes de 1 pixel de grosor y se consideran únicamente los píxeles cuya magnitud es máxima en bordes gruesos; se descartan los píxeles cuya magnitud no alcanzan el máximo.

La fuerza del borde proporcionará una medida del contraste local de la imagen, mientras que la orientación del borde identificará la dirección perpendicular al borde en forma de vector unitario. [7]

- Fase 3 (Histéresis de umbral a la supresión no máxima): En esta fase se eliminan los máximos procedentes de ruido, primero se encuentra la dirección que mejor se aproxima a la dirección normal al borde. Si la fuerza del borde es menor que al menos uno de los dos vecinos a lo largo de esa dirección suprimimos el punto como punto perteneciente a un borde. La salida es un conjunto de bordes conectados de contornos de la imagen, así como la magnitud y orientación.

- Fase 4 (Cierre de contornos abiertos):_La imagen de entrada para esta fase se encuentra

binarizada (1= Borde, 0= No borde). Para cada punto de borde de un extremo abierto se le asigna un código que determina las direcciones de búsqueda para el cierre del contorno: ej:

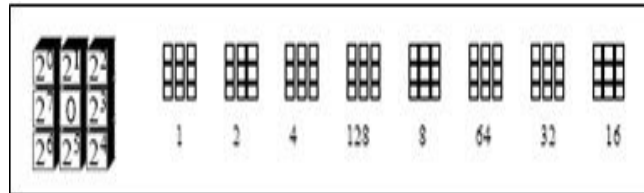


Ilustración 35. Ejemplo de un cierre de contornos abiertos [7]

Para los píxeles a los cuales se les ha asignado un código, se marca como pixel de borde el de máximo gradiente que se encuentre en las direcciones posibles.

Se itera sobre este mismo procedimiento, hasta que todos los contornos se cierren.

Ejemplo del algoritmo Canny:



Ilustración 36. Paralelo entre una imagen recién capturada y después de aplicarle el algoritmo de Canny [19]

* *Algoritmo de Burn (Segmentación de bordes)*: Los segmentos de una imagen constituyen una de las características geométricas más relevantes del procesamiento digital, debido a que suponen la base para una interpretación 3D de la escena [7]. La ejecución del proceso de segmentación se complica debido a la presencia de ruido y a la propia distorsión radial de la imagen. Es necesario extraer de la imagen como extremos del segmento los puntos que mejor definen la línea que se ajusta al borde para implementar una segmentación de

calidad. Para esto se debe desarrollar un proceso de segmentación que tenga como objetivo obtener un conjunto de segmentos (1D) definido por sus coordenadas extremas. El tiempo de procesamiento de la fase de segmentación dependerá linealmente del número de píxeles detectados como bordes en la fase anterior (algoritmo de canny) es aquí donde resulta importante aplicar por lo menos las 3 fases básicas del filtro de Canny.

La segmentación inicia con un barrido de arriba hacia abajo y de izquierda a derecha, para encontrar píxeles que se puedan clasificar dentro de un mismo grupo. El planteamiento principal consiste en etiquetar cada píxel de borde en grupos que soporten segmentos basándose en la similitud de las orientaciones del gradiente. De esta forma, cada píxel de borde será analizado con respecto a sus ocho vecinos en cada una de las direcciones posibles resultantes del gradiente. El grupo de píxeles que mantengan una orientación similar serán etiquetados como pertenecientes a un mismo borde, obteniendo una primera agrupación de los bordes en función de su gradiente. [7]

Los bordes resultantes de esta fase de etiquetación serán filtrados mediante el parámetro de mínima longitud de un borde. Para este filtro se calcula la longitud de los bordes anteriormente etiquetados y se compara con la tolerancia en la longitud fijada. Si la longitud es igual o superior a la tolerancia fijada por el usuario, el borde será transformado a segmento, el cual tendrá como extremos las coordenadas del centro del píxel correspondientes al inicio y fin del borde. Por el contrario, si la longitud es inferior a la tolerancia fijada, el borde será eliminado, no generando el correspondiente segmento.

* *RANSAC+MMCC*: RANSAC (RANmdom SAmple Consensus), es un estimador robusto desarrollado por Fischler y Bolles en 1981, este es un método iterativo basado en la aplicación de una técnica de votación resultante de un muestreo aleatorio; con el fin de calcular y determinar los parámetros de un modelo matemático de un conjunto de datos observados; es decir, el número de observaciones válidas ‘inliers’ y el número de observaciones erróneas ‘outliers’.

En esta parte del presente trabajo, este método y/o técnica se aplica en la determinación de posibles errores groseros que de no ser detectados, conlleva a gravísimos errores en la extracción de líneas.

Los pasos a considerar por parte de RANSAC serían los siguientes [7]:

- i. Selección aleatoria de dos puntos aleatorios para constituir una posible recta candidata.
- ii. Validación de la recta, en función del número de puntos que más menos una cierta tolerancia tengan una variación mínima de su distancia ortogonal a la recta candidata.
- iii. Repetición del paso 1 y 2 un determinado número de veces.
- iv. El máximo resultante del proceso de votación después de un número aleatorio de combinaciones se corresponderá con la recta candidata y por consiguiente los votos favorables ‘inliers’ del máximo contendrán los puntos favorables a constituir una recta. Los segmentos que superen la tolerancia en los tres grupos generados serán considerados como segmentos erróneo ‘outliers’ y serán eliminados.

MMCC. Una vez determinados los errores groseros se llevará a cabo una extracción de las líneas mediante la estrategia de MMCC.

Partiendo de la consideración de que entre x e y existe la relación lineal $y = ax + b$, el objetivo será determinar los parámetros a y b a partir de n -medidas de x e y ; donde a es la pendiente de la recta, es decir, la tangente del ángulo que forma con el eje de abscisas (eje x), y b la ordenada en el origen, es decir la altura a la que corta la recta al eje de ordenadas (eje y). [7]

Este vendría siendo el paso número 5 de esta técnica.

- v. Finalmente, se reajustará una nueva línea mediante una regresión ortogonal apoyada en los puntos clasificados por RANSAC como válidos ‘inliers’, obteniendo una recta ajustada al borde más precisa y fiable.

* *Transformada de Hough para líneas*: La Transformada de Hough es una técnica para la detección de figuras en imágenes digitales. Con la transformada de Hough es posible encontrar todo tipo de figuras como rectas, circunferencias o elipses; con el fin de poder ser expresadas matemáticamente; aunque está diseñada especialmente para encontrar líneas.

La transformada de Hough fue propuesta y patentada en 1962, por Paul Hough, inicialmente esta técnica solo se aplicaba a la detección de rectas en una imagen, más tarde se extendió para identificar cualquier figura que se pudiera describir con unos cuantos

parámetros; más comúnmente circunferencias y elipses. La transformada de Hough, como se usa actualmente, fue inventada por Richard Duda y Peter Hart en 1972, quienes lo llamaron "Transformada de Hough Generalizada". Dana H. Ballard popularizó este método en la comunidad de Visión por Computadora en un artículo publicado en 1981, llamado [generalizando la transformada de Hough para detectar figuras arbitrarias](#).

El objetivo de la transformada de Hough es agrupar los puntos que pertenecen a los bordes de posibles figuras, a través de un procedimiento de votación sobre un conjunto de figuras parametrizadas. La transformada de Hough considera las relaciones globales entre píxeles de borde permitiendo encontrar ciertos patrones en la imagen como líneas y círculos.

Primero se plantea la transformada de Hough para líneas y luego para círculos.

Supongamos que para n puntos de la imagen, se desean encontrar aquellos subconjuntos de puntos que caen en líneas rectas. Una posible solución podría ser en primer lugar encontrar todas las líneas determinadas por cada par de puntos y entonces encontrar todos los subconjuntos de puntos que están cerca de cada recta en particular. Este problema así planteado requiere encontrar $n(n-1)/2 \sim n^2$ rectas y realizar $n(n(n-1))/2 \sim n^3$ comparaciones de cada punto a línea. Este método no será viable salvo en casos triviales.

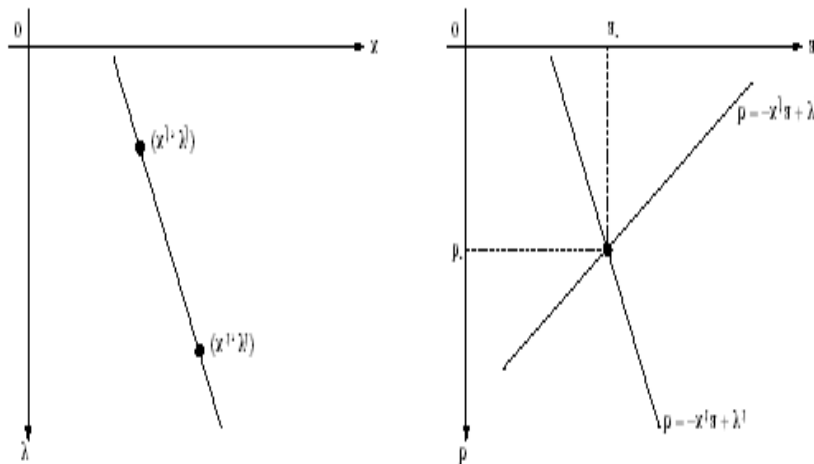


Ilustración 37. (a) Plano xy . (b) Espacio de parámetros. Tomada de Imágenes de Google.

Una alternativa al método de búsqueda exhaustiva es la transformada de Hough. Consideramos un punto (x_i, y_i) y la ecuación de la recta, de la forma pendiente y ordenada al origen,

$$y_i = ax_i + b \quad (1)$$

Por el punto (x_i, y_i) pasan infinitas rectas, pero todas satisfacen la ecuación anterior para diferentes valores de a y b . Sin embargo, escribiendo esta ecuación en la forma

$$b = -x_i a + y_i \quad (2)$$

y considerando el plano ab (también denominado espacio de parámetros) da lugar a una única recta para el par (x_i, y_i) constante. Si ahora consideramos un segundo punto (x_j, y_j) , también va a tener su recta asociada en el espacio de parámetros. Estas dos rectas se cortan en el espacio de parámetros en un punto (a', b') , donde a' es la pendiente y b' la ordenada al origen de la recta que contiene a los puntos (x_i, y_i) y (x_j, y_j) en el plano xy , como se puede ver en la figura 19. De hecho, todos los puntos de esa recta en el plano xy darán lugar a rectas diferentes en el espacio de parámetros que se cortan en un único punto (a', b') .

El atractivo de la transformada de Hough proviene de subdividir el espacio de parámetros en celdas acumuladoras, como se puede ver en la figura 38, donde (a_{\min}, a_{\max}) y (b_{\min}, b_{\max}) son los rangos esperados para la pendiente y la ordenada al origen. La celda de coordenadas (i, j) con un valor de acumulador $A(i, j)$ corresponde al cuadrado asociado con las coordenadas (a_i, b_j) del espacio de parámetros. Inicialmente se ponen todos los acumuladores a cero. Entonces para cada punto (x_k, y_k) de la imagen, permitimos que el parámetro a pueda tomar cualquier valor de entre los a_i permitidos y calculamos b usando la ecuación 2. Los valores resultantes para el parámetro b se redondean hasta los b_j permitidos. Si para un valor a_p resultó un valor b_q se tiene que

$$A(p, q) = A(p, q) + 1 \quad (3)$$

Al final, un valor de M en el acumulador $A(i, j)$ significa que M puntos del plano xy caen sobre la recta $y = a_i x + b_j$. La precisión en la colinealidad de estos puntos depende del número de celdas del espacio de parámetros.

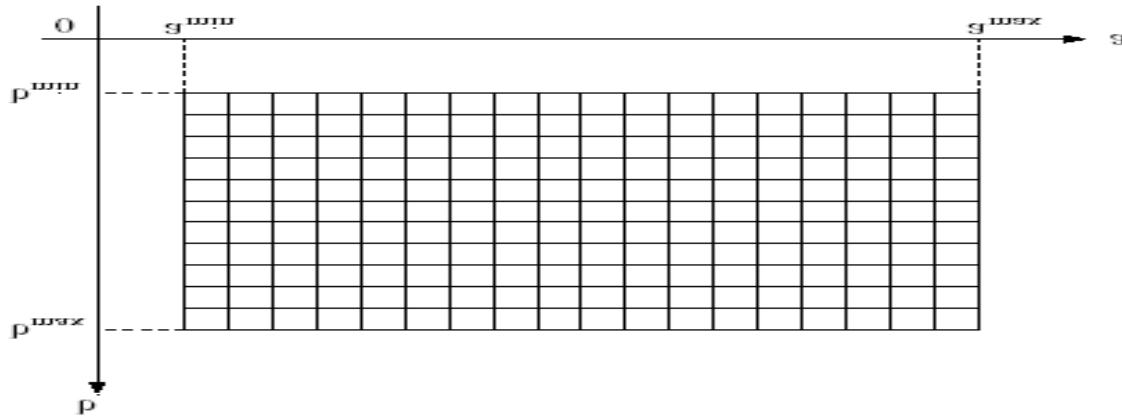


Ilustración 38. Celdas acumuladoras en el espacio de parámetros. [7]

Si subdividimos el eje a en K celdas, para cada punto (x_k, y_k) , obtenemos K valores de b correspondientes a los K posibles valores de A . Si la imagen tiene n puntos, la carga computacional es del orden de nK . La transformada de Hough es lineal en n , y el producto nK es mucho menor que si hubiéramos empleado una búsqueda exhaustiva, a menos que K sea del orden o exceda a n .

Un problema que surge al emplear la ecuación de la recta $y=ax+b$ para representar una línea es que tanto la pendiente como la ordenada al origen pueden llegar a valer infinito, según la línea se hace vertical. Una forma de solventar este problema consiste en utilizar la representación normal de la recta

$$x \cdot \cos \theta + y \cdot \sin \theta = \rho \quad (4)$$

En la ecuación 4 se puede ver el significado de los nuevos parámetros (ρ, θ) . El uso de esta representación para construir la tabla de acumuladores es similar al método explicado para las rectas en la forma pendiente y ordenada al origen. A cada punto del plano xy corresponde ahora una senoide en el plano $\rho\theta$ en lugar de una recta. Al igual que antes, M puntos colineales a la recta $x \cos \theta_j + y \sin \theta_j = \rho_i$ darán lugar a M sinusoides que se cortan en el punto (ρ_i, θ_j) en el espacio de parámetros. Incrementando θ y calculando ρ , obtendremos M entradas en el acumulador $A(i,j)$ correspondiente al par (ρ_i, θ_j) .

En la ilustración 39 podemos ver la tabla de acumuladores del espacio de parámetros en este caso. El rango para el ángulo θ es $\pm 90^\circ$, medido con respecto al eje de abscisas. Se

permiten valores negativos de ρ para rectas por detrás del origen de coordenadas del plano xy . Por ejemplo, una recta horizontal corresponde a un valor de $\theta = 0^\circ$ y un valor de ρ igual a la ordenada al origen, mientras que una recta vertical corresponde a un valor de $\theta = 90^\circ$ y un valor de ρ igual a la abscisa en el origen.

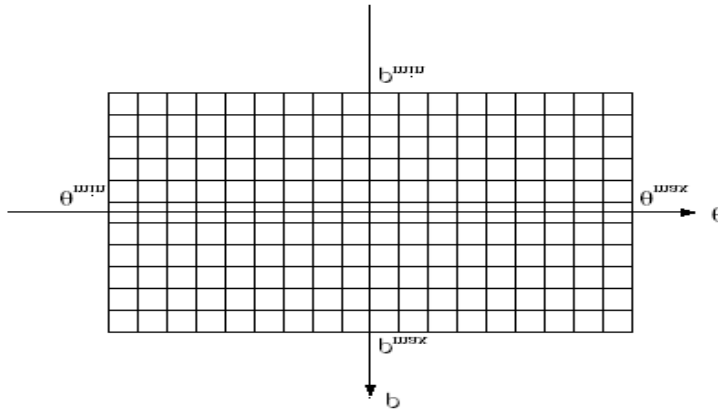


Ilustración 39. Celdas acumuladoras en el espacio $\rho\theta$.

En la ilustración 40 se ilustra con un ejemplo cómo funciona la transformada de Hough. La imagen está deformada por cuatro puntos correspondientes a las esquinas de un cuadrado. Estos cuatro puntos dan lugar a cuatro sinusoides en el espacio $\rho\theta$. Las cuatro sinusoides se cortan en seis puntos (en la figura 2 (b) aparecen ocho puntos, pero hay que recordar que los dos puntos para $\theta = 90^\circ$ son los mismos que los puntos para $\theta = -90^\circ$, por lo que de ocho son seis distintos), correspondientes a las seis rectas posibles que pasan por los cuatro puntos del plano xy , que son a saber, los cuatro lados del cuadrado y las dos diagonales.

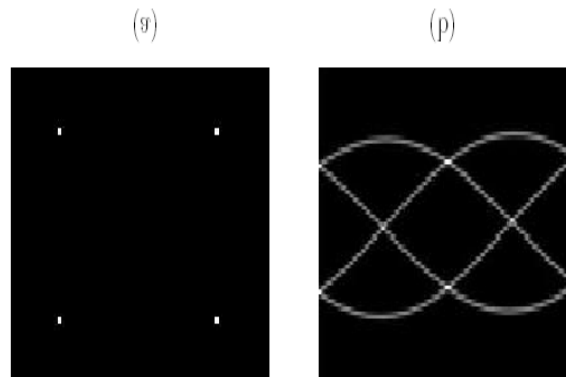


Ilustración 40. (a) Imagen con cuatro puntos de borde. (b) Transformada de Hough mostrando seis puntos de corte correspondientes a las seis rectas que pueden pasar por los cuatro puntos. [7]

Aunque hemos hecho un análisis para el caso de rectas, la transformada de Hough también

es aplicable a cualquier función de la forma

$$g(v, c) = 0 \quad (5)$$

Donde v es un vector de coordenadas y c es un vector de coeficientes. Por ejemplo, puntos que caen en el círculo

$$(x - c_1)^2 + (y - c_2)^2 = c_3^2 \quad (6)$$

Se pueden detectar empleando también la transformada de Hough. En este caso tenemos tres parámetros (c_1, c_2, c_3), lo que dará lugar a un espacio de parámetros de tres dimensiones, con celdas con forma de cubo y acumuladores de la forma $A(i,j,k)$. El procedimiento en este caso es para cada punto del plano xy , para cada c_1 y para cada c_2 , calcular el valor de c_3 y actualizar el acumulador correspondiente a (c_1, c_2, c_3). La complejidad de la transformada de Hough es claramente dependiente del tamaño del espacio de parámetros. [19]

$$(x - c_1)^2 + (y - c_2)^2 = c_3^2 \quad (6)$$

c) Extracción de círculos.

* *Transformada de Hough para círculos:* La transformada de Hough aplicada a cualquier figura geométrica está basada en transformar puntos de la imagen en un espacio de parámetros. Generalmente se realiza detección de bordes a la imagen, y luego se aplica la transformada a esta. De esta forma son menos los puntos que hay que recorrer y por lo tanto más rápido es el algoritmo.

Para detectar los círculos hay que aplicar una ecuación que consta de tres parámetros; uno de los parámetros es utilizado para el radio del círculo y los otros dos para el centro del círculo; lo que implica que el espacio de parámetros sea de dimensión tres.

El procedimiento realizado es muy similar al de la extracción de líneas, los pasos a seguir son los siguientes: [23]

- Elegir las máscaras a aplicar, existen diferentes procesos para obtener las máscaras, planteamos como ejemplo el operador Sobel para obtener las máscaras.
- El operador Sobel calcula el gradiente de la intensidad de una imagen en cada punto (píxel).
- Matemáticamente, el operador utiliza dos kernels de 3×3 elementos para aplicar

convolución a la imagen original para calcular aproximaciones a las derivadas, un kernel para los cambios horizontales y otro para las verticales. Si definimos \mathbf{A} como la imagen original, el resultado, que son las dos imágenes \mathbf{G}_x y \mathbf{G}_y que representan para cada punto las aproximaciones horizontal y vertical de las derivadas de intensidades, es calculado como:

$$\mathbf{G}_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * \mathbf{A} \quad \text{and} \quad \mathbf{G}_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} * \mathbf{A}$$

Ilustración 41. Kernel horizontal y vertical de las derivadas de intensidades. [23]

- En cada punto de la imagen, los resultados de las aproximaciones de los gradientes horizontal y vertical pueden ser combinados para obtener la magnitud del gradiente, mediante:


$$\mathbf{G} = \sqrt{\mathbf{G}_x^2 + \mathbf{G}_y^2}$$

Ilustración 42. Combinación de los gradientes. [23]

- Con esta información, podemos calcular también la dirección del gradiente:

$$\Theta = \arctan\left(\frac{\mathbf{G}_y}{\mathbf{G}_x}\right)$$

Ilustración 38. Dirección del gradiente. [23]

donde, por ejemplo,  es 0 para bordes verticales con puntos más oscuros al lado izquierdo.

- Aplicar las máscaras utilizando convolución discreta (2D)
- Aplicar por separado para obtener 2 gradientes G_x y G_y
- Obtener G con la fórmula $G = \sqrt{G_x^2 + G_y^2}$
- A partir de las matrices de gradientes obtenidas, calcular los ángulos para seno y coseno con las fórmulas: $\cos\theta = G_x/G$ y $\sin\theta = G_y/G$.
- Hallar las coordenadas del centro del círculo aplicando las fórmulas $x_c = x - r\cos\theta$ y $y_c = y - r\sin\theta$

- Hacer esto por cada pixel para obtener las coordenadas centro a partir de un radio establecido.
- Declarar un acumulador de votos, que puede ser un arreglo del mismo tamaño de la imagen en pixeles. El acumulador de votos recoge los potenciales centros de círculos de acuerdo al número de votaciones de los pixeles que lo rodean.
- Agrupar los votos que estén demasiado cerca, que el pixel con mayores votos se coma a los pixeles con menores votos para reducir el ruido y centros falsos.
- Los pixeles con mayores votos son los centros, implementar las rutinas necesarias para descartar falsas alarmas no detectadas en el proceso anterior.
- A partir de los centros, dibujar el contorno de los círculos encontrados, marcar su centro, etiquetar cada círculo y obtener el diámetro total.

Las imágenes necesitan tratarse antes de pasar por el proceso de detección de círculos:

- Binarizar la imagen
- Obtener su contorno

El radio es un parámetro de la simulación y se mide utilizando cualquier software de edición de imágenes, simplemente se mide cuántos pixeles hay de contorno a contorno en el centro, así obtenemos el diámetro, diámetro entre dos nos da el radio estimado en pixeles.

La extracción de características permite el desarrollo de grandes avances, una aplicación muy importante es el estudio de las señales electromiográficas que son una fuente de información bastante apropiada para el control de dispositivos virtuales como la prótesis de mano.

La extracción de características es una de las tareas claves que deben resolver los sistemas de visión artificial que cada vez presentan nuevos campos de aplicación. [27]

3. ANÁLISIS COMPARATIVO

Realce y mejora – Técnicas en el dominio espacial								
Características	Realce punto a punto						Realce con operaciones entre vecinos	
	Escala de grises	Negativo de imagen	Aumento del contraste	Compresión del rango dinámico	Fraccionamiento del nivel de gris	Procesado del histograma	Filtro pasa bajos	Filtro de mediana
Mejorar condiciones de bajo contraste, baja luminosidad o demasiada claridad	x	x	x	x	x			
Explotar el uso del contraste facilita la lectura de una imagen.	x		x			x		
Realzar detalles blancos o grises sobre un fondo oscuro.		x						
Incrementa el rango dinámico en los niveles de gris para mejorar la iluminación.			x					
Se utiliza cuando una imagen procesada excede ampliamente la capacidad del dispositivo de representación				x				
Destaca un rango especificado de niveles de gris para Mejorarlos, como en las masas de agua en imágenes satelitales o					x			

imágenes de rayos X con defectos								
Permite evidenciar ciertas particularidades de la imagen como el tipo de fondo, el contraste y en general si los valores de los niveles de gris están homogéneamente distribuidos o no						x		
Tiene la ventaja de que el valor final del pixel es un valor real presente en la imagen y no un promedio, de este modo se reduce el efecto borroso que tienen las imágenes que han sufrido un filtro de media.								x
Es útil cuando la imagen tiene gran cantidad de ruido y se quiere eliminar							x	x
Reemplaza el valor de cada pixel de la imagen con el valor promedio de sus vecinos incluyéndose el mismo. Esto posee el efecto de eliminar valores de píxeles poco representativos de sus vecinos.							x	

Realce y mejora – Técnicas en el dominio de la frecuencia		
Características	Filtro Ideal	Filtro Butterworth
Elimina ruido y suaviza la imagen para mejorar la calidad.	x	x
Atenúa las frecuencias altas y mantiene sin variaciones las bajas. El resultado en el dominio espacial es equivalente al de un filtro de suavizado, donde las altas frecuencias que son filtradas se corresponden con los cambios fuertes de intensidad. Consigue reducir el ruido suavizando las transiciones existentes.	x	
Atenúa las frecuencias bajas manteniendo invariables las frecuencias altas. Puesto que las altas frecuencias corresponden en las imágenes a cambios bruscos de densidad, este tipo de filtros es usado, porque entre otras ventajas, ofrece mejoras en la detección de bordes en el dominio espacial, ya que estos contienen gran cantidad de dichas frecuencias. Refuerza los contrastes que se encuentran en la imagen.		x

Segmentación – Detección de Bordes				
Características	Roberts	Prewitt	Sobel	Canny
Aplica la derivada de primer orden	x	x	x	x
Aplica la derivada de primer orden y la de segundo orden			x	x
Realiza un rastreo horizontal y vertical	x	x	x	x
Realiza un rastreo horizontal, vertical y hacia otras direcciones	x			
Permite discriminar los valores en que el gradiente no es máximo local				x

Segmentación - Umbralización			
Características	Técnica de los dos picos	Algoritmo ISODATA	Técnica Otsu
Permite encontrar el umbral más bajo entre dos picos	x		
Permite encontrar el segundo pico más alto	x		
Son iterativos		x	
Permite obtener el umbral óptimo de manera automática		x	x
Utiliza distribución de probabilidad, media y varianza			x

Segmentación – Basada en Regiones			
Características	Algoritmo basado en regiones	Crecimiento de regiones	División y fusión de regiones
Se obtiene un mayor conocimiento sobre las regiones que se van a extraer	x		
Aplica el criterio de homogeneidad	x		
Forma regiones más grandes a través de píxeles comunes		x	
El criterio de parada de este algoritmo muchas veces representa un problema		x	
Divide y subdivide la imagen en regiones para facilitar su procesamiento			x
Aplica recursividad			x

Extracción de características – Extracción de puntos			
Características	Puntos de Fuga		Puntos de Esquina
	Intersección de rectas	Minimización del área del triángulo	Métodos de puntos de esquina
Su complejidad computacional es la más sencilla	x		
Su cálculo se da a través de la ecuación del punto pendiente y la ecuación normal.	x		
Permite minimizar las superficies de triángulos que se forman		x	
Establece sumas cuadráticas a través de cuatro direcciones: vertical, horizontal y dos diagonales			x
Aplican derivada de primer orden			x
Aplican Umbralización			x

Extracción de características – Extracción de líneas y círculos					
Características	Canny + Burns	Algoritmo de Canny	Algoritmo de Burns	RANSAC + MMC	Transformada de Hough
Es un proceso multifase jerárquico	x	x	x		
Es más útil en imágenes con geometrías regulares		x			
Aplica el cálculo de la primera derivada		x			
Se le aplica segmentación con un barrido de arriba hacia abajo y de izquierda a derecha			x		
Es un método iterativo				x	
Permite encontrar todo tipo de figuras geométricas diferente a las líneas					x

5. CONCLUSIONES

- En la etapa de realce y mejora de imagen, se identificó que era necesario agrupar las técnicas en dos grupos principales (dominio espacial y dominio de la frecuencia), con el fin de poder distinguir las operaciones matemáticas en el plano espacial y en el espectro frecuencial para su respectiva explicación.
- Se identificaron más beneficios a través de las técnicas aplicadas en el dominio espacial (eliminación del ruido de la imagen, mejora de condiciones de iluminación, bajo contraste, mejoramiento en los niveles de gris para lograr destacar el contenido de la imagen) que en las técnicas aplicadas en el dominio de frecuencia.
- A diferencia de las otras técnicas aplicadas en el procesamiento de la imagen, las de realce y mejora utilizan una complejidad computacional menor a las demás.
- En la etapa de segmentación se identificaron bastantes técnicas, las cuales se clasificaron en 3 grupos (Detección de bordes, Umbralización y basado en regiones) con el fin de poder hacer comparaciones más precisas y realizar una análisis más claro y entendible.
- La extracción de características de acuerdo a la forma en la imagen se hizo a partir de 3 niveles: puntos, líneas y círculos, los que permiten determinar otras formas de extracción de figuras geométricas como elipses, parábolas, rombos, triángulos, rectángulos, cuadrados.
- Con el fin de que el lector tenga una mejor apreciación comparativa de técnicas para hacer procesamiento digital de imágenes, se elaboró un cuadro donde se resumen las principales características frente a las técnicas más usadas.
- Se espera que este documento sirva de apoyo a investigadores o en general a personas interesadas en conocer técnicas para el procesamiento digital de imágenes.

6. BIBLIOGRAFÍA.

- [1] Dr. Rivas Pablo. (2015). Tu centro médico online. Recuperado de:
<http://www.webconsultas.com/pruebas-medicas/tomografia-axial-computarizada-tac-8015>
- [2] Martinez Elena. (2014). Curso de Procesamiento Digital de Imágenes. Universidad Nacional Autónoma de México, Departamento de ciencias de la computación. Recuperado de: http://turing.iimas.unam.mx/~elena/PDI-Mast/Tema_1_ABC.pdf
- [3] M. Monteagudo Valdivia, “La imagen digital, una mirada interna”, Revista Cubana de Informática Medica, Recuperado de:
www.rcim.sld.cu/revista_3/articulos_html/articulo_pedro.htm
- [4] Real Academia Española. 23.^a edición. Madrid: Espasa Libros, S. L. U., 2014. Edición en cartóné, un volumen, formato: 17,5 x 26 cm. Recuperado de: <http://www.rae.es/>
- [5] Blog del grupo 9 de Psicología y TICS, “Sistemas inteligentes al servicio de la psicología aplicada”, 2011, Recuperado de:
<https://g9tics.wordpress.com/2011/11/06/sistemas-inteligentes-definicion-y-concepto/>
- [6] Profesores Academ 12, “Procesamiento digital de imágenes”, Benemerita Universidad Autónoma de puebla, 2015, Recuperado de:
<ftp://ece.buap.mx/pub/profesor/academ75/RENOVACION%20PNPC/Mtria%20en%20Ingenieria%20Electronica/1%29Estructura%20del%20Programa/1%20Plan%20de%20Estudios/1.1%20Plan%20de%20estudios.pdf>
- [7] González Aguilera D. Master de geotecnologías cartográficas en ingeniería y arquitectura, Universidad de Salamanca, Recuperado de:
<http://ocw.usal.es/eduCommons/enseanzas-tecnicas/procesamiento-avanzado-de-imagenes-digitales/contenidos/Tema2.pdf>
- [8] “Técnicas de realce de imágenes”, Marzo 2014, Recuperado de:
<http://dea.unsj.edu.ar/imagenes/recursos/capitulo3.pdf>
- [9] Peguero Núñez Pedro D. (2010), “Realce y restauración de imágenes”, Recuperado de:
https://www.lpi.tel.uva.es/muitic/pim/docus/Realce_y_restauracion.pdf
- [10] Ing. en Automatización y Control Industrial, “Segmentación por umbralización”,

Octubre 2005, Recuperado de:
<http://iaci.unq.edu.ar/materias/vision/archivos/apuntes/Segmentaci%C3%B3n%20por%20umbralizaci%C3%B3n%20-%20M%C3%A9todo%20de%20Otsu.pdf>

[11] Martín M. (2004), Técnicas clásicas de segmentación de imagen, Tomado de:
<http://lmi.bwh.harvard.edu/papers/pdfs/2003/martin-fernandezCOURSE03b.pdf>

[12] García, G. (2013), Procesamiento Audiovisual. *Adquisición y representación de imágenes*. Dept. de Informática y Sistemas, Universidad de Murcia. Tomado de: <http://dis.um.es/profesores/ginesgm>

[13]. García, G. (2013), Procesamiento Audiovisual. *Formatos de almacenamiento de imágenes*. Dept. de Informática y Sistemas, Universidad de Murcia. Tomado de: <http://alereimondo.no-ip.org/OpenCV/uploads/41/tema1.pdf>

[14] B. Aldalu y M. Santamaría, (2012). *Realce de Imágenes*. Departamento de Ingeniería, Departamento de Matemática Universidad Nacional del Sur Bahía Blanca, Argentina. Tomado de: <http://www.aet.org.es/revistas/revista17/AET17-04.pdf>

[15] Dra. La Serna Palomino N. (2009), Técnicas de segmentación en procesamiento digital de imágenes, Facultad de Ingeniería de Sistemas e Informática Universidad Nacional Mayor de San Marcos, Tomado de:
http://sisbib.unmsm.edu.pe/bibvirtual/publicaciones/risi/2009_n2/v6n2/a02v6n2.pdf

[16] Aznar Casanova J. Antonio, “Detección de bordes”, Recuperado de:
<http://www.ub.edu/pa1/node/101>

[17] Gálvez Sebastián, 30 de septiembre del 2014, "Operadores de detección de bordes", Recuperado de: https://www.u-cursos.cl/usuario/b4eb6d37062854338662ba7470704112/mi_blog/r/EL7007_-_Introduccion_al_Procesamiento_Digital_de_Imagenes_-_Tarea_3_-_Sebastian_Galvez.pdf

[18] "Filtros de detectar bordes", Recuperado de: <https://docs.gimp.org/es/plugin-sobel.html>

[19] Ing. en Automatización y Control Industrial, “Segmentación por umbralización”, Octubre 2005, Recuperado de:

<http://iaci.unq.edu.ar/materias/vision/archivos/apuntes/Segmentaci%C3%B3n%20por%20umbralizaci%C3%B3n%20-%20M%C3%A9todo%20de%20Otsu.pdf>

[20] Ceballos T, Noviembre del 2009, “Tratamiento de imágenes”. Recuperado de: <http://imagenesnovenio.blogspot.com.co/2009/11/principales-aplicaciones-de-la.html>

[21] Martín M. (2004), Técnicas clásicas de segmentación de imagen, Tomado de: <http://lmi.bwh.harvard.edu/papers/pdfs/2003/martin-fernandezCOURSE03b.pdf>

[22] Fernández Asunción O. (2013), DETECCIÓN DE BORDES: algoritmo de canny, Universidad nacional de trujillo, Tomado de: <http://oefa.blogspot.com.co/2009/04/deteccion-de-bordes-algoritmo-de-canny.html>

[23] J. C. Espinosa Ceniceros (2013), Detección de círculos (radio conocido), Tomado de: <http://juankenny.blogspot.com.co/2013/03/vc-tarea-4-deteccion-de-circulos-radio.html>

[24] Otero Diego, "Técnicas de umbralización", Universidad del valle, Tomado de: https://campusvirtual.univalle.edu.co/moodle/pluginfile.php/254390/mod_resource/content/0/Clases/thresholding.pdf

[25] Universidad de Sevilla, "Segmentación de imágenes". Tomado de: http://asignatura.us.es/imagendigital/Tema5_SegmentacionRegionesUmbralizacion.pdf

[26] Suárez Bravo Álvaro (2008), “Análisis de métodos de procesamiento de imágenes estereoscópicas forestales”, Universidad Complutense de Madrid, Tomado de: http://eprints.ucm.es/9875/1/An%C3%A1lisis_de_M%C3%A9todos_de_Procesamiento_de_Im%C3%A1genes_Estereosc%C3%B3picas_Forestales_-_%C3%81lvaro_Su%C3%A1rez_Bravo.pdf

[27] Giraldo Eduardo, Alvarez Damián A (2011), “ICA aplicado a la extracción de características en imágenes”, Universidad Tecnológica de Pereira. Tomado de: <http://biblioteca.versila.com/?q=extraccion+de+caracteristicas>

[28] Álvarez Rubiños Román, Septiembre del 2004, “Procesado de la imagen”, Tomado de: <http://sabia.tic.udc.es/gc/Contenidos%20adicionales/trabajos/Imagenyvideo/Procesadoh%20de%20imagen/TrabajoGC.htm#2.2>