

**RECONOCIMIENTO DE PATRONES EN IMÁGENES FIJAS USANDO
WAVELETS**



Trabajo de Grado

Fabián Ramírez Muñoz

Rider Valdiviezo Bolaños

Director: Ing. Harold Romo

Universidad del Cauca

Facultad de Ingeniería Electrónica y Telecomunicaciones

Departamento de Telecomunicaciones

Popayán, febrero de 2007

CONTENIDO

1. INTRODUCCION	1
2. CONCEPTOS BASICOS SOBRE TEORIA WAVELETS Y RECONOCIMIENTO DE PATRONES	5
2.1 CONCEPTOS BASICOS SOBRE TEORIA WAVELETS	5
2.1.1 Preliminares	5
2.1.2 Transformada wavelet (WT)	7
2.1.3 Análisis multirresolución	10
2.1.4 Codificación de sub-bandas	12
2.1.5 Transformada wavelet en imágenes	15
2.2 CONCEPTOS BASICOS SOBRE RECONOCIMIENTO DE PATRONES	18
2.2.1 Áreas básicas del análisis de imágenes	19
2.2.2 Modelo general para el procesamiento de imágenes	19
2.2.2.1 Adquisición	19
2.2.2.2 Preprocesamiento	20
2.2.2.3 Normalización	20
2.2.2.4 Segmentación	23
2.2.2.5 Descripción	23
2.2.2.6 Clasificación	23
2.2.2.7 Base de conocimientos	24
2.2.3 Reconocimiento estadístico de patrones	25
2.2.3.1 Patrones y clases de patrones	25
2.3 Descomposición wavelet en el reconocimiento de patrones	26
3 ALGORITMOS	29
3.1 Introducción	29
3.2 Selección de familias wavelet	30
3.2.1 Identificación de placas de vehículos	30
3.2.2 Reconocimiento de texto	33

3.2.3	Identificación de cultivos	35
3.2.4	Detección de bacilos de tuberculosis	38
3.3	Identificación de placas de vehículos	41
3.3.1	Localización de la placa	42
3.3.2	Normalización de la imagen	44
3.3.2.1	Redimensionamiento	44
3.3.2.2	Corrección de inclinación	45
3.3.3	Segmentación	47
3.3.4	Descripción	48
3.3.5	Clasificación	52
3.4	Identificación de texto plano	54
3.4.1	Localización de renglones, palabras y letras	55
3.4.2	Descripción	57
3.4.3	Clasificación	59
3.5	Identificación de un cultivo	60
3.5.1	Pre-procesamiento de la imagen	60
3.5.2	Identificación del patrón	63
3.6	Detección de tuberculosis	66
3.6.1	Procesamiento inicial	67
3.6.2	Eliminación de zonas ambiguas	68
3.6.3	Identificación de los bacilos Mycobacterium tuberculosis	69
4	CONCLUSIONES Y RECOMENDACIONES	74
4.1	Conclusiones	74
4.2	Recomendaciones y líneas futuras	76
ANEXOS:		
A	TEORIA SOBRE TRANSFORMADA WAVELET	78
B	TEORÍA SOBRE RECONOCIMIENTO DE PATRONES	88
C	MANUAL DE USUARIO	97

1. INTRODUCCION

Los seres humanos confiamos fundamentalmente en el sentido de la vista, pues el 99 % de nuestra información la recogemos a través de nuestros ojos [1]. El procesamiento de imágenes es un proceso que concierne tanto a ordenadores como a algoritmos diversos para un propósito específico, una vez que hemos adquirido imágenes de interés a través de una cámara y / o un escáner [1].

Mucha de la información existente en la actualidad se presenta a las personas en forma de imágenes, las cuales contienen varios objetos que se diferencian entre si porque cada uno tiene características diferentes. Entre esos objetos, algunos comparten un conjunto de características comunes que los identifica y permite agruparlos bajo algún criterio. Se dice entonces, que ese grupo de objetos tiene el mismo patrón y conforman una clase. Un patrón es una descripción cuantitativa o cualitativa de un objeto o clase y una clase es un conjunto de objetos que comparten propiedades similares. A esas características que conforman un patrón se les llamará descriptores.

El reconocimiento de patrones es el estudio de los mecanismos por los cuales se clasifica un objeto como perteneciente a una clase particular con base en sus descriptores. Suena lógico el buscar una forma para diseñar y construir mecanismos que puedan reconocer patrones porque de una forma u otra es uno de los mecanismos naturales que utilizamos a diario gracias a nuestros sentidos. Esos mecanismos se pueden soportar en el uso ordenadores ya que es necesario

que el procesamiento de datos sea rápido para analizar observaciones del mundo real y obtener información de objetos de interés en una imagen.

Una imagen puede ser definida como una función bi-dimensional $f(x,y)$, donde x e y son coordenadas espaciales, y la amplitud de f en cualquier par de coordenadas se denomina intensidad o nivel de gris de la imagen en el punto. Cuando las coordenadas (x,y) y los valores de la amplitud de f son cantidades discretas, estaremos ante una imagen digital. El procesamiento de imágenes digitales concierne al uso del ordenador en el procesamiento de imágenes. Note que una imagen digital está compuesta de un número finito de elementos, cada uno de ellos con una localización y un valor particular, que se llaman elementos pictóricos, elementos unidad de imagen o más comúnmente, píxeles.

A diferencia de los seres humanos, que están limitados a la banda visual del espectro electromagnético (EM), las máquinas de imagen cubren prácticamente todo el espectro EM desde los rayos gamma hasta las ondas de radio. Pueden trabajar también con imágenes generadas por fuentes que los humanos no están acostumbrados a asociar con imágenes. Estas fuentes incluyen las imágenes ultrasónicas, la microscopía electrónica y las imágenes generadas por ordenador. En definitiva, el procesamiento de imágenes digitales concierne a un amplio y variado elenco de campos de aplicación [2].

Desde los años sesenta del siglo pasado, el procesamiento de imágenes digitales se ha convertido gradualmente en una de las áreas de investigación científica más importantes. El desarrollo del procesamiento de imágenes está siendo acelerado

aún más con el rápido avance de las tecnologías relacionadas con la computación en paralelo, el incremento en la capacidad de memoria de los chips, y el sistema de visualización en color de alta-resolución.

El procesamiento digital de imágenes se utiliza con múltiples fines, entre ellos el reconocimiento de patrones, que es el objetivo de este trabajo y para ello se utiliza la transformada wavelet, debido a que esta transformada es particularmente útil para analizar cambios bruscos o discontinuidades que suceden en señales e imágenes a distintas frecuencias y al descomponer una imagen en aproximación y detalles se obtienen más elementos para describir y clasificar los objetos que la componen, que es la idea principal del reconocimiento de patrones. Como se trabaja con imágenes digitales se utiliza la transformada wavelet discreta en las aplicaciones que se esbozan a continuación:

Identificación de placas de vehículos

En esta aplicación se identifica, a partir de la fotografía de un vehículo, los caracteres alfanuméricos que conforman la placa. La clasificación de los caracteres se logra a partir de los coeficientes wavelet.

Reconocimiento de texto

Esta aplicación identifica caracteres alfanuméricos sobre la imagen de un texto, obtenida de un fax o un scanner. La clasificación de los caracteres se logra a partir de los coeficientes wavelet.

Identificación de un cultivo específico

En esta aplicación se identifican distintas texturas en una imagen, es decir, se busca en una imagen la zona o zonas que tengan una textura particular y que se hace corresponder, por ejemplo, con un cultivo conocido. El reconocimiento de las texturas se logra por medio de los coeficientes wavelet.

Detección del bacilo de tuberculosis

Esta aplicación identifica los bacilos de tuberculosis presentes en la muestra de esputo de una persona. Para esto se analizan fotografías microscópicas y con la ayuda de la transformada wavelet se detectan los bacilos basándose en su morfología.

2. CONCEPTOS BÁSICOS SOBRE TEORÍA DE WAVELETS Y RECONOCIMIENTO DE PATRONES

2.1 CONCEPTOS BASICOS SOBRE TEORIA DE WAVELETS

2.1.1 Preliminares

Vamos a realizar éste estudio teórico sobre wavelets alrededor del análisis multirresolución y la transformada rápida discreta wavelet en una señal unidimensional y extendiendo esta idea a señales bidimensionales como son las imágenes. Aunque es inevitable nombrar otras técnicas de análisis de señales que anteceden a las wavelets; sobre todo para entender el ¿por qué? y las ventajas de esta nueva técnica.

El objetivo de este capítulo es hacerse una idea básica del procesamiento digital de señales; *análisis* (descomposición) y *síntesis* (reconstrucción), en particular sobre imágenes usando la transformada wavelet en su versión de *análisis multirresolución*. Por lo tanto se tratará de evitar la utilización de formulas matemáticas que describan los proceso en la medida de lo posible. El soporte teórico tanto conceptual como matemático se encuentra en el anexo A.

En la teoría wavelets se trabaja de manera similar a la teoría de Fourier, la cual analiza una señal en sinusoides de diferentes frecuencias, transformando la señal desde un marco temporal a un marco de frecuencias, lo que conocemos como Transformada de Fourier (Fourier Transform. (FT)) [3]. Para muchas señales el análisis de Fourier es muy útil, debido a que el contenido de frecuencias de la señal es de gran importancia, pero el análisis de Fourier tiene una seria

desventaja. Cuando se realiza la transformación al dominio de frecuencias, se pierde la información temporal, al observar la transformación de Fourier de una señal, no es posible saber en que instante sucedió una frecuencia en particular [3]. Si una señal no cambia mucho en el tiempo (esto es, si la señal es estacionaria) esta desventaja no es muy importante. No obstante, la mayoría de señales interesantes contienen numerosas características no estacionarias o transitorias: tendencias, cambios bruscos, comienzos o finalizaciones de eventos, etc. Estas características son a menudo, la parte más importante de la señal, y el análisis de Fourier no está adecuado para detectar estos cambios.

Para tratar de corregir esta deficiencia Gabor adaptó la transformación de Fourier. Analiza la señal por “secciones pequeñas” mediante la utilización de una función ventana. La adaptación de Gabor, denominada la Transformación de Fourier de corto tiempo (Short Time Fourier Transform (**STFT**)) lleva una señal a una función bidimensional de tiempo y frecuencia. La STFT representa un intento de obtener conjuntamente los marcos temporales y de frecuencias de una señal, aportando alguna información sobre cuándo y en qué frecuencias ocurre [4]. No obstante, la información obtenida mediante esta técnica presenta una precisión limitada, ya que la precisión es función del tamaño de la ventana. Mientras que la información de tiempo y de frecuencia (obtenida por la STFT) puede ser útil, la desventaja es que, una vez se elige un tamaño particular para la ventana temporal, la ventana es la misma para todas las frecuencias. Muchas señales necesitan una aproximación mucho más flexible, donde se pueda variar el tamaño de la ventana para determinar más detalladamente tanto el tiempo como la frecuencia [4].

El principio que esta por debajo del inconveniente anterior es el *Principio de Incertidumbre de Heisenberg*, el cual, en términos de procesamiento de señales establece que no es posible saber la frecuencia y el tiempo exacto de ocurrencia de esta frecuencia en una señal [5], esto es, una señal simplemente no puede ser representada como un punto en el espacio tiempo-frecuencia.

A partir de este inconveniente surge la inquietud sobre si es posible tener una ventana con rango dinámico que permita realizar un análisis de la señal de manera ideal, es decir, tener una buena resolución en tiempo para frecuencias altas, y una buena resolución en frecuencia para analizar las frecuencias bajas. Precisamente como solución a este problema tiene lugar la herramienta matemática llamada transformada wavelet.

2.1.2 Transformada wavelet

La transformada wavelet (Wavelet Transform (**WT**)) descompone una función o señal a analizar en términos de funciones denominadas wavelets que son generadas por una función wavelet de transformación denominada *wavelet madre*. El término wavelet significa “onda pequeña”. La pequeñez se refiere al hecho de que esta función (ventana) es de longitud finita (de soporte compacto) y onda por ser de naturaleza oscilatoria. El término madre da a entender que las funciones con diferentes regiones de actuación que se usan en el proceso de transformación provienen de una función principal o wavelet madre, es decir, la wavelet madre es un prototipo para generar las otras funciones ventana. [4]

La WT comprende básicamente la transformada wavelet continua (Continuous Wavelet Transform (**CWT**)) y la transformada wavelet discreta (Discrete Wavelet

Transform (**DWT**)), herramientas matemáticas que posibilitan el análisis de señales proporcionando información en el dominio del tiempo y en el dominio de la frecuencia. El análisis mediante la CWT se efectúa sobre señales analógicas de energía finita (definida en el teorema de Parseval). [4]

La CWT es la suma de la señal a lo largo de todo el tiempo, multiplicada por versiones de la wavelet (versiones estiradas/comprimidas y adelantadas/retardadas). Este proceso produce coeficientes wavelet, que son función de la escala y la traslación (ver anexo A). Los coeficientes constituyen los resultados de una regresión de la señal original realizada con las wavelets (su magnitud es proporcional a la similitud de la wavelet con “el segmento” de señal analizado). Mientras más estirada sea la wavelet, más larga es la “porción” de la señal con la cual está siendo comparada y, por lo tanto, más gruesas son las características de la señal que está siendo medida por los coeficientes wavelets. Gráficamente, el proceso es el siguiente, ver figura 2.1.

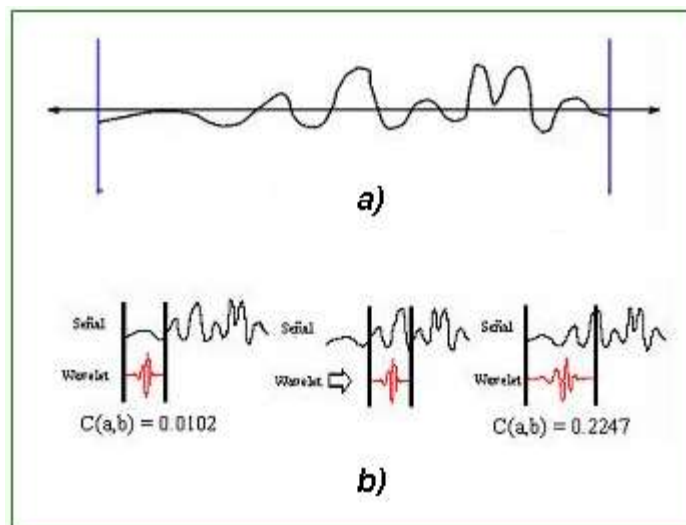


Figura 2.1 Modo de funcionamiento de la CWT, a) Señal a analizar, b) representación del proceso de traslación y cambio de escala sobre la señal [7].

En la figura 2.1 vemos como entre más se parece la señal a la wavelet (con una escala a y en una posición b determinada) más alta es la magnitud del coeficiente wavelet.

Existe una correspondencia entre las escalas y frecuencias de las wavelets, y es la siguiente:

- Baja escala \rightarrow wavelet comprimida \rightarrow cambios rápidos en detalle \rightarrow alta frecuencia.
- Alta escala \rightarrow wavelet estirada \rightarrow cambios lentos, características gruesas \rightarrow baja frecuencia.

Aun cuando la transformada wavelet continua puede evaluarse computacionalmente de manera discretizada, esto no constituye realmente una transformada discreta, sino una serie wavelet o la versión muestreada de la CWT, con la desventaja que la información que entregan es altamente redundante para la reconstrucción de la señal. Esta redundancia significa además un aumento significativo del tiempo de cálculo [4]. Por este motivo se utiliza la transformada wavelet discreta que es capaz de entregar suficiente información tanto para el análisis como para la reconstrucción de una señal con una significativa reducción del tiempo de procesamiento, además, es mucho más fácil de implementar que la CWT. Básicamente, lo que hace la transformada wavelet en el dominio discreto es filtrar una señal en el dominio del tiempo mediante filtros paso bajo y paso alto que eliminan ciertas componentes de alta o baja frecuencia de la señal, el

procedimiento se repite para las señales resultantes del proceso de filtrado anterior (un *análisis multirresolución*).

El cálculo de los coeficientes wavelet en cada escala origina una gran cantidad de trabajo, y genera una gran cantidad de datos. Pero si elegimos solamente un subconjunto de escalas y posiciones con las cuales realizar los cálculos por ejemplo, si escogemos escalas y posiciones basados en potencias de dos (las denominadas escalas y posiciones diádicas) los cálculos serán mucho más eficientes sin mayor pérdida de precisión. Resultando una transformada rápida wavelet discreta (como una función donde se introduce una señal y por donde se obtienen rápidamente coeficientes wavelet) [4].

2.1.3 Análisis multirresolución

Para ser útil la teoría de wavelets debe disponer de algoritmos rápidos para encontrar los coeficientes wavelets y para reconstruir la señal que representan. Existe una familia rápida de algoritmos basados en el análisis multirresolución (Multiresolution Analysis (**MRA**)). El análisis multirresolución consiste en la descomposición de una señal original en una *aproximación (A)* y un *detalle (D)*. La aproximación es entonces descompuesta en un segundo nivel de aproximación y detalle, y el proceso es repetido sucesivamente. Es por esta razón que, en análisis wavelet, se habla de *aproximaciones* y *detalles*. Las aproximaciones son las componentes de gran escala y baja frecuencia de la señal. Los detalles son las componentes de pequeña escala y alta frecuencia. El proceso de filtración, en su nivel más básico (a un solo nivel) para una señal **S** se ve en la figura 2.2 a) Y en la figura 2.2 b) una descomposición en tres niveles.

Para muchas señales, el contenido de baja frecuencia es la parte más importante, es lo que aporta identidad a la señal. El contenido de alta frecuencia, por otro lado, tiene un carácter secundario [4]. Consideremos como ejemplo la voz humana: si eliminamos los componentes de alta frecuencia, la voz suena diferente, pero podemos entender todavía qué se está diciendo. No obstante, si eliminamos suficientes componentes de baja frecuencia, comenzamos a no entender el habla.

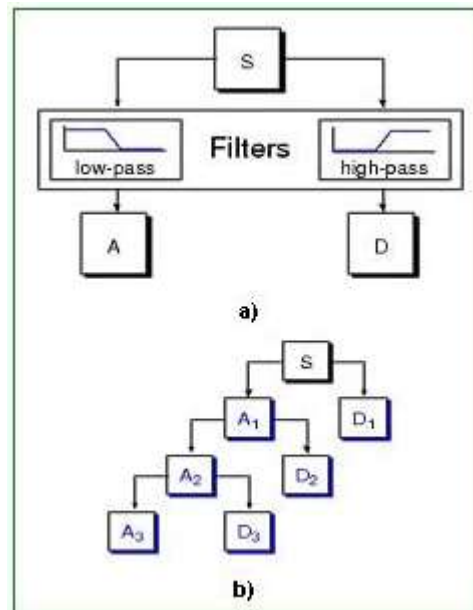


Figura 2.2 a) Modo de descomposición de la señal (S), en que opera el análisis Wavelet. En la descomposición se obtiene una Aproximación (A) y un Detalle (D). b) descomposición de la señal (S) a tres niveles [7]

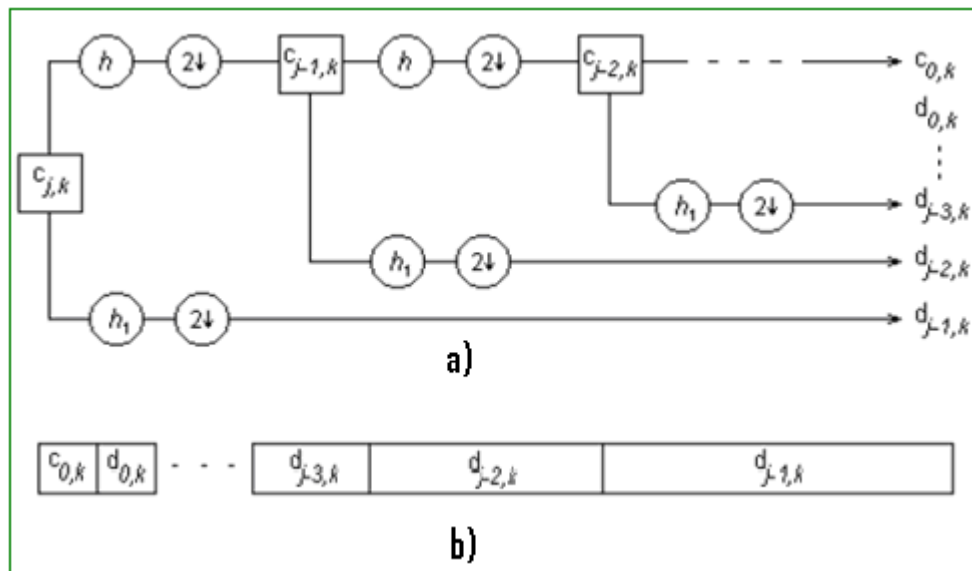
El proceso de descomposición puede ser iterativo, con aproximaciones descompuestas sucesivamente en diferentes niveles, de forma que es posible caracterizar una señal mediante el análisis de sus componentes de menor resolución. Esto es denominado el árbol de descomposición wavelet. La observación de una descomposición de señales de esta manera, aporta una información muy valiosa. Resumiendo, el concepto de descomponer la señal en

aproximación y detalles es el *análisis multirresolución* y el proceso para hacerlo se conoce como *codificación de sub-bandas*.

2.1.4 Codificación de sub-bandas

En la codificación de sub-bandas un filtro paso bajo de media banda elimina todas las frecuencias que están por encima de la mitad de la mayor frecuencia de la señal, una vez que la señal ha pasado por el filtro, la mitad de las muestras se pueden eliminar de acuerdo al *Criterio de Nyquist* [6]. Es decir la tasa de muestreo de Nyquist se puede reducir a la mitad ya que al pasar por el filtro paso bajo la frecuencia máxima de la señal es ahora la mitad de la frecuencia máxima de la señal original. Con este propósito se elimina una de cada dos muestras de la señal, submuestreo por 2 (El submuestreo por un factor “n” reduce el número de muestras de la señal “n” veces) con lo cual se reduce el número de puntos a la mitad puesto que la mitad del número de muestras son redundantes y la escala de la señal se duplica. El filtrado paso bajo reduce a la mitad la resolución (aquí la resolución se entiende como la medida de la cantidad de detalle de la señal) pero no altera la escala.

El procedimiento anterior se denomina codificación de sub-bandas y el algoritmo se le llama *algoritmo piramidal*, y puede repetirse para conseguir una mayor descomposición, en este caso en cada etapa. El filtrado y el submuestreo darán como resultado una disminución a la mitad del número de muestras (resolución en el tiempo dividida) y de la banda de frecuencias abarcada (resolución en frecuencia duplicada). Lo anterior es el proceso de análisis y se esquematiza en la figura 2.3. [6]



**Figura 2.3. a) Esquema clasificación de sub-bandas para el análisis (descomposición)
b) Organización del vector de coeficientes wavelets.**

En la anterior figura, c_{jk} son los valores de la señal a analizar, h_1 representa el filtro paso-alto, h el filtro paso-bajo, $2\downarrow$ simboliza el submuestreo por dos, c los coeficientes de aproximación y d los coeficientes de detalles. Aquí j maneja el nivel de resolución que se desea trabajar y la longitud de la señal debe ser 2^j .

Las frecuencias que son más dominantes en la señal original aparecerán como altas amplitudes en la región de la DWT que incluye esas frecuencias. La diferencia entre la FT y la DWT es que con la DWT no se pierde la localización en el tiempo de estas frecuencias. Sin embargo, la localización en el tiempo tendrá una resolución que dependerá del nivel en que aparezca, de este modo si la información principal contenida en la señal está en altas frecuencias, como sucede a menudo, entonces la localización en el tiempo de estas frecuencias será más precisa, puesto que estarán caracterizadas por un mayor número de muestras. Por otro lado, si la información principal está a muy bajas frecuencias entonces su

localización en el tiempo no podrá ser muy precisa, dado que existirán muy pocas muestras para caracterizar la señal a estas frecuencias.

Las bandas de frecuencia que no son muy dominantes en la señal darán origen a coeficientes de la DWT muy pequeños, los cuales pueden despreciarse sin mayor pérdida de información, pero si con una importante reducción de los datos.

La DWT de la señal original se obtiene concatenando todos los coeficientes, comenzando desde el último nivel de descomposición, La DWT tendrá entonces el mismo número de coeficientes que la señal original.

Un hecho importante de destacar es que al utilizarse un submuestreo sucesivo por dos el número de muestras de la señal debe ser potencia de dos o por lo menos un múltiplo de dos de manera que el procedimiento de cálculo sea eficiente. La cantidad de muestras de la señal determina el número máximo de niveles de descomposición, por ejemplo, si la señal tiene 1024 muestras entonces es posible realizar hasta 10 niveles de descomposición ($2^{10} = 1024$).

El proceso de reconstrucción esto es la *síntesis* es similar al de descomposición, la señal en cada nivel es interpolada por dos, pasada por dos filtros de síntesis, esto se esquematiza en la figura 2.4. [6]

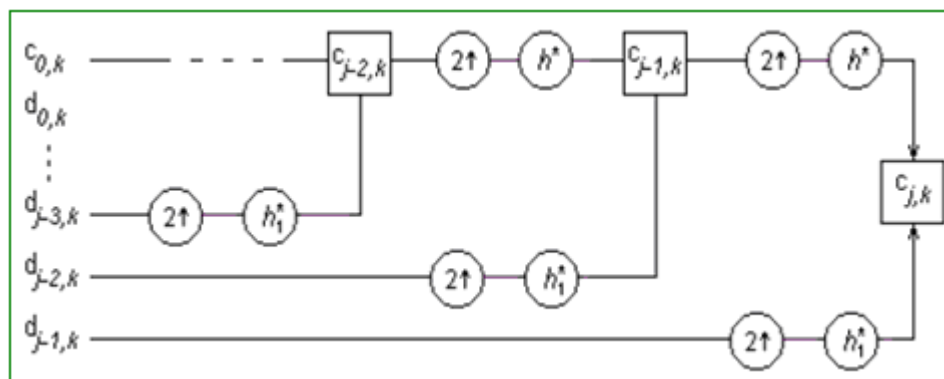


Figura 2.4. Esquema clasificación de sub-bandas para la síntesis (reconstrucción)

En la anterior figura, \mathbf{h}_1^* y \mathbf{h}^* representan los filtros de síntesis, $2\uparrow$ simboliza la interpolación por dos.

2.1.5 Transformada wavelet en imágenes

Computacionalmente, una imagen es una matriz de N^2 píxeles (sin pérdida de generalidad suponemos una imagen cuadrada). Si $N = 2^J$; partiendo de una resolución inicial máxima (de 2^J píxeles por lado), podremos analizarla a $J = \log_2 N$ niveles de resolución distintos y decrecientes; es decir, a J escalas distintas crecientes, que van proporcionando versiones "suavizadas" de la imagen. Para una señal finita s bidimensional de N píxeles por lado, se aplican dos filtros: paso bajo y paso alto, en dos direcciones (columnas y filas), dando lugar a cuatro filtrados posibles en cada escala. Por eso el algoritmo proporciona cuatro matrices de orden $N/2$. En cada escala j , empezando por $s = c^0$; el algoritmo en cascada para el análisis consiste en formar una matriz cuadrada de orden N , ver figura 2.5

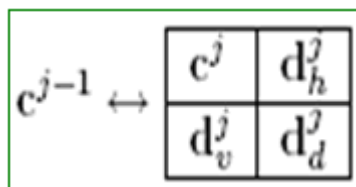


Figura 2.5 Descomposición de la matriz original en las tres imágenes detalle y la imagen de aproximación.

que contiene los detalles de cada punto de la imagen en la escala $j-1$, las aproximaciones c_j y los detalles d_j según las direcciones horizontal, vertical y diagonal:

- c_j : contiene la versión suavizada de la imagen a escala j .

- \mathbf{d}_h^j : contiene los detalles o saltos obtenidos sobre las líneas verticales de la imagen; por tanto, contienen los contornos horizontales en la escala j .
- \mathbf{d}_v^j : contiene los detalles o saltos obtenidos sobre las líneas horizontales; por tanto, proporcionan los contornos verticales en la escala j .
- \mathbf{d}_d^j : contiene los detalles o saltos obtenidos sobre las líneas diagonales; por tanto, dan información de los contornos diagonales en la escala j .

Descomponiendo de forma similar y recursiva la submatriz c^j hasta llegar a la máxima escala J se obtendría su *representación piramidal*. [8] Esquemáticamente lo vemos en la figura 2.6

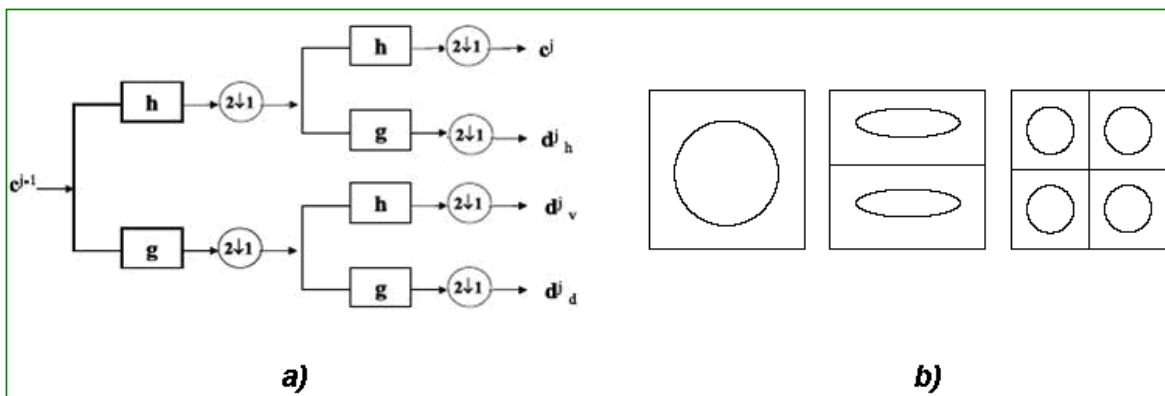
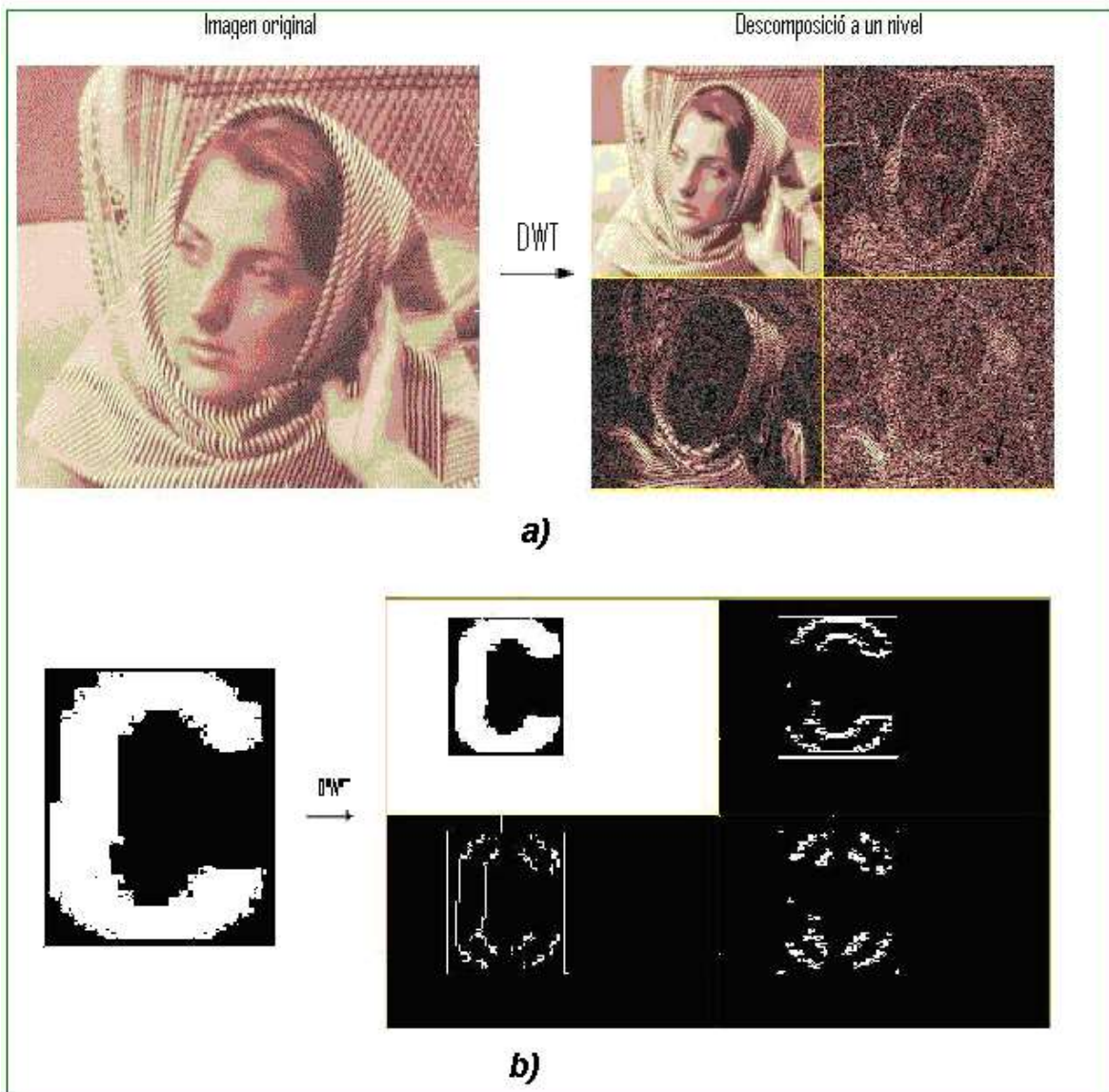


Figura 2.6. a) Etapa de análisis de la imagen mediante los filtros paso bajo h y paso alto g . b) Proceso aplicado a una imagen primero filas luego columnas.

En la anterior figura con la notación $(2\downarrow 1)$ se simboliza el submuestreo por dos, en la imagen filtrada se habrá eliminado una fila cada dos filas y columna cada dos columnas en cada nivel. El proceso de análisis aplicado a una imagen real se ve en la figura 2.7.



**Figura 2.7. Imagen original y la imagen de aproximación y las tres imágenes de detalle.
a) Fotografía de un rostro b) Un caracter de la placa del vehículo [7]**

Al igual que en el caso unidimensional, una vez obtenida la matriz con las imágenes detalle y la imagen tendencia, se puede aplicar a éstas el algoritmo inverso para reconstruir la imagen original. Ahora bien, si se aplica el algoritmo de reconstrucción solo a los coeficientes wavelet de un determinado nivel de resolución haciendo cero el resto de la matriz, se puede reconstruir cualquier nivel de detalle. De la misma manera, si se aplica el algoritmo de reconstrucción sólo a

los coeficientes de aproximación, haciendo cero el resto de la matriz se puede reconstruir la tendencia. Cada imagen detalle resulta de tres componentes independientes en el dominio wavelet que dan cuenta de los detalles de la imagen original. También se puede analizar la imagen en diferentes bandas de frecuencia y reconstruirla usando sólo los coeficientes que pertenecen a una banda en particular. Esquemáticamente lo vemos en la figura 2.8.

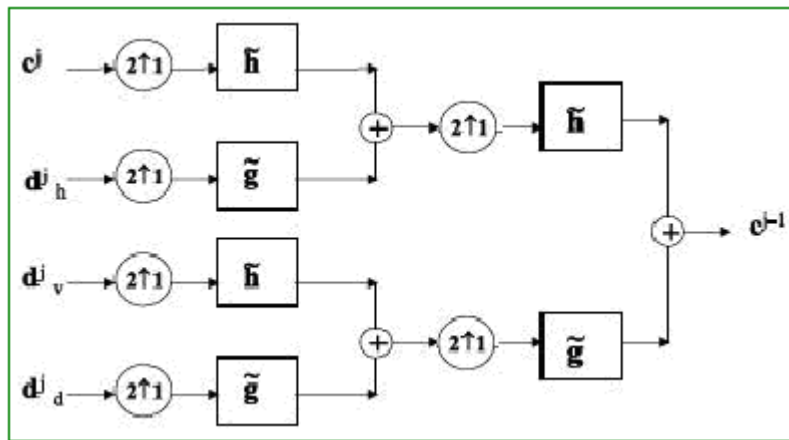


Figura 2.8. Etapa de síntesis de la imagen.

En la figura anterior \tilde{h} y \tilde{g} son filtros de síntesis y $(2^1 1)$ simboliza interpolación por dos. En la imagen reconstruida se habrá insertado una fila entre dos filas y una columna entre dos columnas en cada nivel, los valores de inserción pueden ser ceros o un promedio entre los dos valores en que se hace la inserción.

2.2 CONCEPTOS BASICOS SOBRE RECONOCIMIENTO DE PATRONES

El reconocimiento de patrones se ubica en la parte final del análisis de imágenes, este análisis se divide comúnmente en tres áreas o procesos.

2.2.1 Áreas básicas del análisis de imágenes

Se puede dividir el análisis de imágenes en tres áreas básicas:

1) *Procesado de bajo nivel*: trata de un tipo de funciones a las que se puede considerar como reacciones automáticas, y que no requieren inteligencia por parte del sistema de análisis de imágenes. Esta clasificación incluye actividades desde el proceso mismo de formación de las imágenes hasta las diversas compensaciones aplicables, como la reducción del ruido o la supresión, en la medida de lo posible, del emborramiento de las imágenes.

2) *Procesado de nivel intermedio*: se refiere a la labor de extracción y caracterización de los componentes de la imagen que se obtiene de un proceso a bajo nivel y abarca la segmentación y la descripción.

3) *Procesado de alto nivel*: comprende el reconocimiento y la interpretación, estos dos procesos tienen un gran parecido con lo que normalmente se conoce como conocimiento inteligente [9].

El reconocimiento de patrones es parte del procesado de alto nivel por lo tanto esta será nuestra área de interés.

2.2.2 Modelo general para el procesamiento de imágenes

2.2.2.1 Adquisición: El primer paso del proceso es la adquisición de la imagen, esto es, digitalizarla. En este trabajo se utilizan imágenes a color con codificación RGB, imágenes en tonos de gris e imágenes binarias.

Imagen RGB: Una imagen RGB se compone de tres matrices que representan la intensidad de los colores rojo (R), verde (G) y azul (B), respectivamente. En cada

matriz se especifican espacialmente los valores de intensidad que cada píxel de la imagen tiene en el color correspondiente.

Imagen en tonos de gris: Es una imagen en blanco y negro donde cada píxel tiene una intensidad que está generalmente entre 0 (negro) y 255 (blanco).

Imagen binaria: Es una imagen que solo tiene 2 valores lógicos, normalmente unos (blanco) y ceros (negro).

2.2.2.2 Preprocesamiento: Después de que la imagen digital ha sido obtenida, el siguiente paso es el preprocesamiento. Su función es mejorar la imagen de manera que se incremente la oportunidad de éxito de los siguientes procesos. El preprocesamiento típicamente trata con técnicas para realzar el contraste y remover ruido.

2.2.2.3 Normalización: En algunas ocasiones después del preprocesamiento es necesario normalizar la imagen realizando las siguientes operaciones.

a) Corrección de inclinación. Consiste en modificar la imagen para dejarla simétrica a los ejes x y y . Para esto es necesario obtener primero los ángulos de inclinación de la imagen y esto se logra por medio de la transformada de Hough.

Transformada de Hough

La transformada de Hough es una técnica utilizada para aislar características de forma particular dentro de una imagen. La idea básica es encontrar curvas que puedan ser parametrizadas como líneas rectas, polinomios y círculos.

Se puede analíticamente describir un segmento de línea en varias formas. Sin embargo una ecuación conveniente para describir un conjunto de líneas es la notación paramétrica o normal:

$$\rho = x \cos \theta + y \sin \theta$$

(2.1)

Donde ρ es la longitud de una normal desde el origen hasta la línea y θ es el ángulo de ρ con respecto al eje x. Cualquier línea recta en una imagen es representada en un punto simple (ρ_1, θ_1) en el espacio paramétrico (ρ, θ) .

Puntos colineales (x_i, y_i) con $i = 1, 2, 3 \dots N$ son transformados en N curvas sinusoidales en el plano (ρ, θ) . Figura 2.9

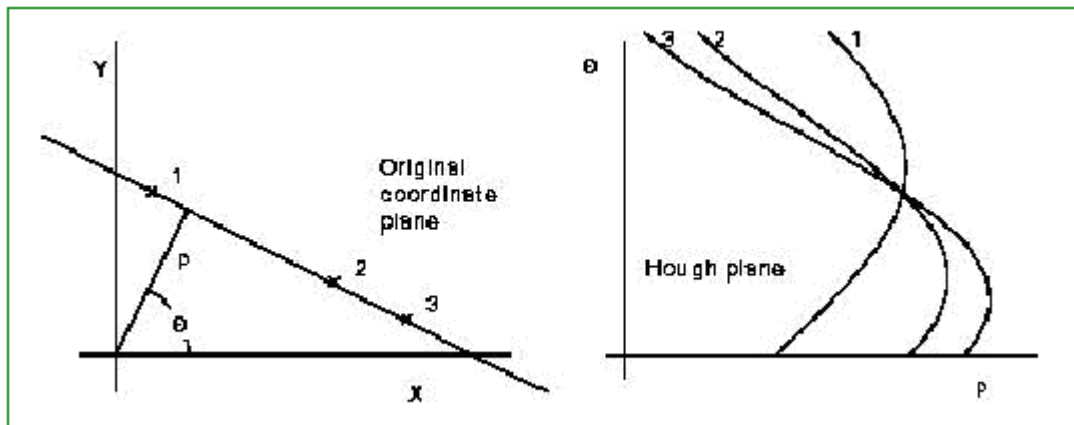


Figura 2.9 Representación gráfica de la transformada de Hough.

Los puntos de intersección de las curvas en el espacio paramétrico, corresponden a los parámetros (ρ_k, θ_k) de las posibles rectas que se encuentran en la imagen. El espacio paramétrico es representado por una estructura rectangular de celdas, llamada arreglo acumulador y cuyos elementos son las celdas acumuladoras

$A(\rho_i, \theta_i)$, las cuales son los rangos esperados de (ρ, θ) . Las celdas acumuladoras con una magnitud superior a un cierto umbral pueden ser consideradas como posibles líneas.

Algoritmo de Hough

Sea E una imagen binaria de tamaño $M \times N$, en la cual cada píxel de la imagen es uno o cero. Sean ρ_d, θ_d vectores que contienen valores discretos del espacio paramétrico ρ, θ ($\rho \in [0, (M+N)^{1/2}]$ y $\theta \in [0, 180]$). Los pasos a seguir para aplicar el algoritmo de Hough son los siguientes:

Descripción del algoritmo

- 1.) Discretize el espacio parámetro ρ, θ .
- 2.) Sea $A(R, T)$ una matriz de contadores inicializada en 0, donde R y T son el número de elementos de ρ y θ respectivamente.
- 3.) Para cada píxel $E(i, j)$ que pertenezca a la línea desde $h = 1, 2, \dots, T$ con $k = 1, 2, \dots, K$.
 - a) Hacer $\rho = i \cos \theta_d(h) + j \sin \theta_d(h)$.
 - b) Halle el intervalo K que sea mas cercano al ρ obtenido.
 - c) Incremente matriz de contadores $A(k, h)$.
- 4.) Encuentre el máximo local (k_p, h_p) tal que sea mayor que un umbral.
- 5.) Con el ρ y θ encontrados se describe la línea en forma polar [10].

Una vez obtenidos los ángulos de inclinación se procede a corregirlos realizando una transformación espacial sobre la imagen.

b) Redimensionamiento. Consiste en modificar el tamaño de la imagen o de los objetos constituyentes cuando sea necesario. Esto se logra por medio de una transformación espacial con las dimensiones y el método de interpolación deseados.

2.2.2.4 Segmentación. Definida en forma amplia, la segmentación particiona una imagen de entrada en sus partes constituyentes u objetos. Generalmente, la segmentación automática es una de las tareas más difíciles en el procesamiento de imágenes. En términos del reconocimiento de caracteres, por ejemplo, el rol de la segmentación es extraer caracteres individuales y palabras del fondo de la imagen.

2.2.2.5 Descripción. Llamada también selección de características, trata con extracción de los rasgos que resulta en alguna información cuantitativa de interés o características que son básicas para diferenciar una clase de objetos con otra. En términos del ejemplo, para el reconocimiento de caracteres, descriptores tales como lagos (huecos) y bahías son rasgos poderosos que ayudan a diferenciar una parte del alfabeto de otra.

2.2.2.6 Clasificación. Este último estado comprende al **reconocimiento** y la **interpretación**. El reconocimiento es el proceso que etiqueta o asigna un nombre a un objeto basándose en la información que proveen sus descriptores. La interpretación involucra la asignación de significado a un conjunto de objetos reconocido.

2.2.2.7 Base de conocimientos. La base de conocimientos mantiene todo el conocimiento que se tiene acerca del problema tratado. Este conocimiento puede ser simple, teniendo solo las regiones de la imagen donde la información de interés se conoce que está localizada, con lo que se limita la búsqueda que conduce a esa información. La base de conocimientos también puede ser compleja, tal como una lista interrelacionada de todos los defectos mayores posibles en un problema de inspección de materiales o una base de datos de imágenes conteniendo imágenes de satélite en alta resolución de una región en conexión con aplicaciones de detección de cambios. Además de guiar la operación de cada modulo del proceso, la base de conocimientos también controla la interacción entre módulos. Esto es representando con flechas de doble punta en la figura 2.9 Con esto se indica que la comunicación entre módulos del proceso generalmente está basada sobre conocimiento previo del resultado que se podría esperar. Este conocimiento no sólo guía la operación de cada módulo, sino también ayuda en las operaciones de retroalimentación entre módulos a través de la base de conocimientos.

Es importante hacer notar que los resultados del procesamiento pueden verse en la salida de cualquier paso de la figura 2.9. También debe notarse que no todas las aplicaciones del procesado de imágenes requieren la complejidad de interacciones de la figura 2.9. Numerosas aplicaciones caen fuera de este esquema, de hecho, no todos los módulos son siempre necesarios. Por ejemplo, el realce de imágenes para interpretación visual humana rara vez va más allá de la etapa de preprocesamiento. En general, las funciones de procesamiento que incluyen reconocimiento e interpretación están asociadas con aplicaciones de

análisis de la imagen en la cual el objetivo es que se extraiga la información de una imagen en forma automática o semiautomática [9].

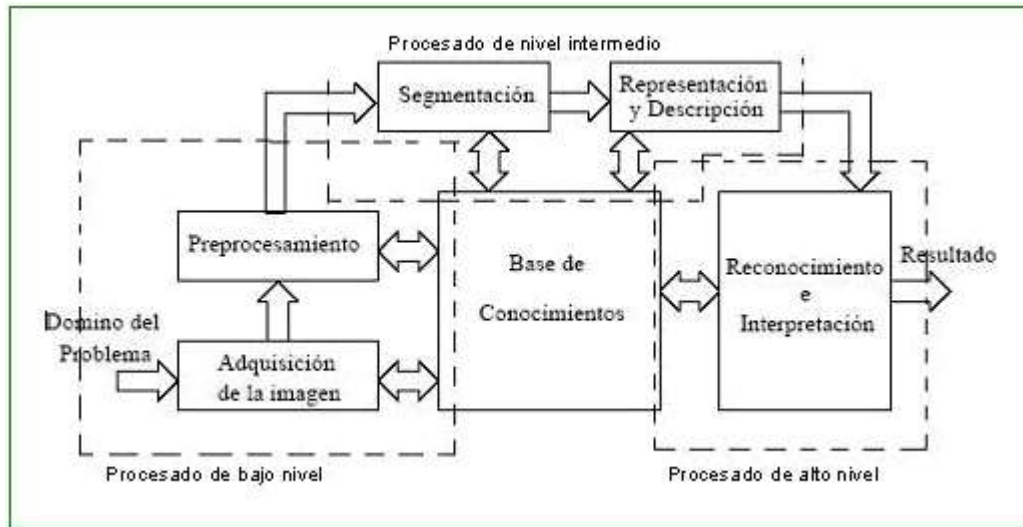


Figura 2.10 Elementos del análisis de imágenes

2.2.3 Reconocimiento Estadístico de Patrones.

2.2.3.1 Patrones y clases de patrones.

Un patrón es una descripción estructural o cuantitativa de un objeto o de alguna otra entidad de interés en una imagen. En general, un patrón está formado por uno o más descriptores. Una clase de patrones es una familia de patrones que comparten algunas propiedades comunes. Las clases de patrones se representan por w_1, w_2, \dots, w_M , donde M es el número de clases.

Las tres representaciones de patrones principalmente utilizadas en la práctica son los vectores (para descripciones cuantitativas), las cadenas y árboles (para descripciones estructurales).

Los patrones vectoriales se representan de la siguiente manera:

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} \quad (2.2)$$

Donde cada componente, \mathbf{x}_i , representa el i -ésimo descriptor y n es el número de descriptores. La naturaleza de los componentes de un patrón vectorial \mathbf{X} depende de la técnica de medida utilizada para describir el propio patrón físico. Tras la selección de un tipo de medida un patrón vectorial pasa a ser la representación completa de cada ejemplo físico.

Los patrones vectoriales pueden generarse de muchas formas. Un concepto clave a tener en cuenta es que la selección de las medidas utilizadas para obtener cada componente de un patrón vectorial tiene una profunda influencia en el rendimiento final de un sistema de análisis de imágenes basado en la técnica de los patrones vectoriales. En este trabajo se utilizó el clasificador de mínima distancia, que es un método de decisión teórica. Este método y otros métodos usados en el reconocimiento de patrones se describen en el Anexo B.

2.3 Descomposición wavelet en el reconocimiento de patrones

Una de las etapas más importantes en el reconocimiento de patrones es la extracción de características o descripción, ya que resalta los aspectos más sobresalientes de cada clase y a partir de esa descripción podemos asignar un patrón desconocido a una clase en particular. Los métodos de extracción de características y los criterios usados para escoger esas características varían de una aplicación a otra.

Una imagen digital está conformada por pequeños cuadros de color llamados píxeles cuyo color queda definido por un valor numérico que depende de la codificación utilizada. Por ejemplo, en escala de grises, los valores están generalmente entre 0 y 255 donde el 0 representa al negro y el 255 al blanco, pasando por 254 tonos distintos de gris (más claro mientras más alto sea el número) llamados niveles de brillo. En este trabajo se utilizan inicialmente imágenes a color con codificación RGB. Esta codificación guarda información en matrices con los datos que conforman cada píxel [11]. Una imagen RGB se compone de tres matrices que representan la intensidad de los colores rojo (R), verde (G) y azul (B), respectivamente. En cada matriz se especifican espacialmente los valores de intensidad que cada píxel de la imagen tiene en el color correspondiente, por lo que cada píxel de color es un componente de la tripleta correspondiente a rojo, verde y azul [12]. Como la herramienta “Wavelet 2D” de matlab solo trabaja con imágenes monocromáticas, posteriormente se transforman esas imágenes a color en imágenes en tonos de gris en unos casos, y en otros se descompone la imagen y se trabajan con las matrices R, G, ó B por separado.

En una imagen en tonos de gris los píxeles difieren unos de otros en forma discreta, es decir, la variación de un píxel a otro no tiene forma suave o continua sino que presenta un cambio abrupto entre niveles discretos, que se hace mayor en los contornos o fronteras de las regiones que conforman la imagen. Como se mencionó en el capítulo 2 una particularidad de la transformada wavelet es que es útil para estudiar esos cambios repentinos y el análisis multirresolución se convierte en una herramienta importante en el procesamiento de imágenes ya que

permite analizar las imágenes para destacar características notorias en diferentes frecuencias. Por esta razón se utiliza en este trabajo la DWT para extraer las características más sobresalientes, de los objetos que conforman la imagen, cuantificarlas y obtener los descriptores.

La DWT aplicada a imágenes proporciona una matriz de coeficientes, conocidos como coeficientes wavelet. La aproximación contiene la mayor parte de la energía de la imagen, es decir, la información más importante, mientras que los detalles tienen valores próximos a cero [13]. Los coeficientes de aproximación indican la tendencia de la imagen mientras que los coeficientes de detalle son una medida de la magnitud de los cambios en las direcciones horizontal, vertical y diagonal por separado. Con lo anterior se concluye que los coeficientes wavelet pueden constituir una herramienta eficaz en la descripción de las imágenes ya que brindan información relevante sobre las características de los objetos que la conforman, información que se puede, mediante algún operador, cuantificar para formar un patrón que describa claramente a un objeto o a una clase.

En todas las aplicaciones desarrolladas en este trabajo se utilizaron los coeficientes wavelet para obtener los descriptores con los que se realiza la clasificación de clases. El comportamiento de una wavelet varía dependiendo de la aplicación, por lo que en cada aplicación se ensayaron todas las familias wavelet provistas por el toolbox wavelet de matlab para escoger la que mejor se ajustara a una necesidad particular. En la sección 3.2 se verá como se seleccionaron las familias wavelet en cada aplicación.

3. ALGORITMOS

3.1 Introducción

El procesamiento digital de imágenes busca obtener información puntual de las imágenes a partir de técnicas y métodos específicos dependiendo de una aplicación particular. El presente trabajo se enfoca en el reconocimiento de patrones específicos en imágenes digitales para identificar objetos de interés.

El procesamiento digital de imágenes tiene inimaginables aplicaciones en diferentes campos de la ciencia y aquí se abordarán algunas de ellas para dar idea de su aplicabilidad. Las imágenes analizadas aquí son fotografías de distintos escenarios en los cuales se desea obtener más información y más precisa que la brindada por la imagen a simple vista.

Estas aplicaciones son:

- Identificación de placas de vehículos.
- Identificación de texto en una imagen.
- Detección de tuberculosis y cuantificación del grado de infección.
- Detección e identificación de vegetación específica.

Es de resaltar la importancia que tiene en todas las aplicaciones el pre-procesamiento de las imágenes. Es decir preparar la imagen antes del reconocimiento como tal.

3.2 Selección de familias wavelet

A continuación se ven algunos resultados de las pruebas realizadas para seleccionar la familia wavelet utilizada en cada aplicación.

3.2.1 Reconocimiento de placas de vehículos

En el reconocimiento de placas de vehículos se utilizan los coeficientes de detalle horizontal, vertical y diagonal de la descomposición wavelet para la extracción de características de las imágenes que quedan tras la etapa de segmentación (sección 3.2.3). Se escogieron las tres matrices de coeficientes de detalle para tener más elementos que describan un objeto, en este caso, un carácter alfanumérico. Por lo tanto se debe buscar la wavelet que más resalte los detalles tras aplicar descomposición, o sea, cuyos coeficientes de detalle, tengan valores altos. Debido a que se necesita buena resolución de los detalles se realiza descomposición en el primer nivel, ya que en niveles superiores la descomposición se hace con una escala mayor y la resolución disminuye. En la siguiente tabla se muestra la suma de los valores absolutos de los coeficientes de detalle horizontal, vertical y diagonal.

WAVELET	DETALLE HORIZONTAL	DETALLE VERTICAL	DETALLE DIAGONAL		TOTAL
DB1	47,5	154,5	23,5		225,5
DB2	74,6	180,6	32,8		288
DB3	88,9	172,4	35,5		296,8
DB5	101,9	214,7	43,2		359,8
DB7	98,4	240,7	52,6		391,7

DB9	122,2	243,7	41,4		407,3
SYM2	74,6	180,5	32,7		287,8
SYM4	86,3	194,9	37,5		318,7
SYM6	91,5	204,7	40,9		337,1
SYM8	96,5	219,6	44,2		360,3
SYM10	101,8	241,3	48,8		391,9
COIF1	83,7	181,2	32,8		297,7
COIF3	93,4	211,5	44		348,9
COIF5	108,7	247,9	55,3		411,9
BIOR 1,1	47,5	154,5	23,5		225,5
BIOR 2,4	70,3	136,9	22,7		229,9
BIOR 3,1	73,5	144	20,9		238,4
BIOR 3,7	73,1	151,3	21,1		245,5
BIOR 5,5	99,6	221,1	57		377,7
RBIO 1,1	47,5	154,5	23,5		225,5
RBIO 2,4	109,6	224,6	57		391,2
RBIO 3,1	106,5	273	111		490,5
RBIO 5,5	85,1	178,6	29,4		293,1

Tabla 3.1 Suma de coeficientes de detalle

En la tabla se mira que la wavelet que produce valores más altos es la biortogonal inversa 3.1 (rbio 3.1) pero la imagen (figura 3.1a) es algo difusa por lo tanto se descartó. La Daubechie 9 (db 9) también tiene valores altos pero la imagen se distorsiona porque aparece una especie de reflexión en los extremos (figura 3.1b)

y también se descartó. Cuando, al hacer descomposición, una wavelet se adapta bien a la forma de los objetos de una imagen particular, las líneas de detalle son bien definidas e intensas y sus coeficientes correspondientes en las matrices respectivas son de gran magnitud. En general, aunque, las wavelet que resaltan mejor los detalles de una imagen particular tienen coeficientes con valores altos y las wavelet que no se adaptan muy bien a la forma de los objetos de la imagen desprecian muchos detalles o los resalta de manera muy tenue produciendo coeficientes de pequeña magnitud, en este caso, el simple criterio basado en la magnitud de los coeficientes no garantiza la conveniencia de utilizar esa familia wavelet por lo que aunque se tiene en cuenta la magnitud de los coeficientes el criterio prevaleciente es el visual, es decir trabajar con la herramienta “wavelet 2D” del toolbox wavelet de matlab y ensayar todas las familias wavelet para ver cual resalta mas claramente los detalles horizontal, vertical y diagonal. Por esa razón se seleccionó la wavelet biortogonal 5.5 (bior 5.5) (figura 3.1c) ya que después de aplicar descomposición se aprecia que resalta en forma clara y definida los detalles horizontal, vertical y diagonal y se valida de alguna manera por el hecho de que la descomposición produce coeficientes de gran magnitud.

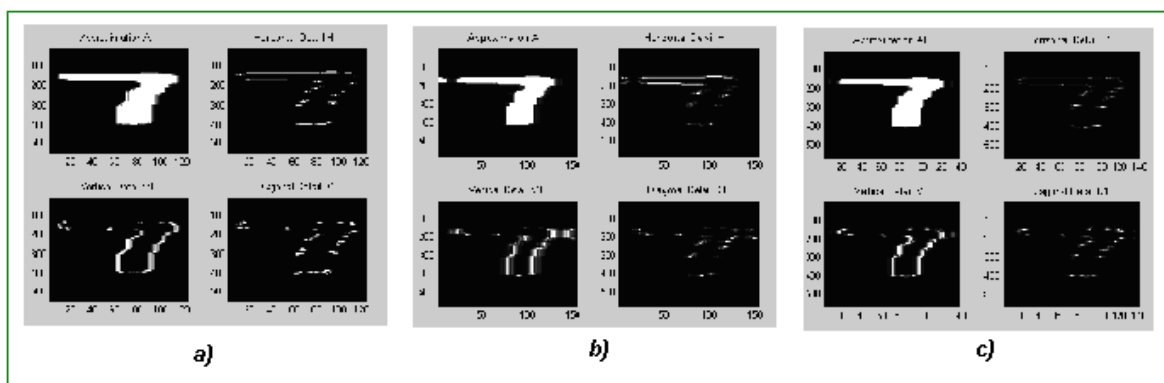


Figura 3.1 Descomposición wavelet de la imagen del dígito 7. a) rbio 3.1, b) db 9, c) bior 5.5

En la etapa de descripción (sección 3.2.4) se explica como se obtienen los descriptores a partir de la matriz de coeficientes de detalle.

3.2.2 Reconocimiento de texto

En el reconocimiento de texto se utilizan los coeficientes de aproximación de la descomposición wavelet. No se utilizan los coeficientes de detalle porque los caracteres de la imagen provienen de un editor de texto y, aunque sufren distorsión al momento de transmitir el fax o de escanear el texto, las formas se conservan y no hay mucha variación entre muestras de un mismo carácter, por lo que es suficiente trabajar con una sola matriz y, además, se agilizan los cálculos. En la figura 3.3 se muestra el resultado de la descomposición utilizando varias wavelet. Al realizar la descomposición wavelet con la Daubechie 1 (db1) la aproximación tiene forma cuadrículada y se descarta (figura 3.2a). Con otras wavelet se mejora la aproximación como con la Daubechie 7 (figura 3.2b) pero la imagen se distorsiona por reflejar la imagen en los extremos. Lo mismo se repite con otras wavelet como la Biortogonal 6.8 (figura 3.2c) y en gran medida la Coiflet 5 (figura 3.2d). Se escogió la Biortogonal inversa 2.2 (rbio 2.2) ya que la aproximación es muy satisfactoria (figura 3.2e)

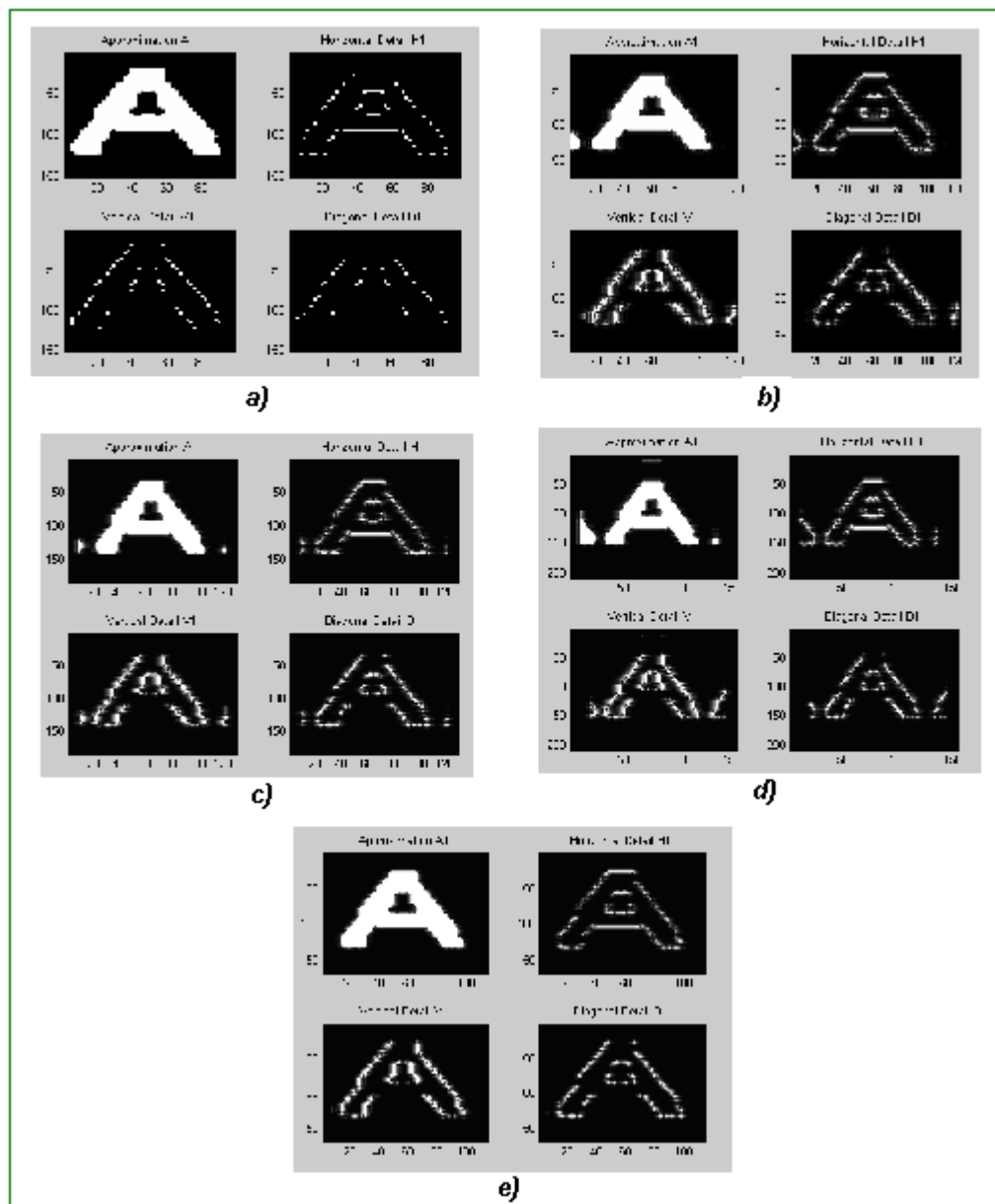


Figura 3.2 Descomposición wavelet a) db1, b) db7, c) bior 6.8, d) coif 5, e) rbio 2.2

En la etapa de descripción (sección 3.3.2) se explica como se obtienen los descriptores a partir de la matriz de coeficientes de aproximación.

3.2.3 Identificación de cultivos.

En la identificación de cultivos, por la naturaleza de las imágenes, se trabaja con coeficientes de detalle vertical ya que, en este caso, esos coeficientes realzan más la diferencia entre sectores de la imagen con diferentes texturas. Un ejemplo de esto se aprecia en la figura 3.4, donde se mira que a partir de los coeficientes de detalle vertical se aprecia mejor la zona que corresponde a cultivo de maíz, área encerrada en la figura 3.3.



Figura 3.3 Imagen con cultivo de maíz

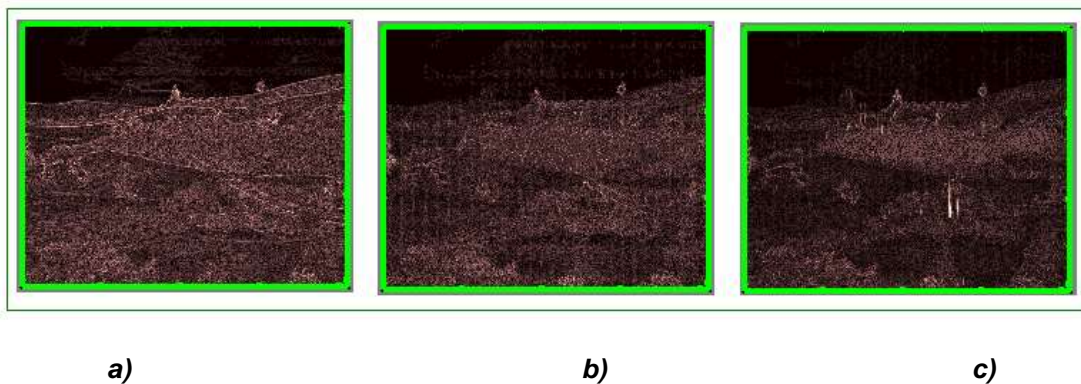


Figura 3.4 Coeficientes de detalle utilizando la wavelet rbio 3.1 a) horizontal, b) diagonal, c) vertical

En las fotografías de esta aplicación, además de la región de interés, se encuentran zonas con vegetación gruesa, como por ejemplo árboles, arbustos grandes o cerca del punto de donde se tomó la fotografía, que tienen rasgos fuertes en casi todas las direcciones y puede influir negativamente en la detección de las regiones buscadas. Es necesario, entonces, ubicar primero esas zonas indeseadas para eliminarlas antes de realizar la detección propiamente dicha. Con este propósito se utiliza la descomposición en segundo nivel ya que los rasgos de esa vegetación gruesa tienen una frecuencia mas baja y es necesario examinarlos con una escala mayor. Luego se analiza la imagen para identificar las regiones de interés, que en este caso son regiones de cultivo de maíz, y ese análisis se hace sobre los coeficientes de detalle vertical del primer nivel de descomposición ya que los detalles más pequeños o delicados presentes en estas regiones tienen frecuencia más alta y es necesario analizarlas con una escala menor para lograr una mayor resolución. La figura 3.20 muestra la detección de vegetación gruesa a partir de los coeficientes de detalle vertical de segundo nivel.

En la figura 3.5 se muestran los detalles verticales de primer nivel, resultado de la descomposición, usando varias familias wavelet. La idea es mirar cuál familia wavelet se ajusta más a los detalles de la imagen, es decir, con cuál familia wavelet son más altos los valores de la correlación entre las señales de la imagen y la wavelet seleccionada, haciendo mas notoria la diferencia entre superficies con diferente textura.

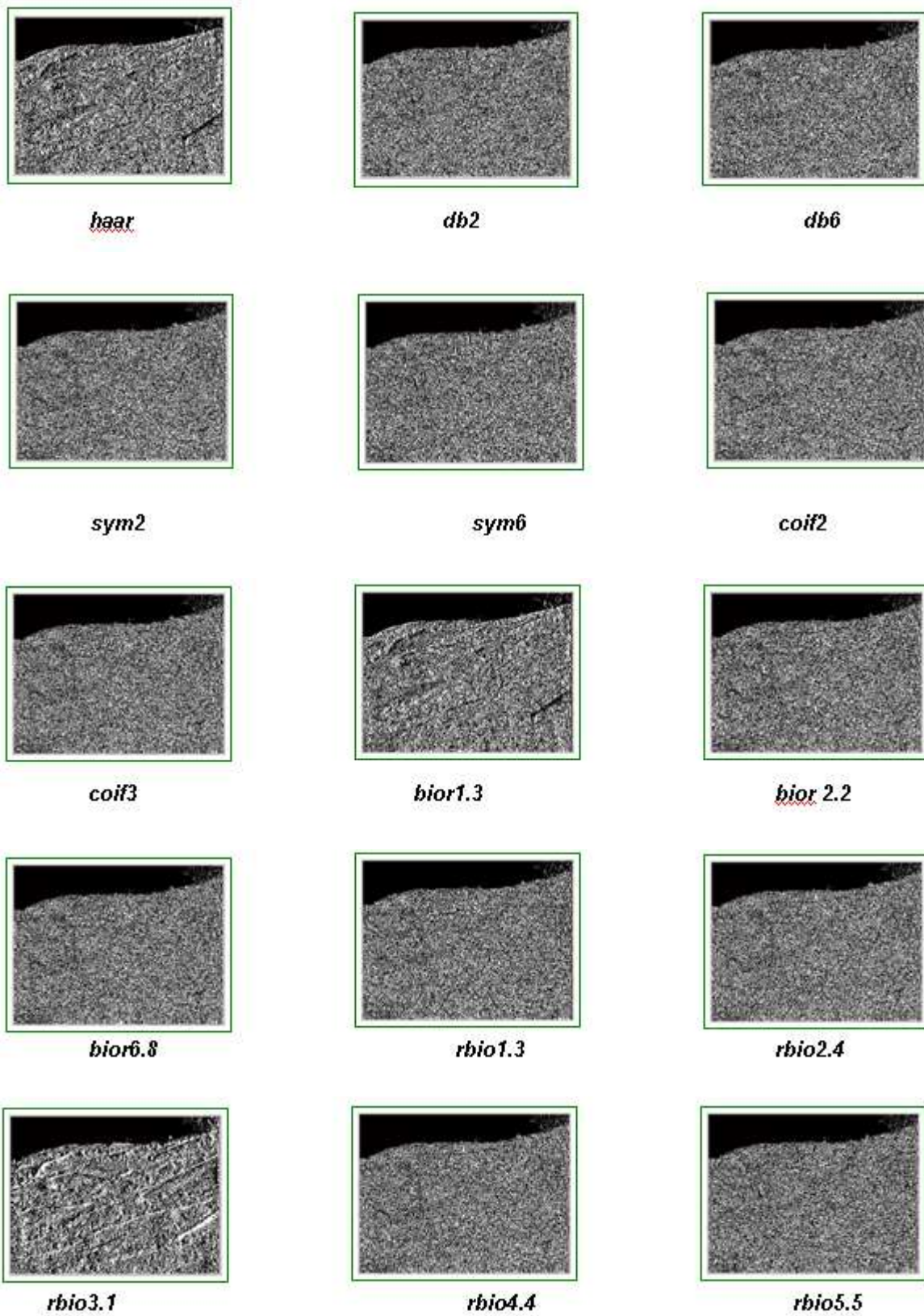


Figura 3.5 Coeficientes de detalle vertical de primer nivel con varias familias wavelet

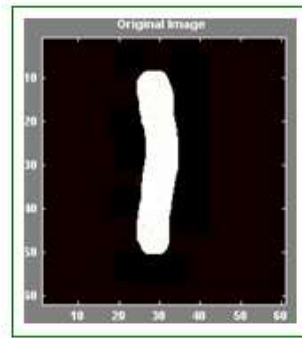
Las wavelet que mejor se ajustan a la imagen son: Haar, Biortogonal 1.3 y Biortogonal inversa 3.1. La que más se ajusta es la Biortogonal inversa 3.1 (rbio 3.1) y es esta familia la que se utiliza para la descomposición de la imagen y la identificación de las zonas buscadas.

En la sección 3.2.2 se explica como se realiza la identificación del cultivo de maíz a partir de los coeficientes de detalle vertical nivel 1, usando la wavelet biortogonal inversa 3.1 (rbio 3.1)

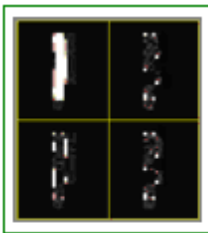
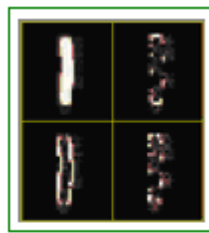
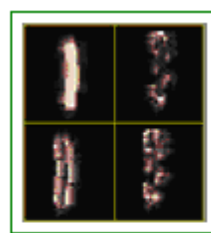
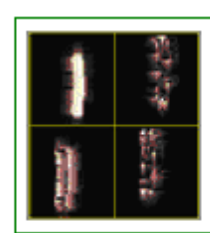
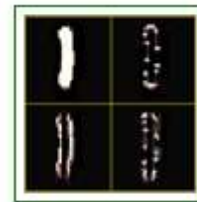
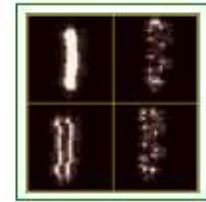
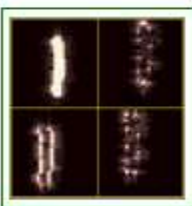
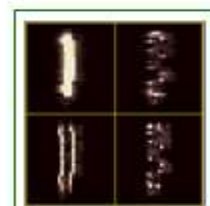
3.2.4 Detección de bacilos de tuberculosis

La detección del bacilo de tuberculosis (*Mycobacterium tuberculosis* ó bacilo de Koch), se hizo basándose en sus características morfológicas para lo cual se utilizó la descomposición wavelet ya que las características morfológicas de los bacilos se pueden describir tanto por las particularidades del contorno del bacilo como por las propiedades del objeto en si, como por ejemplo tamaño, grosor y longitud. Se utilizó la descomposición wavelet ya que por medio de los coeficientes de detalle (horizontal, vertical y diagonal) se pueden describir propiedades del contorno tales como cantidad de detalle horizontal, vertical y diagonal que tienen los bacilos y por medio de los coeficientes de aproximación se pueden describir propiedades como tamaño grosor y longitud.

Se ensayaron todas las familias wavelet provistas por el toolbox wavelet de matlab para escoger la que más se adaptara a la imagen. En la figura 3.6 se aprecia el resultado de la descomposición usando algunas de las familias wavelet que se probaron.



a)

*haar**db2**db5**db8**sym2**sym8**coif1**coif3**coif5**bior 1.1**bior 1.5**bior2.6*



b)

Figura 3.6 a) Imagen binaria de un bacilo b) Descomposición con varias familias wavelet

Debido a que los bacilos tienen poco tamaño, a pesar de utilizar interpolación bicúbica para redimensionarlos, las imágenes de la descomposición no son muy definidas y son algo difusas. Además cuando la wavelet no se ajusta muy bien a las características de la imagen, los valores de la correlación entre la imagen y la wavelet son pequeños y esto se nota en la imagen porque las figuras no son tan blancas y en cambio aparecen algo tenues o amarillentas como ocurre con la Daubechie 5 (db 5) y la Biorotogonal 2.6 (bior 2.6). La que mejor se ajusta a la

forma y tamaño de los bacilos es la Coiflet 1 (coif1), y es esta la que se utiliza para hacer la descomposición wavelet. En la sección 3.5.3 se explica como se realiza la detección de los bacilos a partir de los coeficientes wavelet tanto de aproximación como de detalle.

A continuación se detallará el algoritmo desarrollado en cada aplicación.

3.3 IDENTIFICACION DE PLACAS DE VEHICULOS

El objetivo de esta aplicación es poder reconocer la matrícula de un vehículo a partir de su fotografía como se indica en la figura 3.7. Para esto, se tomaron diversas fotografías a gran cantidad de vehículos y se procesaron siguiendo los pasos que se listan a continuación:

- Localización de la placa.
- Normalización
- Segmentación
- Descripción
- Clasificación



Figura 3.7 Fotografía de la placa de un vehículo

3.3.1 Localización de la placa

Como se mencionó en la sección 2.3 una imagen RGB se compone de tres matrices que representan la intensidad de los colores rojo (R), verde (G) y azul (B), respectivamente. El anterior concepto se utilizó en la localización de la placa ya que estas tienen un color específico (amarillas para vehículos particulares, blancas para vehículos públicos y azules para vehículos oficiales) claramente diferenciado en cada componente de color. El procedimiento es descomponer la imagen original en sus matrices R, G y B y así discriminar zonas que obedecen al color de la placa. Por ejemplo, al restar la componente azul de la fotografía de la imagen en blanco y negro, el color amarillo aparece como una zona totalmente oscura ya que el amarillo no tiene componente azul. Umbralizando y binarizando la imagen se obtienen las zonas correspondientes al color deseado. Luego se realiza una operación "opening" (primero erosión y después dilatación) sobre la imagen para encerrar el área de la placa. De esta manera es fácil distinguir en la imagen original las zonas que obedecen al color de la placa. Este método es muy eficiente pero tiene el inconveniente de detectar otras zonas del mismo color de la placa, por ejemplo cuando el vehículo es del mismo color que la placa, (aunque en realidad este es un caso muy poco frecuente) o cuando en la imagen hay zonas con componentes de color semejantes al de la placa. De todas maneras cuando hay más de una zona seleccionada se procede a analizar por medio de la técnica de componentes conectados, que se discutirá mas adelante, cual de todas las zonas seleccionadas contiene verdaderamente la placa. El análisis se lleva a cabo en tres etapas: Primero, evaluando las dimensiones de cada área para descartar aquellas en las que sus proporciones disten mucho de las proporciones de la

placa. Obviamente se deja un margen ya que las proporciones en las dimensiones de la placa pueden variar dependiendo de varios factores como por ejemplo el grado de inclinación con la que se obtuvo la fotografía. Después del primer descarte, se analiza la zona media de las regiones vigentes para cuantificar los cambios que ocurren entre blanco y negro y viceversa. Estos cambios obedecen al cambio entre los caracteres de la placa (negros) y el fondo (blanco). Si el número de estos cambios es menor o sobrepasa unos márgenes establecidos se descarta. Por último se evalúa el gradiente (calculado previamente) de la región o regiones vigentes, que representa los cambios no solo en la zona media sino en toda la región, terminando así la primera etapa del análisis.

Si después de este procedimiento se descartan todas las zonas seleccionadas, es decir, no se obtiene ninguna zona que contenga la placa, el sistema asume que se trata de una placa pública, es decir, de color blanco y se inicia un nuevo proceso encaminado a localizar placas de otro color. El algoritmo está diseñado así debido a que estadísticamente la gran mayoría de los carros son particulares, o sea, tienen placas de color amarillo. En segundo lugar están los vehículos públicos, que tienen placas blancas y solo un mínimo porcentaje pertenece a vehículos oficiales que tienen placas azules.

Cuando el vehículo es del mismo color que la placa se obtiene una zona muy grande que abarca gran parte del carro y la placa. En este caso, se binariza la región nuevamente pero con diferentes niveles de umbral para resaltar líneas delgadas como las del marco de la placa con el fin de obtener nuevas zonas

aisladas, las cuales se analizan nuevamente descartando las zonas indeseables y dejando finalmente la zona de la placa.

El procedimiento anterior arroja una imagen con un fondo negro donde solo aparece la placa del vehículo como se indica en la figura 3.8.



Figura 3.8 Localización de la placa.

3.3.2 Normalización de la imagen

3.3.2.1 Redimensionamiento

Una vez localizada la placa se crea una nueva imagen que contiene solo la región de interés, es decir, se elimina el fondo para dejar solamente la placa del vehículo. Esta nueva imagen, es normalmente muy pequeña por lo que se hace necesario redimensionarla. De esta manera se obtiene una imagen ampliada, que contiene únicamente la placa del carro, de dimensiones específicas independiente del tamaño de la placa en la imagen original, como se indica en la figura 3.9. Para redimensionar imágenes Matlab usa tres clases de interpolación dependiendo del grado de calidad requerido. En orden creciente de calidad estas son: “nearest”, “bilinear” y “bicubic”. Los procesos posteriores, especialmente la descomposición

wavelet requieren que la imagen analizada sea de una buena calidad por lo que se utilizó la interpolación bicúbica ('bicubic').



Figura 3.9 Área de la placa redimensionada.

3.3.2.2 Corrección de inclinación

La imagen resultante contiene la placa ampliada pero esta puede estar inclinada en alguno de sus ejes o en los dos dependiendo del ángulo con el que fue tomada la fotografía. Para corregir ese ángulo primero hay que conocerlo y la forma de averiguarlo es aplicando la transformada de Hough.

Antes de aplicar la transformada de Hough se deben detectar los bordes en la imagen para luego analizar cuales son las líneas más comunes que indicarán el grado de inclinación de la placa. Se evaluaron varios métodos de detección de bordes siendo el de Canny el que mejor se comportó y por ello es el que se utiliza. El método de Canny encuentra bordes buscando máximos locales del gradiente de la imagen de entrada. El gradiente es calculado usando la derivada de un filtro Gausiano. El método usa dos umbrales para detectar bordes fuertes y débiles, e incluye bordes débiles en la salida solamente si ellos están conectados a bordes

fuerzas. Este método es por lo tanto menos propenso que otros a ser engañados por ruido y más propenso a detectar bordes débiles verdaderos. [14]

La imagen resultante se muestra en la figura 3.10.



Figura 3.10 Detección de bordes con la máscara de canny.

Se implementaron dos rutinas para calcular, por medio de la transformada de Hough, los ángulos de inclinación horizontal y vertical por aparte. Una vez calculados los ángulos de inclinación se procede a corregirlos realizando una transformación espacial bidimensional¹ para obtener una imagen lista para la segmentación como se muestra en la figura 3.11.



Figura 3.11 Corrección de inclinación.

¹ Una transformación geométrica T es una función vectorial que asocia un pixel (x,y) a una nueva posición (x',y') . [15]

3.3.3 Segmentación

Una vez normalizada la imagen se utiliza nuevamente la técnica de componentes conectados para separar los caracteres de la imagen. Téngase en cuenta que para aplicar esta técnica, la imagen debe ser binaria, esto es, representada por una matriz de valores lógicos. La técnica de componentes conectados busca en la imagen zonas donde los píxeles están ligados entre si y a cada zona le asigna una etiqueta. La figura 3.12 ilustra este proceso.

De esta forma a cada carácter de la imagen de la placa se le asigna una etiqueta diferente para luego analizar una por una y, evaluando ciertos parámetros, comprobar si se trata de un caracter o es una zona indeseada. La evaluación consiste en examinar la zona correspondiente a cada etiqueta y chequear si las dimensiones corresponden a las de un caracter alfanumérico. Si el número de zonas es menor que seis y el ancho de alguna zona es muy grande es posible que se trate de dos caracteres pegados.

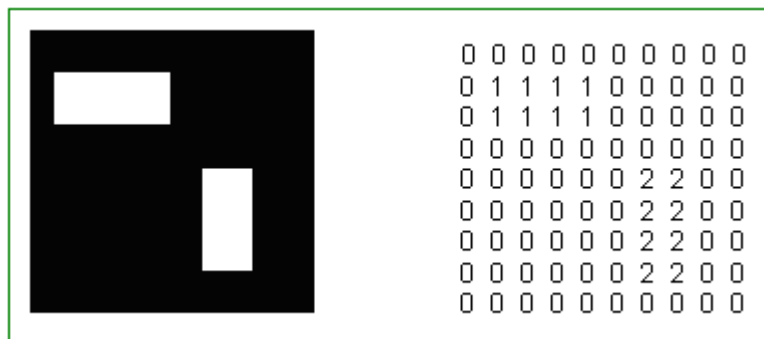


Figura 3.12 Componentes conectados.

Entonces se busca, en el área cuyas dimensiones sobrepasen un umbral, el punto mas débil que corresponderá al segmento donde están unidos los caracteres para

luego eliminarlo. De esta forma se logran separar caracteres ligados. Si la búsqueda es infructuosa se procede a eliminar paulatinamente las regiones más exteriores de la imagen con el fin de suprimir ruido causado por los bordes de la placa como el marco, tornillos, golpes, etc., que en ocasiones están en contacto con los caracteres y estropean la búsqueda. La búsqueda termina cuando se han encontrado las seis regiones que cumplen con los requisitos. Cuando esto sucede, se extrae de la imagen cada región para crear seis imágenes diferentes, cada una con un carácter alfanumérico. Estas imágenes tienen dimensiones pequeñas y no son óptimas para la siguiente fase que es la descomposición wavelet. Por consiguiente es necesario aumentar el tamaño de las imágenes por medio de una interpolación bicúbica², que es la que ofrece mejores resultados. Hecho esto, las imágenes están en condiciones para descomponerlas mediante la transformada wavelet.

3.3.4. Descripción

Como se vio en la sección 2.2.3, un patrón es una descripción estructural o cuantitativa de un objeto o de alguna otra entidad de interés en una imagen y en general, está conformado por uno o más descriptores [9].

En esta etapa de descripción se aplica a cada imagen, o sea a cada carácter alfanumérico, un proceso de extracción de características para obtener los descriptores, que conforman su respectivo patrón vectorial. El espacio de letras que aparecen en las placas contiene 26 elementos mientras que el espacio de números contiene diez. Lo anterior quiere decir que el grupo de letras tiene 26

² Basado en promediar teniendo en cuenta los valores de los 16 píxeles vecinos. [16]

clases diferentes y el grupo de números tiene 10, para un total de 36 clases caracterizadas individualmente por su respectivo patrón, entre las cuales hay que averiguar a cuál clase pertenece un carácter alfanumérico analizado. La forma de saberlo es extrayendo, del carácter desconocido, sus descriptores para luego de aplicar un clasificador, saber a cual clase pertenece. Los descriptores del carácter desconocido se extraen aplicando a cada imagen la transformada wavelet discreta. Como el grupo de letras tiene mas clases (26) que el grupo de números (10) y se sabe que los tres primeros caracteres son letras y los tres restantes son números, se crearon dos rutinas diferentes para descomponer por separado las letras y números. Esto optimiza el proceso ya que se evitan posibles confusiones que se podrían presentar al haber demasiadas clases en un solo grupo. Además, como el número de clases en el grupo de números es pequeño se diferencia más una clase de otra y se puede representar cada clase con menos descriptores haciendo más sencillo el proceso de extracción de características.

Como se dijo en la sección 2.1.5, la transformada wavelet descompone la imagen en aproximación, detalles horizontales, detalles verticales y detalles diagonales. Si se utiliza descomposición multinivel, en cada nivel de análisis se descompone la aproximación del nivel anterior para obtener una nueva aproximación y nuevos detalles horizontales, verticales y diagonales. En general, para cada aplicación que se desarrolle, cada wavelet se comportará diferente dependiendo de varios aspectos tales como calidad de la imagen, tamaño de la imagen, aspecto y tamaño de los objetos que se quieren analizar en la imagen etc. Esto hace que una wavelet que se comporte muy bien en una aplicación particular no tiene porqué hacerlo en otra.

Como se dijo en la sección 2.3.1 la wavelet que mejor se adapta a esta aplicación es la wavelet biortogonal 5.5 (en el primer nivel de descomposición) porque hace un buen realce de los detalles para esta clase de imágenes en particular.

Cada imagen de tamaño $M \times N$ que contiene el caracter alfanumérico se descompone, en el primer nivel, con la wavelet biortogonal 5.5 para obtener la matriz de coeficientes de aproximación, matriz de coeficientes de detalle horizontal, matriz de coeficientes de detalle vertical y matriz de coeficientes de detalle diagonal. Todas estas matrices son aproximadamente de tamaño $M/2 \times N/2$. Cada matriz de coeficientes de detalle horizontal, vertical y diagonal contiene una medida espacial de la magnitud de los cambios presentados en la imagen en las direcciones horizontal, vertical y diagonal respectivamente. La idea es saber como están distribuidos esos cambios en la imagen y para esto se divide cada imagen en tres sectores horizontales como se muestra en la figura 3.13

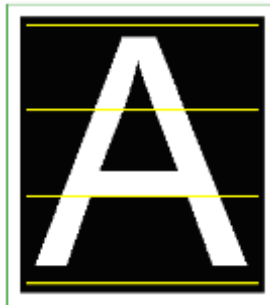


Figura 3.13 Sectorización de la imagen.

Pero como la idea es conocer, de cada imagen, la distribución de los detalles horizontales, verticales y diagonales por separado, el procedimiento a seguir es sectorizar cada matriz de coeficientes de detalle en tres sectores horizontales

como se indica en la figura 3.14 y luego calcular la suma de los valores absolutos de los coeficientes de detalle en cada sector. Se toman los valores absolutos ya que las matrices de detalles contienen tanto valores positivos como negativos y nuestro interés, en este caso, se centra en la magnitud de los coeficientes y no en su signo.

2.36	15.24	-3.10	1.20
2.10	17.42	-2.21	0.87
3.45	5.12	4.58	1.56
4.21	4.97	5.12	2.07
-22.36	18.21	35.11	28.95
-24.60	19.24	28.54	30.04

Figura 3.14 Sectorización de la matriz.

Al hacer esto con las matrices de coeficientes de detalle horizontal, vertical y diagonal, se obtienen para el caso de las letras 9 descriptores, tres por cada matriz de detalle. Para el caso de los números, como se trata de una población más pequeña, se trabaja con las matrices de coeficientes sin sectorizarlas, es decir, con la suma total de los valores absolutos de los coeficientes de cada matriz, por lo que se tienen 4 descriptores correspondientes a los coeficientes de aproximación, detalle horizontal, detalle vertical y detalle diagonal. A pesar de que cada imagen pasó por una etapa de normalización, por fenómenos como inclinación de la placa, desgaste de pintura de la placa, deformidad de la placa etc. se presenta alguna variación del tamaño o grosor del caracter final, que aunque es mínima, afectan los cálculos y por esta razón se optó por medir la contribución de cada sector de la correspondiente matriz en forma porcentual. Es decir, se mide la

contribución en valor absoluto de cada sector con relación a la magnitud de la suma de los valores absolutos de toda la matriz. O sea, en cada caracter se mide que tantos detalles horizontales, verticales y diagonales tiene cada sector en relación con el total de detalles horizontales, verticales y diagonales respectivamente, obteniendo, a partir de los coeficientes wavelet de detalle, nueve descriptores que caracterizan cada caracter.

3.3.5 Clasificación

Cada clase es representada por un punto en el espacio n -dimensional, donde n es la cantidad de descriptores. Una vez caracterizado el caracter se procede a identificar a que clase pertenece, para lo cual se utilizó el clasificador de mínima distancia (descrito en el ANEXO B sección 1.1.1), que consiste en calcular la distancia euclídea entre el caracter de la imagen de prueba y los caracteres de entrenamiento, pero en vez de tomar todas las muestras de la población de entrenamiento, se toman los valores promedio de las mismas. Es decir, se procesó gran cantidad de fotografías de placas de vehículos para obtener la media de los valores de los descriptores (nueve para el caso de las letras y cuatro para el caso de los números) que caracterizan los caracteres. Esos valores se almacenan en una matriz de tamaño $R \times S$ donde R es igual al número de descriptores (9 para letras y 4 para números) y S es la cantidad de clases, es decir, 26 para la población de las letras y 10 para la población de los números. De esta forma, cuando se extraen los descriptores de un nuevo símbolo, o sea, su patrón, si se trata de los tres primeros caracteres de la placa, se calcula la distancia euclídea entre ese punto y cada uno de los 26 puntos del espacio de

clases de letras y si se trata de los tres caracteres restantes se calcula la distancia euclídea entre ese punto y cada uno de los 10 puntos del espacio de clases de los números. En ambos casos, se identifica el punto al cual hay la menor distancia y ese punto indicará la clase a la cual pertenece el caracter en cuestión. La figura 3.15 es una representación del cálculo de la distancia euclídea entre un punto de prueba pp y cuatro clases (A, H, U, Z) salvo que el espacio real es de 9 dimensiones para el caso de las letras y de 4 dimensiones para el caso de los números, pero por la imposibilidad física de representar estos espacios se utiliza un espacio tridimensional con fines ilustrativos. En la figura, la menor distancia está entre el patrón desconocido pp y el patrón promedio de la clase Z, por lo que el patrón desconocido es asignado a la clase Z.

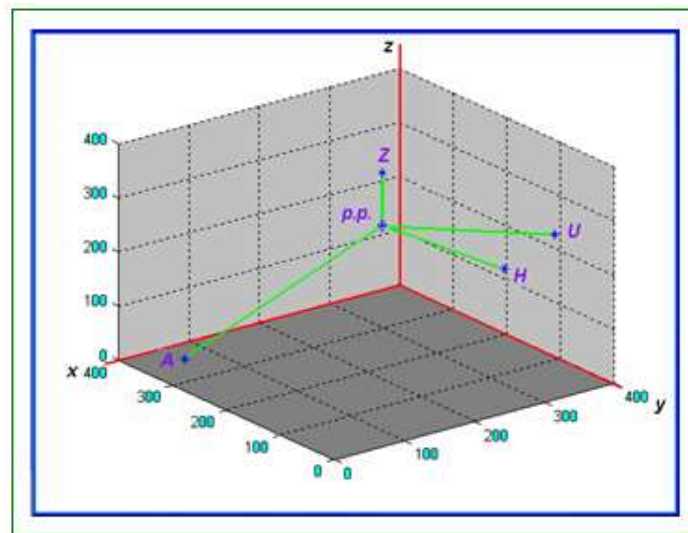


Figura 3.15 Cálculo de la distancia euclídea

De esta manera se obtienen los seis caracteres (3 letras y 3 números) que conforman la placa de los vehículos colombianos. Este dato se almacena en un vector y se despliega como se muestra en la figura 3.16.



Figura 3.16 Interfase mostrando el resultado de la identificación

Para evaluar el rendimiento de esta aplicación se tomaron, al azar, 40 muestras (40 fotografías) y la identificación no fue exitosa en 4 oportunidades, por lo tanto la eficiencia estimada es de 90%. El tiempo promedio que toman las etapas de descomposición wavelet, extracción de características y reconocimiento es de 1.36 segundos.

3.4 IDENTIFICACIÓN DE TEXTO PLANO

Esta aplicación reconoce texto en una imagen, obtenida por ejemplo, de un fax o un scanner, como la mostrada en la figura 3.17. Esto se hace siguiendo las siguientes etapas:

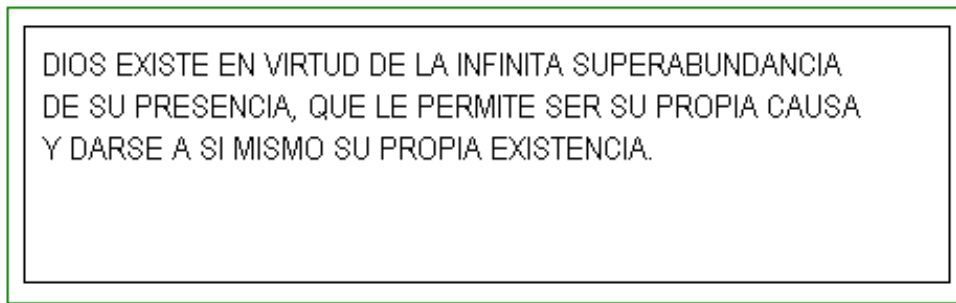


Figura 3.17 Imagen que contiene texto.

3.4.1 Localización de renglones, palabras y letras

Después de leer la imagen se binariza de tal modo que la imagen se representa por dos valores, uno lógico para el fondo (blanco) y cero lógico para las letras (negras). Luego se invierte la imagen para que el fondo quede en negro y las letras en blanco (lo cual es necesario para aplicar más adelante la técnica de componentes conectados). A la imagen se aplica una operación de “opening”, utilizando un elemento estructurante mayor que los espacios entre palabras, con el fin de encerrar las palabras en bloques de tal modo que los renglones aparecen como bloques blancos sobre un fondo negro como se indica en la figura 3.18



Figura 3.18 Renglones detectados.

Se aplica entonces la técnica de componentes conectados para ubicar y enumerar los renglones para luego procesar renglón por renglón. Teniendo la ubicación de cada renglón se toma esa porción de la imagen original (figura 3.19) y se aplica un nuevo “opening” pero con un elemento estructurante más pequeño con el fin de

encerrar en bloques cada palabra para obtener una subimagen del tamaño del renglón de fondo negro con bloques blancos que representan las palabras en cada renglón como se muestra en la figura 3.20

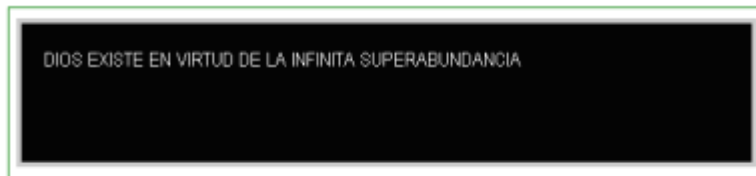


Figura 3.19 Primer renglón identificado.



Figura 3.20 Identificación de palabras en el primer renglón.

Con la técnica de componentes conectados se ubican y enumeran las palabras de cada renglón para procesarlas una por una. Ahora se toma bloque por bloque, o sea, palabra por palabra de la imagen original y se aplica nuevamente un nuevo “opening” con un elemento estructurante, aún más pequeño, encaminado a diferenciar las letras para obtener subimágenes del tamaño de cada palabra con fondo negro y con las letras en blanco como lo muestra la figura 3.21



Figura 3.21 Identificación de letras por palabra.

Nuevamente mediante componentes conectados se enumeran y ubican las letras de cada palabra para luego identificarlas.

De esta forma se va procesando renglón por renglón, palabra por palabra, letra por letra.

3.4.2 Descripción

Una vez obtenidas las imágenes que contienen cada letra se redimensionan para aumentarlas de tamaño y posteriormente se aplica la transformada wavelet para obtener los coeficientes de aproximación y los coeficientes de detalles horizontal, vertical y diagonal, caracterizando cada imagen. Como se dijo en la sesión 2.3.2 la wavelet que mejor se adapta a este tipo de imágenes es la wavelet biortogonal inversa 2.2 (rbio2.2). La idea es analizar cómo está distribuida espacialmente la energía de cada letra por lo que en este caso se trabajará solo con los coeficientes de aproximación, ya que como se dijo en la sección 2 los coeficientes de aproximación contienen la mayor parte de la energía de la imagen en tanto que los coeficientes de detalle tienen valores pequeños. El procedimiento utilizado para obtener las características o los descriptores que conforman el patrón de cada letra a partir de la matriz de coeficientes de aproximación es el siguiente: Dividir la imagen en tres sectores horizontales y tres sectores verticales como se muestra en la figura 3.22

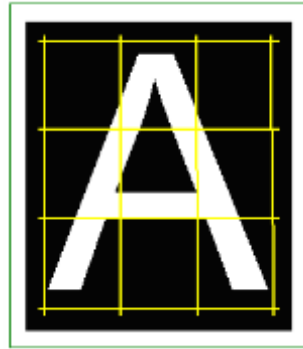


Figura 3.22 Sectorización de la imagen.

Como el interés se centra en los coeficientes de aproximación, una vez obtenida la matriz de coeficientes de aproximación se divide la matriz en tres sectores horizontales y tres verticales, es decir nueve sectores, como se muestra en la figura 3.23 obteniendo entonces nueve descriptores que consisten, cada uno, en la suma de los valores absolutos de los coeficientes wavelet de aproximación de cada sector.

0.00	0.00	30.36	36.72	1.20	0.00
0.01	3.23	29.32	31.52	4.21	0.00
0.00	4.25	28.63	32.54	4.87	0.00
0.95	5.32	26.09	28.67	6.02	0.08
5.86	8.92	14.52	15.08	9.00	6.87
25.63	20.03	18.75	20.56	22.38	24.06

Figura 3.23 Sectorización de la matriz de aproximación.

Estos nueve valores se ubican en un punto en el espacio de dimensión 9, donde cada eje representa una de las nueve características del patrón de cada letra.

3.4.3 Clasificación

Previamente se han almacenado los patrones de cada una de las letras del alfabeto en una matriz que se utiliza para calcular la distancia euclídea más corta entre un punto desconocido, localizado en el espacio de características (non-dimensional), y los puntos que representan cada letra conocida como se ilustró en la figura 3.15, es decir se aplica el método de mínima distancia (Descrito en el ANEXO B sección 1.1.1). De esta forma se reconoce a qué letra pertenece el patrón de una imagen desconocida. Como se dijo inicialmente, cuando se obtiene el segmento de la imagen original que contiene cada letra, se redimensiona la imagen para dejarla de un tamaño estándar, lo que proporciona cierta independencia del tamaño de la fuente.

Una vez reconocida cada letra se la concatena en un vector que contiene las letras reconocidas anteriormente, por lo que el vector al final contendrá el texto de toda la imagen con cada letra en cada posición del vector. Cada vez que se procesa una palabra se coloca un espacio libre en la siguiente casilla del vector lo que permite separar las palabras a la hora de desplegar el contenido del vector.

Para evaluar el rendimiento de esta aplicación se procesó la imagen de un texto con 210 caracteres y se presentaron identificaciones erróneas en tres caracteres, por lo tanto, la eficiencia estimada es de 98.6 %. El tiempo promedio que toman las etapas de descomposición wavelet, extracción de características y reconocimiento es de 0.24 segundos.

3.5 IDENTIFICACIÓN DE UN CULTIVO

Con esta aplicación se logra identificar, a partir de una imagen, un cultivo específico ubicado en medio de otra clase de vegetación. Esto se logra analizando las texturas de las diferentes regiones de la imagen y filtrando, a partir de los coeficientes wavelet, la región de interés.

La idea principal es estudiar las imágenes de una clase específica de vegetación para obtener su patrón característico, con el fin de determinar si ese patrón se encuentra en una imagen dada. Para este fin se analizaron, con las familias wavelet del toolbox wavelet de matlab, varias fotografías donde se encontraba una plantación de maíz. Como se mencionó en la sección 2.3.3, la función wavelet seleccionada fue la wavelet biortogonal inversa 3.1 (rbio3.1) por sus mejores resultados para esta aplicación en particular. El procedimiento seguido para tal fin es el siguiente:

3.5.1 Pre-procesamiento de la imagen

Una vez obtenida la imagen original (RGB) se la transforma en imagen en tonos de gris ya que una imagen RGB está compuesta por tres matrices como se mencionó anteriormente, en cambio, una imagen en tonos de gris tiene una sola matriz que contiene los niveles de intensidad de gris de cada píxel. A esta imagen se le aplica la transformada wavelet para obtener los coeficientes de aproximación y los de detalle. Con la wavelet biortogonal inversa 3.1 (rbio3.1), se hizo una descomposición en nivel 2 ya que en la imagen se presenta a veces un tipo de vegetación gruesa (ya sea por la presencia de vegetación grande como árboles o por la proximidad de alguna otra vegetación en el momento de tomar la fotografía),

cuyos detalles no están definidos en una dirección particular sino que tiene componentes de gran magnitud en casi todas las direcciones y puede interferir con la selección de las zonas que contienen el cultivo de interés. La figura 3.24 representa las zonas muy irregulares, cuyos coeficientes son de gran magnitud, en color amarillo (esta imagen está en falso color y se logró fijando el colormap en 1-jet). Estas son las zonas que hay que eliminar primero antes de entrar a identificar las regiones de interés.

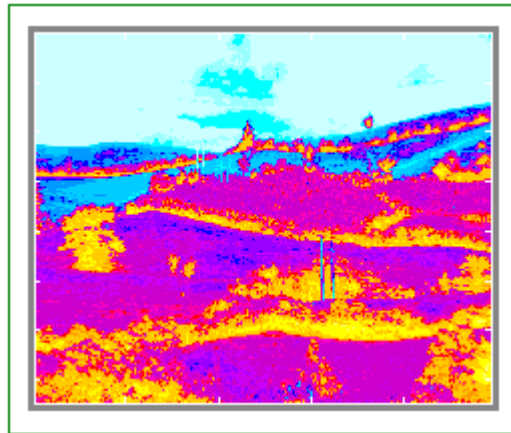


Figura 3.24 Imagen en falso color. Las zonas amarillas representan zonas muy irregulares con coeficientes de gran magnitud.

Es necesario, entonces, eliminar esas zonas indeseadas para evitar ambigüedades y esto se hace detectando esas zonas de vegetación gruesa y descartándolas cuando se haga la selección de las zonas de interés. Como esa vegetación es gruesa, tiene rasgos geométricos grandes (baja frecuencia) y son precisamente esos rasgos los que servirán para identificarla y luego eliminarla. El análisis wavelet permite el uso de zonas grandes cuando se quiere información de baja frecuencia y zonas más pequeñas donde se quiere información de alta frecuencia [7]. Esta es la ventaja de hacer análisis multirresolución ya que nos

permite analizar las imágenes en varias escalas en un solo paso como se muestra en la figura 3.25

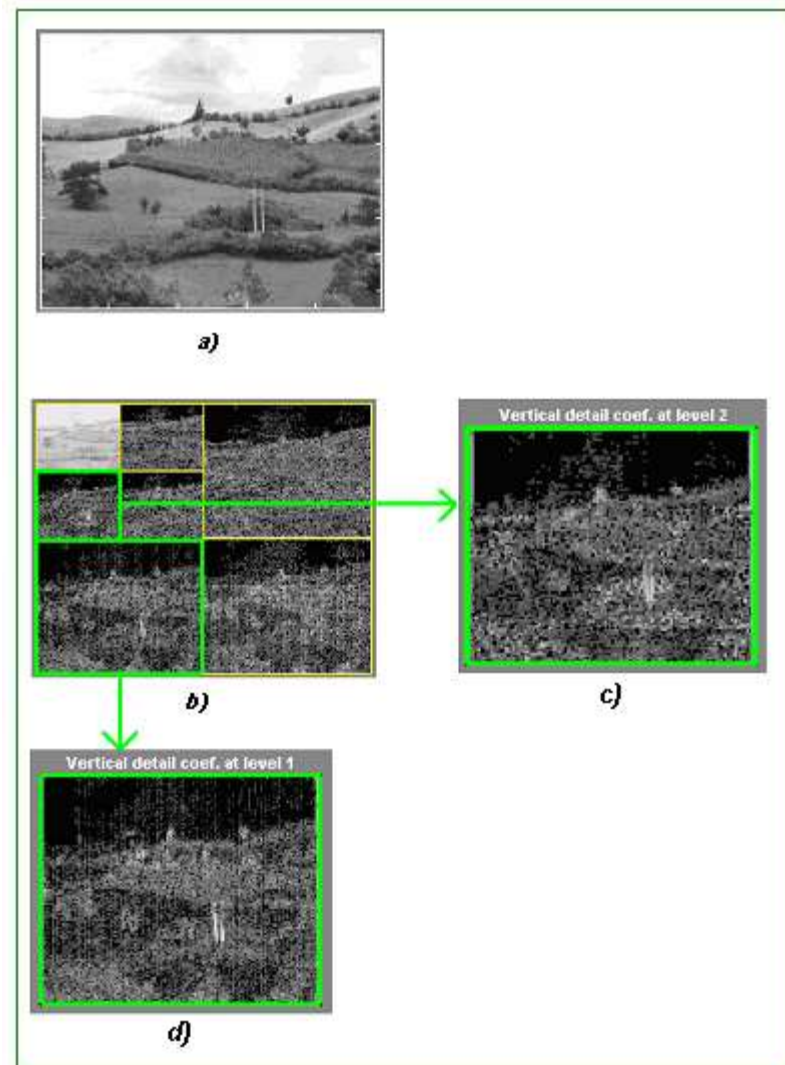


Figura 3.25 a) Imagen original, b) Descomposición nivel 2 con la wavelet bior 3.1 c) Imagen de detalle vertical nivel 2 d) imagen de detalle vertical nivel 1.

Como la vegetación gruesa se identifica mejor en un nivel mayor de descomposición, se utiliza el segundo nivel de descomposición para identificar la vegetación gruesa y el primer nivel para identificar el cultivo deseado. En la figura

3.26a se mira el resultado de la detección de la vegetación gruesa por la magnitud de los coeficientes de detalle vertical nivel 2 y en la 3.26b la imagen anterior después de aplicar la operación “closing”.

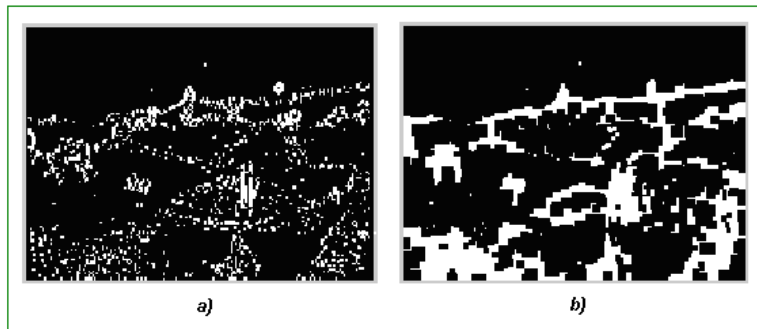


Figura 3.26 a) Detección de vegetación gruesa a partir de coef. de detalle vertical segundo nivel b) Imagen con “closing”

Una vez eliminadas las zonas indeseables (vegetación gruesa), la imagen está lista para buscar, en ella, el patrón de las zonas deseadas.

3.5.2 Identificación del patrón

Una imagen está compuesta por objetos, cada uno con características o patrones diferentes. Esos patrones son los que los diferencian entre si y permiten localizarlos y procesarlos separadamente. En este caso un objeto es una zona de un cultivo determinado que se diferencia del resto de la imagen. Particularmente el cultivo de maíz en las imágenes presenta rasgos sobresalientes en la matriz de coeficientes de detalle vertical del primer nivel, utilizando la wavelet biortogonal 3.1 por lo que se utiliza esta matriz para identificar el patrón perteneciente a tal cultivo. Como se dijo en la sección 3.2.1 cuando una wavelet se adapta bien a las características de los objetos de una imagen, visualmente se aprecian líneas de

detalle definidas e intensas y sus coeficientes son de gran magnitud. Por esta razón se utiliza la magnitud de los coeficientes para identificar las zonas que pertenecen al cultivo de maíz ya que la magnitud de los coeficientes de detalle vertical correspondientes a esas zonas donde las líneas de detalle son más intensas es mayor que en las demás zonas. El siguiente paso es comparar los coeficientes de las zonas correspondientes el cultivo de maíz con los demás coeficientes y establecer el umbral para el cual la magnitud de los coeficientes de las zonas de cultivo de maíz es mayor que el resto. De la misma manera se establece un umbral superior, es decir, un valor para el cual la magnitud de los coeficientes de las zonas de cultivo de maíz es menor que el resto. Esto se aplica para eliminar zonas como por ejemplo postes colocados artificialmente donde los coeficientes de detalle vertical tienen mayor magnitud que los coeficientes de las zonas de maíz. Los umbrales superior e inferior han sido establecidos a criterio de los desarrolladores de este proyecto fruto de la comparación de los coeficientes de las matrices de detalle vertical después de descomponer varias imágenes. Cuando se esta procesando una imagen, al correr la aplicación, se comparan todos los coeficientes de la matriz de detalle vertical para evaluar cuáles están entre los umbrales superior e inferior para dejarlos y descartar los que estén por fuera de esta rango. En la figura 3.27 se ilustra este proceso.

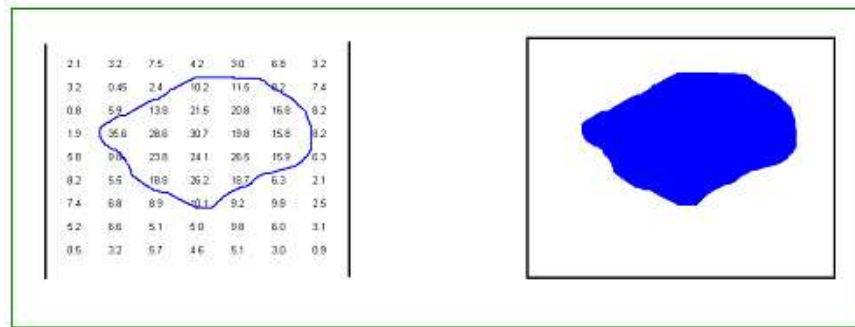


Figura 3.27 Selección áreas a partir de los coeficientes wavelet de detalle vertical

Después de hacer esta selección hay que hacer una etapa de “closing” para unificar las zonas blancas y luego “opening” para eliminar ruido. En la figura 3.28 se mira el resultado.

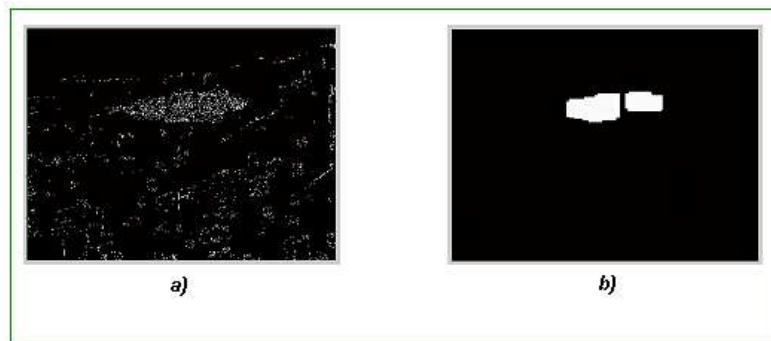


Figura 3.28 a) Identificación de la zona deseada a partir de los coeficientes de detalle vertical nivel 1 con wavelet biortogonal inversa 3.1 (rbio 3.1) b) Imagen anterior después de closing y opening

De esta manera se logra identificar el cultivo de maíz a partir de los coeficientes wavelet de detalle vertical, nivel 1. El resultado se muestra en la figura 3.29

Así mismo se pueden crear aplicaciones que reconozcan, en general, texturas diferentes en una imagen, como por ejemplo diferenciar bosque, concreto, superficies de agua, etc., todo en una sola imagen, a partir del análisis de sus coeficientes.



Figura 3.29 Detección de cultivo de maíz.

Para evaluar el rendimiento de esta aplicación se analizaron, al azar, 6 fotografías que contenían cultivo de maíz y en las 6 fotografías el reconocimiento en forma general fue exitoso, aunque algunas pequeñas regiones, sobre todo en los bordes del cultivo, se pasan por alto. La cantidad de píxeles que no se resaltan, tratándose efectivamente de una zona con cultivo de maíz, es aproximadamente el 10% del área total que debería resaltarse. Por lo tanto se estimó una eficiencia aproximada del 90%. El tiempo promedio que toman las etapas de descomposición wavelet, eliminación de vegetación gruesa y reconocimiento del cultivo de maíz es 3.78 segundos.

3.6 DETECCIÓN DE TUBERCULOSIS

Esta es una aplicación interesante ya que analizando una fotografía del esputo de una persona, como la indicada en la figura 3.30, identifica si hay infección de tuberculosis marcada por la presencia del bacilo *Mycobacterium tuberculosis*.

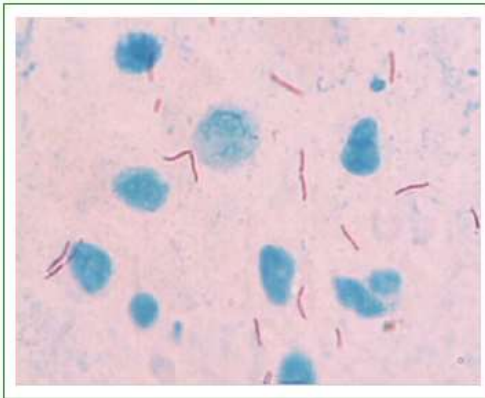


Figura 3.30 Fotografía de muestra de esputo.

3.6.1 Procedimiento inicial

El procedimiento para detectar la tuberculosis consiste en obtener muestras de esputo de los pacientes y luego aplicarles un colorante que resalta la presencia del bacilo *Mycobacterium tuberculosis* y luego se analiza esta imagen para determinar el grado de infección del paciente. El procedimiento aplicado aquí es el mismo solo que se hace automáticamente. La imagen RGB, que es una fotografía de la muestra de esputo de una persona, se la descompone en sus tres componentes (rojo, verde y azul) ya que el color es muy importante en la detección del bacilo. Particularmente en la componente azul, el bacilo sobresale más que en las otras dos componentes. Las tres componentes de una imagen RGB se muestran en la figura 3.31.

Se encontró cierta dificultad debido a que los resultados eran dependientes de la cantidad y calidad de colorante utilizado por el laboratorio. Además, en la muestra de esputo hay presencia de leucocitos y en ocasiones parte de esos leucocitos presentan propiedades de color semejantes al bacilo.

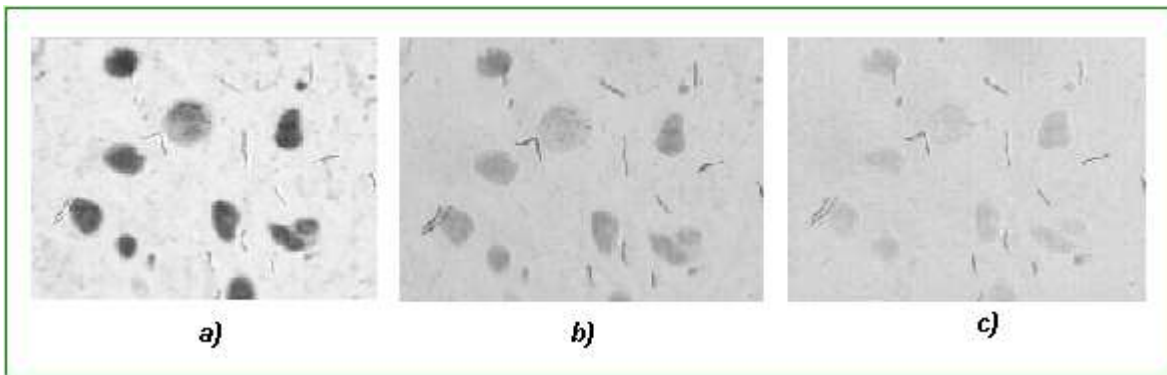


Figura 3.31 Componentes de una imagen RGB. a) roja, b) verde, c) azul.

Para solventar estos inconvenientes se implementaron dos soluciones para aislar tanto lo relacionado con el colorante como la ambigüedad presentada con los leucocitos. Este mecanismo consta de dos pasos que son:

1. Determinar la cantidad de colorante presente en la imagen para tomarlo como referencia a la hora de identificar el bacilo.
2. Localizar los leucocitos y variar la intensidad de color de tal forma que no interfiera en la detección del bacilo.

Los pasos, entonces, para detectar la presencia de tuberculosis son los siguientes:

3.6.2 Eliminación de zonas ambiguas

El objetivo de esta etapa es identificar la presencia de leucocitos en la imagen para lo cual se analiza la componente roja de la imagen, ya que en esta componente sobresalen y es fácil localizarlos. Una vez localizados se aumenta un poco el valor de esos píxeles logrando que aparezcan menos pigmentados, como se aprecia en la figura 3.32, evitando que algunas zonas de los leucocitos se confundan con los bacilos.

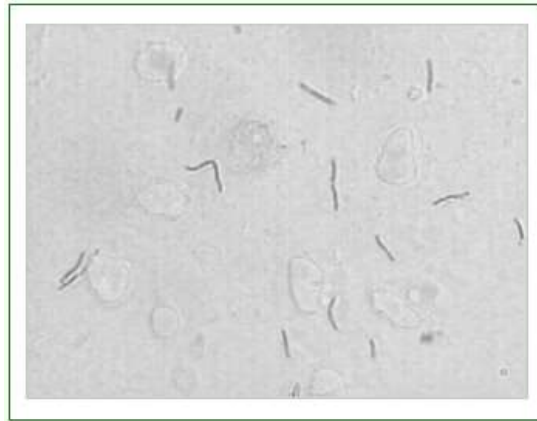


Figura 3.32 *Componente azul modificada.*

3.6.3 Identificación de los bacilos *Mycobacterium tuberculosis*

Los bacilos se resaltan más en la componente azul de la imagen por lo que se los localiza a partir de esta. En esta etapa cobra importancia el haber determinado anteriormente el nivel de pigmentación de la imagen tras aplicar colorante. Con este valor conocido se fija un valor de umbral por debajo del cual se encuentran los píxeles correspondientes a los posibles bacilos. Posteriormente se entra a discriminar entre las zonas seleccionadas cuáles pertenecen en realidad al bacilo *Mycobacterium tuberculosis*. Esta discriminación consiste en evaluar las características morfológicas del área seleccionada y aplicar la transformada wavelet para analizar sus coeficientes y descartar las áreas que no corresponden a las particularidades de los bacilos. El análisis discriminatorio es el siguiente:

Se toma la imagen y se forman nuevas imágenes que contienen cada región seleccionada y para cada una se calculan los siguientes parámetros:

- Alto y ancho de la sección de la imagen que contiene la figura en cuestión.
- Alto y ancho de la figura misma.

- Relación entre las dimensiones de la figura.
- Número de cambios que ocurren entre blanco y negro y viceversa al recorrer cada imagen horizontal y verticalmente por su zona media (esto indica si hay mas de una figura en la imagen, por ejemplo en el caso de que hayan dos bacilos unidos en algún punto).

Posteriormente se le aplica a cada imagen la transformada wavelet para estudiar los coeficientes de aproximación y de detalle de cada una. Como se dijo en la sección 2.3.4, la wavelet que mejor se adaptó a este tipo de imágenes es la coiflet1 y se hace descomposición en el primer nivel. Con estos coeficientes se calculan los siguientes parámetros:

- Suma de valores absolutos de coeficientes de aproximación. Este parámetro nos da idea del tamaño de la figura seleccionada.
- Suma de los valores absolutos de coeficientes de detalle horizontal, vertical y diagonal.
- Relación entre los dos valores anteriores.

En el caso de los bacilos, los valores de los parámetros anteriores están comprendidos entre ciertos rangos específicos. Por esa razón cuando se están analizando los elementos que componen la imagen se evalúan los valores de los parámetros anteriores uno a uno y si al final todos esos valores están comprendidos en unos rangos predefinidos, se asume que el objeto analizado es un bacilo. Este proceso de selección se logra implementando funciones de decisión que evalúan si los valores de los parámetros mencionados arriba están

entre los rangos validos ó no y cuando una condición ó más no se cumple, se asume que se trata de un objeto extraño y se descarta. Finalmente se obtiene una imagen, únicamente, con los bacilos detectados.

Una vez detectados los bacilos se crea una imagen binaria con fondo negro, y en blanco las regiones que corresponden a la ubicación de los bacilos, como se muestra en la figura 3.33.



Figura 3.33 Detección de bacilos *Mycobacterium tuberculosis*.

Adicionalmente se hace una aproximación del grado de infección, cuantificando el área blanca que corresponde a la presencia de bacilos y se da este valor en términos porcentuales en relación al área total de la imagen. Es decir, se dice que porcentaje del área total esta ocupada por los bacilos. En la imagen de la figura 3.33, por ejemplo, se encontró presencia del bacilo en un 0.8280 por ciento.

Para evaluar el rendimiento de esta aplicación se analizaron, al azar, 5 placas y se encontró que, en promedio, de 11 bacilos presentes en la imagen original 1 bacilo se detecta erradamente (no se detecta o se reconoce como bacilo un objeto distinto). Por lo tanto la eficiencia estimada es mayor a 90%. El tiempo promedio

que toman las etapas de descomposición wavelet, extracción de características y reconocimiento es de 0.27 segundos

Cabe anotar que para esta aplicación se implementó otro algoritmo, utilizando otra técnica diferente, con el fin de comparar el rendimiento. La otra técnica utilizada fue el de correspondencia por correlación, es decir, utilizar una subimagen o plantilla, que contenga el objeto que se desea encontrar, para recorrer la imagen y calcular la correlación entre la plantilla y la sección de la imagen correspondiente. La técnica de correspondencia por correlación se explica en el Anexo B sección 2.2. En cada paso, del recorrido de la plantilla, se calcula el valor de la correlación y el máximo valor corresponderá a la región de la imagen que contenga el objeto buscado, en este caso el bacilo de Koch.

La correspondencia por correlación no solo identifica el bacilo sino que también da la localización del bacilo pero, en este caso la ubicación de los bacilos no es relevante porque solo importa si hay presencia del bacilo sin importar dónde este se encuentre y en cambio esta información adicional implica mayor costo computacional y por ende mayor tiempo de ejecución. El algoritmo implementado solo considera una posición del bacilo y sin embargo el tiempo de ejecución es muy elevado. Para mejorar esto se puede aumentar el tamaño de los saltos de la plantilla pero, aunque se gana en velocidad se pierde eficiencia porque la identificación de los objetos de interés no va a ser tan precisa. Además para que el algoritmo sea eficiente se deben tener en cuenta todas las posiciones en las que

se encuentren los bacilos y esto haría al algoritmo menos eficiente porque aumentaría mucho más el tiempo de ejecución.

4. CONCLUSIONES Y RECOMENDACIONES

En este capítulo se dan las consideraciones finales del proyecto indicando algunas generalidades de la transformada wavelet, el reconocimiento de patrones utilizando la transformada wavelet y las aplicaciones desarrolladas aquí.

4.1 CONCLUSIONES

- Una de las principales ventajas de las wavelets frente a los métodos clásicos, como la transformada de Fourier, es que en el segundo caso se maneja una base de funciones bien localizada en frecuencia pero no en tiempo, mientras que la mayoría de las wavelets interesantes presentan una buena localización en tiempo y en frecuencia. En el caso de las imágenes esto es importante ya que se presentan zonas tanto suaves y homogéneas como zonas con cambios abruptos, especialmente en los bordes de los objetos que la componen y por esta razón las wavelets y el análisis de multirresolución constituyen una potente herramienta para afrontar problemas fundamentales en el tratamiento de imágenes.
- El análisis multirresolución permite estudiar una imagen en distintas resoluciones y extraer características propias de cada nivel de resolución, dependiendo del tipo de problema sobre el que se esté trabajando, para describir y reconocer los objetos de interés en una imagen.

- En la identificación de placas de vehículos y la identificación de texto en una imagen, la transformada wavelet fue una herramienta eficiente para obtener los descriptores de los caracteres alfanuméricos permitiendo diferenciar bien los patrones de cada caracter.
- En el procesamiento de imágenes médicas la transformada wavelet es útil para identificar patrones de objetos de interés, en este caso la detección de bacilos de tuberculosis a partir de propiedades morfológicas.
- En la identificación de patrones a partir de imágenes fijas es vital el pre-procesamiento de las imágenes, o sea, manipular de cierta forma las imágenes adecuándolas a los procedimientos posteriores de detección de patrones como tal para aumentar las probabilidades de éxito de los pasos posteriores.
- El análisis de texturas permite identificar objetos de interés en una imagen y sus propiedades. Para este propósito, la transformada wavelet es especialmente útil ya que permite diferenciar y analizar las diferentes regiones de una imagen a partir de los diferentes detalles a diferentes niveles de resolución. Lo anterior se utilizó para identificar una clase de cultivo particular en medio de otro tipo de vegetación.

- En la actualidad son muchos los campos de la ciencia y la tecnología donde las imágenes son de vital importancia por brindar mucha información. De ahí la importancia de expandir el estudio de técnicas y aplicaciones relacionadas con el procesamiento y análisis de imágenes en diversos campos como la biomedicina, teledetección, etc.

4.2 RECOMENDACIONES Y LINEAS FUTURAS

- En la era de la información es de vital importancia el desarrollo de aplicaciones que procuren mejorar las condiciones de vida de las personas, específicamente en el campo de la salud. Así como en este trabajo se utilizaron técnicas como la transformada wavelet y el procesamiento de imágenes para hacer una aplicación para detectar tuberculosis, estas y otras técnicas se pueden usar para crear aplicaciones que apoyen la detección de cualquier otra enfermedad, básicamente para facilitar la prestación de servicios médicos a las personas que viven alejadas de las ciudades.
- El análisis de texturas tiene aplicabilidad en muchos campos, como por ejemplo, muchas etapas de procesos industriales como detección de fisuras en materiales, chequeo de rugosidad, identificación de objetos etc. Con este fin, la transformada wavelet es una herramienta muy útil y vale la pena extender su aplicabilidad promoviendo el desarrollo de aplicaciones que optimicen etapas de procesos industriales.

- Por último queda reconocer y alabar el interés del departamento de telecomunicaciones por abordar el estudio y la investigación en el campo del procesamiento de señales e imágenes con herramientas prometedoras como la transformada wavelet.

ANEXO A

TEORÍA SOBRE TRANSFORMADA WAVELETS

1. Transformada de Fourier

En la teoría de wavelets se trabaja de manera similar a la teoría de Fourier, la cual establece que una señal se puede descomponer en una serie de funciones sinusoidales. Recordando, la transformada de Fourier (FT) de una señal $\mathbf{x}(t)$ esta definida por: [3]

$$FT(f) = \int_{-\infty}^{+\infty} \mathbf{x}(t)e^{-i2\pi ft} dt \quad (\text{A.1})$$

La transformada de Fourier trabaja bien si la señal $\mathbf{x}(t)$ esta compuesta de unos cuantos componentes estacionarios. Sin embargo, algún cambio repentino en el tiempo en una señal $\mathbf{x}(t)$ no estacionaria, no es reflejado en el espacio de frecuencias. Para contrarrestar estas desventajas se ha modificado la transformada de Fourier dando origen a la Transformada de Fourier de Tiempo Corto (STFT- *Short Time Fourier Transform*) también conocida como la Transformada de Gabor [6].

2. Transformada de Fourier de Tiempo Corto

La STFT usa una función $\mathbf{g}(t)$ como ventana de análisis que se mueve sobre la señal para determinar el espectro. La expresión que nos permite este análisis es [6]:

$$STFT(b, f) = \int_{-\infty}^{+\infty} \mathbf{x}(t)\mathbf{g}^*(t - b)e^{-i2\pi ft} dt \quad (\text{A.2})$$

donde el parámetro b indica la posición en la que se hace el análisis g^* denota el complejo conjugado.

Ahora bien, el soporte de la ventana constituye un parámetro de gran importancia ya que establece un compromiso entre resolución en tiempo y resolución en frecuencia, de tal manera que al establecer una ventana angosta se analiza una pequeña parte de la señal teniéndose una buena resolución en tiempo pero pobre en frecuencia ya que solo se conoce una pequeña fracción del espectro total de la señal. Por el contrario, si la ventana es muy ancha se tiene una buena resolución en frecuencia pero pobre resolución en tiempo. En el caso extremo una ventana de ancho infinito corresponde a la transformada de Fourier clásica. Por lo tanto un defecto de la STFT es que no puede entregar una buena resolución tanto en tiempo como en frecuencia de manera simultánea ya que el soporte de la ventana es fijo. El principio que está por debajo del inconveniente anterior es el Principio de Incertidumbre de Heisenberg, el cual, en términos de procesamiento de señal establece que es imposible saber la frecuencia exacta y el tiempo exacto de ocurrencia de esta frecuencia en una señal, es decir, una señal simplemente no puede ser representada como un punto en el espacio tiempo-frecuencia [5].

A partir de este inconveniente surge la inquietud sobre si es posible tener una ventana con rango dinámico que permita realizar un análisis de la señal de manera ideal, es decir, tener una buena resolución en tiempo para frecuencias altas, y una buena resolución en frecuencia para analizar las frecuencias bajas. Precisamente como solución a este problema tiene lugar la herramienta matemática llamada transformada wavelet.

3. Transformada wavelet (WT)

La transformada wavelet descompone una función o señal a analizar en términos de funciones denominadas wavelets que son generadas por una función wavelet de transformación denominada *wavelet madre*. El término wavelet significa “onda pequeña”. La pequeñez se refiere al hecho de que esta función (ventana) es de longitud finita (compactamente soportada) y el término onda se refiere a la condición que esta función es de naturaleza oscilatoria [4]. El término madre da a entender que las funciones con diferentes regiones de actuación que se usan en el proceso de transformación provienen de una función principal o wavelet madre, es decir, la wavelet madre es un prototipo para generar las otras funciones ventanas.

La WT comprende básicamente la transformada continua wavelet (CWT) y la transformada discreta wavelet (DWT), herramientas matemáticas que posibilitan el análisis de señales proporcionando información en el dominio del tiempo y en el dominio de la frecuencia. El análisis mediante la CWT se efectúa sobre señales analógicas de energía finita (definida en el teorema de Parseval) [3]. Sin embargo, el procesamiento digital de señales, tal como su nombre lo indica, se aplica sobre muestras de datos digitales en tiempos discretos.

Básicamente, lo que hace la transformada Wavelet en el dominio discreto es filtrar una señal en el dominio del tiempo mediante filtros paso bajo y paso alto que eliminan ciertas componentes de alta o baja frecuencia de la señal, el procedimiento se repite para las señales resultantes del proceso de filtrado anterior.

4. Transformada continua wavelet

La transformada Wavelet continua (CWT) fue desarrollada como una técnica alternativa a la STFT como una manera de superar el problema de resolución. El análisis wavelet se realiza de manera similar al análisis STFT, en el sentido que la señal es multiplicada por una función (función wavelet) de manera similar a la función ventana en la STFT, y la transformada se calcula separadamente para distintos segmentos de la señal en el dominio del tiempo. Así como lo hace la STFT, esta transformada utiliza una función ventana con la cual se analiza por intervalos una señal, sin embargo, esta ventana es de ancho variable. La CWT intenta expresar una señal $x(t)$ continua en el tiempo, mediante una expansión de términos o coeficientes proporcionales al producto interno entre la señal y diferentes versiones escaladas y trasladadas de una función prototipo $\psi(t)$, conocida como wavelet madre. Entre este tipo de wavelets una de las más conocidas es la Haar representada en la figura 2.1 [3], siendo ésta la más antigua y simple de todas, no obstante, debido a su forma es muy utilizada para analizar señales que presentan cambios abruptos.

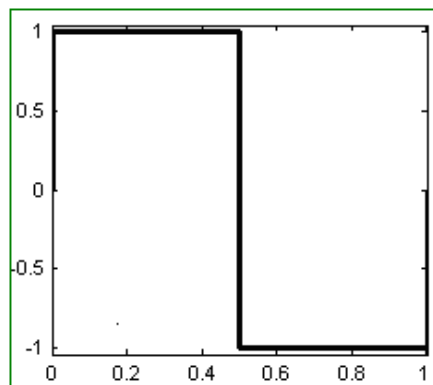


Figura 2.1. Función Wavelet Haar

Algunos ejemplos de otras wavelets madre se muestran en la figura A.2.

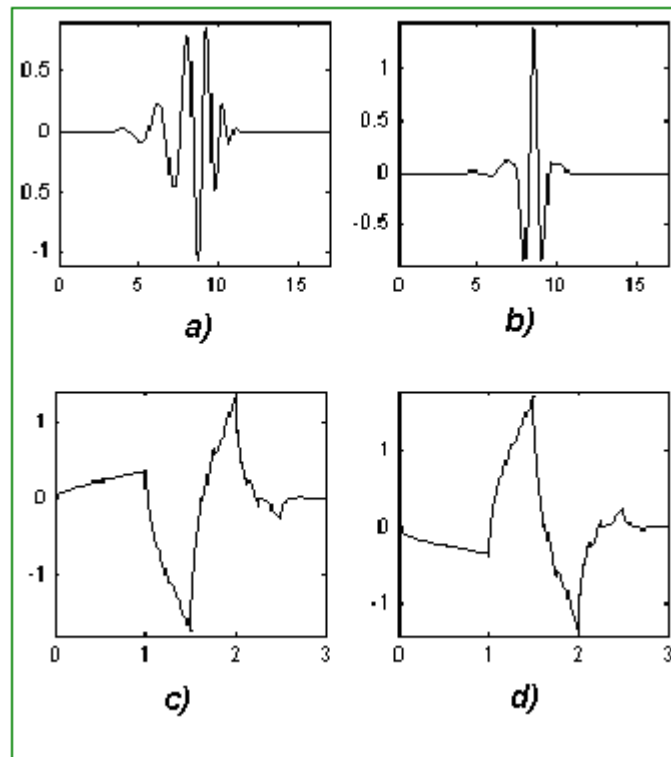


Figura A.2. Funciones wavelets más comunes: a) Daubechies9 b) Coiflet3 c) Symlet2 d) Daubechies2

Partiendo de que tanto la señal $\mathbf{x}(t)$ como la función wavelet madre $\boldsymbol{\psi}(t)$ son de energía finita, entonces se define la CWT de una señal $\mathbf{x}(t)$ como:

$$\mathbf{CWT}(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} \mathbf{x}(t) \boldsymbol{\psi}\left(\frac{t-b}{a}\right) dt, \quad a, b \in \mathbb{R} \text{ y } a \neq 0 \quad (\text{A.3})$$

Donde \mathbf{a} determina el ancho o soporte de la función ventana y \mathbf{b} determina la ubicación en el dominio del tiempo de la misma [4]. De acuerdo a esto se identifica a la variable \mathbf{a} como *variable de escala*, la cual permite comprimir (reducir la escala) o dilatar (subir la escala) la función $\boldsymbol{\psi}$, en otras palabras, establece el grado de resolución con el cual se analiza la señal, mientras que la variable \mathbf{b} se llama *variable de traslación*.

De acuerdo a la definición de la CWT puede concluirse que más que una representación tiempo-frecuencia, es una representación tiempo-escala, siendo escala lo opuesto de frecuencia, ya que el término de frecuencia está reservado a la transformada de Fourier. En tanto que la información global de la señal, es decir, las bajas frecuencias, se analizan con altas escalas y se obtiene buena resolución en frecuencia, mientras que los detalles de la señal, correspondientes a las altas frecuencias, se analizan con bajas escalas, con lo que se obtiene buena resolución en tiempo. En este sentido, se dice que la escala y la frecuencia tienen una relación inversamente proporcional. La representación gráfica de la CWT se realiza en un plano denominado plano tiempo-escala, mostrado en la figura A.3.

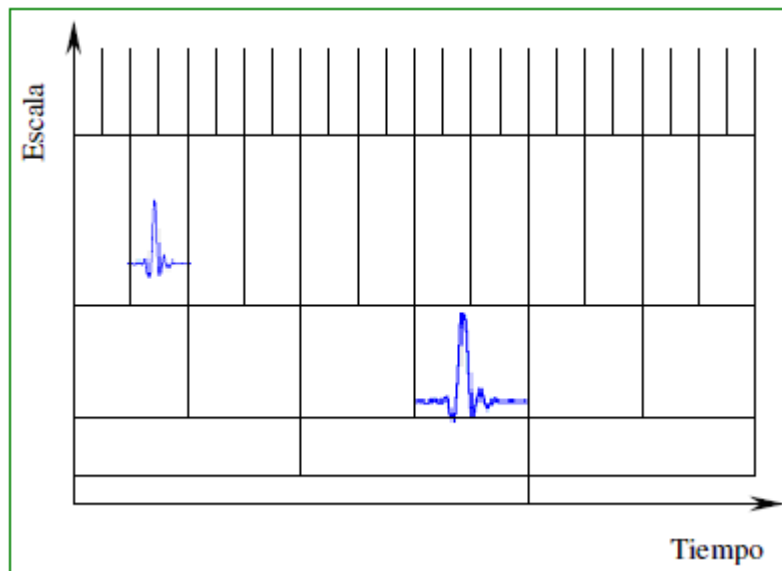


Figura A.3. Plano tiempo -escala en la CWT

Desde un punto de vista intuitivo, la CWT consiste en calcular un índice de semejanza entre la señal que está siendo analizada, y la wavelet, tal como se muestra en la figura A.4.

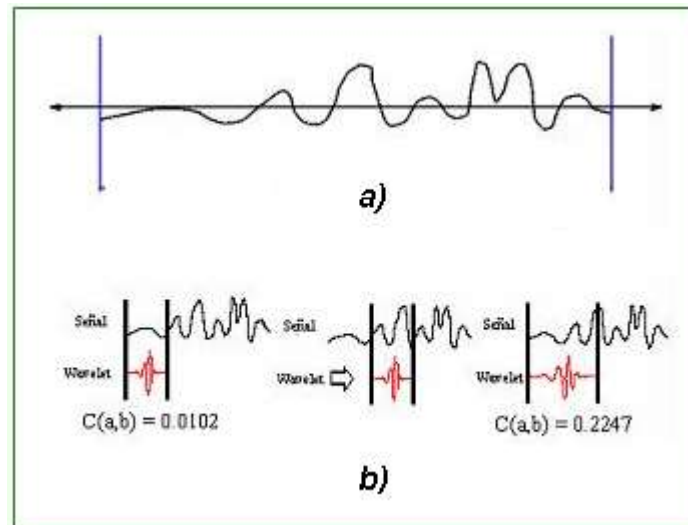


Figura A.4 Modo de funcionamiento de la CWT, a) Señal a analizar, b) representación del proceso de translación y cambio de escala sobre la señal [7].

5. La escala [4]

El parámetro escala, en el análisis wavelet, es similar a la escala usada en los mapas, como en el caso de éstos, las altas escalas corresponden a vistas globales no detalladas (de la señal) y las bajas escalas corresponden a vistas en detalle. En términos de frecuencia, las bajas frecuencias (altas escalas) corresponden a la información global de una señal (que usualmente se extienden al total de la señal), mientras que las altas frecuencias (bajas escalas) corresponden a información detallada de un patrón oculto en la señal (que usualmente dura un tiempo relativamente).

El escalamiento es una operación matemática, que dilata o comprime a una función. Si $f(t)$ es una función dada, $f(at)$ corresponde a una versión contraída (comprimida) de $f(t)$ si $a > 1$ y expandida (dilatada) si $a < 1$. Sin embargo, en la definición de la WT, el término escalamiento es usado en el denominador, y por lo

tanto, se cumple lo opuesto de la sentencia enunciada, en este caso, escalas $a > 1$ dilatan la señal, mientras que escalas $a < 1$ comprimen la señal.

6. Traslación

El término traslación se usa con el mismo sentido que fue usado en la STFT y está relacionado con la localización de la ventana a medida que ésta se desplaza a través de la señal. Obviamente, este término corresponde a la información del tiempo en el dominio transformado.

La variable b controla la ubicación de la función en el espacio de tiempo permitiéndonos deslizar $\psi(t)$ sobre el intervalo de tiempo en el que se haya definido $x(t)$. Un punto importante es que la función wavelet se traslada cubriendo toda la señal para cada valor de a , es decir, si la escala escogida es pequeña habrá más traslaciones de $\psi(t)$ que si la escala escogida es grande. Por lo tanto, la variable b nos da la cantidad por la cual $\psi(t/a)$ ha sido trasladada en el dominio del tiempo [4]. El proceso de cálculo de la CWT se puede describir mediante los siguientes pasos [3]

1. Tomar una wavelet madre y compararla con un intervalo al principio de la señal original.
2. Dados dos valores a y b , calcular un coeficiente $C_{a,b}$ que representa la correlación entre la wavelet y la sección de la señal bajo análisis. Entre mayor sea el valor del coeficiente, mayor es la similitud, de ahí que los resultados dependen de la wavelet elegida para hacer dicho análisis.

3. Desplazar la wavelet en el sentido positivo del eje temporal, y repetir los pasos anteriores hasta que se haya cubierto la totalidad de la señal.
4. Escalar la wavelet en el tiempo y repetir los pasos 1 a 3.

Un ejemplo de este proceso se representa en la figura A.4., en donde se puede observar como entre más se parece la señal a wavelet (con una escala a y en una posición b determinada) más alta es la magnitud del coeficiente wavelet.

Es importante anotar que para que el análisis wavelet de una señal sea posible y se logre una perfecta reconstrucción de la misma a partir de la transformada, la función ψ debe cumplir con la condición de admisibilidad [3], la cual implica que su valor medio debe ser igual a cero (la integral en todo su dominio es cero, ya que hay valores positivos y negativos), en otras palabras, se requiere que sea una onda. Además, teniendo en cuenta que la función wavelet analiza la señal por intervalos de tiempo de acuerdo al tamaño de la ventana, se dice que ésta es de soporte compacto, es decir, es una onda definida sobre un intervalo de tiempo finito, de ahí su nombre wavelet u ondita.

7. Transformada discreta wavelet

La diferencia con la CWT radica principalmente en la forma en que los parámetros de desplazamiento y escala son discretizados.

A nivel teórico se habla de la CWT, donde dicha continuidad radica en el hecho que la variable de escala y la variable de traslación varían en forma continua. Sin embargo a nivel práctico, es decir, a nivel computacional, se hace necesario

discretizar la transformada, y la solución más lógica es asignar valores discretos a las dos variables mencionadas, dando lugar así a la transformada wavelet discreta DWT. La forma más común de discretizar los valores de a y b es utilizar una red diádica [5], es decir, $a = 2^{-j}$ y $b = k 2^{-j}$, de tal manera que el conjunto de funciones wavelets determinadas por:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad a, b \in \mathbb{R}, a \neq 0 \quad (\text{A.4})$$

adquiere la forma:

$$\Psi_{j,k}(t) = 2^{\frac{j}{2}} \psi(2^j t - k) \quad j, k \in \mathbb{Z}. \quad (\text{A.5})$$

Este conjunto de funciones se conoce como la versión diádica discretizada de la función wavelet o como conjunto de expansión wavelet, y es el resultado de traslaciones y escalamientos de la función wavelet madre ψ . El factor $2^{\frac{j}{2}}$ se denomina constante de normalización y se hace necesario para cumplir con la condición de ortonormalidad.

De otro lado, la función wavelet madre ψ lleva asociada consigo una función escala $\varphi(t)$, de tal manera que con una de estas funciones o con ambas es posible aproximar cualquier función $f(t) \in L^2(\mathbb{R})$ (espacio de Hilbert), de la siguiente forma [3]:

$$f(t) = \sum_k \sum_j c_{j,k} \varphi(t) + \sum_j \sum_k d_{j,k} \Psi(t) \quad j, k \in \mathbb{Z} \quad (\text{A.6})$$

En esta expresión se ha aproximado f en términos de la función escala $\varphi(t)$ y de la función wavelet ψ .

ANEXO B

1. Enfoques de Reconocimiento de Patrones

En el presente anexo se trata con detalle los métodos de decisión teórica ya que estos se utilizan en el desarrollo del proyecto y se mencionan otros enfoques de reconocimiento de patrones dando una breve descripción de ellos solo para tener referencia, ya que no se utilizan en el desarrollo del proyecto.

1.1 Métodos de decisión teórica.

Los métodos de decisión teórica para el reconocimiento se basan en la utilización de funciones de decisión (o discriminantes).

Si $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]^T$ (donde T indica el vector transpuesto) representa un patrón vectorial de dimensión n . Para M clases de patrones $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_M$, el problema básico en el reconocimiento de patrones por decisión teórica consiste en encontrar M funciones de decisión $\mathbf{d}_1(\mathbf{X}), \mathbf{d}_2(\mathbf{X}), \dots, \mathbf{d}_M(\mathbf{X})$, asociadas a las M clases que tengan la propiedad de que, si un patrón \mathbf{X} pertenece a la clase \mathbf{w}_i , entonces

$$\mathbf{d}_i(\mathbf{X}) > \mathbf{d}_j(\mathbf{X}) \quad j = 1, 2, \dots, M; \quad j \neq i \quad (\text{B.1})$$

En otras palabras, un patrón desconocido \mathbf{X} pertenece a la i -ésima clase de patrones si, al sustituir la \mathbf{X} en todas las funciones de decisión, $\mathbf{d}_i(\mathbf{X})$ toma el mayor valor numérico. En caso de empate, los conflictos se resuelven arbitrariamente.

La *frontera de decisión* que separa la clase \mathbf{w}_i de la \mathbf{w}_j viene dada por valores de \mathbf{X} para los que $\mathbf{d}_i(\mathbf{X}) = \mathbf{d}_j(\mathbf{X})$.

La práctica común consiste en hacer que la frontera de decisión sea la función $\mathbf{d}_{ij}(\mathbf{X}) = \mathbf{d}_i(\mathbf{X}) - \mathbf{d}_j(\mathbf{X}) = \mathbf{0}$. De esta forma, $\mathbf{d}_{ij}(\mathbf{X}) > \mathbf{0}$ para los patrones de la clase \mathbf{w}_i y $\mathbf{d}_{ij}(\mathbf{X}) < \mathbf{0}$ para los patrones de la clase \mathbf{w}_j .

1.1.1 Clasificador de mínima distancia

Si cada clase de patrones está representada por un vector prototipo (o medio):

$$\mathbf{m}_j = \frac{1}{N_j} \sum_{\mathbf{X} \in \mathbf{w}_j} \mathbf{X} \quad j = 1, 2, \dots, M \quad (\text{B.2})$$

donde N_j es el número de patrones vectoriales de la clase \mathbf{w}_j y la suma se realiza para todos los vectores. Una forma de determinar la pertenencia a una clase de un patrón vectorial desconocido \mathbf{X} consiste en asignarlo a la clase del prototipo más próximo. Si se utiliza la distancia euclídea para determinar el grado de proximidad, se reduce el problema del cálculo de las medidas de distancia:

$$D_j(\mathbf{X}) = \|\mathbf{X} - \mathbf{m}_j\| \quad j = 1, 2, \dots, M \quad (\text{B.3})$$

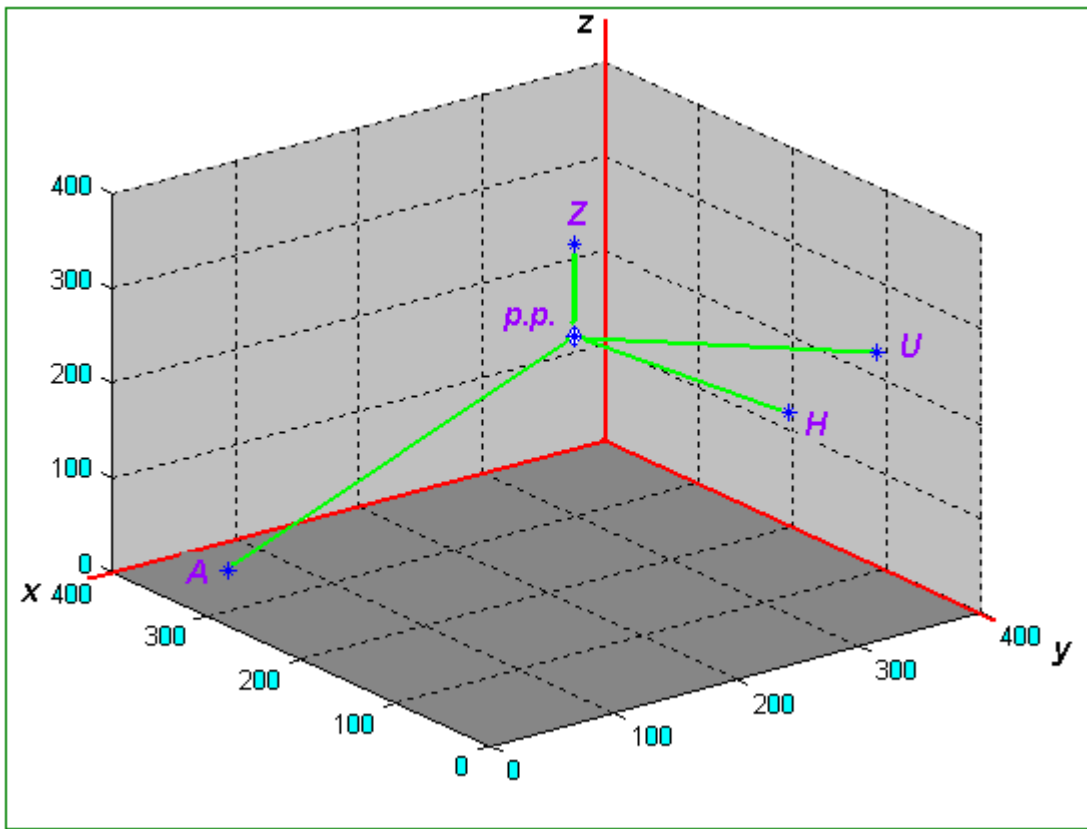


Figura B.1 Cálculo de la distancia euclídea, espacio tridimensional. La menor distancia está entre el patrón desconocido *pp* y el patrón promedio de la clase *Z*, por lo que el patrón desconocido es asignado a la clase *Z*.

Si $D_i(\mathbf{X})$ es la menor distancia, entonces se asigna la \mathbf{X} a la clase w_i . Ver figura B.1. Esto es, en esta formulación, la menor distancia implica el mejor emparejamiento o correspondencia. Esto es equivalente a evaluar las funciones

$$d_j(\mathbf{X}) = \mathbf{X}^T \mathbf{m}_j - \frac{1}{2} \mathbf{m}_j^T \mathbf{m}_j \quad j=1,2,\dots,M \quad (\text{B.4})$$

y asignar \mathbf{X} a la clase w_i si $d_j(\mathbf{X})$ toma el mayor valor numérico. Esta formulación concuerda con el concepto de función de decisión.

La frontera de decisión entre las clases w_i y w_j para un clasificador de mínima distancia es:

$$d_{ij}(X) = d_i(X) - d_j(X) = X^T(m_i - m_j) - \frac{1}{2}(m_i - m_j)^T(m_i - m_j) = 0 \quad (\text{B.5})$$

La superficie dada por esta ecuación es el bisector (pasa por el punto medio) perpendicular a la línea que une m_i y m_j . Para $n = 2$, el bisector perpendicular es una línea, para $n = 3$ es un plano, y para $n > 3$ es un hiperplano.

En la práctica, el clasificador de mínima distancia funciona bien cuando la distancia entre los promedios es grande en comparación con las variaciones de los elementos de cada clase con respecto a la media.

1.2 Reconocimiento Lógico Combinatorio de Patrones

Este enfoque se basa en la idea de que la modelación del problema debe ser lo más cercana posible a la realidad del mismo, sin hacer suposiciones que no estén fundamentadas. Uno de los aspectos esenciales del enfoque es que las características utilizadas para describir a los objetos de estudio deben ser tratadas cuidadosamente [9].

1.3 Reconocimiento Sintáctico de Patrones

Este enfoque se basa en encontrar las relaciones estructurales que guardan los objetos de estudio, utilizando la teoría de lenguajes formales. El objetivo es construir una gramática que describa la estructura del universo de objetos [9].

1.4 Redes Neuronales.

Este enfoque supone que tiene una estructura de neuronas interconectadas que se estimulan unas a otras, las cuales pueden ser “entrenadas” para dar una cierta respuesta cuando se le presentan determinados valores [9].

1.5 Reconocimiento Estadístico de Patrones.

Este enfoque se basa en la teoría de probabilidad y estadística y supone que se tiene un conjunto de medidas numéricas con distribuciones de probabilidad conocidas y a partir de ellas se hace el reconocimiento. Este enfoque es el que utilizaremos en nuestro trabajo ya que tenemos un conocimiento previo de del conjunto de descriptores de las clases [9].

2. Métodos de reconocimiento de caracteres más usados.

2.1 Un caso de Clasificador Sintáctico.

El clasificador sintáctico se basa en la utilización de un conjunto de primitivas que describen a los caracteres. Las primitivas se etiquetan y se generan enunciados preposicionales que se utilizan para crear un conjunto de características que distinguen a cada caracter. La descripción de los caracteres se almacena en el sistema de forma que cuando se quiere clasificar un caracter desconocido se buscan las primitivas y los enunciados almacenados que lo representan.

Ali y Yarman presentan un método de reconocimiento de caracteres basado en este clasificador [10], donde dividen el proceso de clasificación en tres componentes: El detector de primitivas, el control global y el comparador de descripciones. Así, dada una imagen que contiene caracteres, el primer componente utiliza un procedimiento de erosión morfológica para encontrar las primitivas, tomado esas mismas primitivas como elementos estructurantes. El componente de control global maneja el proceso de búsqueda y de agrupación (o partición) de las primitivas detectadas, mientras que el componente de comparación encuentra otras primitivas ubicadas en posiciones relativas

adecuadas para completar un símbolo. En la figura B1 se presenta un ejemplo de definición y detección de primitivas para la letra A. En la Figura B.2 de izquierda a derecha: La imagen del caracter, el conjunto de primitivas P1, P2 y P3 y finalmente el resultado de erosionar la imagen original con las tres primitivas.

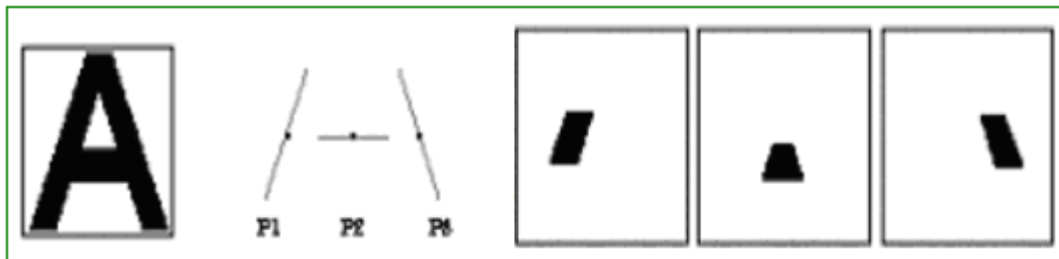


Figura B.2. Un ejemplo de definición y detección de primitivas [10].

2.2 Correspondencia por correlación

El siguiente es el concepto básico de la correlación: Si se tiene una sub-imagen $w(x, y)$ de tamaño $J \times K$ dentro de una imagen $f(x, y)$ de tamaño $M \times N$, donde $J \leq M$ y $K \leq N$. La correlación entre $w(x, y)$ y $f(x, y)$ en un punto (i, j) está dada por:

$$c(\mathbf{s}, \mathbf{t}) = \sum_x \sum_y f(x, y) w(x - \mathbf{s}, y - \mathbf{t}) \quad (\text{B.6})$$

donde $s = 0, 1, \dots, M-1$; $f = 0, 1, \dots, N-1$, y la sumatoria se calcula para la región de la imagen donde se solapan w y f . Se asume que el origen de la imagen f está en la esquina superior izquierda. La correlación es el proceso de mover el patrón o sub-imagen w a través del área de la imagen y calcular el valor C en esa área. Esto envuelve una multiplicación en cada píxel en el patrón por cada píxel que es trasladado en la imagen y entonces sumando los resultados sobre todos los píxeles del patrón. El máximo valor de C indica la posición donde w se parece más a f [9]. En la figura B.3 se representa esta situación.

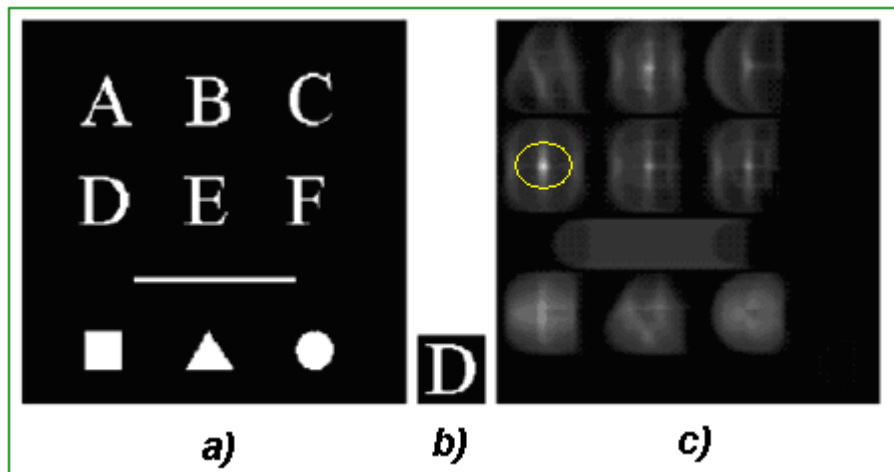


Figura B.3. a) Imagen b) Sub-imagen d) Correlación.

Los valores de correlación no son precisos en los bordes de la imagen. La correlación básica es muy sensible a cambios de amplitud en la imagen, tales como intensidad. Por ejemplo, si la intensidad de la imagen f es doblada, también serán doblados los valores de C .

2.3 Un caso de Clasificador con Redes Neuronales Artificiales.

Las redes neuronales artificiales (RNA) son un método de aprendizaje automático inspirado en un modelo computacional simplificado del cerebro humano. La unidad básica de las RNA son las neuronas, éstas reciben entradas y las combinan de cierta forma ejecutando una operación no lineal para generar un resultado que es enviado como salida a otras neuronas.

La figura B.4 muestra la representación de una neurona artificial. Las entradas a la red se representan con los símbolos x_i . Cada entrada se multiplica por una conexión con peso w_i . En el caso más simple se obtiene una suma de los productos $x_i \cdot w_i$ que se pasa a la función de transferencia para limitar la amplitud

de la salida generada. La función de activación más común es la función sigmoideal por su importante propiedad de ser derivable.

El aprendizaje en las redes neuronales artificiales se logra actualizando los pesos de la red en un proceso de entrenamiento para minimizar la diferencia entre la salida de la red y el valor deseado de cada uno de los ejemplos de entrenamiento.

La actualización de pesos continúa hasta que los resultados de la red aplicados a un conjunto de ejemplos de prueba son satisfactorios.

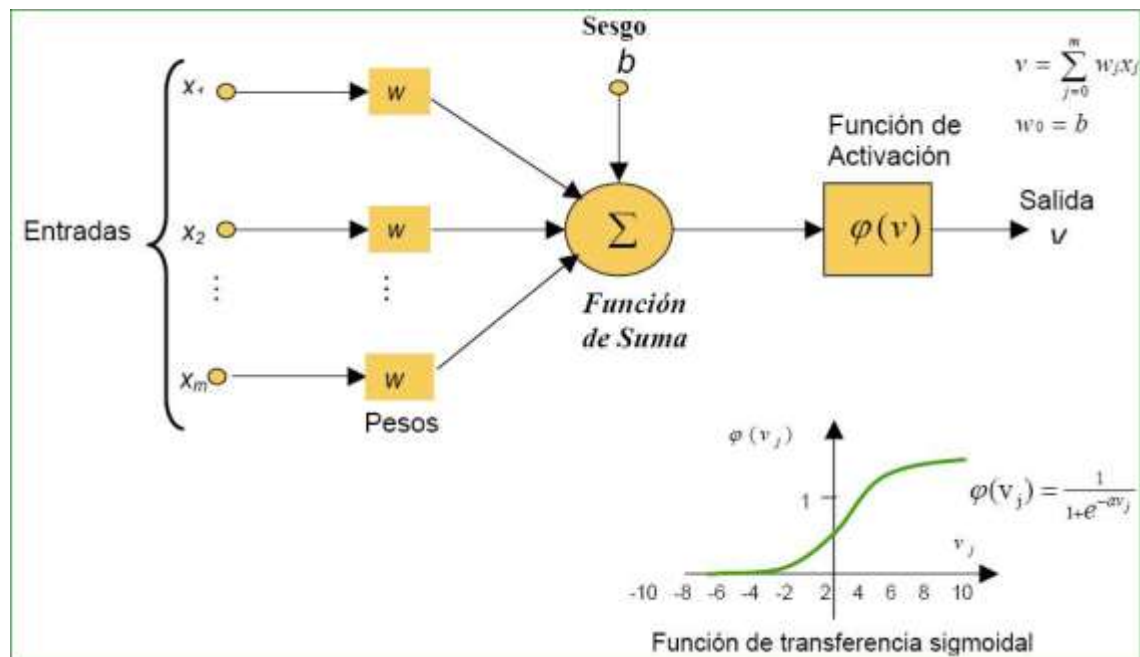


Figura B.4. La unidad básica de las redes neuronales y la función de activación sigmoideal [10].

En las aplicaciones modernas incluidas las de reconocimiento óptico de caracteres (OCR), se utilizan estructuras en varias capas de Unidades Básicas donde entre las capas de entrada y salida pueden existir varias capas ocultas. Las líneas de comunicación entre las neuronas son muy importantes en las redes neuronales ya que determinan la retroalimentación entre capas o la fuerza de conexión entre las

neuronas que generan los estados de activación o inhibición de otras neuronas [10].

2.4 HISTOGRAMA SOBEL

Para la detección de bordes en imágenes el operador de Sobel es uno de los más famosos. Su principio fundamental es calcular la diferencia de intensidad en la vecindad de cada píxel en las direcciones vertical y horizontal, lo cual se interpreta como la derivada de la función en el punto. Con el operador de Sobel puede obtenerse información de la magnitud y la dirección del cambio en cada punto [10].

ANEXO C

MANUAL DE USUARIO

Los algoritmos desarrollados en el proyecto “Reconocimiento de patrones en imágenes fijas usando wavelets” se hicieron sobre MATLAB 7. Por lo tanto se debe instalar MATLAB 7 ó los programas no correrán correctamente. El sistema de reconocimiento presenta al usuario una interfaz amigable (figura C1) donde se le permite seleccionar la aplicación que desea ejecutar. Las aplicaciones son:

- Identificación de placas de vehículos.
- Detección de tuberculosis.
- Identificación de un cultivo específico (maíz).
- Reconocimiento de texto.



Figura C1. Interfaz

Los pasos a seguir para realizar el procesamiento de las imágenes son los siguientes:

Seleccionar la aplicación

Seleccione la aplicación en el menú **Tipo de Aplicación**. Por ejemplo si se desea ejecutar la aplicación '**Identificación de placas de vehículos**' se hace clic en esa aplicación (figura C2).

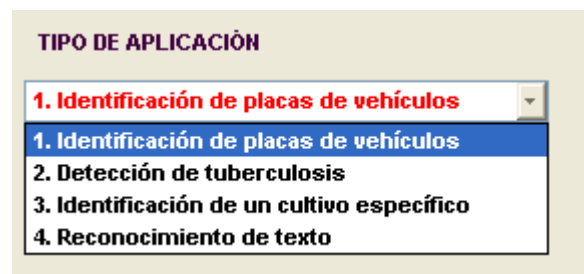


Figura C2. Selección de la aplicación.

Cargar imagen

Al hacer clic en la opción **File** y seleccione **Cargar imagen** como se muestra en la figura C3.

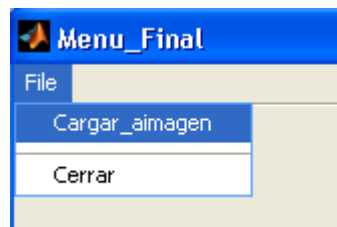


Figura C3. Opción cargar imagen

Al hacer esto se abrirá un navegador donde se encuentran las carpetas que contienen las imágenes, figura C4.

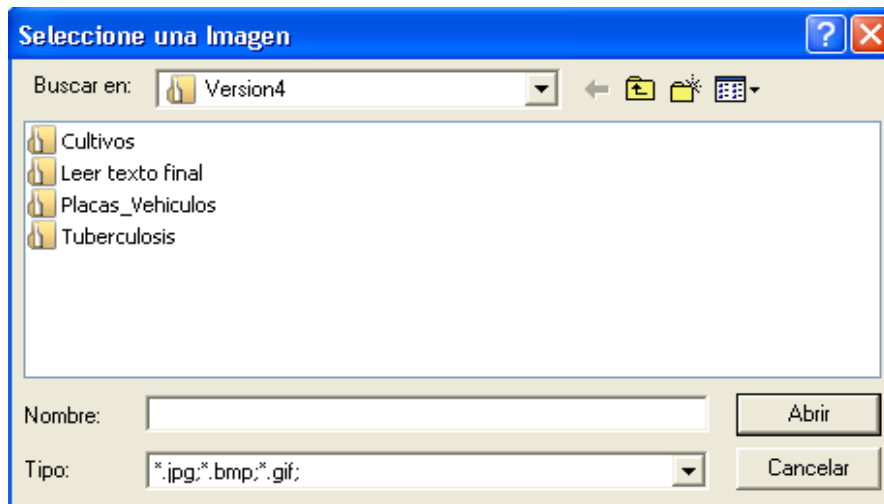


Figura C4. Selección de carpeta

Dependiendo de la aplicación seleccione la carpeta y luego la imagen y esta se cargará en la interfaz principal como se muestra en la figura C5.



Figura C5. Imagen cargada.

Iniciar reconocimiento

Cuando se ha cargado la imagen se hace clic en el botón **Iniciar reconocimiento** y comienza el procesamiento de la imagen, mostrando las imágenes correspondientes a las diferentes etapas del procesamiento, y finalmente se

muestran los resultados del reconocimiento. La figura C6 muestra, por ejemplo, el resultado del reconocimiento de la placa del vehículo de la imagen car176.jpg



Figura C6. Resultado del reconocimiento.

Para las demás aplicaciones se procede de manera similar. Es decir, se selecciona la aplicación que se desee ejecutar y luego se carga la imagen que se quiera procesar. En cada aplicación se ejecutan los programas respectivos y se muestran los resultados correspondientes. A continuación se muestra el resultado del reconocimiento de patrones en las demás aplicaciones. Como las etapas de selección de la aplicación y selección de la imagen a procesar se hacen de igual manera que en el caso anterior, solo se muestra los resultados finales.

Detección de tuberculosis

En la figura C7 se mira el resultado de la detección de tuberculosis. Después de seleccionar la aplicación y de cargar la imagen correspondiente, al pulsar **iniciar reconocimiento** el programa comienza procesar la imagen eliminando zonas

indeseadas y luego detectando los bacilos *Mycobacterium tuberculosis* presentes en la imagen. En la interfaz se muestra la imagen original, la imagen resultante de la detección y debajo de esta, aparece un número que representa, en términos porcentuales, el porcentaje de infección del paciente equivalente a la razón entre el área ocupada por los bacilos y el área total de la imagen.



Figura C7. Resultado de la detección de tuberculosis.

Identificación de cultivo

En la figura C8 se muestra el resultado de la identificación de cultivo de maíz. Al pulsar **iniciar reconocimiento** el programa comienza a procesar la imagen eliminando zonas indeseadas e identificando el cultivo de maíz. En este caso se muestra la imagen original y otra imagen en donde se resaltan las zonas que el sistema ha identificado como cultivo de maíz.



Figura C8. Identificación de cultivo de maíz.

Reconocimiento de texto

En la figura C9 se muestra la identificación de texto en una imagen. Al pulsar **iniciar reconocimiento** el programa comienza a buscar, inicialmente, renglones, luego palabras y finalmente letras para proceder con reconocimiento de éstas. Aquí se muestra la imagen original y en el cuadro inferior se muestra el texto que el sistema ha identificado.



Figura C9. Identificación de texto en una imagen.

Finalmente cuando se desee salir de las aplicaciones haga clic sobre la opción **File** y luego en **Cerrar** (Figura C10) o cierre la ventana principal.



Figura C10. Salir de la aplicación.

REFERENCIAS

- [1] <http://alojamientos.us.es/gtocoma/pid/introduccion.html> 05/07/2006.
- [2] <http://www.upv.es/frechet/wavelets/imagenes/presentacion.htm> 20/05/2006
- [3] IGLESIAS Lorenzo J. "Diseño de un relé neuronal de protección para líneas aéreas de AT con procesado de señal mediante la transformada wavelet". Departamento de Ingeniería Electrónica. Universidad Politécnica de Cataluña. 2003.
- [4] LUQUE MARIN José A. Capítulo 9. "Análisis de los ciclos en el contenido en materia orgánica del sedimento del Lago de Sanabria". Subcapítulo 25. "Análisis Wavelet: un nuevo método para análisis de señales". Departamento de Geoquímica. Universidad de Barcelona. Mayo de 2003.
- [5] FAUNDEZ Pablo, FUENTES Álvaro. "Procesamiento Digital de Señales Acústicas utilizando Wavelets". Instituto de Matemáticas UACH. Chile.
- [6] DE CASTRO FERNANDEZ Rosa Ma. "Análisis de la Teoría de Ondículas Orientadas a las Aplicaciones en Ingeniería Eléctrica: Fundamentos". Escuela Técnica Superior de Ingenieros Industriales, Departamento de Ingeniería Eléctrica. Madrid Julio de 2002.
- [7] Wavelets: A New Tool For Signal Analysis, (Wavelet Toolbox – Matlab 7.0 2004).
- [8] DOMÍNGUEZ JIMÉNEZ Maria E, VIDAI Gabriela S. "La transformada wavelet una introducción". Departamento de Matemática Aplicada a la Ingeniería Industrial. Universidad Politécnica de Madrid. 2005.

[9] FUKUNAGA K, Introduction to statical pattern recognition. Orlando, San Diego: Academic Press, 1972.

[10] OSPINA Emmanuel, URREA Juan, "Implementación de la transformada de Hough para la detección de líneas para un sistema de visión de bajo nivel", Ingeniería Electrónica, Universidad Nacional de Colombia Sede Manizales. Diciembre de 2002.

[11] RESNIKO H.L. y WELLS, R. Jr. Wavelet analysis, the scalable structure of information. New York, EUA: Springer-Verlag 1998.

[12] GONZALES Rafael, WOODS Richard y EDDINS Steven. Digital Image Processing using Matlab. Addison - Wesley, 2004.

[13]<http://coco.ccu.uniovi.es/immed/compresion/descripcion/spiht/discreta/discreta.htm> 11/2005

[14] Edge (Image processing Toolbox User's Guide – Matlab 7.0 2004).

[15] <http://www.tsc.uc3m.es/~jcid/cursotdi/fourier/rotaci/index.html> 26/01/2007

[16] <http://www.fotonostra.com/glosario/interpolacion.htm> 26/01/2007

[17] <http://www.upv.es/frechet/wavelets/imagenes/presentacion.htm> 06/08/2006

[18] CARRASCO OCHOA Jesús A. "Reconocimiento de Patrones". Instituto Nacional de Astrofísica Óptica y Electrónica. Lisboa, Portugal. Septiembre de 2000.

[19] LÓPEZ LEMUZ Rafael. “Un Sistema de Visión para el Registro Automático de Vehículos basado en la Lectura de Números VIN”. Instituto Nacional de Astrofísica, Óptica y Electrónica, 2002.