

# Capítulo 3

## Marca de agua en el dominio transformado

### 3.1 *Introducción*

Una imagen digital, función bidimensional que asigna a cada par de variables espaciales un valor de intensidad de luz, se puede representar en un dominio distinto al espacial. En otro dominio se tiene una distribución de la energía de la imagen digital diferente a la distribución que se tiene en el dominio espacial. Un aspecto importante de la transformación de imágenes es la posibilidad de realizar el proceso inverso mediante la transformación inversa.

Las transformadas más utilizadas en la representación de imágenes son la transformada de Fourier, la transformada coseno discreta, la transformada *wavelet*, la transformada *contourlet*, entre otras. En el procesamiento digital de imágenes resulta útil poder representar una imagen digital en diferentes dominios. Por ejemplo, la transformada coseno discreta es adoptada por el estándar de compresión con pérdidas JPEG (*Joint Photographic Experts Group*) por su capacidad de representar una señal con pocos coeficientes; es decir, la energía de la señal está compactada en un número pequeño de coeficientes de la transformada. De manera similar, la transformada *wavelet* es empleada en el estándar JPEG-2000 de compresión de imágenes con pérdidas mediante el uso del algoritmo EZW (*Embedded Zerotree Wavelet*).

En este trabajo explotamos la descomposición *contourlet* de la imagen en diferentes bandas de frecuencia y utilizamos técnicas de espectro disperso para la inserción de la marca de agua digital con el propósito de hacerla más robusta. Para expandir su espectro, la marca de agua binaria es multiplicada por una secuencia pseudoaleatoria, generando una señal de mayor banda [12], es decir, tiene más componentes en el dominio de la frecuencia.

Expandir el espectro de una señal aumenta su robustez ante el ruido y en nuestro caso, aumenta la robustez de la marca de agua digital. Una vez incrustada la marca de agua en los coeficientes *contourlet* se realiza la síntesis de la imagen para obtener la imagen marcada.

### 3.2 Transformada Discreta de Fourier

La transformada de Fourier es una transformada reversible que permite representar una señal como la sumatoria de funciones exponenciales complejas de diferente frecuencia. Para el caso de señales discretas, la transformada discreta de Fourier está dada por:

$$X(k) = \frac{1}{N} \sum_{n=0}^{N-1} f(n) e^{-i2\pi nk/N} \quad (3.1)$$

$$x(n) = \sum_{k=0}^{N-1} F(k) e^{i2\pi nk/N}$$

donde  $n = \{1, 2, \dots, N - 1\}$ ,  $k = \{1, 2, \dots, N - 1\}$ .

Dado que una imagen digital es función de dos variables espaciales discretas, se extiende la transformada discreta de Fourier y su transformada inversa al caso de dos dimensiones:

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-i2\pi \left( \frac{ux}{M} + \frac{vy}{N} \right)} \quad (3.2)$$

para  $u = 1, 2, \dots, M - 1$ ,  $v = 1, 2, \dots, N - 1$ .

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{i2\pi \left( \frac{ux}{M} + \frac{vy}{N} \right)}$$

para  $x = 1, 2, \dots, M - 1$ ,  $y = 1, 2, \dots, N - 1$ .

La desventaja de la transformada de Fourier es que al pasar al dominio de la frecuencia no se puede conocer la información referente al dominio espacial. Podemos conocer la distribución de la energía de la señal en el espectro, pero no sabemos cuál es su

localización en el espacio. Eso no sería problema si la imagen no presentara variaciones de frecuencia al recorrerla en el espacio. Sin embargo la mayoría de las señales, como las imágenes, tienen componentes de diferentes frecuencias en distintas regiones espaciales. Para el caso de señales unidimensionales en función del tiempo, la transformada de Fourier no es una buena representación de éstas si sus componentes frecuenciales cambian en el tiempo; a este tipo de señales sería mejor poder caracterizarlas en el tiempo y la frecuencia. Dicha caracterización se puede lograr mediante la STFT (*Short Time Fourier Transform*). En el caso de señales unidimensionales en función del tiempo, esta transformada consiste en dividir la señal en intervalos de tiempo en los que se asume que es estacionaria; es decir, intervalos de tiempo en los cuales sus componentes en frecuencia permanezcan constantes, y hacer la transformada de Fourier en cada uno de esos intervalos. Para ello se multiplica la función base de la transformada de Fourier por una ventana temporal de longitud finita (su longitud también se conoce como soporte) que se va recorriendo a lo largo de la señal que se va a transformar.

$$STFT(\tau, f) = \int_{-\infty}^{\infty} x(t)g(t - \tau)e^{-i2\pi ft} dt \quad (3.3)$$

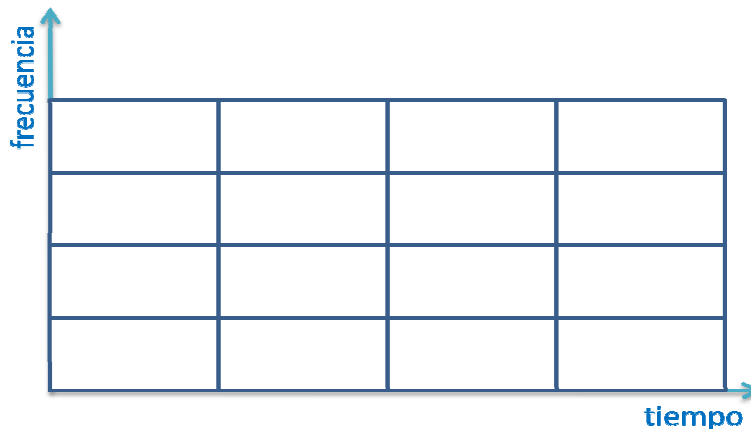
Sin embargo se tiene una restricción en la concentración de la energía de la señal en el dominio del tiempo y la frecuencia dada por el principio de incertidumbre de Heisenberg. En mecánica cuántica este principio indica la incertidumbre existente entre el momento lineal de una partícula y su posición. En el caso de la STFT el principio de incertidumbre implica que mientras más estrecha sea la ventana en el tiempo mejor será la localización de la señal en el tiempo pero mayor será la incertidumbre en la frecuencia y viceversa. Es decir, sólo se puede relacionar ciertas bandas de frecuencia con intervalos de tiempo [13].

$$\sigma_t \sigma_\omega \geq \frac{1}{2} \quad (3.4)$$

A mayor resolución en el tiempo menor será la resolución en la frecuencia. La transformada de Fourier es el caso límite en el cual la ventana temporal es infinita y es posible localizar una sola componente frecuencia de la señal pues se tiene la máxima

resolución en la frecuencia. Dennis Gabor introdujo dichas ventajas temporales utilizadas en la STFT y las llamó átomos de Gabor [13].

La STFT divide el plano tiempo-frecuencia en rectángulos del mismo tamaño como se muestra en la figura 3.1. La relación entre la anchura en el tiempo y altura en la frecuencia de cada rectángulo está determinada por el tamaño de la ventana temporal utilizada y el principio de incertidumbre. Ventanas anchas en el tiempo nos dan buena resolución en la frecuencia pero mala resolución en el tiempo; ventanas angostas en el tiempo brindan buena resolución en el tiempo pero mala resolución en la frecuencia.



**Figura 3.1** División del plano tiempo-frecuencia hecha por la STFT. La resolución es la misma en cualquier región del plano tiempo-frecuencia.

Aunque la STFT nos permite conocer qué componentes espectrales de la señal están presentes en cada intervalo de tiempo existe el problema de escoger el ancho de la ventana temporal que nos brinde la mejor discriminación de las componentes presentes en cada señal.

### 3.3 Transformada *wavelet*

La transformada *wavelet* permite descomponer las señales en versiones de diferente resolución temporal y frecuencial. Para ello se parte de una función  $\varphi(t)$  conocida como

función de escalamiento la cual puede ser trasladada en el tiempo por el parámetro  $k$  y escalada en el tiempo por el parámetro  $j$ .

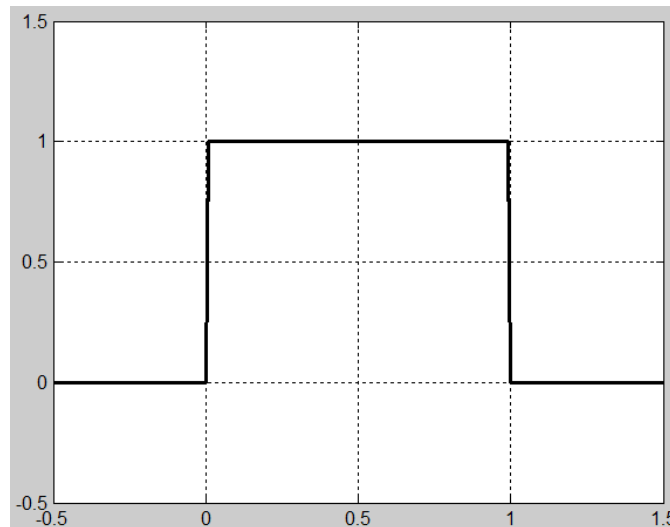
$$\varphi_{j,k}(t) = 2^{\frac{j}{2}}\varphi(2^j t - k) \quad (3.5)$$

donde  $j, k \in \mathbb{Z}$ .

Las funciones de escalamiento pertenecen al espacio vectorial  $L^2$  de funciones cuya integral de su cuadrado es finita. Cada función de escalamiento también pertenece a un subespacio vectorial:  $V_j: \{\varphi_{j,k}(t)\}$ . Y además:

$$V_j \subset V_{j+1} \subset L^2 \quad (3.6)$$

Lo anterior implica que una función de escala  $j$  puede crearse a partir de funciones de escalamiento de mayor resolución  $j + n$  mas no es posible lo inverso. Esta es la base del análisis multiresolución.



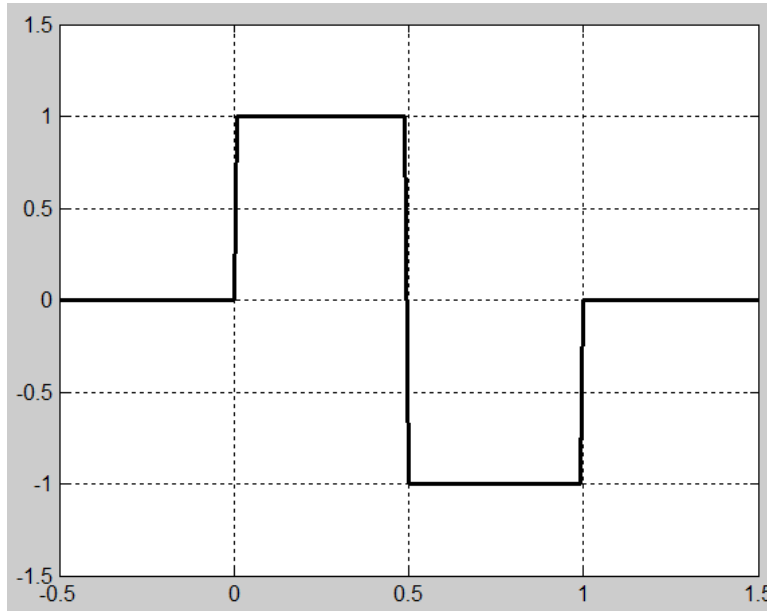
**Figura 3.2 Función de escalamiento Haar.**

Sin embargo las características de una señal pueden ser descritas mejor mediante un conjunto de funciones que abarquen la diferencia de los espacios vectoriales. Este conjunto de funciones se conocen como *wavelets* [15]. Cada función *wavelet* obtenida a partir de la

combinación lineal de funciones de escalamiento es conocida como “*wavelet madre*” y está dada por:

$$\psi(t) = \sum_n h_\psi(n)\sqrt{2}\varphi(2t - n) \quad (3.7)$$

donde  $n \in \mathbb{Z}$ .



**Figura 3.3** “*wavelet madre*” Haar.

Versiones escaladas y trasladadas de la función “*wavelet madre*” pueden ser obtenidas a partir de la siguiente ecuación:

$$\psi_{j,k}(t) = 2^{\frac{j}{2}}\psi(2^j t - k) \quad (3.8)$$

Una vez definido el origen de la función *wavelet*, la transformada *wavelet* continua en una dimensión está dada por [13]:

$$Wf(\tau, s) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{s}} \psi^* \left( \frac{t-\tau}{s} \right) dt \quad (3.9)$$

donde  $\tau$  es la posición de la *wavelet* en el tiempo y  $s$  su escala. La escala es inversa a la frecuencia.

Para cada valor de  $\tau$  y  $s$  se realiza la integral del producto de la función  $f(t)$  y la función *wavelet* escalada y trasladada. Componentes de la señal presentes en la escala  $s$  y posición en el tiempo  $\tau$  darán como resultado un valor grande para  $Wf(\tau, s)$ . La transformada *wavelet* nos permite localizar componentes de la señal original en el tiempo y frecuencia con soportes temporales de diferente tamaño: soportes grandes en el tiempo ofrecen buena resolución en la frecuencia pero mala en el tiempo, y soportes pequeños en el tiempo ofrecen buena resolución en el tiempo pero mala en la frecuencia.

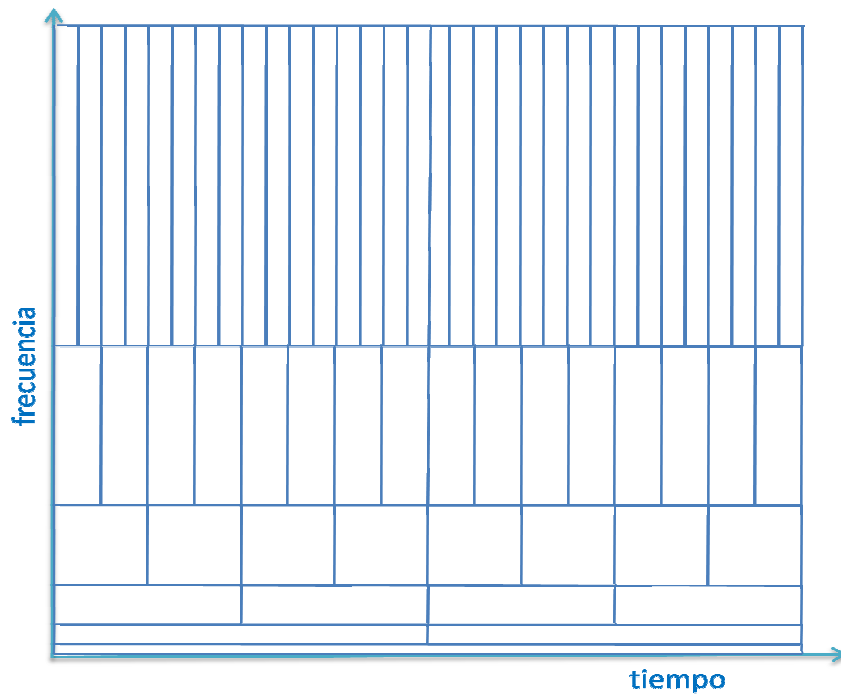


Figura 3.4 División del plano tiempo-frecuencia hecho por la transformada *wavelet*.

### 3.3.1 Transformada discreta *wavelet* en 1-D

En realidad las computadoras trabajan con un número finito de muestras de señales discretas en el tiempo. En [15] se muestra que la transformada discreta *wavelet* en una dimensión se obtiene a partir de la convolución de la respuesta al impulso de los filtros *wavelet*  $h_\psi$  y  $h_\phi$  con la secuencia original  $s(n)$ , seguidas por un submuestreo.  $h_\phi$  es un filtro pasa bajas y  $h_\psi$  un filtro pasa altas. Por lo tanto, la descomposición *wavelet* discreta se puede hacer de manera iterativa utilizando un banco de filtros de análisis *wavelet*.

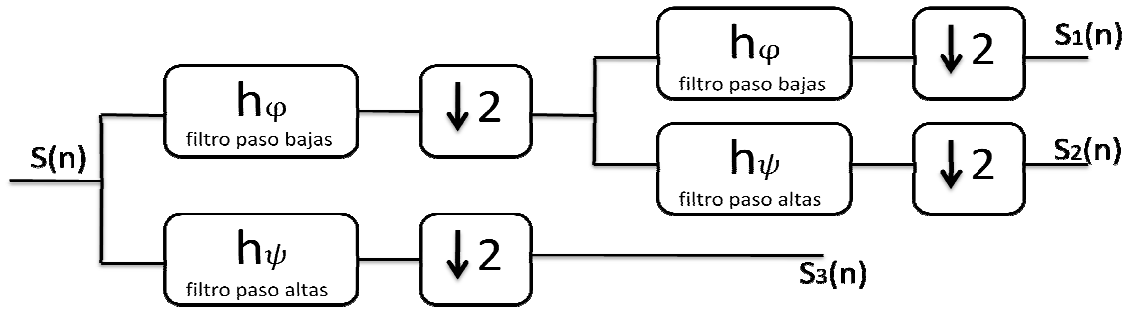


Figura 3.5 Banco de filtros *wavelet* 1-D.

Luego de realizar el filtrado tenemos la mitad del ancho de banda de la señal en cada canal. Es por ello que podemos hacer un submuestreo por 2 sin perder información. En el primer nivel de descomposición, la secuencia original de tamaño  $M$  es descompuesta en dos secuencias de tamaño  $\frac{M}{2}$  cada una.

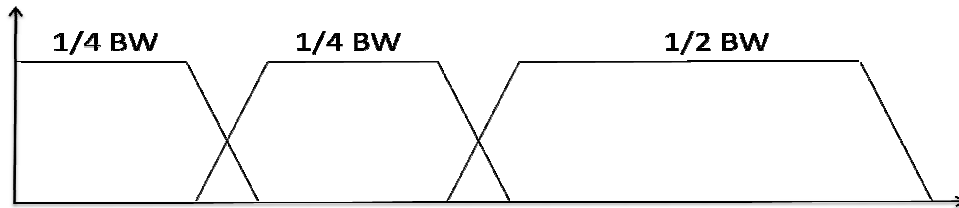


Figura 3.6 Respuesta al impulso del banco de filtros pasa banda. Se divide el ancho de banda original  $BW$ .

La síntesis *wavelet* se realiza de manera inversa a la descomposición. Cada canal pasa banda sobremuestreado es filtrado con su respectivo filtro de síntesis. De manera gráfica se muestra el proceso en la figura 3.7.

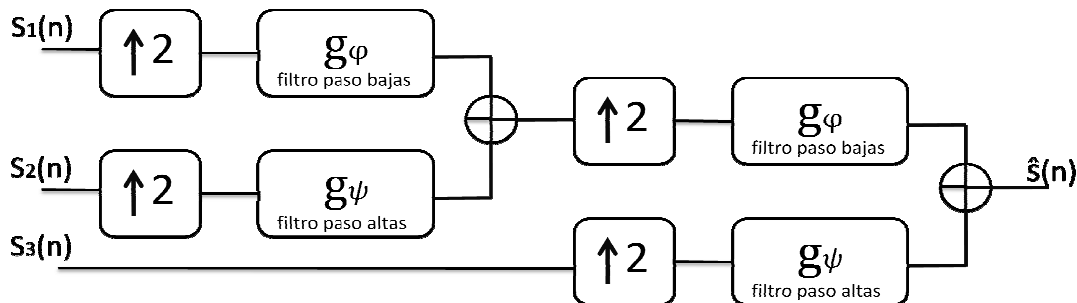


Figura 3.7 Síntesis *wavelet* 1-D.



3.3.2 Transformada discreta wavelet en 2-D

Para el caso de señales bidimensionales, como las imágenes digitales, la transformada discreta wavelet se realiza aplicando el filtro paso bajas  $h_\varphi$  y el paso altas  $h_\psi$  en la dirección horizontal  $x$  y vertical  $y$  de la imagen.

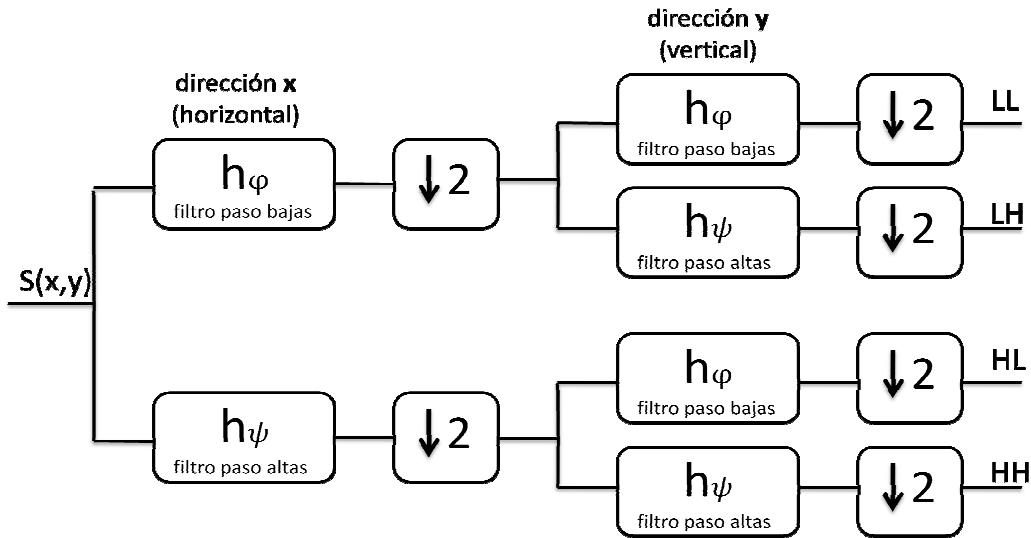


Figura 3.8 Análisis wavelet en 2-D. “L” denota banda de paso baja, “H” denota banda de paso alta.

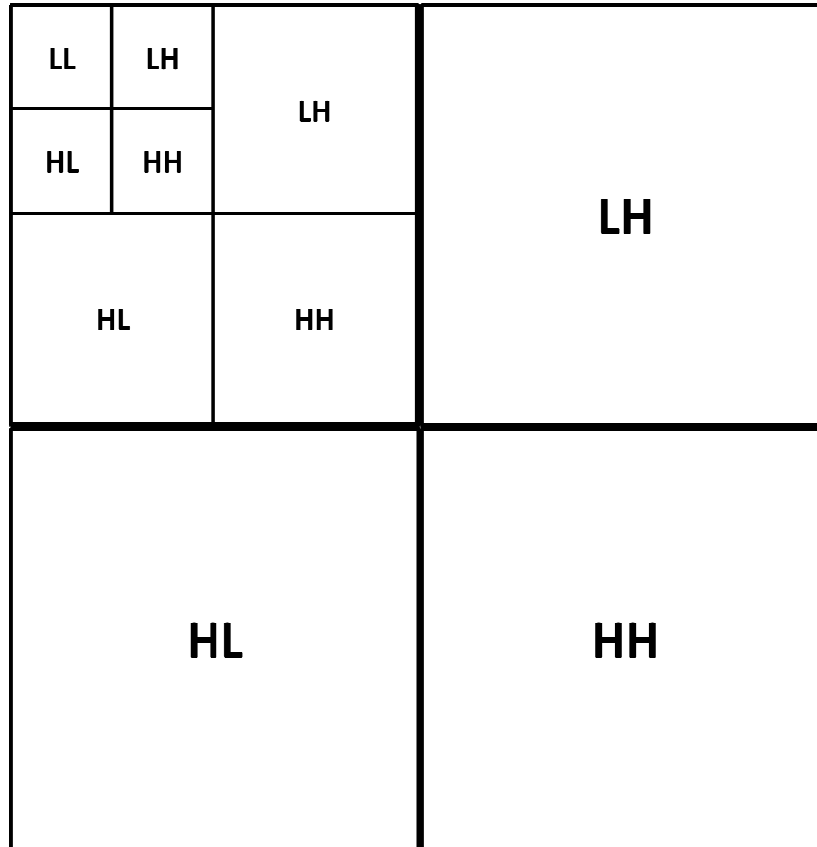
Al igual que en el caso unidimensional, luego del filtrado, las muertas redundantes hacen posible realizar el decimado por 2 sin perder información.

La DWT aplicada a imágenes en cada nivel de descomposición nos da 4 versiones reducidas en resolución de la imagen original, cada una caracterizada por la combinación de los filtros paso bajas y paso altas. Dicha combinación permite extraer los bordes horizontales, verticales y diagonales de la imagen como se muestra en la tabla 3.1.

LL	baja-baja
LH	bordes horizontales
HL	bordes verticales
HH	bordes diagonales

Tabla 3.1

Como las componentes de baja frecuencias son abundantes en las imágenes, las descomposiciones subsecuentes se realizan usualmente en la versión LL.



**Figura 3.9** Análisis *wavelet* en imágenes.

La transformada *wavelet* permite capturar los detalles horizontales, verticales y diagonales de una imagen, sin embargo no es buena para capturar contornos en otras direcciones.



Figura 3.10 Descomposición *wavelet* de la imagen "Bárbara" a nivel 3.

### 3.4 Transformada *contourlet*

Supóngase que se desea construir un contorno suave como el que se muestra en la figura 3.11. Se puede hacer a partir de *wavelets* 2-D pero con la limitante de utilizar "formas base" rectangulares. En resoluciones cada vez más finas se observa que se requiere un gran número de rectángulos para capturar el contorno y resulta evidente la deficiencia de las *wavelets* en 2-D. Sin embargo, si se utilizan rectángulos de diferentes tamaños que puedan ser orientados en diferentes direcciones se puede representar dicha curva suave a partir de un pequeño número de rectángulos inclusive en resoluciones finas. Esto es precisamente lo que realiza la Transformada *contourlet* como se verá más adelante.

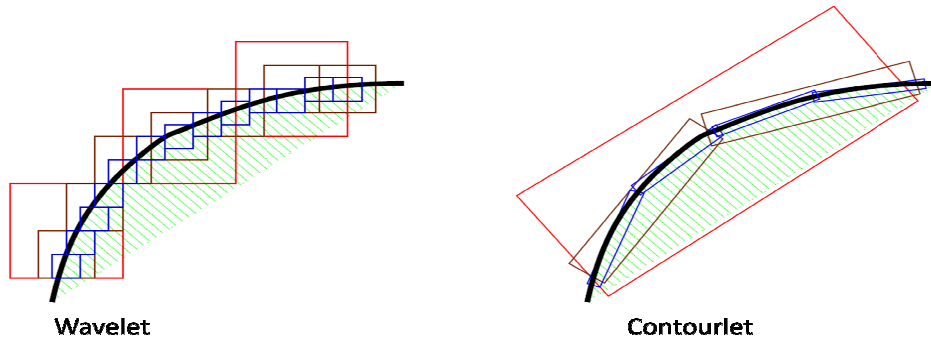


Figura 3.11 Aproximación del contorno suave a diferentes resoluciones [16].

Existe un conjunto de características deseables en la representación de imágenes, las cuales se mencionan enseguida [17]:

**Multiresolución:** Se debe de poder aproximar las imágenes desde una versión “tosca” hasta una versión de resolución fina.

**Localización:** Los elementos base de la imagen deben estar bien localizados en el dominio espacial y de la frecuencia.

**Muestreo crítico:** La representación debe realizarse con la menor redundancia posible.

**Direccionalidad:** Las funciones base deben de poder ser orientadas en un gran número de direcciones diferentes.

**Anisotropía:** Las funciones base deben tener diferentes formas y relación de aspecto para poder capturar los contornos.

A diferencia de la representación *wavelet* que sólo cumple con las tres primeras características, la representación *contourlet* cumple con las cinco características.

La transformada *contourlet* realiza una descomposición direccional multiresolución. Utiliza la pirámide Laplaciana para hacer el análisis multiresolución y posteriormente el canal pasa-banda es descompuesto en diferentes direcciones. Enseguida se mostrará más a detalle cómo se realiza la descomposición *contourlet*.

### 3.4.1 Pirámide Laplaciana

La pirámide Laplaciana [18] es utilizada en la transformada *contourlet* para realizar la descomposición multiresolución de la imagen. Primero la imagen es filtrada con un filtro Gaussiano pasobajas seguida de un submuestreo obteniendo como resultado una versión de menor resolución de la imagen original (aproximación). Es posible dicha reducción en la frecuencia de muestreo pues previamente se ha reducido la frecuencia máxima de la imagen con el filtrado pasobajas. Sobremuestreando la aproximación se obtiene una predicción de la imagen original. A partir de la diferencia entre la imagen original y la predicción se obtiene una versión pasobanda o de error que contiene los detalles de la imagen. Se puede seguir la descomposición de manera iterativa con la versión pasobajas. La pirámide Laplaciana se construye con las versiones pasobanda que se van obteniendo.

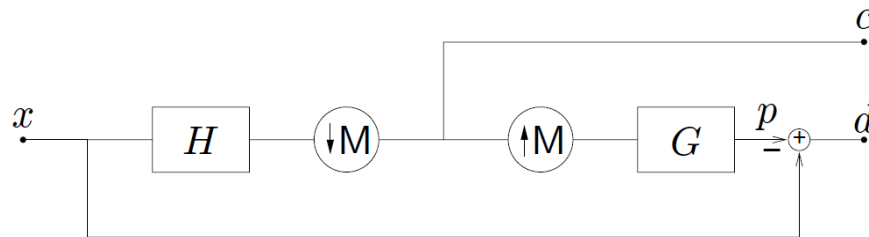


Figura 3.12 Esquema de la pirámide Laplaciana [17]. “*H*” es el filtro pasobajas de análisis y “*G*” el de síntesis. “ $\downarrow M$ ” representa el submuestreo y “ $\uparrow M$ ” el sobremuestreo.

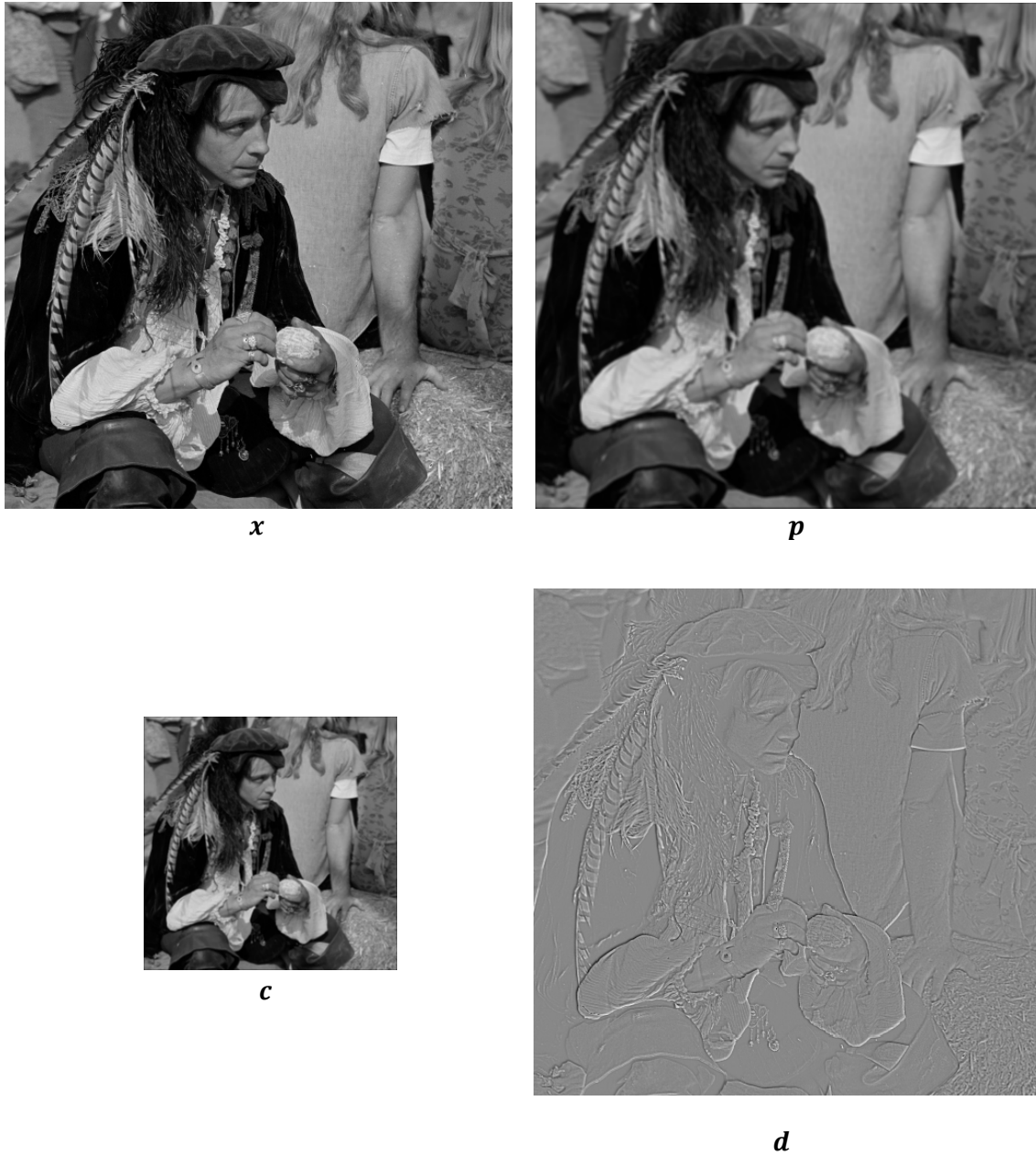
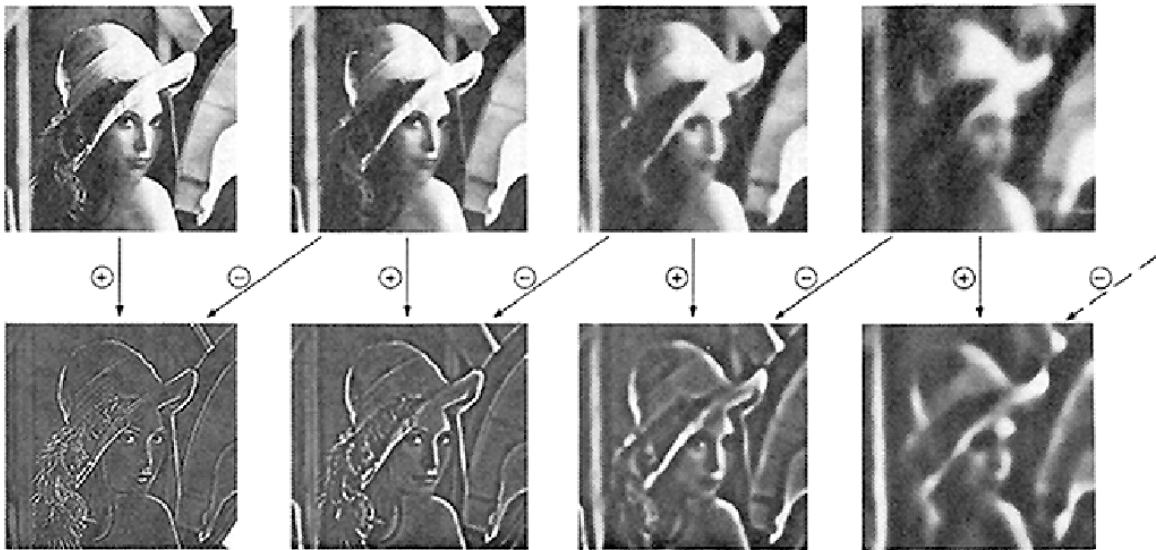


Figura 3.13 Se obtiene la aproximación “ $c$ ” a partir de la imagen original “ $x$ ” filtrada paso bajas por “ $H$ ” y submuestreada por “ $M$ ”. Sobremuestreando “ $c$ ” y luego de un filtro de síntesis se obtiene la predicción “ $p$ ”. La diferencia entre la imagen original “ $x$ ” y la predicción “ $p$ ” nos da los detalles “ $d$ ”.

La Pirámide Laplaciana genera sólo una señal pasobanda en cada nivel de descomposición.



**Figura 3.14** Primeros 4 niveles de de la pirámide Gaussiana y Laplaciana [18]. Pirámide Gaussiana columna superior, pirámide Laplaciana columna inferior.

La reconstrucción se realiza a partir de la suma de la aproximación y la versión pasobanda.



**Figura 3.15** Síntesis Laplaciana.  $\hat{x}$  es la imagen reconstruida [17].

### 3.4.2 Multiresolución

El análisis realizado con la pirámide Laplaciana se trata de un sistema multiescala. El filtro paso bajas  $G$  en cada nivel de la pirámide Laplaciana corresponde con la función de escalamiento  $\varphi(t) \in L^2$ . La función  $\varphi(t)$  puede ser escalada por el parámetro  $j$  y “recorrida” por el parámetro  $n$  [17].

$$\varphi_{j,n} = 2^{-j}\varphi(2^{-j}t - n) \quad (3.10)$$

donde:  $j \in \mathbb{Z}, n \in \mathbb{Z}^2$ .

Sea el conjunto de funciones  $\{\varphi_{j,n}\}$  una base ortonormal del subespacio  $V_j$ . El análisis multiresolución está representado por una secuencia de subespacios vectoriales anidados  $\{V_j\}_{j \in \mathbb{Z}}$  cuya intersección es el conjunto vacío y unión es  $L^2$ .

$$\dots V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \dots \quad (3.11)$$

$$\bigcup_{j \in \mathbb{Z}} V_j = L^2$$

$$\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$$

Es decir, cada espacio vectorial  $V_j$  está contenido dentro de un espacio vectorial de mayor resolución  $V_{j-n}$ , lo cual implica que cualquier vector dentro del espacio  $V_j$  se puede representar como la combinación lineal de los vectores de un espacio  $V_{j-n}$  de resolución superior.

$$\varphi_{0,n} = \varphi(t) = 2 \sum_n g[n] \varphi(2t - n) = \sum_n g[n] \varphi_{-1,n} \quad (3.12)$$

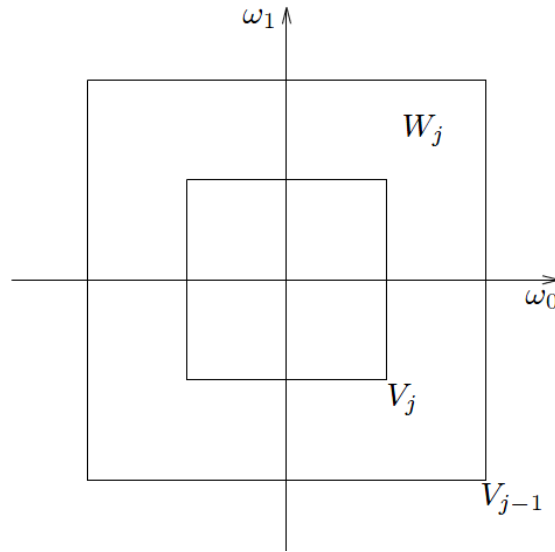
donde:  $n \in \mathbb{Z}^2$  es el parámetro de corrimiento y  $g[n]$  es el coeficiente asociado a cada  $n$ .

Como se muestra en la ecuación 3.12, la función de escalamiento  $\varphi_{0,n}$  en el subespacio  $V_0$  corresponde con la suma de versiones ponderadas por  $g[n]$  y trasladadas de la función de escalamiento  $\varphi_{-1,n} = 2\varphi(2t - n)$  en el subespacio  $V_{-1}$ .

Sea  $W$  la diferencia entre dos subespacios consecutivos. La unión del subespacio  $V_j$  y  $W_j$  da como resultado el subespacio de escala  $j - 1$ :

$$V_{j-1} = V_j \oplus W_j \quad (3.13)$$





**Figura3.16 Subespacios multiescala generados por la pirámide Laplaciana [17].**

Por lo tanto,  $W_j$  es el complemento ortogonal de  $V_j$ . Esto implica que dentro de  $W_j$  existen funciones  $\psi_{j,n}$  que cubren dicho subespacio diferencia y que se pueden construir a partir de la combinación lineal de funciones de escalamiento de su siguiente subespacio:

$$\psi^{(i)}(t) = 2 \sum_n f_i[n] \varphi(2t - n) \quad (3.14)$$

$$\psi_{j,n}^{(i)} = 2^{-j} \psi^{(i)}(2^{-j}t - n) \quad (3.15)$$

donde:  $j \in \mathbb{Z}, n \in \mathbb{Z}^2$ .

Sin embargo en este caso los coeficientes  $f_i[n]$  por ser el complemento de  $g[n]$  corresponden a un filtro paso altas, cada uno asociado a cada función  $\psi^{(i)}(t)$ .

El desarrollo anterior se relaciona con la pirámide Laplaciana de la siguiente manera. Supóngase que  $f$  es una función en el subespacio vectorial  $V_{j-1}$  representada a partir de la combinación lineal de las funciones base de  $V_{j-1}$ :

$$f(t) = \sum_n c^{(j-1)}[n] \varphi_{j-1,n}(t) \quad (3.16)$$

donde  $c^{(j-1)}[n]$  son los coeficientes.

Dado que las funciones de escalamiento  $\varphi_{j-1,n}$  son funciones ortogonales base de  $V_{j-1}$  los coeficientes se calcula a partir de:

$$c^{(j-1)}[n] = \langle f, \varphi_{j-1,n} \rangle \quad (3.17)$$

El producto interno de  $f$  en escala  $j - 1$  con funciones base de la siguiente escala  $j$  está dado por:

$$c^{(j)}[n] = \langle f, \varphi_{j,n} \rangle = \sum_k c^{(j-1)}[k]g[k - Mn] \quad (3.18)$$

$$d_i^{(j)}[n] = \langle f, \psi_{j,n}^{(i)} \rangle = \sum_k c^{(j-1)}[k]f_i[k - Mn] \quad (3.19)$$

donde  $k \in \mathbb{Z}^2$  y  $M$  es la matriz diagonal de submuestreo.

Se observa que las ecuaciones 3.18 y 3.19 corresponden a un filtrado multitasa pues la secuencia de entrada  $c^{(j-1)}$  se convoluciona con el filtro paso bajas para obtener  $c^{(j)}[n]$  y con el filtro paso altas para obtener  $d^{(j)}[n]$ . El cambio de escala se debió al submuestro  $M$  implícito dentro del proceso de filtrado. Por lo tanto, la salida de la pirámide Laplaciana para la secuencia de entrada  $c^{(j-1)}$  es  $\{c^{(j)}[n], d^{(j)}[n]\}$ .

Lo anterior aplicado al análisis de imágenes mediante la pirámide Laplaciana implica que una imagen de tamaño  $N \times N$  de escala  $j - 1$  introducida en las ecuaciones 3.18 y 3.19 nos dará como resultado dos imágenes  $c$  y  $d$  de tamaño  $\frac{N}{2} \times \frac{N}{2}$  y escala  $j$ , siendo  $c$  la versión pasobajas y  $d$  los detalles, es decir la versión paso altas. Este es el análisis multiresolución que se realiza con la pirámide Laplaciana.

### 3.4.3 Banco de filtros direccionales

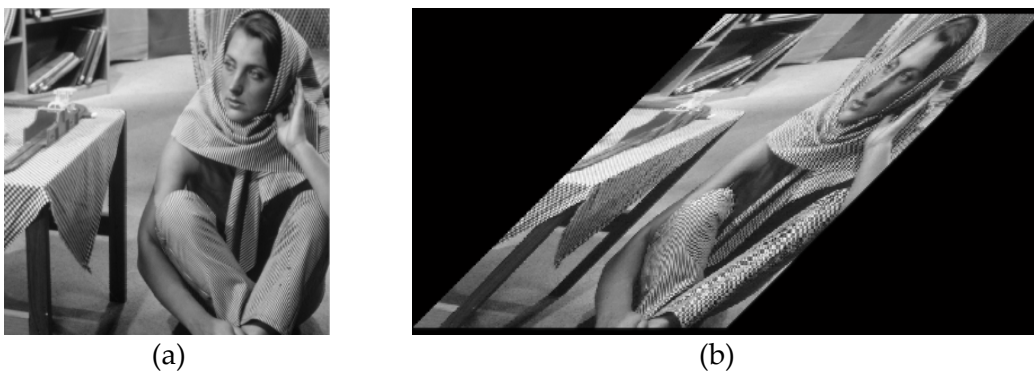
El análisis direccional que realiza la transformada *contourlet* se efectúa mediante un banco de filtros direccionales DFB (*Directional Filter Bank*), proceso que se puede iterar para la descomposición en diferentes niveles.

En sistemas multidimensionales multitasas las operaciones de muestreo se definen sobre matrices. En el caso del banco de filtros direccionales que implementa la transformada *contourlet* se tratan de matrices *quincunx*. Reciben ese nombre por asemejar una estructura geométrica en forma de diamante con cuatro aristas y un quinto punto central. Dichas matrices son utilizadas para realizar un mapeo lineal de las coordenadas de cada pixel que compone la imagen a una nueva posición dando como resultado la rotación de la misma. En la ecuación 3.20 se muestra la forma en que se realizan estas transformaciones lineales. También son conocidas como operaciones de remuestreo pues es posible realizar el submuestreo o sobremuestreo de la imagen a partir de dichas matrices.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (3.20)$$

donde  $(x, y)$  son las coordenadas de cada pixel de la imagen antes de la operación lineal y  $(x', y')$  las coordenadas luego de aplicar el operador lineal.

Para poder realizar la descomposición direccional se requiere de matrices que efectúen operaciones de rotación.



**Figura 3.17** Ejemplo de imagen remuestreada. (a) Imagen de “Bárbara”. (b) “Bárbara” remuestreada por la matriz  $R = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ .

En realidad la matriz de muestreo *quincunx*  $Q$  utilizada en el banco de filtro *quincunx* QFB

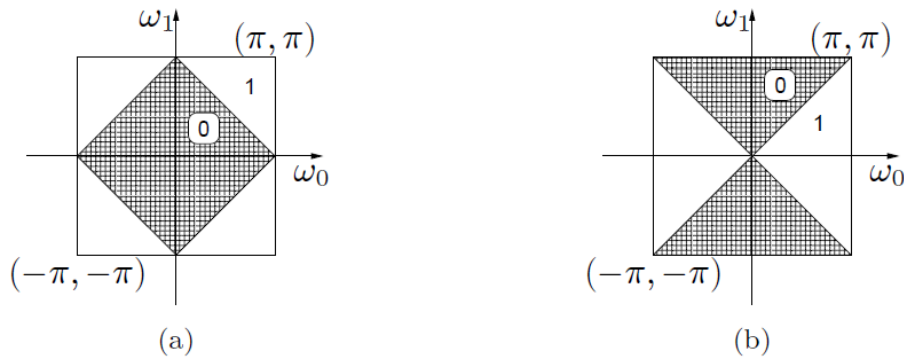
(*Quincunx Filter Bank*) puede ser  $Q_0 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$  ó  $Q_1 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ , cada una

correspondiente a la rotación de la imagen de entrada  $-45^\circ$  y  $+45^\circ$  respectivamente [17], correspondiendo el valor positivo al sentido de giro de las agujas del reloj.



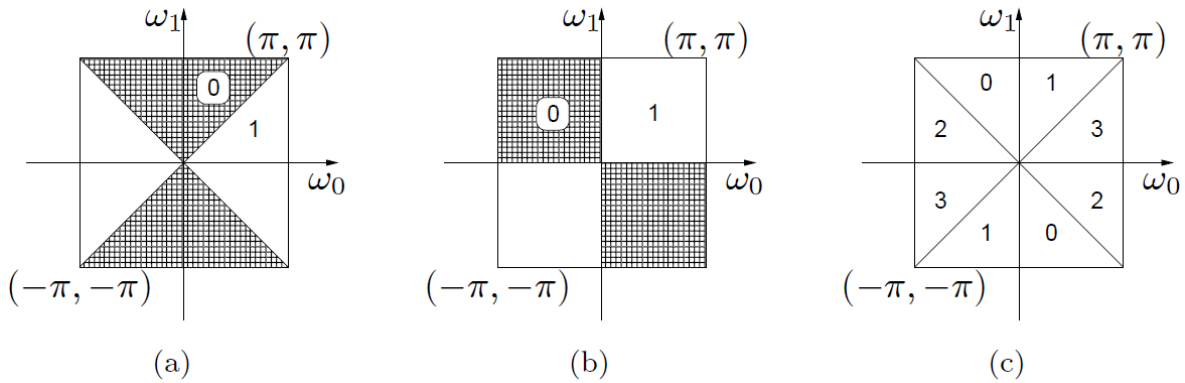
**Figura 3.18** Imagen rotada por matriz de muestreo *quincunx*  $Q_0$ .

El banco de filtros *quincunx* se puede utilizar para dividir el espectro en un canal paso altas y otro paso bajas utilizando un par de filtros en forma de diamante o en un canal horizontal y vertical utilizando un par de filtros de abanico, como se muestra en la figura 3.19.



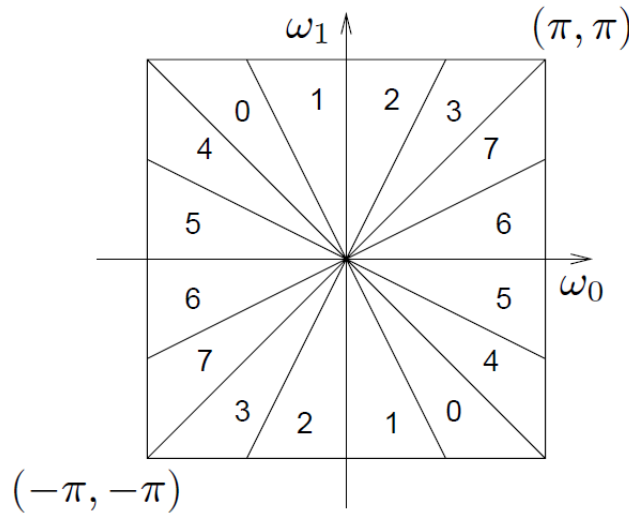
**Figura 3.19** Respuesta en frecuencia de (a) filtro en forma de diamante (b) filtro en forma de abanico. [17]

El banco de filtros direccionales propuesto por Minh Do y Martin Vetterli [17] se basa en el uso de un banco de filtros *quincunx* con filtros de abanico combinados con operaciones de rotación hechas mediante remuestreo logrando como resultado un filtro con respuesta en frecuencia dividida en cuadrantes, como se muestra en la figura 3.20 (b). Combinando este filtro con el filtro en forma de abanico se obtiene la partición de frecuencia mostrada en la figura 3.20 (c).



**Figura 3.20 (a) Filtro de abanico. (b) Filtro de cuadrantes. (c) cuatro direcciones de descomposición [17].**

El resultado es una descomposición direccional en  $l$  niveles que genera  $2^l$  subbandas con una partición en frecuencia como la que se muestra en la figura 3.21.



**Figura 3.21 Partición direccional en frecuencia para  $l=3$  [17].**

Las matrices de submuestreo empleadas en la descomposición direccional son matrices diagonales cuyos elementos dependen del nivel de descomposición  $l$ .

$$S_k^{(l)} = \begin{cases} \text{diag}(2^{l-1}, 2) & \text{para } 0 \leq k < 2^{l-1} \\ \text{diag}(2, 2^{l-1}) & \text{para } 2^{l-1} \leq k < 2^l \end{cases} \quad (3.21)$$

Supóngase que el DFB utiliza filtros ortogonales con  $l$  niveles. La familia de filtros direccionales  $g_k^{(l)}[n]$ :

$$\{g_k^{(l)}[m - S_k^{(l)}n]\}_{0 \leq k < 2^{l-1}, n \in \mathbb{Z}^2} \quad (3.22)$$

es una base ortonormal del espacio  $l^2(\mathbb{Z}^2)$ , donde  $k = 0, \dots, 2^{l-1} - 1$  corresponde con las direcciones dentro de  $[-45^\circ, +45^\circ]$ ;  $k = 2^{l-1}, \dots, 2^l - 1$  corresponde con las direcciones comprendidas dentro del intervalo  $[+45^\circ, +135^\circ]$ .

Aplicando el banco de filtros direccionales a los subespacios multiresolución  $V_j$ :

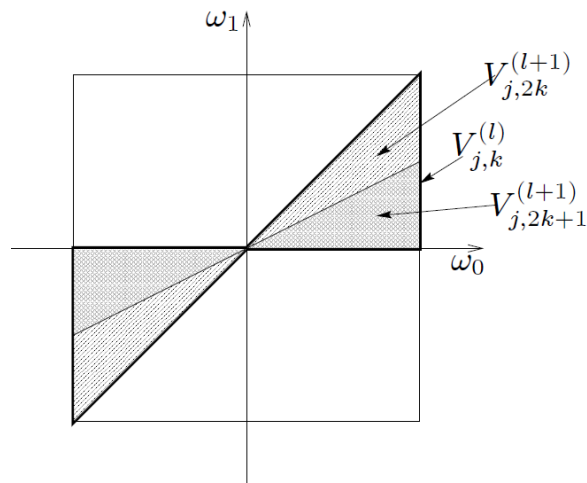
$$\theta_{j,k,n}^{(l)}(t) = \sum_{m \in \mathbb{Z}^2} g_k^{(l)}[m - S_k^{(l)}n] \varphi_{j,m}(t) \quad (3.23)$$

La familia  $\{\theta_{j,k,n}^{(l)}(t)\}_{n \in \mathbb{Z}^2}$  es una base ortonormal del subespacio direccional  $V_{j,k}^{(l)}$  para  $k = 0, \dots, 2^l - 1$ . Estos subespacios son ortogonales y además:

$$V_{j,k}^{(l)} = V_{j,2k}^{(l+1)} \oplus V_{j,2k+1}^{(l+1)} \quad (3.24)$$

$$V_j = \bigoplus_{k=0}^{2^l-1} V_{j,k}^{(l)} \quad (3.25)$$

Es decir, el subespacio multiresolución  $V_j$  se divide en los subespacios direccionales  $V_{j,k}^{(l)}$ .



**Figura 3.22** Subespacios multidireccionales generados por el banco de filtros direccionales [17].

Aplicando la descomposición de la familia de filtros direccionales  $g_k^{(l)}[n]$  al subespacio  $W_{j+1}$  definido previamente, sea:

$$\mu_{j,2n+k_i}(t) = \psi_{j+1,n}^{(i)} \quad (3.26)$$

$$\rho_{j,k,n}^{(l)}(t) = \sum_{m \in \mathbb{Z}^2} g_k^{(l)}[m - S_k^{(l)}n] \mu_{j,m}(t) \quad (3.27)$$

La familia  $\{\rho_{j,k,n}^{(l)}\}_{n \in \mathbb{Z}^2}$  es un marco ajustado del subespacio  $W_{j+1,k}^{(l)}$  es decir, es una base de dicho subespacio. Además:

$$W_{j+1,k}^{(l)} = W_{j,2k}^{(l+1)} \oplus W_{j,2k+1}^{(l+1)} \quad (3.28)$$

$$W_{j+1} = \bigoplus_{k=0}^{2^l - 1} W_{j+1,k}^{(l)} \quad (3.29)$$

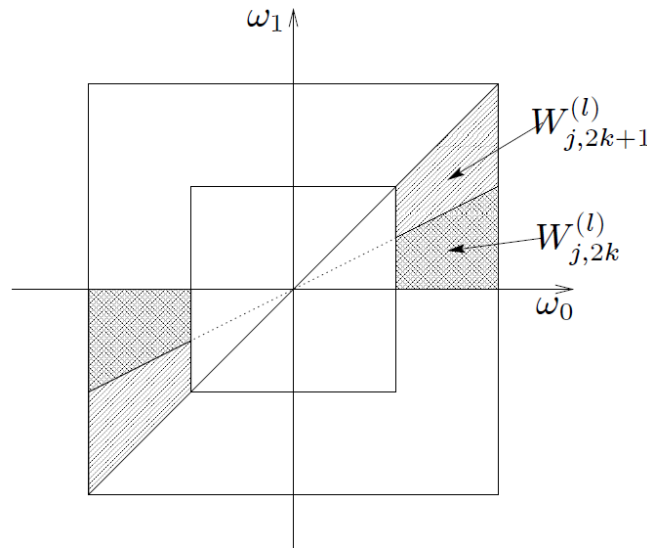
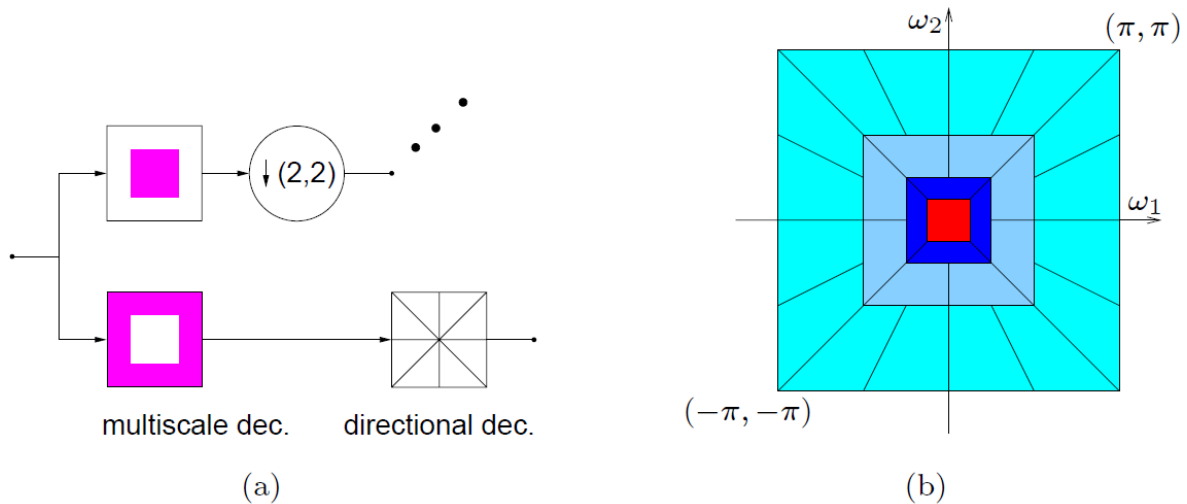


Figura 3.23 Subespacios generados por la descomposición multiescala multidirección [17].

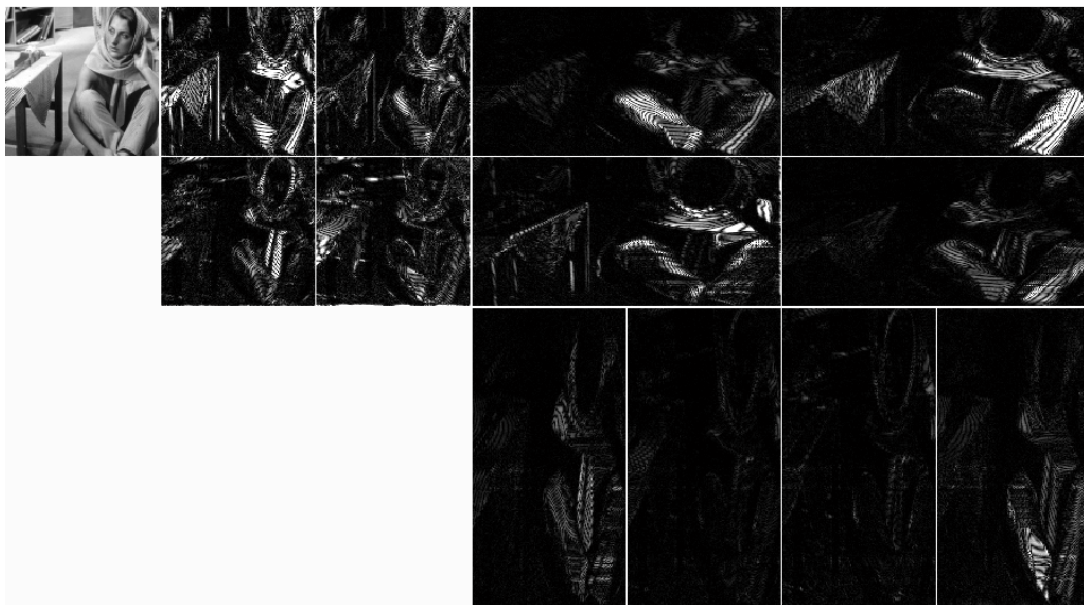
### 3.4.3 Descomposición Direccional y Multiescala

Realizando el análisis direccional al canal paso banda de cada nivel de la pirámide Laplaciana se obtiene la descomposición direccional piramidal a través de bancos de filtros

direccionales piramidales PDFB (*Piramidal Directional Filter Bank*). Se realiza el análisis direccional al canal paso banda por ser este el que contiene las componentes de alta frecuencia que representan direccionalidad en las imágenes. El resultado es una descomposición direccional en múltiples escalas.



**Figura 3.24** Descomposición piramidal direccional. (a) Se realiza la descomposición piramidal; la descomposición direccional es aplicada solamente al canal paso altas. (b) División en frecuencia resultante: el número de direcciones aumenta con la frecuencia.



**Figura 3.25** Descomposición *contourlet* de la imagen "Bárbara" con 2 niveles de descomposición. 4 direcciones en el primer nivel y 8 direcciones en el segundo nivel.



Como se observa en la figura 3.25 la transformada *contourlet* realiza una descomposición multiresolución y direccional. Los coeficientes de mayor valor son los que se muestran en blanco. La transformada *contourlet* captura las componentes de alta frecuencia en diferentes direcciones y permite identificar texturas presentes en la imagen por lo que es adecuada para capturar contornos.

### 3.5 Espectro disperso

Basados en [18] y [1], adoptamos una técnica aditiva de espectro disperso para la inserción de la marca de agua en el dominio *contourlet*. Cada bit de la marca de agua codificada (ver 2.10.4) es multiplicado por un código binario pseudoaleatorio arreglado en forma matricial y añadido en alguna de las bandas resultantes a partir del análisis *contourlet*.

$$Y' = Y + \gamma \sum_{i=1}^h w_i P_i \quad (3.30)$$

donde  $\bar{w} = (w_1, w_2, \dots, w_h)$  es la marca de agua codificada de tamaño  $h$ , nótese que los “0s” se han cambiado por -1,  $w_i \in \{-1, 1\}$ ,  $i = \{1, 2, \dots, h\}$ ,  $Y$  es una matriz de tamaño  $m \times n$  cuyos elementos contienen los coeficientes *contourlet* de la banda en donde se incrusta la marca de agua,  $Y'$  es la matriz de tamaño  $n \times m$  con los coeficientes de la banda marcada,  $\{P_1, P_2, \dots, P_h\}$  es el conjunto de matrices (idealmente no correlacionadas entre sí) que contienen los “chips” de las secuencias pseudoaleatorias binarias, cada matriz es de tamaño  $n \times m$ , es decir, del tamaño de la banda en donde se incrusta la marca de agua y  $\gamma$  es un parámetro de escala que controla la intensidad de la marca de agua. Las secuencias pseudoaleatorias son creadas a partir de una clave que sirve como semilla para generarlas, por lo que en la recuperación se requiere de la clave para poder generar las secuencias correctas para su detección. De esta forma se logra expandir cada bit de la marca de agua codificada en toda la banda donde se desea incrustarla. Finalmente se obtiene la imagen marcada a partir de la síntesis *contourlet*.

La técnica de espectro disperso es utilizada por brindar a la marca de agua una mayor robustez, pues al esparcir la señal en toda una banda de la descomposición *contourlet* el ruido tendría que estar presente en toda la banda para poder producir bits en error en la detección. Sin embargo, no sólo el ruido producido por ataques afecta a la marca de agua, también la misma banda *contourlet* interfiere con la marca de agua aditiva. Existe una técnica llamada *espectro disperso mejorado* que reduce la interferencia entre la banda y la marca de agua, como se describirá más adelante. Explotando la descomposición multiescala y multidirección de la transformada *contourlet* junto con el conocimiento del sistema visual humano, como se verá en el siguiente capítulo, se puede conseguir mayor robustez y menor perceptibilidad de la marca de agua.