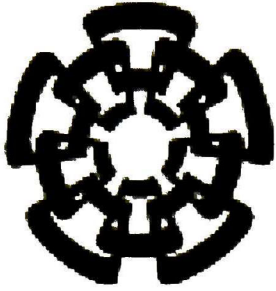


XX(104961.1)

CLASIF.	
ADQUIS.	SSI-210
FECHA	27 NOV-2002
PROGEO.	TESIS-02
	\$



CINVESTAV

Centro de Investigación y de Estudios Avanzados del IPN
Unidad Guadalajara

**CINVESTAV I.P.N.
SECCION DE INFORMACION
Y DOCUMENTACION**

TRANSFORMADA CUATERNIÓNICA WAVELET: TEORÍA Y APLICACIONES

**TESIS QUE PRESENTA
JUAN FRANCISCO ALVARADO CASAS**

**PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS**

**EN LA ESPECIALIDAD DE
INGENIERÍA ELÉCTRICA**

**CINVESTAV
I.P.N.
ADQUISICION
DE LIBROS**

Guadalajara, Jal. Agosto de 2002

TRANSFORMADA CUATERNIÓNICA WAVELET: TEORÍA Y APLICACIONES

Tesis de Maestría en Ciencias
Ingeniería Eléctrica

Por:

Juan Francisco Alvarado Casas

Ing. en Sistemas Computacionales
Instituto Tecnológico y de Estudios Superiores de
Occidente

Becario de CONACYT, expediente No. 157954

Director de Tesis:
Dr. Eduardo José Bayro Corrochano.

CINVESTAV del IPN Unidad Guadalajara, Junio del 2001

Agradecimientos

Agradezco a mis padres, a mi asesor, a mis amigos,

Un agradecimiento especial al Conacyt.

Becario de Conacyt No. 157954.

Índice General

1	Introducción	3
1.1	Motivación .	5
1.1.1	Objetivos	5
1.2	Estado del Arte	5
1.3	Estructura de la Tesis	6
2	Fundamentos	9
2.1	El Álgebra Cuaterniónica .	9
2.1.1	El ángulo/fase cuaterniónico	15
2.2	Enlace Tiempo-Frecuencia	18
2.2.1	La Transformada de Fourier	18
2.2.2	La Transformada de Fourier por ventanas	20
2.2.3	La Transformada Wavelet	23
3	Transformada Cuaterniónica Wavelet	31
3.1	Filtros Cuaterniónicos de Gabor	31
3.1.1	Fase local Cuaterniónica	33
3.2	Análisis de imágenes usando la TQW	35
3.3	Complejidad Algorítmica de la TQW	39
4	Aplicaciones	41
4.1	Estimación del Flujo Optico	42
4.1.1	El Problema	42
4.1.2	Métodos para resolver el problema de estimación de disparidad	43
4.1.3	Estimación usando la Fase cuaterniónica	43
4.1.4	Estimación de la confianza	44
4.2	Método utilizando la TQW y la Fase cuaterniónica	46
4.3	Experimentos	50
4.3.1	Datos sintéticos	50
4.3.2	Datos reales	53
4.4	Comparación con otros métodos y discusión	55
5	Conclusiones	61

Apéndices	63
.1 Complementos Matemáticos	63
.2 Imágenes de la Transformada Cuaterniónica Wavelet	64
.2.1 Flujo óptico del cubo Rubik.	71

Índice de Figuras

1.1	La primer wavelet la función de Haar.	4
2.1	La energía de G es medida por su ancho de banda $\Delta\omega$ y la amplitud máxima A de los lóbulos laterales, localizados en $\omega = \pm\omega_0$	22
2.2	Gráficos de cuatro ventanas g con soporte $[-1/2, 1/2]$	23
2.3	Cobertura tiempo-frecuencia con una wavelet de Daubechies	26
2.4	Izquierda, Piramide Wavelet de una imagen $f(x,y)$. Derecha, Diagrama de flujo del analisis del nivel $j - 1$ al nivel j	28
2.5	a). Función Wavelet, b) Función de escalamiento. Ambas funciones complejas.	29
3.1	Filtro de Gabor Cuaterniónico con $c_1 = c_2 = 3$ y $\sigma = \pi/6$	34
3.2	La función $f(x, y) = \cos(x)\cos(y)$ con $(x, y) \in [0, 4\pi] \times [0, 3\pi[$ izquierda y $(x, y) \in [0, 2\pi] \times [0, \pi[$ derecha. Figura Generada por Bülow [Bul99]	35
3.3	Dependencia de la tercer componente de fase ψ en la estructura local de la imagen. Figura Generada por Bülow [Bul99]	36
3.4	Piramide Wavelet de una imagen en 2 niveles.	37
3.5	Función $\sin(xu + yv)$ con $(u, v) = (\frac{\pi}{6}, 0), (\frac{\pi}{6}, \frac{\pi}{6}), (0, \frac{\pi}{6})$	38
3.6	Transformada cuaterniónica de la imagen mostrada en la figura (3.5).	39
4.1	Diagrama de la estimación del flujo óptico en un nivel	48
4.2	función $\sin(xu)\sin(yv)$ con $u = v = \frac{\pi}{6}$. a. Imagen original 80×80 . b. resultado de trasladar -3 pixeles en la dirección x la imagen original.	49
4.3	Disparidad horizontal y vertical para el nivel 1. El color negro corresponde a valores cero.	51
4.4	Mascara de confianza para el nivel 1. El Blanco es usado. El negro no.	51
4.5	Estimación del flujo óptico en el nivel 1 sin usar la mascara de confianza.	53
4.6	Estimación del flujo óptico una vez eliminadas las regiones con poca energía o cero.	54
4.7	Estimación final del Flujo óptico.	55
4.8	Estimación del flujo óptico en el nivel 1 de una rotación.	56
4.9	Imagenes del cubo Rubik usadas en los experimentos.	57
4.10	Mascara de Confianza nivel 1.	58
4.11	Estimación del flujo óptico nivel 1.	59

4.12 Flujo óptico del cubo Rubik (Final).	60
1 Estimación del flujo óptico en el nivel 1	70

NOTACIÓN

\mathbb{N}	Conjunto de números naturales
\mathbb{Z}	Conjunto de números enteros.
\mathbb{R}	Conjunto de números reales.
\mathbb{C}	Conjunto de números complejos.
\mathbb{H}	Conjunto de números cuaternionicos.
$f(x)$	función de x .
$F(x)$	Trasformada de Fourier de $f(x)$
$WF(x)$	Transformada wavelet de $f(x)$.
$F^q(x)$	Transformada Cuaterniónica Wavelet de $f(x)$.
ψ	Función wavelet.
ϕ	Función de escalamiento.
$\langle f, g \rangle$	Producto punto de dos funciones.
$\ f\ $	Norma de una función
$\exp(x)$	Funcion exponencial de x .
q	Cuaternion
h	Filtro pasa-bajos
g	Filtro pasa-altos
H_i^q	Filtro pasa-bajos cuaterniónico en el nivel i
G_i^q	Filtro pasa-altos cuaterniónico en el nivel i
$A_j f$	Aproximación de una función $f(x, y)$
$D_j f$	Diferencia entre dos aproximaciones A_{j-1} y A_j
$O(n)$	Orden de complejida algorítmica n .

Capítulo 1

Introducción

Basta de filósofos que interpreten al mundo! se necesitan filósofos que lo transformen.

Las transformaciones son esenciales para la vida, el cuerpo humano es una máquina compleja que subsiste por medio de transformaciones, la sinapsis, la respiración, la digestión, la vista, el oído, todo esta relacionado con transformaciones. En el mundo exterior también nos enfrentamos a muchas clases de transformaciones, desde transformaciones en el clima hasta en estructuras políticas y sociales. Vivimos rodeados de transformaciones y el campo de la ciencia no se puede quedar atrás. La palabra transformación proviene del latín *transformatio* – *ōnis* y es la acción y efecto de transformar. Transformar proviene de *transformāre* que se refiere a cambiar la forma de una cosa. En la ciencia el concepto va mas haya de un simple cambio de forma, implica cambio de espacio, de representación, de dimensión.

En el procesamiento de señales, una transformada es una nueva representación de la señal. Existen muchas bases para hacer una transformación Una de las mas famosas es la transformada de Fourier que lleva una señal del espacio del tiempo al de la frecuencia. Como la transformada de Fourier existen muchas otras clases de transformaciones. Estas nuevas transformaciones son mas locales. Un evento permanece conectado al tiempo en que ocurre, en lugar de transformar de una descripción de la señal en el tiempo por una descripción pura en la frecuencia, los nuevos métodos encuentran un balance entre el tiempo y la frecuencia. Es como una partitura musical, con frecuencias específicas en tiempos específicos. Donde el principio de la incertidumbre de Heisenberg se interpone en el camino a la perfección. Este principio se refiere al hecho de perder precisión en el tiempo cuando ganamos en la frecuencia. Los músicos no pueden cambiar entre frecuencias instantáneamente, pero nuestros ojos y oídos logran asignar un lugar así como una frecuencia. Las transformada wavelet busca cumplir el mismo propósito.

Aun cuando el concepto de wavelet no es una idea brillante y nueva, es el resultado del trabajo realizado en diferentes áreas de la ciencia, como son las matemáticas, la física y la ingeniería. La conexión existente entre estas áreas ha creado un flujo de nuevas ideas muy enriquecedor para la construcción de nuevas bases o transformadas. Por esta fusión de ideas, es difícil hablar de un “padre” de la teoría wavelet, tal vez

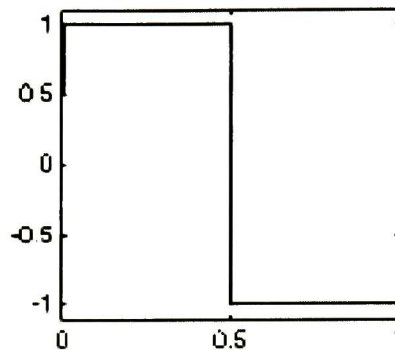


Figura 1.1: La primer wavelet la función de Haar.

mencionar a una sola persona como el fundador sería injusto para aquellos que han contribuido al desarrollo de esta teoría.

Estableceremos la historia del desarrollo de la teoría wavelet. La palabra wavelet fue utilizada por primera vez en el análisis de señales por el matemático Alfred Haar en el año de 1909, designando así a una función discontinua que se asemeja a una función escalon. Esta función se muestra en la figura (1.1). Toda la teoría detrás de la transformada wavelet viene de mucho mas atrás, del siglo XIX, del trabajo desarrollado por el matemático francés Joseph B. Fourier. La transformada wavelet (TW) fue olvidada en los libros por muchos años, las razones, similares a las de la transformada de Fourier fueron, el costo de implementación y la complejidad algorítmica. En el año de 1989 la TW resurge al ser programable y útil para la solución de problemas, gracias al desarrollo de un algoritmo basado en un banco de filtros por Yves Meyer y Stéphane Mallat, desde esa fecha se han desarrollado muchas bases, tal vez de las mas importantes son las wavelets de Daubechies, Meyer, Morlet, Shannon, y también están las Symlets, Coiflets, Biorthogonales, Gaussianas y función $\text{sinc}(x)$, por mencionar algunas.

La TW es aplicable en muchas áreas del procesamiento de señales. El FBI tiene mas de 30 millones de conjuntos de huellas digitales y el numero se incrementa a diario. Estas huellas necesitan ser digitalizadas y almacenadas para utilizarlas en la identificación de un sospechoso. La huella digital del sospechoso se envía a una terminal central donde se busca una correspondencia. Es posible que la transformada de Fourier sea útil y exitosa en esta aplicación. Pero con una compresión de 20:1 los patrones de la huella no podrían ser encontrados. Las líneas estarían quebradas en bloques de 8 por 8. La transformada de Fourier por ventanas introduciría muchos mas bloques. La transformada wavelet generan un panorama mejor. En 1994 el director del FBI eligió usar wavelets. Este es un ejemplo, pero la TW también es aplicada en discriminación de texturas sobre tejidos para el control de calidad en líneas de producción o para el detección de tumores. Las áreas de aplicación de la TW es muy amplio y se desarrollan

nuevas muy rápidamente.

1.1 Motivación

Extender a espacios de mayor dimensión o separar un problema en partes cuando no se puede encontrar una solución en el espacio actual es una tarea común en muchas áreas de la ciencia. Con esto en mente, la extensión de la transformada wavelet en términos de cuaterniones, busca descomponer una señal 2D en sus simetrías, de esta forma buscamos obtener una herramienta poderosa para el análisis de señales en 2 o 3 dimensiones. Resolviendo problemas donde otras herramientas no son útiles o mejorando la solución de métodos existentes.

Hacemos énfasis en el uso de filtros de Gabor. Los filtros de Gabor están bien localizados simultáneamente en el dominio del espacio y el de la frecuencia, además, son filtros linealmente invariantes al desplazamiento, lo que los hace muy convenientes de usar, ya que pueden ser aplicados convolucionando la señal con el filtro. En este trabajo se aprovechan las propiedades de los filtros cuaternionicos de Gabor.

1.1.1 Objetivos

Los objetivos que buscamos cumplir en esta tesis son:

1. Utilizar las propiedades de los cuaterniones como una forma de ampliar la transformada wavelet.
2. Utilizar las simetrías resultantes de la transformación cuaterniónica wavelet para obtener mas elementos en el análisis de señales, y utilizar estos elementos para la solución de problemas o la optimización de métodos existentes.
3. Explorar e identificar las ventajas y desventajas de la Transformada Cuaterniónica Wavelet (TQW) en el análisis de señales en comparación con la transformada wavelet real y compleja.
4. Usar la transformada cuaterniónica wavelet en experimentos, demostrando que es útil para la solución de problemas.

1.2 Estado del Arte

La transformada wavelet a tomado mucho auge en los últimos años y el número de aplicaciones en las que se aplica es cada vez mas grande. Los únicos trabajos sobre la transformada cuaterniónica wavelet son [Tra96] y [Tra01]. Un trabajo relacionado, que es citado en los anteriores y trata con transformaciones wavelet mas allá de los números complejos es [Mit94].

En [Mit94], Mitrea no habla de la wavelet cuaterniónica, pero nos parece importante mencionar este trabajo aquí por ser el primer trabajo que trata de extender la transformada wavelet a un plano mas complejos. Mitrea estudia las Clifford Wavelets, llevando la transformada a un punto de vista de números hipercomplejos empleando la wavelet de Haar.

Traversoni en [Tra96] desarrolla la transformada cuaterniónica wavelet usando un kernel cuaterniónico de Haar, la wavelet de Haar es la mas simple y presenta algunas deficiencias en aplicaciones reales. Esto no presenta un problema en el articulo porque no se exponen ningun ejemplo de la TQW. Sin embargo, este es el primer articulo en el que se describe la transformada cuaterniónica wavelet. Este articulo es teórico y no busca mas que la definición de la TQW de una forma totalmente matemática. El articulo en general es complicado y es dificil llevara la teoría a aplicaciones por el nivel de abstracción con el que se trabaja.

En [Tra01] Traversoni habla del análisis de imágenes utilizando la TQW, logrando esclarecer un poco el uso de la transformada wavelet cuaterniónica como herramienta para el procesamiento de señales. Sin embargo, se sigue usando un kernel de Haar, limitando la aplicación de la TQW a aplicaciones donde la wavelet de Haar siga dando resultados óptimos. Se muestra un pequeño ejemplo y no se habla de la forma en que fue implementado, ni se muestran resultados intermedios.

El panorama en materia de wavelet cuaterniónica esta poco explorado, son pocas las referencias existentes y estas se encuentran enteramente dedicadas al área de las matemáticas. Los artículos con muy abstractos y no se encuentra una aplicación orientada a un problema real que demuestre que la transformada cuaterniónica wavelet es programable y útil.

1.3 Estructura de la Tesis

El capitulo 2 trata todos los aspectos básicos relacionados con la transformada cuaterniónica wavelet, y esta enfocado para lectores que no tengan una idea clara del álgebra de cuaterniones, las transformaciones y los conceptos esenciales de filtros. El capitulo comienza con un breve estudio del álgebra de cuaterniones, se presentan las definiciones mas importantes o relevantes de esta álgebra para el desarrollo de la TQW, como la definición de la función exponencial para cuaterniones y el ángulo o fase cuaterniónica. Se describe la transformada de Fourier y sus teoremas principales como el de convolución de señales, por ser la base sobre la cual se desarrolla toda la teoría de la transformada wavelet. También, en este capitulo se describe la transformada de Fourier por ventanas, tratando de llevar al lector por el proceso natural e histórico a través del cual se llevo al desarrollo de la teoría wavelet. Toda esta experimentación empleando diferentes transformaciones es un resultado de la búsqueda por tratar de superar el problema de incertidumbre de Heisenberg, ver teorema (28) en el apéndice. En esta búsqueda llegamos a la definición de la transformada wavelet en la sección

(2.2.3), donde también se describe el análisis multiresolución que forma parte medular de la transformada discreta wavelet y del algoritmo de Mallat. Finalmente se dan los conceptos básicos de la transformada compleja wavelet y se expone uno de los kernels usados en esta transformación.

El capítulo 3 es la espina dorsal de esta tesis, se describen los conceptos teóricos de la transformada cuaterniónica y la forma como la TQW fue desarrollada. Se comienza por definir los filtros cuaterniónicos de Gabor en una y dos dimensiones, es el kernel que se usa en las aplicaciones. Se define la fase local cuaterniónica como la fase angular de la respuesta de un filtro cuaterniónico de Gabor. En la sección (3.2) se presenta un algoritmo similar al de Mallat para calcular la TQW por medio de un banco de filtros cuaternionico. Al final de este capítulo se hace un análisis de complejidad del algoritmo propuesto para el cálculo de la TQW. Este capítulo es fundamental para entrar al capítulo de aplicaciones.

El capítulo 4 muestra los resultados de la TQW en el análisis de imágenes. En este capítulo, se estudia el problema de estimación del flujo óptico. Primero se hace la descripción del problema de la estimación y se exponen varios métodos para resolverlo. Después se describe la estimación por medio de la fase cuaterniónica que es parte importante del método usado para resolver el problema del flujo óptico en esta tesis. En este capítulo se describe una forma de medir la confianza que es una medida para la fiabilidad de la estimación de la disparidad. La parte mas importante de este capítulo es la sección (4.2) donde se muestra el método completo para estimar el flujo óptico usando la TQW y la fase cuaterniónica. En la sección 4.3 se muestra muy brevemente algunos de los experimentos realizados en el transcurso de esta tesis, se eligieron los experimentos mas representativos. Desafortunadamente estos experimentos generan muchas imágenes que no pueden ser mostradas con el esquema de una tesis. Primero se realizo la estimación sobre datos sintéticos para tener características controladas y así poder encontrar errores en los algoritmos y métodos usados, una vez que se perfeccionaron los algoritmos se aplicó a varios ejemplos reales. El ejemplo mostrado aquí es tomado de una secuencia de imágenes de un cubo rubik montado sobre un plato giratorio. Este ejemplo es común en aplicaciones de estimación de movimiento y esa es la razón para mostrarlo aquí. Haciendo mas fácil la comparación con otros métodos en la ultima sección de este capítulo.

En el capítulo 5, exponemos las conclusiones y comentarios finales de este trabajo.

Capítulo 2

Fundamentos

Este capítulo inicia con un breve repaso del álgebra de cuaterniones, que es un álgebra de cuatro dimensiones no conmutativa. Este tema es extenso, por esta razón, en este capítulo solo se muestran las definiciones y teoremas que son de importancia para el desarrollo de la tesis. Muchos de los teoremas que usamos en esta sección están demostrados o referidos a demostraciones en textos especializados en la materia. Esta sección termina con la definición de fase angular cuaterniónica, esta definición es importante por el uso que hacemos de la fase cuaterniónica en las aplicaciones. La representación del álgebra que se muestra en este capítulo es la usada a lo largo de toda la tesis.

Las transformaciones juegan un papel muy importante en el campo del procesamiento de imágenes; por eso son el tema de la segunda parte de este capítulo. Se revisan en este capítulo la transformada de Fourier, la Transformada de Fourier por ventanas, la transformada wavelet continua y discreta, por último la transformada compleja wavelet.

La teoría de la transformada de Fourier y su significado para el procesamiento de señales, se puede encontrar en muchos libros, por lo tanto en esta sección solo se presentan los teoremas principales, las demostraciones son cortas y solo se presentan cuando es estrictamente necesario. Hacemos un repaso de la transformada wavelet por ventanas como la primer aproximación para eludir el problema de incertidumbre, sacando la transformada de Fourier de pequeñas partes de la señal. Después se expone el siguiente paso lógico, la transformada wavelet, que usa ventanas de tamaño variable, y con kernels que tiene diferentes propiedades. Por último se presenta un esbozo de la transformada compleja wavelet. El tema de transformaciones es sumamente interesante y aun cuando este capítulo es teórico, las aplicaciones de la TQW son interesantes y aplicables a muchos campos de la ciencia.

2.1 El Álgebra Cuaterniónica

Aquí se hace una revisión de las principales propiedades y definiciones del álgebra de cuaterniones y se introduce en la sección 2.1.1 una representación cuaterniónica de la

magnitud/fase. Para una revisión completa de la teoría de cuaterniones ver [Per97]. El Álgebra de cuaterniones fue introducida por W. R. Hamilton en 1843 [Ham66]. Es una álgebra de cuatro dimensiones denotada por \mathbb{H} en honor a Hamilton. La principal motivación para introducir el álgebra de cuaterniones es extender el significado geométrico de los números complejos en el plano 3D a un sistema de números hiper-complejos.

Los cuaterniones son números complejos generalizados de la forma

$$q = q_0 + iq_1 + jq_2 + kq_3$$

donde q_0, q_1, q_2, q_3 son números reales e i, j, k son unidades imaginarias generalizadas que obedecen las siguientes reglas

Definición 1 (Cuaterniones) Consideremos el conjunto de números $q = q_0 + iq_1 + jq_2 + kq_3$ y los elementos i, j, k obedeciendo las siguientes reglas de multiplicación

$$ij = -ji = k \quad y \quad i^2 = j^2 = k^2 = -1 \quad (2.1)$$

junto con la adición y la asociatividad generamos una álgebra asociativa no conmutativa, denotada por \mathbb{H} y llamada álgebra de cuaterniones.

Para los componentes de un cuaternion q algunas veces escribimos

$$q_0 = \mathcal{R}q, \quad q_1 = \mathcal{I}q, \quad q_2 = \mathcal{J}q, \quad q_3 = \mathcal{K}q$$

explícitamente estas operaciones pueden expresarse como

$$\mathcal{R}q = \frac{1}{4}(q - iqi - jqj - kqk) \quad (2.2)$$

$$\mathcal{I}q = \frac{1}{4}(q - iqi + jqj + kqk) \quad (2.3)$$

$$\mathcal{J}q = \frac{1}{4}(q + iqi - jqj + kqk) \quad (2.4)$$

$$\mathcal{K}q = \frac{1}{4}(q + iqi + jqj - kqk) \quad (2.5)$$

Los símbolos i, j y k son unidades imaginarias. La tabla completa de multiplicación de \mathbb{H} está dada por la tabla (2.1)

	1	i	j	k
1	1	i	j	k
i	i	-1	k	-j
j	j	-k	-1	i
k	k	j	-1	-i

Tabla 2.1 Multiplicación en \mathbb{H}

La adición y multiplicación de dos cuaterniones q y r está definida como sigue:

$$q + r = (q_0 + iq_1 + jq_2 + kq_3) + (r_0 + ir_1 + jr_2 + kr_3) \quad (2.6)$$

$$= (q_0 + r_0) + i(q_1 + r_1) + j(q_2 + r_2) + k(q_3 + r_3) \quad (2.7)$$

$$\begin{aligned} qr &= (q_0 + iq_1 + jq_2 + kq_3)(r_0 + ir_1 + jr_2 + kr_3) \\ &= (q_0r_0 - q_1r_1 - q_2r_2 - q_3r_3) \\ &\quad + i(q_0r_1 + q_1r_0 + q_2r_3 - q_3r_2) \\ &\quad + j(q_0r_2 + q_2r_0 - q_1r_3 + q_3r_3) \\ &\quad + k(q_0r_3 + q_3r_0 + q_1r_2 - q_2r_1) \end{aligned} \quad (2.8)$$

La regla de multiplicación (2.8) es resultado de la tabla de multiplicación (2.1), es importante remarcar que la multiplicación de cuaterniones *no es conmutativa*.

Una interpretación de la multiplicación definida por (2.8) puede ser encontrada en la siguiente definición

Definición 2 (Cuaterniones Puros) *Sea q un cuaternion y sea q_0 y $\mathbf{q} := iq_1 + jq_2 + kq_3$ la parte escalar y vectorial de q , respectivamente, entonces si la parte escalar q_0 es igual a cero entonces q es un cuaternion puro.*

El producto de multiplicar dos cuaterniones puros \mathbf{q} y \mathbf{r} es

$$\begin{aligned} \mathbf{qr} &= (-q_1r_1 - q_2r_2 - q_3r_3) \\ &\quad + i(q_2r_3 - q_3r_2) \\ &\quad + j(q_3r_1 - q_1r_3) \\ &\quad + k(q_1r_2 - q_2r_1) \end{aligned} \quad (2.9)$$

De (2.9) se sigue que la parte escalar de qr es igual a menos el producto punto de dos vectores tridimensionales y la parte vectorial es igual al producto cruz de vectores. Entonces, el producto cuaterniónico combina los dos productos fundamentales definidos para vectores tridimensionales. Luego, podemos escribir $qr = -q \cdot r + q \times r$.

Podemos definir, como en el álgebra de números complejos, la operación de conjugación para cuaterniones.

Definición 3 (Conjugación) *El endomorfismo*

$$- : \mathbb{H} \rightarrow \mathbb{H} \quad (2.10)$$

$$q \mapsto \bar{q} = -\frac{1}{2}(q + iq_i + jq_j + kq_k) \quad (2.11)$$

es la conjugación en \mathbb{H} .

De acuerdo con la definición (3) el conjugado de un cuaternion $q = q_0 + iq_1 + jq_2 + kq_3$ está dado por

$$\bar{q} = q_0 - iq_1 - jq_2 - kq_3. \quad (2.12)$$

La norma de q se define como

$$|q| = \sqrt{q\bar{q}} = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2}. \quad (2.13)$$

para cualquier $q \in \mathbb{H} \setminus \{0\}$ un único inverso multiplicativo existe y puede ser demostrado que es $q^{-1} = \frac{\bar{q}}{|q|^2}$ ([Per97]).

El conjunto \mathbb{H} con la adición y la multiplicación como se definió anteriormente forma un campo *torcido o skew*. Con la multiplicación de números reales \mathbb{H} se convierte en una álgebra con división asociativa sobre \mathbb{R} . La asociatividad puede ser probada por validación directa de las 27 igualdades $(e_l e_m) e_n = e_l (e_m e_n)$ donde $l, m, n \in \{1, 2, 3\}$ con $e_1, e_2, e_3 \in \{i, j, k\}$ ([Edd88]).

Sin embargo, se puede demostrar que \mathbb{H} es isomorfa a un álgebra de 2×2 matrices complejas que prueba directamente la asociatividad de \mathbb{H} .

También es posible considerar \mathbb{H} como un espacio lineal real de cuatro dimensiones con bases $\{1, i, j, k\}$. Usando notación vectorial podemos asignar a cada elemento de la base un vector natural de \mathbb{R}^4 : $1 = (1, 0, 0, 0)^\top$, $i = (0, 1, 0, 0)^\top$, $j = (0, 0, 1, 0)^\top$, y $k = (0, 0, 0, 1)^\top$.

El producto escalar Euclidiano en este espacio lineal es $\langle q, r \rangle = \frac{1}{2}(q\bar{r} + r\bar{q}) = q_0r_0 + q_1r_1 + q_2r_2 + q_3r_3$. Ahora podemos definir la noción de ortogonalidad para cuaterniones

Definición 4 (Ortogonalidad) Decimos que dos cuaterniones q y r son ortogonales cuando el producto escalar $\langle q, r \rangle = 0$.

Así, la norma de un cuaternion definida por la ecuación (2.13) puede ser reescrita usando el producto escalar como $|q| = \sqrt{\langle q, q \rangle}$. También introducimos el producto escalar para funciones evaluadas cuaterniónicamente.

Definición 5 (Producto escalar y norma de funciones cuaterniónicas) Sea V el espacio vectorial de funciones integrables $f : \mathbb{R}^n \rightarrow \mathbb{H}$. Entonces, por la formula

$$\langle f, g \rangle = \frac{1}{2} \int_{\mathbb{R}^n} (f(\mathbf{x})\bar{g}(\mathbf{x}) + g(\mathbf{x})\bar{f}(\mathbf{x})) d^n \mathbf{x}$$

introducimos un producto escalar en V . La L^2 -norma de una función $f \in V$ está dada por

$$\|f\| = \sqrt{\int_{\mathbb{R}^n} |f(\mathbf{x})|^2 d^n \mathbf{x}} \quad (2.14)$$

Mostramos como el campo de los números complejos puede ser introducido dentro de \mathbb{H} . Un subconjunto de \mathbb{H} que consiste sólo de cuaterniones de la forma $q = q_0 + iq_1 + j0 + k0$ es isomorfo a \mathbb{C} , estos cuaterniones son llamados *i-cuaterniones*, análogamente podemos definir *j-cuaterniones* y *k-cuaterniones* siendo estos también isomorfos a \mathbb{C} . Generalmente lo siguiente es verdad:

Teorema 6 *Cualquier subconjunto de \mathbb{H} de la forma $\mathbb{R}\{\mathbf{1}, \mathbf{n}\} = \{q | q = a + \mathbf{n}b\} \subset \mathbb{H}$, donde \mathbf{n} es un cuaternion puro, por ejemplo, $\mathbf{n} = in_1 + jn_2 + kn_3$ y $\sqrt{n_1^2 + n_2^2 + n_3^2} = 1$ es isomorfo a \mathbb{C} .*

Prueba. *Solamente necesitamos mostrar que \mathbf{n} se comporta como la unidad imaginaria de los números complejos i .*

$$\begin{aligned} n^2 &= (in_1 + jn_2 + kn_3)^2 \\ &= -n_1^2 - n_2^2 - n_3^2 + i(n_2n_3 - n_3n_2) + j(n_3n_1 - n_1n_3) + k(n_1n_2 - n_2n_1) \\ &= -1 \end{aligned}$$

lo que completa la prueba ■

La operación de conjugación en $\mathbb{C} : z \implies z^*$ también es llamada una involución del álgebra. Por ejemplo, cumple las siguientes dos propiedades: sea $z, w \in \mathbb{C} \implies (z^*)^* = z$ y $(wz)^* = w^*z^*$. En \mathbb{H} existen tres involuciones no triviales [Cher95]:

$$\begin{aligned} \alpha(q) &= -iqi = q_0 + iq_1 - jq_2 - kq_3 \\ \beta(q) &= -jqj = q_0 - iq_1 + jq_2 - kq_3, \text{ y} \\ \gamma(q) &= -kqk = q_0 - iq_1 - jq_2 + kq_3 \end{aligned} \tag{2.15}$$

las cuales utilizaremos para extender la noción de simetría Hermitiana de funciones complejas a cuaterniónicas. Una función $f : \mathbb{R}^n \rightarrow \mathbb{H}$ es llamada simétrica Hermitiana si $f(\mathbf{x}) = f^*(-\mathbf{x})$ para todo $\mathbf{x} \in \mathbb{R}^n$.

Definición 7 (Simetría Hermitiana Cuaterniónica) *Una función $f : \mathbb{R}^2 \rightarrow \mathbb{H}$ es llamada cuaterniónica Hermitiana si:*

$$F(-x, y) = \beta(f(x, y)) \quad \text{y} \quad f(x, -y) = \alpha(f(x, y)) \quad , \tag{2.16}$$

para cada $(x, y) \in \mathbb{R}^2$

Usaremos la función exponencial para cuaterniones en la definición de la transformada cuaterniónica wavelet. Antes de definir la función exponencial introducimos el siguiente lema.

Lema 8 Para cualquier cuaternion q la secuencia $\{a_n\}$ con

$$a_n = \sum_{k=0}^{\infty} \frac{q^k}{k!}, \quad q \in \mathbb{N} \quad (2.17)$$

converge a algún $r \in \mathbb{H}$

Prueba. Desde que \mathbb{H} es un álgebra normalizada tenemos que $|q^k| = |q|^k$. Sabemos que si $|q|$ es real entonces las series $\sum_{k=0}^{\infty} \frac{|q|^k}{k!} = \sum_{k=0}^{\infty} \frac{|q^k|}{k!}$ convergen para $n \rightarrow \infty$. Entonces, la secuencia $\{a_n\}$ converge absolutamente que es lo que implica el lema. ■

Definición 9 (Función exponencial) La función exponencial para cuaterniones $\exp: \mathbb{H} \rightarrow \mathbb{H}$ esta definida por la serie

$$\exp(q) = \sum_{k=0}^{\infty} \frac{q^k}{k!}, \quad q \in \mathbb{H} \quad (2.18)$$

Es importante remarcar que la existencia de $\exp(q)$ esta garantizada para cualquier $q \in \mathbb{H}$ por el lema (2.17)

Un dato importante es la ecuación de adición

$$\exp(q + r) = \exp(q) \exp(r) \quad (2.19)$$

que es verdadera para los números reales y complejos, generalmente no lo es para los cuaterniones q y r . En vez de (2.19) usaremos una relación mas general

$$\exp(h(q, r)) = \exp(q) \exp(r) \quad (2.20)$$

donde en general $h(q, r) \neq q + r$. De hecho $h(q, r)$ contiene el conmutador de q, r y conmutadores de ordenes superiores. La forma de $h(q, r)$ está dada por la fórmula de Campbell-Hausdorff (ver ejemplo [Hei90]).

La forma de calcular $\exp(q)$ para aplicaciones es la siguiente. Escribimos el cuaternion q en la forma $q = q_0 + \mathbf{q}$, donde q_0 y \mathbf{q} denotan la parte escalar y vectorial de q , respectivamente. Podemos evaluar $\exp(q)$ como

$$\begin{aligned} \exp(q) &= \exp(q_0 + \mathbf{q}) \\ &= \exp(q_0) \exp(\mathbf{q}) \\ &= \exp(q_0) \exp\left(\frac{\mathbf{q}}{|\mathbf{q}|} |\mathbf{q}|\right) \\ &= \exp(q_0) \left(\cos(|\mathbf{q}|) + \frac{\mathbf{q}}{|\mathbf{q}|} \sin(|\mathbf{q}|) \right) \end{aligned} \quad (2.21)$$

En el primer paso se utilizo la propiedad del escalar real conmutando con cualquier cuaternion. En el último paso empleamos la fórmula de Euler

$$\exp(i\phi) = \cos(\phi) + i \sin(\phi)$$

es válido si reemplazamos i por un cuaternion puro \mathbf{n}

$$\exp(\mathbf{n}\phi) = \cos(\phi) + \mathbf{n} \sin(\phi) \quad (2.22)$$

Esto por que $\mathbb{R}\{1, \mathbf{n}\}$ es isomorfo a \mathbb{C} como lo establece el teorema (6). Representando \mathbf{n} en la forma polar $\mathbf{n} = i \cos(\phi) \sin(\theta) + j \sin(\phi) \sin(\theta) + k \cos(\theta)$ tenemos la representación polar de q que es $\tilde{q} = (|q|, \phi, \theta, \psi)$. Los componentes q_i de q se pueden obtener de \tilde{q} por

$$\begin{aligned} q_0 &= |q| \cos(\psi) \\ q_1 &= |q| \cos(\phi) \sin(\theta) \sin(\psi) \\ q_2 &= |q| \sin(\phi) \sin(\theta) \sin(\psi) \\ q_3 &= |q| \cos(\theta) \sin(\psi) \end{aligned} \quad (2.23)$$

Sin embargo, como demostraremos después, esta representación polar de los cuaterniones no cubre con nuestras necesidades. Por lo tanto, introducimos otra definición para el ángulo/fase cuaterniónico en la sección 2.1.1.

2.1.1 El ángulo/fase cuaterniónico

En el capítulo 3 usamos el concepto señal-fase, que es introducido aquí, para una explicación mas detallada de los conceptos descritos aquí revisar [Bul99]. En el álgebra de cuaterniones no existe una sola fase angular, por lo tanto, introducimos una definición de esta basada en ángulos de Euler.

Para un número complejo $z = a + bi$ el argumento esta definido como $\text{atan2}(b, a)$, donde atan2 es la función arco-tangente dependiente del signo. Por ejemplo, si z es escrito en la forma $z = r \exp(i\phi)$ con $r \geq 0$ el argumento de z es ϕ . Para la función que regresa el argumento de cualquier número complejo introducimos la notación $\text{arg}(\cdot)$, tal que $\text{arg}(z) = \phi$. Como sabemos el álgebra de cuaterniones contiene tres subcampos complejos con unidades imaginarias ortogonales. Para su uso posterior, definimos tres funciones argumento, cada una tomando un subcampo complejo:

Definición 10 Sea $q = a + bi + cj + dk$ un cuaternion. Entonces definimos las siguientes funciones:

$$\text{arg}_i(q) = a \tan 2(b, a) \quad (2.24)$$

$$\text{arg}_j(q) = a \tan 2(c, a) \quad (2.25)$$

$$\text{arg}_k(q) = a \tan 2(d, a) \quad (2.26)$$

Estas funciones proyectan el cuaternion q en un subcampo complejo como indican los índices i, j y k y evalúan el ángulo-fase resultante del número complejo.

Tomamos otra representación de la fase cuaterniónica, que esta mas relacionada con el concepto clásico de fase, de una forma que quedará clara mas adelante

Teorema 11 Cada cuaternion q puede ser representado de la forma

$$q = |q| e^{i\phi} e^{k\psi} e^{j\theta} \quad \text{con } (\phi, \theta, \psi) \in [-\pi, \pi] \times [-\pi/2, \pi/2] \times [-\pi/4, \pi/4]. \quad (2.27)$$

Antes de probar el teorema se presentan dos lemas. El primero representa el mapeo entre cuaterniones unitarios y una matriz de rotación. El segundo especifica la forma de la cobertura de $SO(3)$ por S^3

Lema 12 Sea $q = a + bi + cj + dk$ un cuaternion, y sea \mathcal{M} el mapeo de cuaterniones a matrices reales de 3×3 de la siguiente forma:

$$\mathcal{M}(q) = \begin{pmatrix} a^2 + b^2 - c^2 - d^2 & 2(bc - ad) & 2(db + ac) \\ 2(bc + ad) & a^2 - b^2 + c^2 - d^2 & 2(cd - ab) \\ 2(bd - ac) & 2(cd + ab) & a^2 - b^2 - c^2 + d^2 \end{pmatrix} \quad (2.28)$$

entonces

$$\mathcal{M} : \{q \in \mathbb{H}, |q| = 1\} \rightarrow SO(3) \quad (2.29)$$

es un homomorfismo y $\mathcal{M}(q)$ es la matriz que representa la misma rotación que el cuaternion q .

Prueba. ver [Kor68] ■

La matriz $\mathcal{M}(q)$ es llamada la matriz de Rodríguez.

Lema 13 Sea q_1 y q_2 cuaterniones unitarios y

$$\mathcal{M}(q_1) = \mathcal{M}(q_2) \quad (2.30)$$

entonces

$$q_1 = q_2 \quad \text{o} \quad q_1 = -q_2 \quad (2.31)$$

Prueba. Como S^3 es un grupo de cobertura de $SO(3)$ le corresponden exactamente dos cuaterniones diferentes por cada matriz de rotación. Usando (2.28) se puede verificar que para todo $q \in S^3$ $\mathcal{M}(q) = \mathcal{M}(-q)$. ■

Prueba del Teorema. 11 Cada cuaternion q puede ser escrito como $q = |q|r$ donde r es un cuaternion unitario. De acuerdo al lema (12) cada cuaternion unitario representa una rotación en \mathbb{R}^3 entonces la matriz $R = \mathcal{M}(r)$ puede ser factorizada con tres rotaciones

$$R = R_x(2\phi)R_z(2\psi)R_y(2\theta), \quad (2.32)$$

Donde R_x , R_z y R_y representan rotaciones sobre los ejes de coordenadas, los ángulos 2ϕ , 2ψ y 2θ son los ángulos de Euler [Kor68]. Note que se utiliza el doble del ángulo. Si $q_1 = \exp(i\phi)$, $q_2 = \exp(j\theta)$, $q = \exp(k\psi)$ entonces es fácil ver que

$$\mathcal{M}(q_1) = R_x(2\phi), \quad \mathcal{M}(q_2) = R_z(2\psi), \quad \mathcal{M}(q_3) = R_y(2\theta). \quad (2.33)$$

Como \mathcal{M} es un homomorfismo, por el lema (12), tenemos que

$$\mathcal{M}(q_1)\mathcal{M}(q_3)\mathcal{M}(q_2) = \mathcal{M}(q_1q_3q_2)R_x(2\phi), = \mathcal{M}(r). \tag{2.34}$$

■

Teorema 14 Prueba. De el lema (13) tenemos que $r = q_1q_3q_2$ o $r = -q_1q_3q_2$, que nos lleva directamente a la representación que buscamos cuando mucho con un signo menos, que puede ser eliminado reemplazando ϕ por $\phi + \pi$, esto nos deja

$$q = \pm|q| \exp(i\phi) \exp(k\psi) \exp(j\theta). \tag{2.35}$$

Lo que prueba que cada cuaternion puede ser expresado en la forma polar propuesta.

■

La segunda parte de la prueba muestra que los intervalos dados para ϕ , θ y ψ son suficientes y mínimos para expresar cualquier cuaternion. La prueba puede verse en [Bul99] pág. 18,19,20.

Definición 15 Sea $q \in \mathbb{H}$ un cuaternion. La fase angular de q esta definida por

$$\arg(q) = (\phi, \theta, \psi) \in [-\pi, \pi] \times [-\pi/2, \pi/2] \times [-\pi/4, \pi/4]. \tag{2.36}$$

de acuerdo al teorema (11).

La fórmulas para la fase cuaterniónica se resumen en la tabla (2.1.1).

$q = a + bi + cj + dk, \quad q = 1$					
$\psi = -\frac{\arcsin(2bc-ad)}{2}$					
si $\psi \in]-\frac{\pi}{4}, \frac{\pi}{4}[$ $\phi = \frac{\arg_i(q\beta(\bar{q}))}{2}$ $\theta = \frac{\arg_j(\alpha(\bar{q})q)}{2}$	Si $\psi = \pm\frac{\pi}{4}$ elegir <hr style="width: 100%;"/> <table style="width: 100%; border: none;"> <tr> <td style="width: 50%; border: none;">$\phi = 0$</td> <td style="width: 50%; border: none;">$\theta = 0$</td> </tr> <tr> <td style="width: 50%; border: none;">$\theta = \frac{\arg_j(\gamma(\bar{q})q)}{2}$</td> <td style="width: 50%; border: none;">$\phi = \frac{\arg_i(\gamma(\bar{q}))}{2}$</td> </tr> </table>	$\phi = 0$	$\theta = 0$	$\theta = \frac{\arg_j(\gamma(\bar{q})q)}{2}$	$\phi = \frac{\arg_i(\gamma(\bar{q}))}{2}$
$\phi = 0$	$\theta = 0$				
$\theta = \frac{\arg_j(\gamma(\bar{q})q)}{2}$	$\phi = \frac{\arg_i(\gamma(\bar{q}))}{2}$				
Si $e^{i\phi}e^{k\psi}e^{j\theta} = -q$ y $\phi \geq 0$ $\phi \rightarrow \phi - \pi$	si $e^{i\phi}e^{k\psi}e^{j\theta} = -q$ y $\phi < 0$ $\phi \rightarrow \phi + \pi$				

Tabla 2.1.1 Cálculo de la fase cuaterniónica de un cuaternion dado en representación cartesiana

2.2 Enlace Tiempo-Frecuencia

El tiempo y la frecuencia están íntimamente relacionados, una muestra de esto es el principio de incertidumbre de Heisenberg, ver teorema (28) en el apéndice, que establece que ganar precisión en frecuencia lleva a perder precisión en el tiempo y viceversa. El análisis de señales cuenta con muchas técnicas matemáticas para *transformar* nuestra vista de una señal basada en el tiempo por una señal basada en la frecuencia. Tal vez la mas conocida de estas herramientas es el análisis de Fourier. Sin embargo si estamos interesados en el análisis de señales transitorias; como por ejemplo, una manzana localizada en la orilla derecha de una imagen o una palabra pronunciada en un momento en particular, entonces la transformada de Fourier no es suficiente.

Existen diferentes aproximación para lidiar con este tipo de señales, una de las primeras fue por parte del físico Dennis Gabor (1946) que adapto el análisis de Fourier para analizar solo pequeñas secciones de la señal (ventana) por vez. Así el análisis de Fourier por ventanas, como es llamada la adaptación de Gabor, provee información acerca de cuando y que frecuencias están presentes en una señal. Desafortunadamente esta información tiene precisión limitada por el tamaño de la ventana como lo establece el principio de incertidumbre.

El análisis wavelet es el siguiente paso lógico: analizar con ventanas de un tamaño variable. Así el análisis wavelet utiliza el concepto de multiresolución, analizado en (2.2.3), que le permite el uso de intervalos largos para frecuencias bajas e intervalos cortos para frecuencias altas. Esta es una forma relativamente nueva de representar una señal, sin embargo el análisis wavelet es una alternativa al análisis de Fourier, mas que un reemplazo, para diseñar y entender las wavelets seguimos utilizando el análisis de Fourier en lo que se refiere al enlace entre el tiempo y la frecuencia.

2.2.1 La Transformada de Fourier

En 1822 El matemático francés Jean Baptiste Joseph Fourier mostró que cualquier función periódica puede ser expresada como una suma infinita de funciones exponenciales complejas. Muchos años después, sus ideas fueron generalizadas primero para funciones no periódicas y después para señales discretas. Sólo después de esta generalización la Transformada de Fourier (TF) se convirtió en una herramienta adaptable a cálculos. Después en 1965 apareció un algoritmo para calcular de una manera efectiva la TF llamado transformada rápida de Fourier y desde entonces se volvió aun mas popular.

La mayoría del material de esta sección se presenta sin prueba. Esto porque las pruebas son largas y se alejan del objetivo de la tesis. Sin embargo, la notación es parte importante del trabajo estableciendo las bases necesarias para la transformada de Fourier por ventanas, además se da un panorama del camino histórico a la transformada cuaterniónica wavelet.

La transformada de Fourier está definida como

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt \quad (2.37)$$

la cual podemos interpretar que mide “cuantas” oscilaciones de la frecuencia ω existen en f . Si f pertenece al espacio de funciones integrables entonces la integral converge y es acotada, entonces tenemos el siguiente teorema para la transformada inversa de Fourier.

Teorema 16 (Transformada inversa de Fourier) *Si f y F pertenecen al espacio de funciones integrables entonces*

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{i\omega t} d\omega \quad (2.38)$$

La fórmula (2.38) descompone la señal f en una suma de ondas senoidales $e^{i\omega t}$ de amplitud $F(\omega)$. Usando esta fórmula se puede mostrar que la hipótesis de que $F(\omega)$ pertenece al espacio de funciones integrables implica que f debe de ser continua. Por lo tanto, la reconstrucción (2.38) no esta probada para funciones discontinuas.

La propiedad mas importante de la transformada de Fourier para el procesamiento de señales es el teorema de convolución. Que es otra forma de expresar que las ondas senoidales $e^{i\omega t}$ son eigenvectores de operadores de convolución.

Teorema 17 (Convolución) *Sean f y h dos funciones pertenecientes al conjunto de funciones integrables. La función $g = h * f$ está en el mismo conjunto de funciones integrables y*

$$G(\omega) = H(\omega)F(\omega). \quad (2.39)$$

Prueba.

$$G(\omega) = \int_{-\infty}^{+\infty} e^{-i\omega t} \left(\int_{-\infty}^{+\infty} f(t-u)h(u)du \right) dt \quad (2.40)$$

como $|f(t-u)||h(u)|$ es integrable en \mathbb{R}^2 podemos aplicar el teorema de Fubini (A.26) y el cambio de variable $(t, u) \rightarrow (v = t - u, u)$ nos lleva a

$$\begin{aligned} G(\omega) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{[-i\omega(u+v)]} f(v)h(u)dudv \\ &= \left(\int_{-\infty}^{\infty} e^{-i\omega v} f(v)dv \right) \left(\int_{-\infty}^{\infty} e^{-i\omega u} h(u)du \right) \end{aligned}$$

verificando (2.39). ■

La respuesta de un sistema lineal invariante en el tiempo $Lf = g = f * h$ puede ser calculado por su transformada de Fourier como $G(\omega) = F(\omega)H(\omega)$ con la fórmula de Fourier inversa

$$g(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(\omega)e^{i\omega t} d\omega, \quad (2.41)$$

que nos lleva a

$$Lf(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega)F(\omega)e^{i\omega t} d\omega \quad (2.42)$$

Cada componente de frecuencia $e^{i\omega t}$ de amplitud $F(\omega)$ es amplificada o atenuada por $H(\omega)$. Por lo tanto la convolución es llamada *filtrado de frecuencia*, y H es la función de transferencia del filtro. La tabla (2.2.1) resume las propiedades mas importantes de la transformada de Fourier, la mayoría de estas fórmulas se prueban con un cambio de variable en la integral de Fourier.

Propiedad	Función	Transformada de Fourier
	$f(t)$	$F(\omega)$
Inversa	$F(t)$	$2\pi f(-\omega)$
Convolución	$f_1 * f_2(t)$	$F_1(\omega)F_2(\omega)$
Multiplicación	$f_1(t)f_2(t)$	$\frac{1}{2\pi}F_1 * F_2(\omega)$
Traslación	$f_1(t - u)$	$e^{-i\omega u}F(\omega)$
Modulación	$e^{-i\xi t}f(t)$	$F(\omega - \xi)$
Escalamiento	$f\left(\frac{t}{s}\right)$	$ s F(s\omega)$
Derivada del tiempo	$f^{(p)}(t)$	$(i\omega)^p F(\omega)$
Derivada Frecuencia	$(-it)^{(p)}f(t)$	$F^{(p)}(\omega)$
Conjugado Complejo	$f^*(t)$	$F^*(-\omega)$
Simetría Hermitian	$f(t) \in \mathbb{R}$	$F(-\omega) = F^*(\omega)$

Tabla 2.2.1 Propiedades de la transformada de Fourier

2.2.2 La Transformada de Fourier por ventanas

Una familia de formas de onda que están bien localizadas en tiempo y en frecuencia son llamadas átomos tiempo-frecuencia. En 1946, Gabor introdujo los *átomos de Fourier* por ventanas, para medir variaciones de frecuencia en sonido. Una función que selecciona una región de la señal que puede ser considerada como estacionaria se le llama *ventana*. Una ventana real y simétrica $g(t) = g(-t)$ es trasladada por u y modulada por la frecuencia ξ :

$$g_{u,\xi}(t) = e^{i\xi t}g(t - u) \quad (2.43)$$

Donde $\|g\| = 1$ está normalizada tal que $\|g_{u,\xi}\| = 1$ para cualquier $(u, \xi) \in \mathbb{R}^2$. La Transformada de Fourier por Ventanas (TFV) resultante de funciones que pertenecen al espacio de funciones con energía finita es

$$Sf(u, \xi) = \langle f, g_{u,\xi} \rangle = \int_{-\infty}^{\infty} f(t)g(t-u)e^{-i\xi t} dt. \quad (2.44)$$

Esta transformada también es llamada Transformada de Fourier de tiempo corto, porque la multiplicación por $g(t-u)$ localiza la integral de Fourier en la región $t = u$. La función ventana g es localizada primero al principio de la señal $t = 0$. Supongamos que el ancho de esta ventana es T , segundos en este instante la función se sobrepone los primeros $T/2$ segundos, después multiplicamos la función ventana y la señal al hacer esto estamos seleccionando sólo los primeros $T/2$ segundos de la señal con el peso apropiado de la señal (por ejemplo, si la ventana es un rectángulo de amplitud 1 el producto será igual a la señal), tomamos este producto y le aplicamos la TF como si estuviéramos tomando la TF de una señal cualquiera.

El resultado de esta transformación es la TF de los primeros $T/2$ segundos de la señal, si esta porción de la señal es estacionaria entonces obtenemos un representación verdadera de la frecuencia en los primeros $T/2$ segundos de la señal. El siguiente paso es trasladar la ventana g a una nueva localidad t_1 , multiplicar por la señal y tomar la TF del producto, este procedimiento continua hasta el final de la señal, desplazando siempre la ventana en intervalos de t_1 segundos.

El problema con la TFV tiene que ver con el ancho de g . Para ser técnicamente correctos, este ancho de la ventana es conocido como *soporte de la ventana*

Soporte Finito En aplicaciones numéricas, g tiene que tener soporte compacto (la ventana tiene que ser angosta). El teorema (2.37) prueba que la transformada de Fourier G tiene soporte infinito necesariamente. Es una función simétrica con un lóbulo principal en $\omega = 0$, que decae a cero con oscilaciones. La figura (2.1) muestra este comportamiento. Para maximizar la resolución en frecuencia de esta transformación, debemos concentrar la energía de G cerca de $\omega = 0$. Esta expansión de energía es medida por tres parámetros importantes:

- La raíz cuadrada media del ancho de banda $\Delta\omega$ que está definido por

$$\frac{|G(\Delta\omega/2)|^2}{|G(0)|^2} = \frac{1}{2}$$

- La amplitud máxima A de los primeros lóbulos laterales localizados en $\omega = \pm\omega_0$ en la figura (2.1), son medidos en decibeles:

$$A = 10 \log_{10} \frac{|G(\omega_0)|^2}{|G(0)|^2}$$

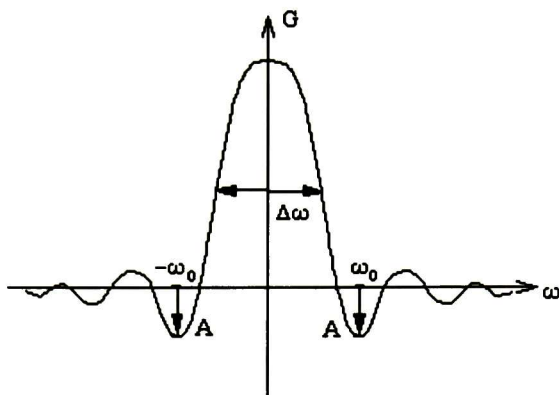


Figura 2.1: La energía de G es medida por su ancho de banda $\Delta\omega$ y la amplitud máxima A de los lóbulos laterales, localizados en $\omega = \pm\omega_0$

- El exponente polinomial p , que da el decaimiento asintótico de $|G(\omega)|$ para frecuencias altas:

$$|G(\omega)| = O(\omega^{-p-1}).$$

La tabla (2.2.2) muestra los valores de estos tres parámetros para diferentes ventanas g cuyos soportes están restringidos a $[-1/2, 1/2]$. La figura (2.2) muestra los gráficos de estas ventanas.

Nombre	$g(t)$	$\Delta\omega$	A	p
Rectángulo	1	0.89	-13db	0
Hamming	$0.54 + 0.45\cos(2\pi t)$	1.36	-43db	0
Gaussiano	$\exp(-18t^2)$	1.55	-55db	0
Hanning	$\cos^2(\pi t)$	1.44	-32db	2
Blackman	$0.42 + 0.5\cos(2\pi t) + 0.08\cos 4\pi t$	1.68	-58db	2

Tabla 2.2.2 Parámetros de frecuencia de cinco ventanas g con soporte restringido a $[-1/2, 1/2]$. Normalizadas tal que $g(0) = 1$ pero $\|g\| \neq 1$.

Como podemos observar, al usar la TFV nos enfrentamos al problema de resolución, el conflicto es la selección de la ventana adecuada. Una ventana angosta nos da buena resolución en tiempo pero mala resolución en frecuencia; mientras una ventana ancha tiene buena resolución en frecuencia pero pobre resolución en tiempo. Además, las ventanas anchas pueden violar la condición de régimen estacionario. La selección de la ventana entonces es dependiente de la aplicación. Sin embargo, el uso de un esquema multiresolución es útil para lidiar con el problema de incertidumbre de una manera efectiva, como veremos en la siguiente sección.

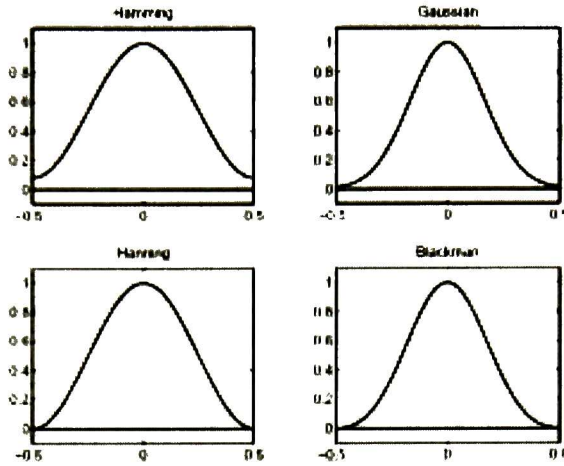


Figura 2.2: Gráficos de cuatro ventanas g con soporte $[-1/2, 1/2]$

2.2.3 La Transformada Wavelet

La primer vez que se uso el término “wavelet” u ondoleta fue en 1909, en la tesis de Alfred Haar. Por razones de popularidad continuaremos usando el término wavelet. Los conceptos relacionados con la transformada wavelet han sido contribuciones de muchos campos incluyendo la ingeniería, física y matemáticas. La transformada wavelet consiste en descomponer una señal o imagen en un conjunto de *aproximaciones*.

Wavelet Real

La transformada wavelet descompone una señal en wavelets trasladadas y dilatadas. Una wavelet es una función ψ con promedio cero:

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0 \quad (2.45)$$

Esta está normalizada $\|\psi\| = 1$ y centrada en un vecindario $t = 0$. Una familia de wavelets es obtenida escalando ψ por un factor s y trasladándola por u :

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right)$$

Todas estas modificaciones corresponden a transformaciones afines. Estos átomos permanecen normalizados: $\|\psi_{u,s}\| = 1$. La transformada wavelet de f con tiempo u y escala s es

$$Wf(u, s) = \langle f, \psi_{u,s} \rangle = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{s}} \psi^*\left(\frac{t-u}{s}\right) dt \quad (2.46)$$

La transformada wavelet puede ser escrita como un producto convolución

$$Wf(u, s) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{s}} \psi^* \left(\frac{t-u}{s} \right) dt = f * \bar{\psi}_s(u) \quad (2.47)$$

con

$$\bar{\psi}_s(t) = \frac{1}{\sqrt{s}} \psi^* \left(\frac{-t}{s} \right) \quad (2.48)$$

y la transformada de Fourier de $\bar{\psi}_s(t)$ es

$$\bar{\Psi}_s(\omega) = \sqrt{s} \hat{\psi}^*(s\omega). \quad (2.49)$$

Como $\hat{\psi}(0) = \int_{-\infty}^{+\infty} \psi(t) dt = 0$, $\hat{\psi}$ tiene la forma de la función de transferencia de un filtro pasa-banda. La convolución (2.47) calcula la transformada wavelet con un filtro pasa-banda. Así, podemos calcular la transformada wavelet por medio de filtros, como veremos mas adelante en este capitulo.

La transformada wavelet para dos dimensiones proyecta la función de la imagen $f(x, y)$ en una familia de funciones que son las dilataciones y traslaciones de la wavelet madre (*mother wavelet*). La TW en 2D de la función $f(x, y)$ con una wavelet madre ψ esta definida por

$$Wf(u, v, s) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \psi(x, y) dx dy \quad (2.50)$$

Una transformación wavelet como la definida en (2.50) puede ser interpretada como la descomposición de una imagen $f(x, y)$ en un conjunto de canales de frecuencia con el mismo ancho de banda en una escala logarítmica, produciendo exactamente una representación espacio-escala de la imagen.

Función de escalamiento

Cuando $Wf(u, s)$ es conocido sólo para $s < s_0$, para recobrar f necesitamos información complementaria de $Wf(u, s)$ para $s > s_0$. Esto es obtenido introduciendo una función de escalamiento ϕ que es una adición a wavelets de escalas mayores que 1. El modulo de la transformada de Fourier esta definida por

$$|\hat{\phi}(\omega)|^2 = \int_1^{+\infty} |\hat{\psi}(s\omega)|^2 \frac{ds}{s} = \int_{\omega}^{+\infty} \frac{|\hat{\psi}(\xi)|^2}{\xi} d\xi \quad (2.51)$$

y la fase compleja de $\hat{\phi}(\omega)$ puede ser elegida arbitrariamente. Uno puede verificar que $\|\phi\| = 1$ y que derivando de la ecuación de admisibilidad, teorema (27) tenemos que

$$\lim_{\omega \rightarrow 0} |\phi(\omega)|^2 = C_{\psi}, \quad (2.52)$$

así la función de escalamiento puede entonces ser interpretada como la respuesta al impulso de un filtro pasa-bajas

En 1986 apareció un nuevo método para hacer análisis y síntesis de señales ese método es conocido como Análisis Multiresolución (AMR), como su nombre lo implica, analiza una señal a diferentes frecuencias con diferentes resoluciones. Cada componente espectral no está resuelto igualmente como era el caso en la transformada de Fourier por ventanas. Denotemos

$$\phi_s(t) = \frac{1}{\sqrt{s}}\phi\left(\frac{t}{s}\right) \quad \text{y} \quad \phi_s(t) = \phi_s^*(-t). \tag{2.53}$$

La aproximación de baja frecuencia de f en la escala s es

$$Lf(u, s) = \left\langle f(t), \frac{1}{\sqrt{s}}\phi\left(\frac{t-u}{s}\right) \right\rangle = f * \phi_s(u). \tag{2.54}$$

Análisis Multiresolución

El Análisis Multi-Resolución (AMR) esta diseñado para dar una buena resolución en tiempo y mala en frecuencia para frecuencias altas y buena resolución en frecuencia y mala en tiempo para frecuencias bajas. Esta aproximación tiene sentido cuando la señal a analizar tiene componentes de alta frecuencia en periodos cortos y componentes de bajas frecuencia por periodos largos. Afortunadamente, las señales encontradas en aplicaciones practicas generalmente son de este tipo.

A una resolución determinada de una señal o imagen, la función de escalamiento $\phi(2^j t - k)$ es la base para un conjunto de señales. El nivel es establecido por j , y el periodos de tiempo en cada nivel por 2^{-j} . Los nuevos detalles en el nivel j están representados por las wavelets $W(2^j t - k)$. Por lo tanto la función suavizada de la señal mas los detalles se combinan en multiresolución de la señal al nivel de detalle $j + 1$. La aproximación viene de la función de escalamiento, los detalles de la wavelet.

$$\begin{array}{l} \text{señal en el nivel } j \text{ (aproximación local)} \searrow \\ + \\ \text{Detalles en el nivel } j \text{ (Detalles locales)} \nearrow \end{array} \quad \text{señal en el nivel } j + 1$$

Esta es la idea de multiresolución para una señal. Cuando aplicamos para todas las señales, tenemos multiresolución para espacio de funciones:

$$\begin{array}{l} V_j = \text{espacio de escalamiento en el nivel } j \searrow \\ \oplus \\ W_j = \text{Espacio wavelet en el nivel } j \text{ (Detalles locales)} \nearrow \end{array} \quad V_{j+1} = \text{espacio de escala en nivel } j + 1$$

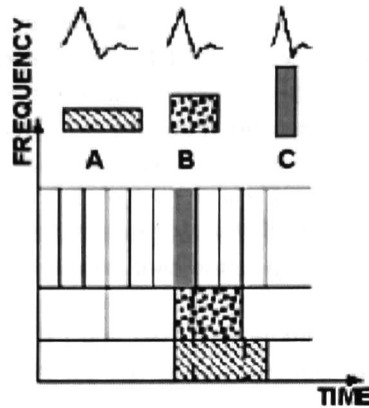


Figura 2.3: Cobertura tiempo-frecuencia con una wavelet de Daubechies

Esta idea de multiresolución es absolutamente básica en la teoría de wavelets. La idea básica es dividir la señal en diferentes escalas de resolución. El plano tiempo-escala toma el lugar para wavelets que el plano tiempo-frecuencia tiene para filtros.

En la figura (2.3) podemos observar la cobertura tiempo frecuencia de una ventana de diferentes tamaños (A,B,C) donde A es una ventana con buena resolución en frecuencia y mala en tiempo para frecuencias bajas, B es una ventana mas angosta que A y por lo tanto tiene mejor resolución en tiempo pero peor resolución en frecuencia que A. Por ultimo, C es una ventana mas angosta y por lo tanto nos da muy buena resolución en tiempo pero mala en frecuencia.

La multiresolución divide la frecuencia en octavas, de ω a 2ω , en vez de bandas uniformes de ω a $\omega + \Delta\omega$. Como se puede observar en la figura (2.3) el espacio tiempo-frecuencia esta particionado en rectángulos de área constante.

Transformada Discreta Wavelet

Para un caso especial de funciones, la redundancia de la transformada continua wavelet (2.50) puede ser aclarado por la discretización del factor de escala s y la traslación (u, v) , en particular un espacio diádico

$$s = \frac{1}{2^j} \quad \text{y} \quad (u = k, v = l) \quad \text{con} \quad (j, k, l) \in \mathbb{Z}^3 \quad (2.55)$$

En el espacio diádico, denotemos $A_j f$ como la aproximación de una función $f(x, y)$ en la escala $s = \frac{1}{2^j}$. Denotemos a $D_j f$ como la diferencia entre dos aproximaciones A_{j-1} y A_j , por ejemplo:

$$D_j f = A_{j-1} f - A_j f, \quad j = 1, 2, \dots, n \quad (2.56)$$

donde A_0 es un operador identidad

Existen dos formas inversas de definir la noción de escala s y de nivel j para una pirámide multiresolución. Para una imagen cuadrada de tamaño $2^n \times 2^n$, definimos el nivel $j = 0$ y la escala $s = 2^{-j} = 1$ como el primer nivel, donde el tamaño original de la imagen $2^n \times 2^n$. el siguiente nivel mas burdo es $j = 1$ y $s = 2^{-1} = \frac{1}{2}$. El nivel mas alto de la pirámide, el nivel mas burdo, es $j = n$ y $s = 2^{-n}$, correspondiente a un solo punto. En la practica solo consideramos un número limitado de niveles $j = 0, 1, 2, \dots, n$, para algún n que elegimos como el nivel mas burdo que corresponde a la escala mas pequeña $\frac{1}{2^n}$.

Mallat probó en ([Mal89]) que para una análisis multiresolución, existen tres componentes para los detalles entre dos aproximaciones $A_{j-1}f$ y A_jf

$$D_jf = A_{j-1}f - A_jf = D_{j,1}f + D_{j,2}f + D_{j,3}f \quad (2.57)$$

por lo tanto el análisis multiresolución de una función $f(x, y)$ puede ser escrito como

$$\begin{aligned} f(x, y) &= A_1f + D_{1,1}f + D_{1,2}f + D_{1,3}f \\ &= A_2f + D_{2,1}f + D_{2,2}f + D_{2,3}f + D_{1,1}f + D_{1,2}f + D_{1,3}f \\ &\quad \vdots \\ &= A_nf + \sum_{j=1}^n [D_{j,1}f + D_{j,2}f + D_{j,3}f] \end{aligned} \quad (2.58)$$

cada aproximación $A_jf(x, y)$ y cada componente de detalle $D_{j,p}f(x, y)$ pueden ser totalmente caracterizados con una función de 2D $\phi(x, y)$ y su función wavelet asociada $\psi_p(x, y)$

$$A_jf(x, y) = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} a_{j,k,l} \phi_{j,k,l}(x, y), \quad D_{j,p}f(x, y) = \sum_{k=-\infty}^{+\infty} \sum_{l=-\infty}^{+\infty} d_{j,p,k,l} \psi_{j,p,k,l}(x, y) \quad (2.59)$$

donde

$$\phi_{j,k,l}(x, y) = \frac{1}{2^j} \phi\left(\frac{x-k}{2^j}, \frac{y-l}{2^j}\right), \quad (j, k, l) \in \mathbb{Z}^3, \quad (2.60)$$

$$\psi_{j,p,k,l}(x, y) = \frac{1}{2^j} \psi_p\left(\frac{x-k}{2^j}, \frac{y-l}{2^j}\right) \quad (2.61)$$

y

$$a_{j,k,l}(x, y) = \langle f(x, y), \phi_{j,k,l}(x, y) \rangle, \quad d_{j,p,k,l} = \langle f(x, y), \psi_{j,p,k,l}(x, y) \rangle \quad (2.62)$$

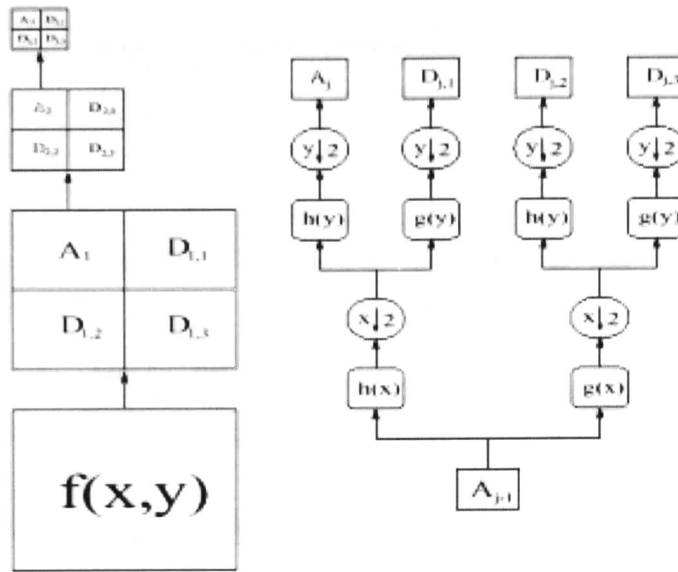


Figura 2.4: Izquierda, Piramide Wavelet de una imagen $f(x,y)$. Derecha, Diagrama de flujo del análisis del nivel $j - 1$ al nivel j

para una análisis multiresolución separable, la función de escalamiento $\phi(x,y)$ y la función wavelet $\psi_p(x,y), p = 1, 2, 3$ puede ser escrito como

$$\begin{aligned}
 \phi(x,y) &= \phi(x)\phi(y), \\
 \psi_1(x,y) &= \phi(x)\psi(y), \\
 \psi_2(x,y) &= \psi(x)\phi(y), \\
 \psi_3(x,y) &= \psi(x)\psi(y).
 \end{aligned}
 \tag{2.63}$$

donde ϕ es una función de escala 1D, ψ es la función wavelet de 1D asociada con ϕ .

Sea h y g la respuesta al impulso del filtro ϕ y ψ respectivamente, los coeficientes $a_{j,k,l}$ y $d_{j,p,k,l}, p = 1, 2, 3$, pueden ser calculados con un procedimiento iterativo, el conocido algoritmo de Mallat [Mal89]. La pirámide wavelet para una imagen $f(x,y)$ y el proceso de construcción es ilustrado en la figura (2.4).

Transformada Compleja Wavelet

La transformada wavelet no es invariante a la traslación ni a la rotación, lo que implica que la información de fase no es representada con facilidad. Explorando la información de fase en imágenes y siguiendo con el concepto de multiresolución, las wavelets-complejas representan una traslación en el espacio como una rotación en el dominio de la fase compleja.

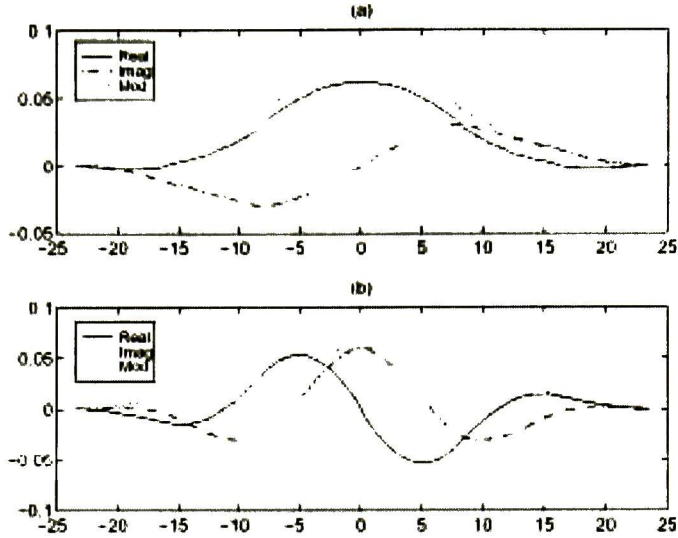


Figura 2.5: a). Función Wavelet, b) Función de escalamiento. Ambas funciones complejas.

La wavelet compleja que exponemos aquí fue diseñada por Margarey y Kingsbury [Mag95] también se puede consultar para más referencias [Lin97] donde se proponen una extensión de las wavelet Daubechies usando números complejos. Existen muchas familias de wavelet complejas, y no sería extraño que aparecieran otras nuevas. Las presentadas aquí, tienen la intención de mostrar solo las bases de este tipo de wavelets.

Aquí en lugar de usar un Filtro con respuesta al impulso (FIR) real usaremos un (FIR) con fase lineal compleja en cuadratura y lo asociamos con las bases de la wavelet. Siguiendo con la idea de la pirámide y un análisis en 2D, los filtros de escalamiento y wavelet g y h son un par de ventanas moduladas complejas de longitud par

$$g_k = w(k + 0.5)e^{i\omega(k+0.5)} \quad (2.64)$$

$$h_k = w(k + 0.5)e^{i\tilde{\omega}(k+0.5)} \quad (2.65)$$

donde $k = -n, -n + 1, \dots, n - 1$ y w es una Gaussiana de longitud $2n$. El soporte más compacto es de longitud 4, por lo tanto $n = 2$. La frecuencias de modulación ω y $\tilde{\omega}$ deben de ser complementarias: $\omega + \tilde{\omega} = \pi$, para cubrir el rango de frecuencias de $[0, \pi]$. Como sabemos ϕ es un filtro pasa-bajas y ψ uno pasa-altas, entonces se tiene que cumplir que $\omega > \tilde{\omega}$. Los filtros (2.64) y (2.65) son comúnmente llamados MKC wavelet (Magarey and Kingsbury Complex wavelet).

Usando los filtros de wavelet y escalamiento complejos en 1D que se muestran en (2.64) y (2.65), el análisis wavelet en 2D puede ser implementado como se muestra en

(2.63). Los filtros resultantes son predominan en el primer cuadrante del dominio de la frecuencia. Como las imágenes reales discretas contienen información significativa en el primer y segundo cuadrante, se usan los filtros complejos conjugados (\bar{h}, \bar{g}) para así producir un conjunto reflejado de wavelet conteniendo la información del segundo cuadrante. En la figura (2.5) se muestran los filtros de wavelet y escalamiento MKC.

Capítulo 3

Transformada Cuaterniónica Wavelet

Este capítulo es la parte medular de la tesis, primero entramos al tema de los filtros cuaterniónicos de Gabor, estos filtros son empleados en un modelo multiresolución para llevar a cabo la transformada cuaterniónica wavelet (TQW). Se muestran también algunas propiedades de dichos filtros, como es la descomposición en filtros complejos para facilitar el proceso de cálculo de la TQW en un AMR. En esta sección se define la fase local cuaterniónica como la fase angular de la respuesta de un filtro de Gabor conforme a las reglas presentadas en el capítulo 2.

Se describe a detalle el proceso para realizar el análisis de una imagen con la TQW. Se usó el algoritmo de Mallat como modelo principal del algoritmo y se extendió para cuaterniones. Los filtros de Gabor no pueden recuperar toda la energía de la señal, por lo tanto se usa una función de escalamiento para generar un banco de filtros con los que se va generando la pirámide multiresolución. Se presenta la forma de los filtros wavelet y de escalamiento usados además de un pequeño ejemplo de la TQW en acción sobre una imagen sintética sencilla.

Finalmente se presenta un análisis de complejidad del algoritmo para el cálculo de los coeficientes de la TQW. Esta sección es fructífera, porque la discusión puede llevar a la generación de un trabajo futuro interesante, donde el objetivo sea reducir la complejidad del algoritmo.

3.1 Filtros Cuaterniónicos de Gabor

Los filtros de Gabor han demostrado ser una herramienta eficiente para el procesamiento y análisis de imágenes como es segmentación y clasificación de texturas, detección de bordes y estimación de la fase local y frecuencia para *imagen matching*. Los filtros de Gabor tienen la principal ventaja de estar óptimamente localizados en el dominio de la frecuencia y el espacio simultáneamente. En otras palabras, estos filtros cumplen la relación de incertidumbre como una igualdad.

Comenzaremos por una breve descripción de los filtros complejos de Gabor que están definidos como filtros LSI (Linear Shift-Invariant)

Definición 18 (Filtro complejo de Gabor 1D) *Un filtro de Gabor complejo de una dimensión es un filtro LSI con respuesta al impulso*

$$h(x; N, u_0, \sigma) = g(x; N, \sigma) \exp(i2\pi u_0 x), \quad (3.1)$$

donde $g(x; N, \sigma)$ es la función Gaussiana

$$g(x; N, \sigma) = N \exp\left(-\frac{x^2}{2\sigma^2}\right). \quad (3.2)$$

Los parámetros del filtro de Gabor son una constante de normalización N , La frecuencia central u_0 y la varianza de σ de la función Gaussiana. Sin embargo, a menos que haya lugar a dudas, no escribiremos estos argumentos explícitamente.

Excepto cuando lo indiquemos específicamente usamos la normalización $N = (\sqrt{2\pi\sigma^2})^{-1}$ tal que $\int_{\mathbb{R}} g(x) dx = 1$. La parametrización dada en la definición (18) no cumple con las propiedades de wavelets de escalamiento. La siguiente parametrización busca obtener un parámetro para hacer el escalamiento.

$$h(x) = g(x) \exp(i\omega_0 x) \quad (3.3)$$

$$= g(x) \exp\left(\frac{icx}{\sigma}\right). \quad (3.4)$$

en (3.3) simplemente reemplazamos la frecuencia central u_0 por la frecuencia angular $\omega_0 = 2\pi u_0$ esto puede convertirse en (3.4) definiendo

$$c = \omega_0 \sigma \quad (3.5)$$

y substituyendo ω_0 por c/σ . De esta forma tenemos una parametrización útil para wavelets, porque el filtro de Gabor no cambia su forma cuando σ es variada mientras mantenemos c . Así, todos los filtro de Gabor con el mismo valor de c pueden ser derivados de un solo filtro mediante el escalamiento, en términos de wavelets a este filtro es la *Mother Wavelet*. La parametrización de este filtro de Gabor es la familia de Wavelets de Gabor.

Ahora definimos los filtro de Gabor complejos en 2D, ya que los filtros cuaterniónicos de Gabor están definidos para 2D únicamente.

Definición 19 (Filtro complejo de Gabor 2D) *Un filtro complejo de Gabor de dos dimensiones es un filtro LSI con respuesta al impulso*

$$H(x; u_0, \sigma, \varepsilon, \phi) = g(x', y') \exp(2\pi i(u_0 x + v_0 y)) \quad (3.6)$$

con

$$g(x, y) = N \exp\left(-\frac{x^2 + (\varepsilon y)^2}{\sigma^2}\right) \quad (3.7)$$

donde ε es el *aspect ratio*. Las coordenadas (x', y') son derivadas de (x, y) por una rotación sobre el origen a través del ángulo

$$\phi : \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (3.8)$$

de nuevo, tomaremos la normalización tal que $\int_{\mathbb{R}} g(x, y) dx dy = 1$; por ejemplo $N = \frac{\varepsilon}{2\pi\sigma^2}$.

Los filtros de Gabor son filtros pasabanda. La frecuencia central de un filtro complejo de Gabor esta dada por $f = \sqrt{u_0^2 + v_0^2}$ y su orientación es $\theta = a \tan(v_0/u_0)$. En la mayoría de los casos es conveniente elegir $\theta = \phi$ tal que la orientación del seno complejo coincida con la orientación de uno de los ejes principales de la función Gaussiana. Ahora, en analogía a los filtro complejos de Gabor introducimos los filtros cuaterniónicos de Gabor.

Definición 20 (Filtro Cuaterniónico de Gabor) *Un filtro cuaterniónico de Gabor es una función base de una Gaussiana*

$$h^q(x, u_0, \sigma, \varepsilon) = g(x, \sigma, \varepsilon) \exp(i2\pi u_0 x) \exp(j2\pi v_0 y) \quad (3.9)$$

$$= g(x, \sigma, \varepsilon) \exp(i\omega_1 x) \exp(j\omega_2 y) \quad (3.10)$$

$$= g(x, \sigma, \varepsilon) \exp\left(i\frac{c_1\omega_1 x}{\sigma}\right) \exp\left(j\frac{c_2\varepsilon\omega_2 y}{\sigma}\right). \quad (3.11)$$

donde la ultima parametrización es obtenida siguiendo un procedimiento similar al de la definición para filtros complejos de Gabor en una dimensión. Aquí no usamos una Gaussiana rotada.

Para frecuencias positivas u_0 y v_0 la mayoría de la energía se concentra en el primer cuadrante. Por lo tanto, convolucionando una imagen con un filtro cuaterniónico de Gabor resulta en una aproximación de una señal analítica cuaterniónica. No esta de mas comentar que un filtro cuaterniónico de Gabor y por lo tanto la TQW son aplicables solo para señales de 2 o 3 dimensiones.

En la figura (3.1) se presentan un ejemplo de un filtros cuaterniónicos wavelet. Esta puede ser llamada la wavelet madre cuaterniónica.

3.1.1 Fase local Cuaterniónica

Ahora definiremos la Fase local cuaterniónica de una señal real de dos dimensiones (imagen), como la fase angular de la respuesta un filtro cuaterniónico de Gabor. La fase angular es evaluada de acuerdo a las reglas dada en la tabla (2.1.1).

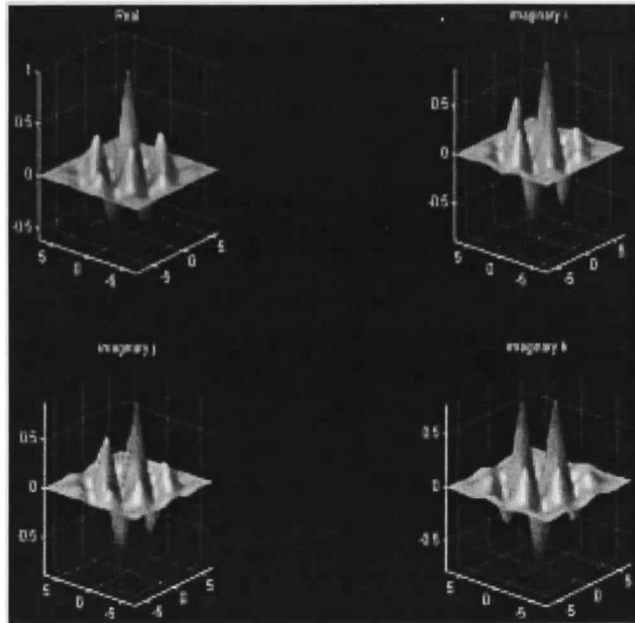


Figura 3.1: Filtro de Gabor Cuaterniónico con $c_1 = c_2 = 3$ y $\sigma = \pi/6$

En 1D podemos hacer la declaración: La estimación de la fase local y la posición espacial son equivariantes. ([Orf96]). En general la fase local de una señal varía monótonicamente sobre el vecindario de 2π . Existen solo algunos puntos de la señal con baja o cero energía en los que esta equivariancia no puede ser encontrada. Un ejemplo simple es la función coseno $\cos(x)$; si aplicamos un filtro de Gabor sintonizado correctamente para la estimación de la fase local ϕ de esta función, encontraremos que es casi igual a la posición espacial: $\phi(x) \approx x$ para $x \in [0, 2\pi]$. Esto nos lleva a una interpretación de la fase local cuaterniónica.

Haremos un ejemplo similar al de una dimensión reemplazando $\cos(x)$ por $\cos(x)\cos(y)$. Las primeras dos componentes de la fase local ϕ y θ se aproximan a la posición espacial: $\phi(x) \approx x$ y $\theta(x) \approx y$ o para $(x, y) \in [0, 2\pi] \times [0, \pi]$. En general estas dos componentes de la fase local son equivariantes con la posición espacial. La razón del intervalo $[0, 2\pi] \times [0, \pi]$ que viene de la definición de la fase angular para cuaterniones unitarios, puede ser entendido más fácilmente de la figura (3.2).

Mientras la posición espacial puede ser recuperado únicamente de la estructura local de la señal sin el intervalo $[0, 2\pi] \times [0, \pi]$, ocurrirán ambigüedades si el intervalo es extendido. El función completa $\cos(x)\cos(y)$ puede ser construida con copias de tamaño $2\pi \times \pi$. Considerando en este ejemplo que la tercer componente de la fase local es siempre cero: $\psi = 0$. El significado de esta componente de fase se vuelve obvio si variamos la estructura de la señal de prueba de la siguiente forma:

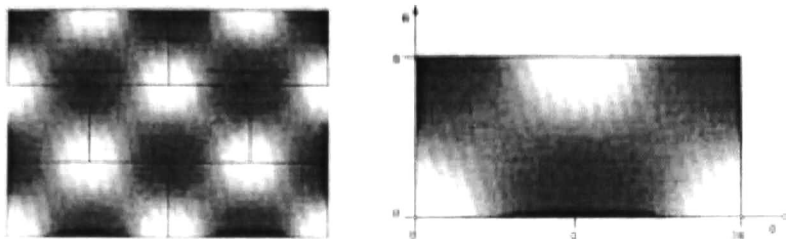


Figura 3.2: La función $f(x, y) = \cos(x)\cos(y)$ con $(x, y) \in [0, 4\pi] \times [0, 3\pi]$ [izquierda y $(x, y) \in [0, 2\pi] \times [0, \pi]$ [derecha. Figura Generada por Búlow [Bul99]

la función $\cos(\omega_1 x_1) \cos(\omega_2 x_2)$ puede ser escrita como la suma

$$\cos(\omega_1 x_1) \cos(\omega_2 x_2) = \frac{1}{2}(\cos(\omega_1 x_1 + \omega_2 x_2) + \cos(\omega_1 x_1 - \omega_2 x_2)).$$

si consideramos combinaciones lineales de la forma

$$f(x, \lambda) = (1 - \lambda)(\cos(\omega_1 x_1 + \omega_2 x_2) + \lambda \cos(\omega_1 x_1 - \omega_2 x_2)) \quad (3.12)$$

podemos considerar que ψ es constante con cada copia o región donde λ sea fija, y varia monóticamente con el valor de $\lambda \in [0, 1]$. este comportamiento es mostrado en la figura (3.3). El gráfico resulta de convolucionar versiones muestreadas de $f(x, \lambda)$ para diferentes valores de λ con un filtro cuaterniónico de Gabor y evaluar la fase ψ resultante de la respuesta del filtro de acuerdo a la tabla (2.1.1).

Otro método es evaluar la fase ψ de $f(x, \lambda)$ calculando la señal analítica cuaterniónica $f_A^q(x, \lambda)$ y su fase ψ analíticamente. De esto resulta

$$\psi(x; \lambda) = -\frac{1}{2} \arcsin \left[\frac{2(1 - 2\lambda)}{1 + (2\lambda - 1)^2} \right]$$

Las primeras dos componentes de fase ϕ y θ no cambian su significado, mientras λ varia. Solo para los valores $\lambda = 0$ y $\lambda = 1$, $\psi = \pm \frac{\pi}{4}$, la estructura se degenera a una estructura intrínsecamente unidimensional. Por tanto, la posición espacial no puede ser recobrada de la estructura local. Esto ocurre por la singularidad en la evaluación de la fase angular de un cuaternion para $\psi = \pm \frac{\pi}{4}$. en este caso solo $\phi \mp \theta$ puede ser evaluada. El grado de libertad restante puede ser eliminado estableciendo $\theta = 0$.

3.2 Análisis de imágenes usando la TQW

Cada filtro cuaterniónico de Gabor corresponde a dos filtros complejos de Gabor, usaremos esta propiedad para hacer el calculo de transformada cuaterniónica wavelets.

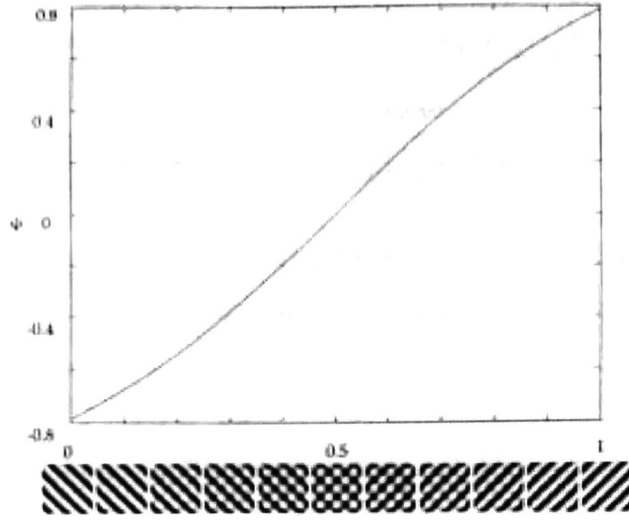


Figura 3.3: Dependencia de la tercer componente de fase ψ en la estructura local de la imagen. Figura Generada por Bülöw [Bul99]

Existe una relación sencilla entre los filtros cuaterniónicos y complejos de Gabor. Cada componente de un filtro complejo de Gabor con $\varepsilon = 1$ puede ser escrito como la suma de dos componentes de un filtro de Gabor cuaterniónico.

$$\begin{aligned}
 h_e(x, y) &= g(x, y) \cos(\omega_1 x + \omega_2 y) \\
 &= g(x, y) (\cos(\omega_1 x) \cos(\omega_2 y) - \sin(\omega_1 x) \sin(\omega_2 y)) \\
 &= h_{ee}^q(x, y) - h_{oo}^q(x, y)
 \end{aligned} \tag{3.13}$$

$$\begin{aligned}
 h_o(x, y) &= g(x, y) \sin(\omega_1 x + \omega_2 y) \\
 &= g(x, y) (\cos(\omega_1 x) \sin(\omega_2 y) - \sin(\omega_1 x) \cos(\omega_2 y)) \\
 &= h_{oe}^q(x, y) + h_{eo}^q(x, y)
 \end{aligned} \tag{3.14}$$

del mismo filtro cuaterniónico podemos obtener otro filtro complejo

$$\begin{aligned}
 h_e(x, y) &= g(x, y) \cos(\omega_1 x - \omega_2 y) h_{ee}^q(x, y) + h_{oo}^q(x, y) \\
 &= h_{ee}^q(x, y) + h_{oo}^q(x, y)
 \end{aligned} \tag{3.15}$$

$$\begin{aligned}
 h_o(x, y) &= g(x, y) \sin(\omega_1 x - \omega_2 y) \\
 &= h_{oe}^q(x, y) - h_{eo}^q(x, y)
 \end{aligned} \tag{3.16}$$

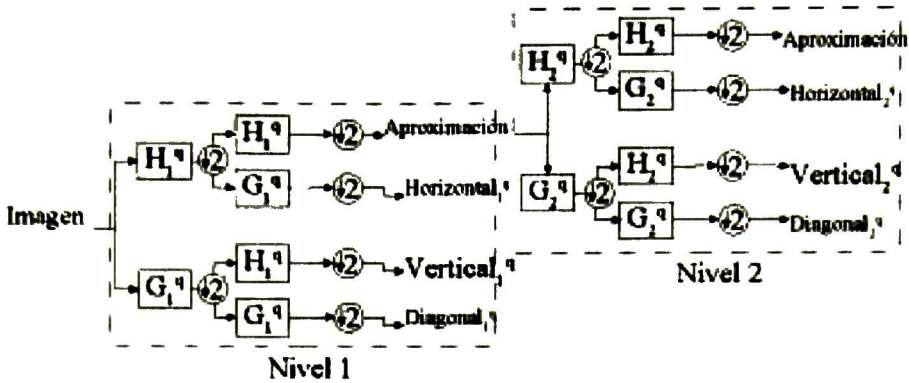


Figura 3.4: Pirámide Wavelet de una imagen en 2 niveles.

aquí h_{mn}^q es la componente cuaterniónica del filtro de Gabor donde $m, n \in (e, o)$ donde e se refiere a la parte par y o a la impar, de esta forma h_{eo}^q es la componente par-impar del filtro cuaterniónico.

El proceso multiresolución que usamos para generar la transformada cuaterniónica wavelet es similar al algoritmo de Mallat discutido en el capítulo 2. Solo que ahora usaremos dos filtros cuaterniónicos H_i^q y G_i^q donde i es la escala del filtro. Así H_1^q es el filtro wavelet de escala 1 también conocido como *mother wavelet*, algo similar sería para G_1^q llamado *Father wavelet*.

El proceso para realizar el análisis wavelet cuaterniónico se describe a continuación:

1. Se toma la señal de entrada en el nivel n y es convolucionada con los filtros H_n^q y G_n^q , sobre las filas.
2. La respuesta de los filtros H_n^q y G_n^q es convolucionada ahora sobre las columnas con H_n^q y G_n^q .
3. De las respuestas de estos filtros se toma una fila y se elimina la siguiente, al igual con las columnas, se toma la primera y se desecha la que sigue, este proceso es llamado en la literatura diezmar o decimar y se denota con el símbolo $(\downarrow 2)$.
4. La parte real de la aproximación en el nivel n es tomada como entrada para el nivel $n + 1$ y se repite el proceso desde el paso 1.

Este proceso en teoría se puede repetir n veces, y generar una pirámide de n niveles. En la práctica el número de niveles está determinado por el tamaño de la imagen, al ir dividiendo la imagen en dos, las imágenes del nivel $n + 1$ tienen la mitad del tamaño de la imagen de entrada al nivel n , en esta forma es inevitablemente detenerse donde la imagen contenga solo un pixel el nivel más alto y burdo de la pirámide.

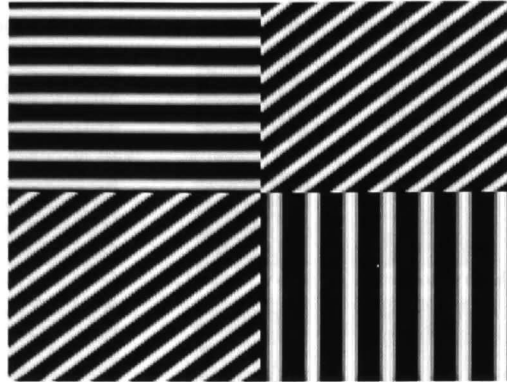


Figura 3.5: Función $\sin(xu + yv)$ con $(u, v) = (\frac{\pi}{6}, 0), (\frac{\pi}{6}, \frac{\pi}{6}), (0, \frac{\pi}{6})$

Desafortunadamente esto también está fuera de la realidad. Para aplicaciones esto no es útil, porque un pixel no contiene información suficiente para hacer un análisis, ni para ser la base de alguno, por lo general se usan de 3 a 5 niveles en la pirámide dependiendo, como mencionamos antes, del tamaño de la imagen. El proceso se muestra en la figura (3.4) para dos niveles.

Es importante remarcar que las respuestas de los filtros son cuaterniónicas y se manejan por separado cada una de sus componentes, Por lo tanto, la aproximación que aparece en la figura (3.4) tiene una parte *real*, y tres imaginarias *i, j* y *k*, cada componente tiene la mitad del tamaño de la imagen de entrada en ese nivel. Esto se muestra en la figura (3.6). Esta figura es el resultado de aplicar el procedimiento descrito anteriormente a la figura (3.5). Se pueden apreciar claramente como los detalles horizontales, verticales y diagonales de la imagen son extraídos con la transformación. además de obtener cuatro imágenes por detalle, descomponiendo así la señal en sus partes cuaterniónicas. Al final eso es lo que se busca con la extensión de la transformada.

La imagen original (3.5) tiene un tamaño de 80×80 píxeles y las imágenes mostradas en la figura (3.6) tiene un tamaño de 40×40 , aquí por razones de presentación se muestran escaladas.

La utilización de la convolución para obtener la transformada está justificada por el teorema de convolución de señales.

Teorema 21 (Teorema de convolución (TQW)) Sean f, g y h señales de dos dimensiones reales y sea F^q, G^q y H^q sus Transformadas wavelet cuaterniónicas.

$$f(x) = (g * h)(x) \iff F^q = G_e^q(u)H^q(u) + G_o^q(u)\beta(H^q(u)). \quad (3.17)$$

donde β denota es uno de los tres automorfismos del álgebra de cuaterniones como se definió en (2.15), G_e y G_o son las componentes pares e impares de G con respecto al segundo argumento.

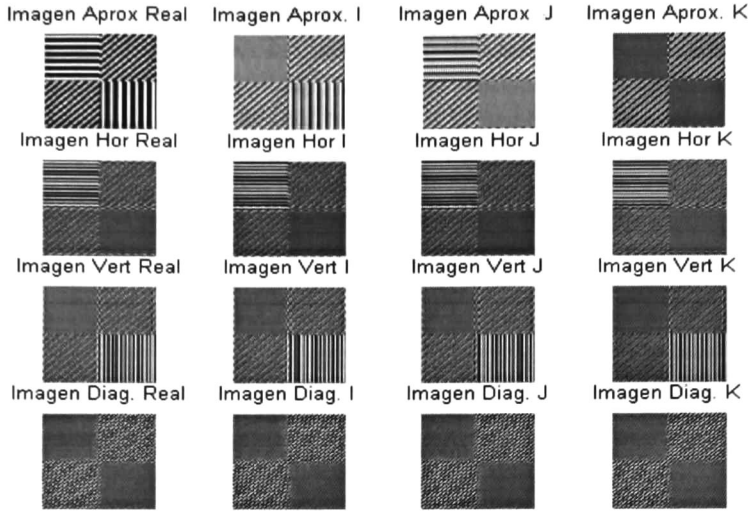


Figura 3.6: Transformada cuaterniónica de la imagen mostrada en la figura (3.5).

Prueba. tomando la definición para la transformada wavelet continua de dos dimensiones y usando los kernel cuaterniónicos definidos anteriormente tenemos

$$\begin{aligned}
 F^q(u) &= \int_{\mathbf{R}^2} e^{-2\pi i x u} \int_{\mathbf{R}^2} (g(x')h(x-x'))d^2x' e^{-2\pi j y v} d^2x \\
 &= \int_{\mathbf{R}^2} e^{-2\pi i x' u} g(x') H^q(u) e^{-2\pi j y' v} d^2x' \\
 &= \int_{\mathbf{R}^2} e^{-2\pi i x' u} g(x') \cos(-2\pi y' v) H^q(u) d^2x' \\
 &\quad + \int_{\mathbf{R}^2} e^{-2\pi i x' u} g(x') \sin(-2\pi y' v) \beta(H^q(u)) d^2x' \\
 &= G_e^q(u) H^q(u) + G_o^q(u) \beta(H^q(u))
 \end{aligned} \tag{3.18}$$

■

3.3 Complejidad Algorítmica de la TQW

En la TQW utilizamos un filtro cuaterniónico de Gabor, al usarlo en la computadora tenemos que usar una mascara discreta del filtro de la forma $h = [h_{m,n}]$, $m, n \in (1, 2, \dots, M)$ donde $h_{m,n}$ es de la forma mostrada en (3.11). De esta forma tenemos una matriz de $M \times M$ y el origen esta localizado en el centro de la matriz, por lo tanto, es

ventajoso la elección de M impar, para tener un pixel en el centro de la mascara. El numero de multiplicaciones necesarias por la convolución de una imagen $N \times N$ con una mascara $M \times M$ es de forma directa $O(N^2 M^2)$ como la mascara y la imagen decrecen en tamaño para el nivel dos de la pirámide tendremos que el orden es $O(\frac{N^2}{2}, \frac{M^2}{2})$, y para el siguiente nivel la mitad de lo anterior, por la regla del máximo el orden de toda la pirámide sigue siendo $O(N^2 M^2)$. Podemos reducir este limite si la mascara es separable $h = (h_c * h_r)$, donde h_c y h_r es un vector columna y un vector fila respectivamente de longitud M . La operación de convolución es asociativa y puede hacer el filtrado como

$$F = f * (h_c * h_r) = (f * h_c) * h_r \quad (3.19)$$

Entonces el numero de multiplicaciones se reduce al orden $O(N^2 M)$ Para revisar como las componentes de un filtro de Gabor pueden ser construidas como la suma de las componentes de un filtro cuaterniónico de Gabor ver ([Bul99]). Como los filtro cuaterniónicos de Gabor son separables, *se abre la posibilidad de implementar la convolución con filtros complejos de Gabor de forma separada*. Esto solo se puede hacer cuando la proporción de forma (*aspect ratio*) no es igual a uno. La razón es que la función Gaussiana con $\varepsilon \neq 1$ ya no es separable. Sin embargo en este caso la reducción de la complejidad algorítmica aun es posible. La función Gaussiana rotada en $\varepsilon \neq 1$ es

$$g(x, y; \varepsilon, \sigma, \phi) = \exp \left(-\frac{(x \cos \phi + y \sin \phi)^2}{2\sigma^2} - \frac{(-x \sin \phi + y \cos \phi)^2}{2(\varepsilon\sigma)^2} \right) \quad (3.20)$$

$$= \exp \left(-\frac{x^2 \cos^2 \phi + 2xy \sin \phi \cos \phi + y^2 \sin^2 \phi}{2\sigma^2} - \frac{x^2 \sin^2 \phi + 2xy \sin \phi \cos \phi + y^2 \cos^2 \phi}{2(\varepsilon\sigma)^2} \right) \quad (3.21)$$

entonces un filtro de Gabor con orientación ϕ y proporción de forma ε puede ser escrito como

$$h(x, y; \varepsilon, \sigma, \phi) = \exp \left(-\frac{2xy \sin \phi \cos \phi}{2\sigma^2} + \frac{2xy \sin \phi \cos \phi}{2(\varepsilon\sigma)^2} \right) (A(x, y) + B(x, y)), \quad (3.22)$$

donde A y B con funciones separables. Desafortunadamente no es posible descomponer la función exponencial en la ecuación (3.22) en una suma de funciones separables.

No obstante, el método puede ser aplicado, cuando la función Gaussiana sea separable, aun si el *aspect ratio* difiere de 1. Este es el caso, cuando la rotación solo es aplicada a los términos de modulación del filtro de Gabor y no a la función ventana. Este resultado no es de gran importancia ya que en las aplicaciones no usaremos filtros con *aspect ratio* diferente de 1.

Capítulo 4

Aplicaciones

El objetivo es demostrar como la transformada cuaterniónica wavelet puede ser utilizada para la solución de problemas. Las aplicaciones de la transformada wavelet son muy variadas, la mayoría son interesantes e implican un gran desafío. Así nos enfrentamos a la disyuntiva de la selección del problema y la solución de este con un algoritmo nunca antes probado.

La elección de la aplicación se basó en diferentes razones, primero el ambiente donde se desarrollo esta tesis, fue en el marco del laboratorio de visión computacional, esto restringe la selección de la aplicación a problemas de análisis de imágenes, por los conocimientos previos sobre análisis de imágenes y era inevitable la selección de una aplicación en esta materia. Además se buscaba una aplicación desarrollada empleando la transformada compleja wavelet, para hacer una comparación mas fácil, rápida y precisa del método con algo similar y cercano a la TQW, también para tratar de eliminar posibles obstáculos en la implementación. Buscábamos además, una aplicación donde los resultados obtenidos pudieran ser usados en los proyectos del laboratorio. Considerando todas estas razones y algunos otros parámetros no muy formales como el interés propio o el impacto a las audiencias, se eligió la estimación del flujo óptico.

En este capítulo se presenta una aproximación para resolver el problema de estimación del flujo óptico. Primero se define el problema que se quiere resolver, después se presentan diferentes tipos de métodos existentes en la literatura para resolver este problema. Se definen dos medidas mas, el calculo de la disparidad por medio de la fase cuaterniónica y la mascara de confianza. Usando estas dos medidas y el algoritmo para calcular la TQW definido en el capitulo anterior se genera un método para estimar el flujo óptico.

Una vez definido el método se presentan dos tipos de experimentos, los sintéticos y los reales. Los experimentos con datos sintéticos tienen como principal objetivo la comprobación de los métodos. Los experimentos con datos reales buscan establecer un punto de comparación con los métodos existentes.

4.1 Estimación del Flujo Optico

La estimación de movimiento es un tema que ha sido resuelto con wavelet complejas por Magarey ([Mag97]). Para la estimación de movimiento la principal característica de wavelet es que una traslación en la imagen resulta en un cambio en fase en los coeficientes wavelet. midiendo los cambios de fase es posible inferir el movimiento en la imagen. El principal obstáculo para la estimación del movimiento es que la fiabilidad de la estimación depende del contenido de la imagen. Por ejemplo, es sencillo determinar el movimiento de un punto en la imagen, pero es mucho más difícil detectar el movimiento de una pieza blanca sobre un fondo blanco. Definamos formalmente el problema que se desea atacar.

4.1.1 El Problema

Dadas dos imágenes f_1 y f_2 que difieren con un desplazamiento local (por ejemplo $f_1(x) = f_2(x + d(x))$) extraer el vector de flujo óptico o de desplazamiento d . Este problema es llamado estimación de disparidad, y d es llamado el vector de desplazamiento o campo de desplazamiento. Las componentes del campo de disparidad serán escritas como $d = (d_x, d_y)^T$ donde los índices indican la disparidad horizontal x , y vertical y .

Tenemos que establecer ciertas restricciones de forma que este problema tenga solución. Primeramente, el rango de d tienen que ser pequeño comparado con el dominio de f_1 y f_2 . De otra manera, estructuras visibles en f_1 pueden estar fuera del dominio de f_2 . Por ejemplo, f_1 y f_2 pueden no tener relación visible. En el caso que tratamos aquí, donde f_1 y f_2 son imágenes, esta restricción significa que ambas imágenes deben representar la misma escena, donde d es debido a diferentes condiciones de visión.

Otra restricción es impuesta por la ocurrencia de estructuras intrínsecamente unidimensionales. Un desplazamiento sobre tal estructura es localmente no perceptible. En este caso solo es posible encontrar la componente del vector de desplazamiento perpendicular a la estructura de una dimensión, si no hay más información *a priori*. Este es el llamado problema de la apertura.

En aplicaciones prácticas las dos imágenes pueden ser dos cuadros consecutivos de la secuencia de imágenes. Por ejemplo, $f_1(x) = f(x, t_1)$ y $f_2 = f(x, t_2)$. En este caso, un intervalo de tiempo $\Delta t = t_2 - t_1$ se encuentra entre el momento que se toma f_1 y f_2 y la solución d puede escribirse como $d = v\Delta t$ donde v es la velocidad de la imagen o flujo óptico. El campo de movimiento es la proyección en 2D del movimiento real en la escena en 3D, mientras la velocidad de la imagen es al entrada del problema de definición de la estructura del movimiento, que tiene la finalidad de encontrar la estructura de la escena de 3D en una secuencia de imágenes. La noción de estructura puede tener diferentes significados. Por ejemplo en ([Fau95]) se introduce una estratificación de visión en 3D y analiza cuanto información es necesaria para encontrar estructuras proyectiva, afina y Euclidiana de la escena.

Para llegar a hacer la estimación del flujo óptico primero se tenemos que resolver el problema de estimación de la disparidad, una vez que se tiene la correspondencia de los puntos de las imágenes f_1 y f_2 podemos encontrar el vector de desplazamiento \mathbf{d} de la escena. Veamos entonces los métodos existentes para resolver el problema de estimación de disparidad. Explicando porque usamos la fase cuaterniónica para esta estimación.

4.1.2 Métodos para resolver el problema de estimación de disparidad

Estableceremos muy brevemente las principales aproximaciones que existen para resolver este problema. Una clase de método usado es el basado en las características. En este método se inicia con la extracción de las características de una imagen como son las aristas, esquinas curvatura máxima local y demás. Generalmente estas características están esparcidas, haciendo posible encontrar sus correspondencia en ambas imágenes buscando solo en un numero pequeño de candidatos. Refiriéndonos a pequeño numero de candidatos comparado con el numero de pixeles en la imagen. Una desventaja de este método es que nos lleva a un mapa de disparidad esparcido, donde los valores de disparidad intermedios tienen que ser interpolados.

El otro método es el basado en regiones. Estos métodos extraen una región de la primera imagen y se busca su correspondiente en la segunda imagen. Esto se hace maximizando la correlación cruzada o minimizando la suma de la distancia cuadrada de los valores de gris. Este tipo de métodos también es llamado de correspondencia por bloques.

Otro importante método usado para estimar la disparidad es el basado en la fase que sera explicado y extendido a fase cuaterniónica en la siguiente sección.

4.1.3 Estimación usando la Fase cuaterniónica

El método de estimación de disparidad usando la fase local comienza con la suposición de que dos imágenes dadas f_1 y f_2 están relacionadas por

$$f_1(x) = f_2(x + \mathbf{d}(x)), \quad (4.1)$$

donde $\mathbf{d}(x)$ es el vector de desplazamiento buscado. Suponiendo que la fase varia linealmente, el desplazamiento $\mathbf{d}(x)$ es

$$d_x(x) = \frac{\phi_2(x) - \phi_1(x) + n(2\pi + k)}{2\pi u_{ref}}, \quad d_y(x) = \frac{\theta_2(x) - \theta_1(x) + m\pi}{2\pi v_{ref}}, \quad (4.2)$$

con frecuencias de referencias (u_{ref}, v_{ref}) que no son conocidas a priori. Aquí $\phi(x)$ y $\theta(x)$ son las primeras dos componentes de fase local cuaterniónicas medidas por un

filtro cuaterniónico de Gabor. Además, $n, m \in \mathbb{Z}$ son escogidas tal que d_x y d_y tomen valores en el rango admitido. Dependiendo de m , k es definido como

$$k = \begin{cases} 0, & \text{si } m \text{ es par} \\ 1, & \text{si } m \text{ es impar} \end{cases}$$

Algunas elecciones se deben hacer de los valores de las frecuencias de referencia (u_{ref}, v_{ref}) para poder estimar la disparidad. Que tan buena sea la aproximación de d depende estrechamente de la elección de las frecuencias de referencia u_{ref} y v_{ref} . Existen dos modelos para elegir las frecuencias. El primero es el llamado modelo constante donde u_{ref} y v_{ref} son elegidas como las frecuencias centrales del filtro de Gabor aplicado. Para el caso complejo el modelo constante ha sido usado en [Mak95] y [San88]. El segundo modelo es el local. Este modelo se basa en la suposición de que la fase local de puntos correspondientes en las dos imágenes tomaran el mismo valor $\phi_1(x) = \phi_2(x + d)$, un estimado de d es obtenida aproximando ϕ_2 por una serie de Taylor de primer orden sobre x :

$$\phi_2(x + d) \approx \phi_2(x) + (d \cdot \nabla)\phi_2(x). \quad (4.3)$$

Aquí usamos la notación $\phi = (\phi, \theta)$. Resolviendo (4.3) para d obtenemos el estimado de disparidad en el modelo local. En nuestros experimentos usamos la suposición adicional de que ϕ varia sobre la dirección x y θ sobre la y . Usando esta suposición estimamos la disparidad usando (4.2) con las frecuencias de referencia:

$$u_{ref} = \frac{1}{2\pi} \frac{\partial \phi_1}{\partial x}(x), \quad v_{ref} = \frac{1}{2\pi} \frac{\partial \theta_1}{\partial y}(x). \quad (4.4)$$

en las localidades donde u_{ref} y v_{ref} son iguales a cero, la ecuación (4.4) no esta definida. Estas localidades serán excluidas de la evaluación por medio de una mascara de confianza que se introduce después en este capitulo.

Los metodos de fase compleja usando el modelo local pueden ser refinados tomando el promedio entre las derivadas de la fase en la primera y segunda imagen, o el peso promedio de las derivadas donde el peso promedio es la energía de la respuesta al impulso del filtro en las imágenes. Estos refinamientos pueden ser aplicados también en el caso cuaterniónico. El uso de filtros cuaterniónicos de Gabor es eficientes por que el filtro es separable.

4.1.4 Estimación de la confianza

En regiones donde la energía de la imagen filtrada es baja la fase local no puede ser estimada monótonamente, al igual que en puntos donde la fase local puede ser cero, las estimación de la disparidad es imposible, porque la disparidad es igual a la diferencia de fase dividida por la frecuencia local.

Por estas razones se entiende la necesidad de tener una medida de confianza, que indique que tan segura es hacer la estimación en un punto. Se crea una mascara de confianza binaria que indica en que partes de la imagen es posible o imposible haber la estimación. En el caso complejo esta medida esta con la energía de la respuesta del filtro.

En el caso cuaterniónico necesitamos dos medidas de confianza, una para la fiabilidad de la componente de fase cuaterniónica ϕ y otro para la componente θ . Un resultado interesante es que estas dos medidas son idénticas. De (2.1.1) sabemos que las dos primeras componentes de fase cuaterniónica estas definidas para casi cualquier punto como

$$\phi(\mathbf{x}) = \arg_i(k^q(\mathbf{x})\beta(k^q(\mathbf{x}))), \quad \theta(\mathbf{x}) = \arg_j(\alpha(k^q(\mathbf{x}))k^q(\mathbf{x})). \quad (4.5)$$

Donde \arg_i y \arg_j son como se definió en (2.24) y como explicamos antes, estas funciones proyectan el cuaternion q en un subcampo complejo como indican los índices i, j y k y evalúan el ángulo-fase resultante del número complejo. k^q es la respuesta al filtro cuaterniónico. Usamos tres nuevas funciones que evalúan la magnitud de esos números complejos.

Definición 22 Sea $q = a + bi + cj + dk$ un cuaternion. Entonces el modulo parcial de q esta definido como

$$\text{mod}_i(q) = \sqrt{a^2 + b^2}, \quad \text{mod}_j(q) = \sqrt{a^2 + c^2}, \quad \text{y} \quad \text{mod}_k(q) = \sqrt{a^2 + d^2}. \quad (4.6)$$

entonces, $\text{mod}_i(q) \exp(i \arg_i(q)) = a + bi$ es la proyección de q sobre la primer subcampo complejo en representación polar. La proyección sobre los otros dos supcampos esta dada analogamente.

Con esta definición es posible adaptar las medidas de confianza conocidas para filtros complejos. Una aproximación es construir una mapa binario de confianza con un umbral de energía en la respuesta al filtro. La medida es hecha por

$$\text{Conf}(\mathbf{x}) = \begin{cases} 1 & \text{si } |k(\mathbf{x})| > \tau \\ 0 & \text{en otro caso} \end{cases}$$

el análogo para el caso cuaterniónico es

$$C_h(k^q(\mathbf{x})) = \begin{cases} 1 & \text{si } \text{mod}_i(k^q(\mathbf{x})\beta(k^q(\mathbf{x}))) > \tau \\ 0 & \text{en otro caso} \end{cases} \quad (4.7)$$

$$C_v(k^q(\mathbf{x})) = \begin{cases} 1 & \text{si } \text{mod}_j(\alpha(k^q(\mathbf{x}))k^q(\mathbf{x})) > \tau \\ 0 & \text{en otro caso} \end{cases} \quad (4.8)$$

Usaremos la siguiente mascara de confianza:

Definición 23 *Teniendo las respuestas al filtro cuaterniónico k_1^q y k_2^q de dos imágenes f_1 y f_2 el mapa de confianza binario esta dado por*

$$Conf_h(\mathbf{x}) = C_h(k_1^q(\mathbf{x}))C_h(k_2^q(\mathbf{x})) \quad (4.9)$$

$$Conf_v(\mathbf{x}) = C_v(k_1^q(\mathbf{x}))C_v(k_2^q(\mathbf{x})) \quad (4.10)$$

Donde, $Conf_h$ y $Conf_v$ son las medidas de confianza para la disparidad horizontal y vertical, respectivamente.

Sin embargo, no es necesario usar estas dos medidas, porque resultan ser idénticas. Como lo demuestra el siguiente teorema general.

Teorema 24 *La identidad*

$$\text{mod}_i(q\beta(\bar{q})) = \text{mod}_j(\alpha(\bar{q})q) \quad (4.11)$$

es verdad para cualquier cuaternion q .

Prueba. *El teorema puede ser verificado calculando ambos lados de la ecuación y haciendo una comparación componente a componente. ■*

Corolario 25 *Para las respuestas al filtro cuaterniónico k_1^q y k_2^q de cualquiera dos imágenes, la medida de confianza horizontal y vertical como están definidas en (4.9) y (4.10) son idénticas:*

$$Conf_h(\mathbf{x}) = Conf_v(\mathbf{x}) \quad (4.12)$$

Prueba. *Del teorema (24) tenemos que las dos funciones definidas en (4.7) y (4.8) son idénticas. Usando este hecho en la definición (23) se prueba el corolario. ■*

4.2 Método utilizando la TQW y la Fase cuaterniónica

El método que utilizamos para la estimación del flujo óptico consiste en 3 etapas, la primera es la transformación cuaterniónica wavelet de las dos imágenes en un concepto multiresolución, la segunda etapa es la estimación de la fase cuaterniónica en cada uno de los niveles de la pirámide construida en la etapa anterior, la tercera y ultima etapa es la estimación del flujo óptico con la información obtenida en los pasos anteriores desde el nivel mas burdo hasta el mas fino. A continuación describiremos cada una de estas etapas en el proceso de estimación del flujo óptico usando la TQW y la fase cuaterniónica.

TQW de las imágenes

En esta etapa empleamos los conceptos establecidos en el capítulo 3 para la creación de una pirámide wavelet de resolución cuaterniónica, usamos dos filtro cuaterniónicos de Gabor h^q y g^q que tienen la siguiente forma

$$h^q = g(x, \sigma_1, \varepsilon) \exp\left(i \frac{c_1 \omega_1 x}{\sigma_1}\right) \exp\left(j \frac{c_2 \varepsilon \omega_2 y}{\sigma_1}\right) \quad (4.13)$$

$$g^q = g(x, \sigma_2, \varepsilon) \exp\left(i \frac{\tilde{c}_1 \tilde{\omega}_1 x}{\sigma_2}\right) \exp\left(j \frac{\tilde{c}_2 \varepsilon \tilde{\omega}_2 y}{\sigma_2}\right) \quad (4.14)$$

con $\sigma_1 = \frac{\pi}{6}$ y $\sigma_2 = \frac{5\pi}{6}$ para que los filtros estén en cuadratura y con $c_1 = \tilde{c}_1 = 3$, $\omega_1 = ?$ y $\omega_2 = ?$, $\varepsilon = 1$. Con estos valores generamos una mascara cuaterniónica de 40×40 pixeles, esta mascara se va reduciendo en factores de dos al igual al mismo tiempo que aumentamos la frecuencia, para general toda la familia cuaterniónica wavelet a usar.

En la estimación del flujo óptico tenemos dos imágenes relacionadas por un desplazamiento, como se describió antes en este capítulo, por lo tanto, generamos dos pirámides wavelet cuaterniónicas de cuatro niveles cada una con los filtro definidos arriba. La elección de este número de niveles tiene que ver con la observación y el análisis al trabajar con las pirámides wavelet, los niveles superiores a cuatro despliegan muy poca información relevante, y por contener menos información y generan ruido, estos niveles resultan mas perjudiciales que ventajosos en el análisis. Este resultado no es general, estamos conscientes que para imágenes mas grandes el uso de 5 niveles o más seria útil y tal vez hasta necesario, pero con imágenes muy grandes, el problema de la complejidad de la TQW se convierte en una restricción, como se expuso en el capítulo 3, por lo tanto nuestros experimentos se limitaron a imágenes de tamaño igual o menor a 100×100 pixeles, y usando un kernel de 40×40 así en el nivel 4 de la pirámide a lo mas contamos con imágenes de 6×6 , otro nivel con el tamaño máximo de la imagen es, como ya lo dijimos, ineficiente e inútil.

En cada nivel de la pirámide tenemos 16 imágenes, como generamos una pirámide por cada imagen contamos con 32 imágenes por nivel, esto se puede observar en la parte superior de la figura (4.1). Todas estas imágenes se necesitan almacenar, en total el numero de imágenes existentes después de la transformada cuaterniónica wavelet es de 128, esto genera un costo de almacenamiento que es necesario tomar en cuenta si se desean hacer construir pirámides multiresolución de mas niveles y con imágenes mas grandes.

Calculo de la fase cuaterniónica

Para cada nivel de la transformada calculamos la fase cuaterniónica, usando la tabla (2.1.1), tomando el cuaternion $q = a + ib + jc + dk$ y definiendo como a, b, c, d las imágenes obtenidas de la TQW, el resultado son cuatro imágenes de fase por nivel, la magnitud y tres ángulos de fase, esto se hace para la aproximación y cada uno de los

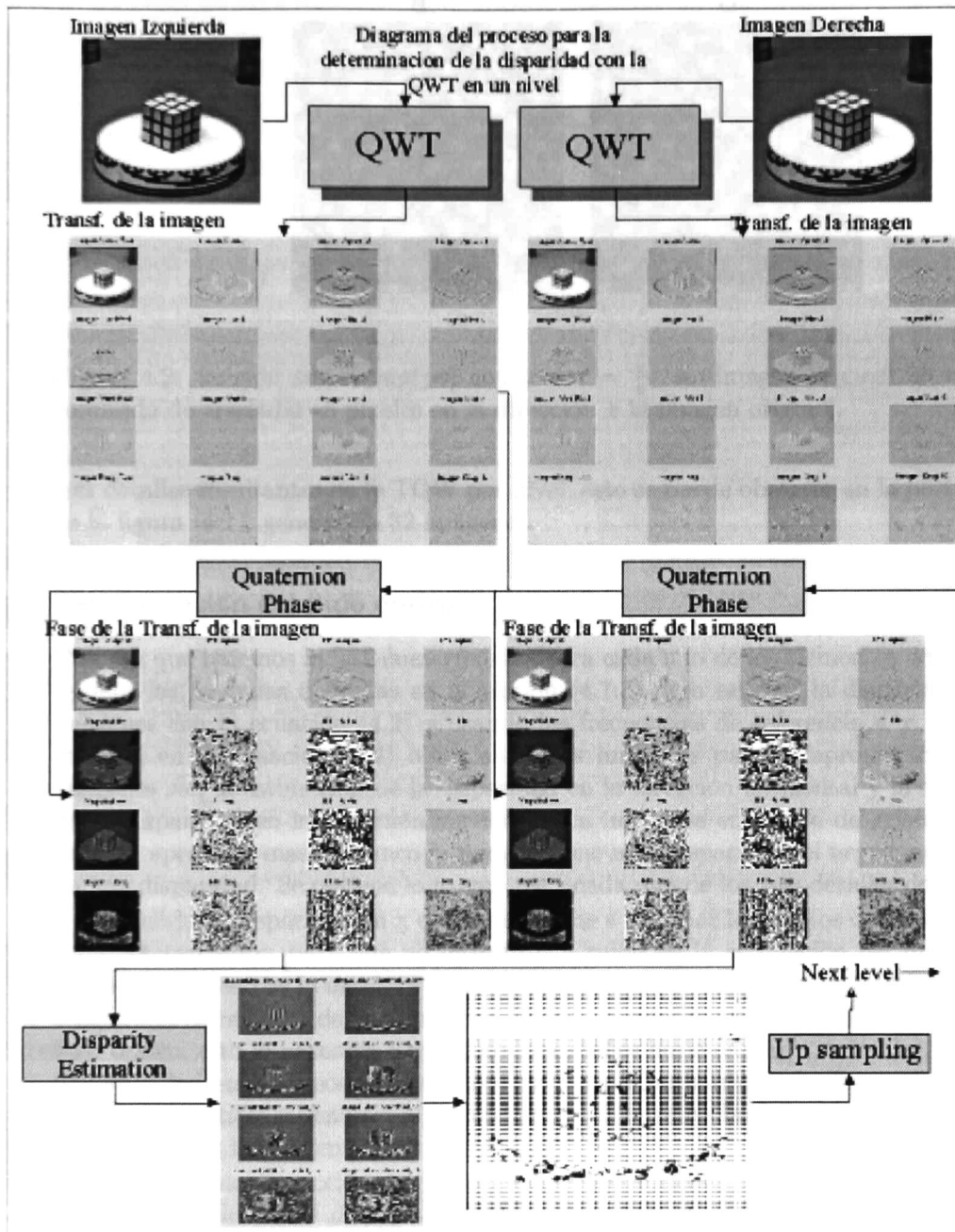


Figura 4.1: Diagrama de la estimación del flujo óptico en un nivel

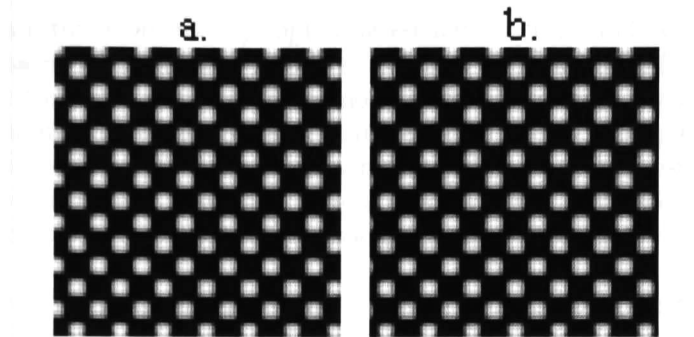


Figura 4.2: función $\sin(xu)\sin(yv)$ con $u = v = \frac{\pi}{6}$. a. Imagen original 80×80 . b. resultado de trasladar -3 pixeles en la dirección x la imagen original.

tres detalles resultantes de la TQW por nivel, esto se puede observar en la parte media de la figura (4.1), generando 32 imágenes.

Determinación del flujo óptico

Una vez que tenemos la fase cuaterniónica para cada uno de los elementos de la TQW, usamos las formulas definidas en la sección (4.1.3) para estimar la disparidad de las imágenes con la ecuación (4.2) y usando las frecuencias de referencia u y v como se muestra en la ecuación (4.4) obteniendo dos imágenes para la aproximación, estas imágenes son la estimación de la disparidad en la dirección horizontal y la estimación de la disparidad en la dirección vertical, para imágenes en escala de grises, entre el valor se aproxime mas al blanco la imagen tiene mas disparidad, el negro muestra que no hay disparidad. Se obtiene lo mismo para cada uno de los tres detalles de la TQW.

Teniendo la disparidad en x e y , procedemos a eliminar los puntos con poca energía o no determinados, para que no introduzcan ruido en la estimación del flujo óptico, para esto empleamos la mascara de confianza definida en la sección (4.1.4), una vez que tenemos la estimación de la disparidad filtrada, procedemos a hacer la estimación del flujo óptico, esto lo hacemos empleando los detalles horizontales y verticales, tomamos la disparidad en la dirección horizontal del detalle horizontal y la disparidad vertical del detalle vertical y generamos un vector que determina la dirección y magnitud del movimiento de la imagen en ese punto.

Hacemos una corrección del flujo óptico en esta misma etapa, empleando la información contenida en el detalle diagonal, tomamos la mascara de confianza junto con la estimación del flujo óptico de la imagen y si estos dos valores existen y estan en relacion con lo estimado por el detalle diagonal, o sea, si el movimiento estimado en cierto punto no es predominantemente vertical o horizontal, entonces usamos el valor de la componente diagonal para refinar la estimación.

Una vez refinada la estimación del flujo óptico la escalamos empleando el método de vecindario. Se toma el valor de un punto y se copia al vecindario de la imagen escalada. Esta estimación reescalada es usada como entrada al siguiente nivel inferior de la pirámide. En el siguiente nivel seguimos el mismo procedimiento, solo que al final de este, en el refinamiento del flujo óptico comparamos, además, con la estimación del flujo óptico del nivel anterior. Si existen pixeles que contienen información en el nivel anterior y en el actual no, copiamos la información de este punto en la estimación actual.

De esta forma procedemos hasta llegar al primer nivel de la pirámide. La estimación se va refinando desde el ultimo nivel de la pirámide, que es el mas burdo, hacia el mas fino que es el primero.

Un método similar se propone en [Wik92], en este trabajo se utiliza un método de detección de bordes junto con el par de imágenes originales correspondientes se construye una pirámide de resolución, donde el numero de niveles de las pirámides depende de la disparidad máxima del par de imágenes estereo.

4.3 Experimentos

El algoritmo descrito anteriormente fue aplicado primero sobre datos sintéticos, donde la disparidad era conocida, para comprobar y corregir posibles errores en el mismo. Se hicieron muchas y diversas pruebas, aquí presentamos solo un ejemplo de estas pruebas, que es el que tiene el menor error en la estimación de la disparidad. Una vez que el algoritmo era eficiente lo aplicamos a imágenes reales, de secuencias, la secuencia del “cubo rubik” y la del “taxi de Hamburgo”, estas imágenes son comunes como ejemplo en otros trabajos de estimación de flujo óptico, haciendo sencilla su comparación, el uso extensivo de estas imágenes se debe a que cumplen con muchas características que son buscadas en esta clase de problemas, estas características se exponen mas adelante.

4.3.1 Datos sintéticos

Aquí usaremos una imagen generada sinteticamente, es la imagen generada por la función $\sin(xu)\sin(yv)$ de dos dimensiones con $u = v = \frac{\pi}{6}$, los primeros experimentos se basaron en movimientos en la dirección x e y únicamente, una vez corregidos algunos errores que se encontraron en los algoritmos y la aplicación de estos se procedió a experimentos con rotaciones y traslaciones, aquí mostraremos un ejemplo detallado del desplazamiento en el eje x. El ejemplo de rotación es mucho mas corto ya que el procedimiento es el mismo, solo mostramos las imágenes resultantes de la estimación.

Las imágenes que se muestran en la figura (4.2) representan un movimiento horizontal de 3 pixeles hacia la izquierda de la imagen. De esta forma es fácil saber si nuestra estimación esta obteniendo resultados correctos. La TQW se presenta solo para la figura (4.2.a) en el apéndice (.2). En base a las dos pirámides y a la información de

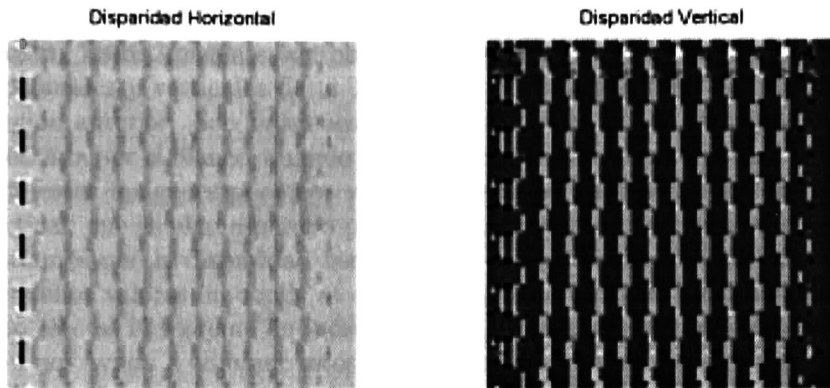


Figura 4.3: Disparidad horizontal y vertical para el nivel 1. El color negro corresponde a valores cero.

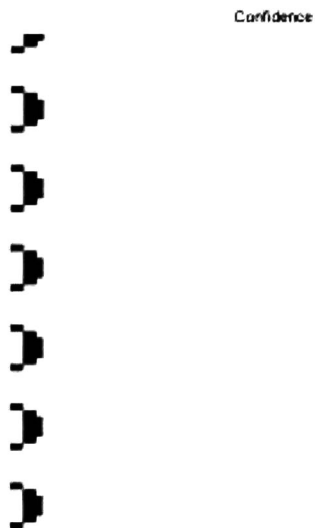


Figura 4.4: Mascara de confianza para el nivel 1. El Blanco es usado. El negro no.

fase obtenemos primero la disparidad horizontal y vertical que existe entre el par de imágenes para cada nivel de la pirámide. En la figura (4.3) se muestra la disparidad horizontal y vertical para el nivel mas fino de la pirámide, cada una de estas imágenes tiene 40×40 pixeles. Estas imágenes son obtenidas de la información de fase de las componentes horizontales y verticales de la TQW. Son el resultado final del proceso descrito en la sección anterior. Son imágenes en escala de grises donde la mayor disparidad es representada por el blanco, cuando no existe disparidad se encuentra el color negro. Analizando estas imágenes podemos ver como la disparidad horizontal es mas grande que la vertical, que es algo esperado por el desplazamiento que hicimos, en un caso ideal esperaríamos que la disparidad vertical seria cero, desafortunadamente, no es así, existen pequeños valores en donde se encuentran dos crestas de las senoidales. La justificación de esto es el teorema de incertidumbre. Afortunadamente, el proceso de multiresolución nos ayuda a eliminar estos errores, como lo veremos mas adelante.

El siguiente paso es filtrar cada imagen con la mascara de confianza, que se presenta en la figura (4.4). En esta figura podemos observar algunos lunares negros en la parte izquierda de la imagen, estos puntos son los que no entraran para la estimación del flujo óptico. Si observamos la figura (4.3) veremos en esta misma región en la disparidad horizontal y vertical áreas grandes con valor 0, que es lo que elimina la mascara de confianza.

Para apreciar mejor la importancia de la confianza veamos el flujo óptico antes de aplicar la mascara de confianza. La figura (4.5) muestra el flujo óptico. En muchas ocasiones la presentación del flujo óptico antes de la aplicación de la mascara de confianza puede causar confusión, aun cuando los vectores parecen mostrar el movimiento correcto, existen vectores en toda la imagen que resultan ser minimizados por los valores de la disparidad que tienen poca energía o son cero.

Una vez aplicada la mascara de confianza sobre la disparidad, se obtiene una imagen con mucha mas información, como lo muestra la figura (4.6). En el apéndice se muestran cada una de las estimaciones del flujo óptico por nivel.

Una vez que tenemos todos vectores del flujo óptico procedemos a la acumulación de la información. Desde el nivel 4 que es el mas burdo, se procede a escalar y acumular con la estimación del nivel 3. La nueva estimación del nivel 3 es ahora y refinada con los datos del nivel 2, y así hasta tener una sola imagen en el nivel mas fino. La estimación de final de este ejemplo se presenta en la figura (4.7). Esta estimación es muy parecida a la de la figura (4.6), pero si observamos bien, encontraremos que contiene mas información en ciertos puntos (compare la parte superior de ambas imágenes). Esta información es la ganancia que obtuvimos de usar el concepto de multiresolución.

Aun cuando la mayor información se presenta en el nivel 1, los niveles superiores son importantes para evitar el problema de la incertidumbre, en este ejemplo los resultados son excepcionales, y los demás niveles proporcionan información verdadera y útil, aunque no primordial, podríamos decir que tenemos información de mas, esto es por nuestra imagen original que cumple con las características necesarias para tener una estimación perfecta y es un movimiento simple. Logrando además que la mascara

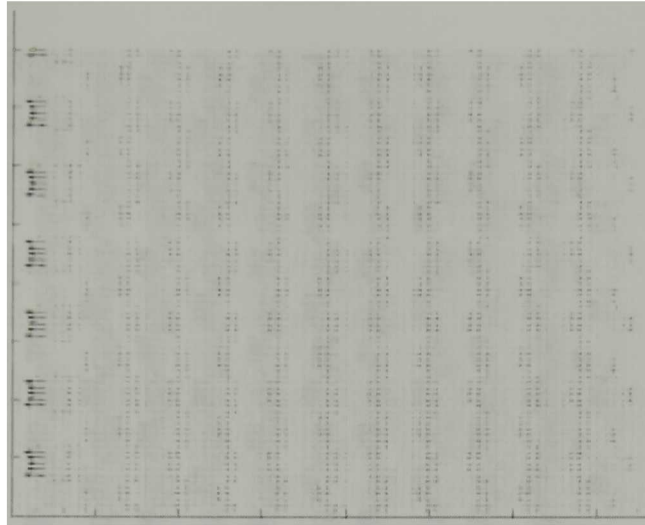


Figura 4.5: Estimación del flujo óptico en el nivel 1 sin usar la mascara de confianza.

de confianza elimine casi todos los valores errados o con ruido.

En los casos de datos reales, se observa información que logra pasar la mascara de confianza, mas aun si tenemos esquinas o bordes muy finos, esto genera ruido. La multiresolución nos ayuda a determinar de una forma mas eficiente, hacia donde se presenta el movimiento realmente y eliminar ciertas ambigüedades. Movimientos mas complejos como el de la una rotación requiere usar un esquema de este tipo mas claro. Observe la imagen en la figura (4.8). que corresponde a una rotación de la figura (4.2.a) los puntos del centro contiene errores estos son eliminados por los niveles siguientes de la pirámide.

4.3.2 Datos reales

Las imágenes usadas para el primer experimento son mostradas en la figura (4.9). Estas dos imágenes son parte de una secuencia, son imágenes tomadas a intervalos de 1 de segundo y el plato sobre el que se encuentra el cubo gira 1 grado por segundo. El movimiento del cubo cuando se observa desde arriba es en el sentido contrario a las manecillas del reloj.

Usamos el mismo procedimiento que se describe en la sección anterior y que se ejemplifica arriba en los datos sintéticos. Ya conociendo el procedimiento solo nos resta comentar sobre las imágenes. La mascara de confianza, muestra un poco de ruido en las esquinas superiores, estos puntos de la imagen no se están moviendo, sin embargo si existen datos en estas regiones serán usados para la estimación del flujo óptico, ya que la mascara no los desecha. Un detalle interesante es la exclusión por parte de la

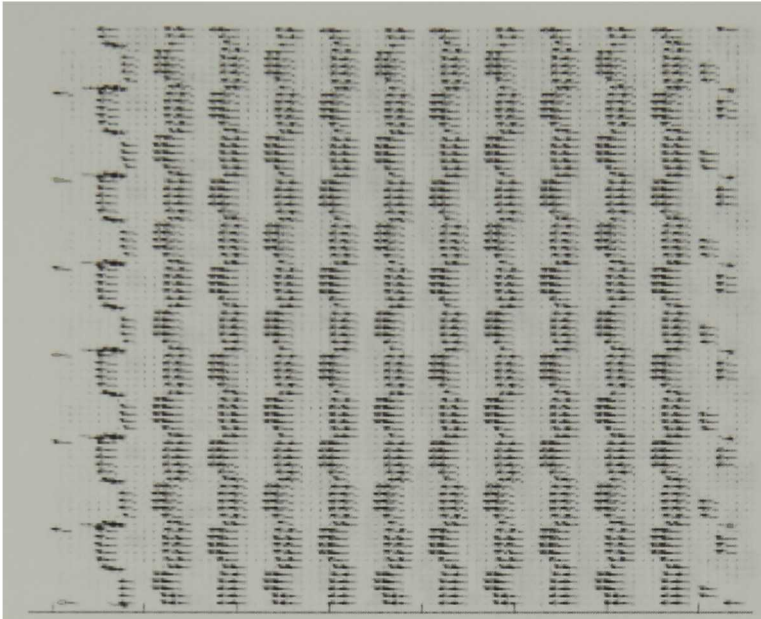


Figura 4.6: Estimación del flujo óptico una vez eliminadas las regiones con poca energía o cero.

mascara de confianza de la parte superior del plato, esto nos ayuda mucho a tener una imagen mas limpia, y así poder identificar mejor las dos partes que se están moviendo en la imagen.

Todas las imágenes de confianza y flujo óptico son mostradas en el apéndice (.2.1), excepto las que se muestran aquí. En la figura (4.11) podemos ver algunos vectores de ruido como los que están la parte superior, y algunos otros que muestran direcciones erradas, como algunos en el borde del plato. Estos puntos son eliminados con la acumulación de la información de los otros niveles.

En la figura (4.12) observamos que a diferencia de las imágenes sintéticas, la estimación mejora mucho por el esquema multiresolución, en este caso podemos observar vectores que no estaban presentes en la estimación del nivel 1. Observe los bordes del plato, y la parte superior del cubo, esta es información que no obtendríamos si solo usáramos la estimación de flujo óptico por medio de fase. Con una vista rápida podría parecer que la imagen final del flujo óptico muestra menos información que la del nivel 1, pero esto es un error, la estimación al nivel 1 esta, tal vez mas saturada de vectores en ciertos puntos, pero muchos de ellos se encuentran en direcciones erradas.

La estimación final del flujo óptico presenta vectores en todos los puntos en movimiento de la imagen, lo que la hace mejor, el problema es que no existe mucha densidad de vectores en estos puntos, lo que provoca que la vista sea menos homogénea. Esta

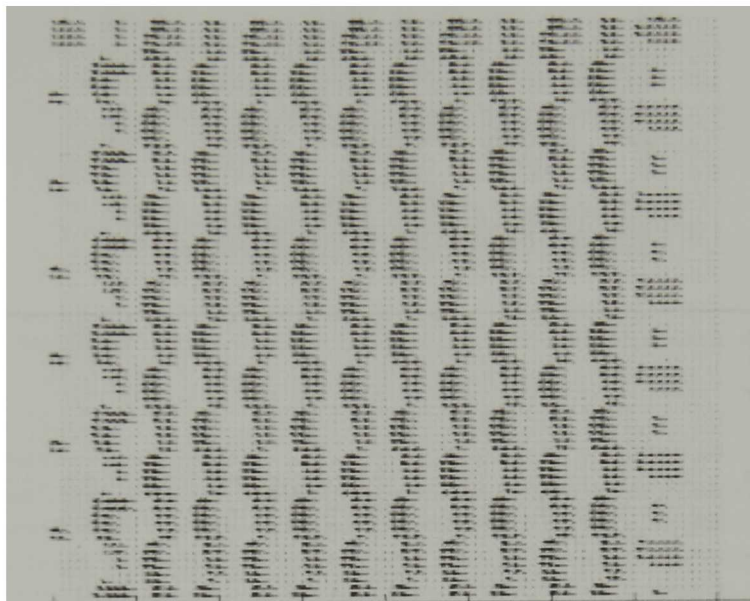


Figura 4.7: Estimación final del Flujo óptico.

dispersión provoca que la figura (4.12) se vea mas correcta que la final. En la practica no se utilizan estas imágenes mas que para presentación.

4.4 Comparación con otros métodos y discusión

El objetivo principal de este capítulo es la demostración de que la TQW puede ser usada para la solución de problemas en aplicaciones del mundo real. Mostramos resultados correctos y verídicos obtenidos del análisis de la TQW y la fase cuaterniónica. Si se comparan nuestros resultados con los mostrados en [Fle90], [Wik92], [Mag95], [Mag97], [Mal89] en su capítulo de aplicaciones, observaremos que no estamos por encima de estos resultados lo cual es una decepción, porque no podemos decir que el uso de la TQW resulta en mejores estimaciones o mejor análisis; pero tampoco tenemos resultados por debajo de la media, lo que demuestra que la transformada es útil y aplicable.

En [Fle90], se dice que la estimación basada en fase esta limitada a los componentes normales al vector de disparidad de una imagen orientada en la imagen. De acuerdo a esta visión, la estimación de la disparidad esta en 2D puede ser solo estimada de medidas locales, tomando en cuenta la suposición de que la imágenes intrínsecamente unidimensional, en cualquier punto. Sin embargo, aun si una señal puede ser localmente aproximada de forma correcta por una señal unidimensional parece no haber

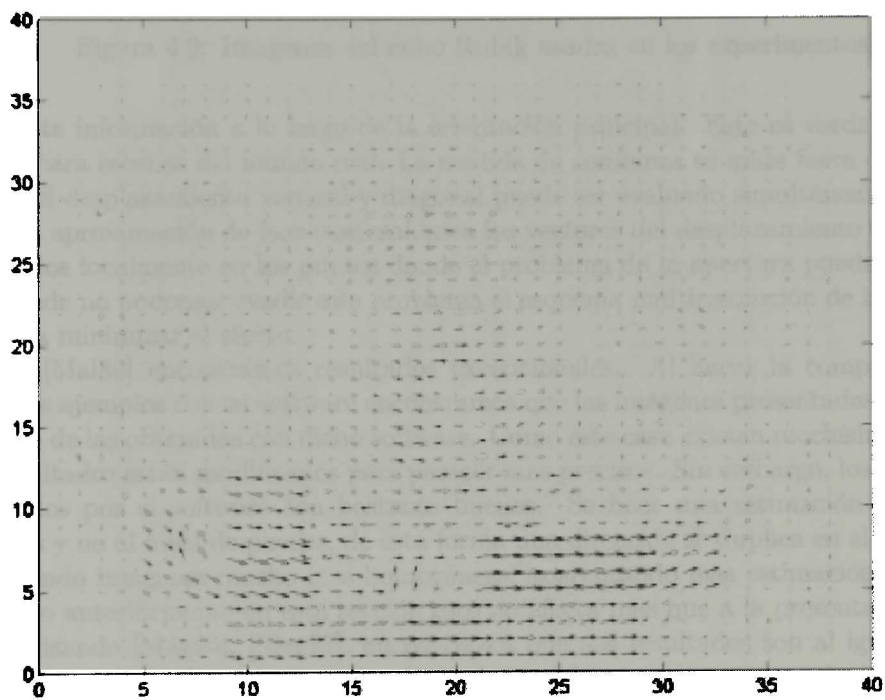


Figura 4.8: Estimación del flujo óptico en el nivel 1 de una rotación.

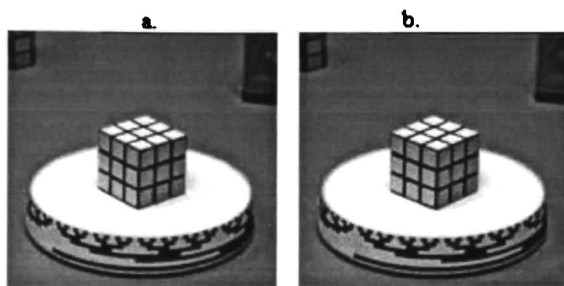


Figura 4.9: Imágenes del cubo Rubik usadas en los experimentos.

suficiente información a lo largo de la orientación principal. Esto es verdad especialmente para escenas del mundo real. La medida de confianza se mide fuera de regiones donde el desplazamiento vertical y diagonal puede ser evaluado simultáneamente. Usando la aproximación de fase cuaterniónica los vectores del desplazamiento totales son evaluados localmente en los puntos donde el problema de la apertura puede ser evadido, donde no podemos evadir este problema el esquema multiresolución de la TQW no ayuda a minimizar el efecto.

En [Mal89] encontramos resultados excepcionales. Al hacer la comparación de nuestros ejemplos con su software encontramos que las imágenes presentadas en el libro difieren de las obtenidas con dicho software. Como este caso existen muchísimos, donde los resultados están modificados para parecer mas precisos. Sin embargo, los resultados obtenidos por el software son bastante buenos. Se hace una estimación a nivel de bloques y no al nivel de pixeles, de esta forma una dirección se duplica en el vecindario obteniendo imágenes mucho mas homogéneas, aparentando mas estimación. Como se comento anteriormente en esta sección esto no afecta mas que a la presentación.

Revisando [Mag95], [Mag97], encontramos que sus resultados son al igual que los de los demás muy buenos. Se usan elipse de confianza que se van reduciendo hasta obtener una estimación confiable, y diferencias elevadas al cuadrado acumulativas. Este método tiene la cualidad de aceptar desplazamientos mas grandes de la imagen que los descritos en esta tesis.



Figura 4.10: Mascara de Confianza nivel 1.

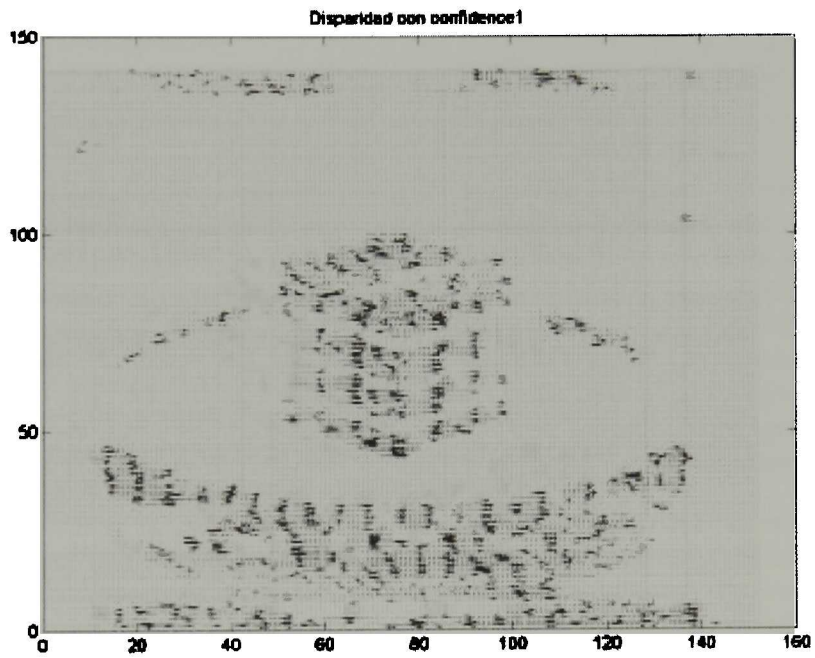


Figura 4.11: Estimación del flujo óptico nivel 1.

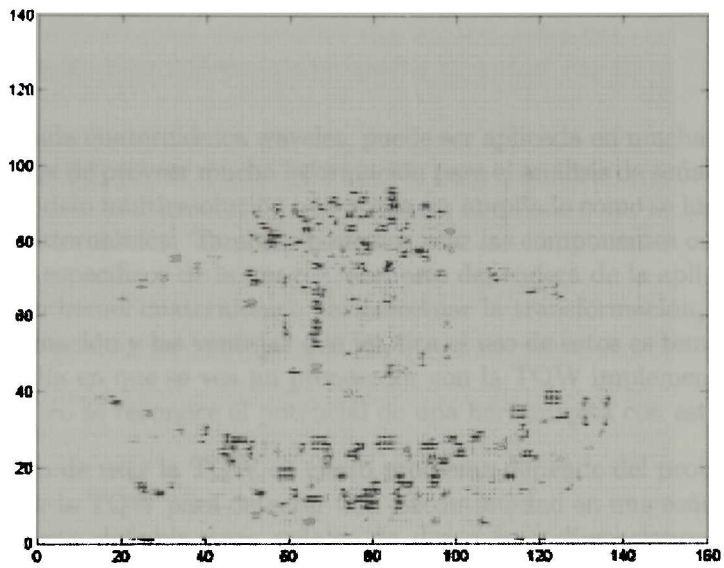


Figura 4.12: Flujo óptico del cubo Rubik (Final).

Capítulo 5

Conclusiones

La transformada cuaterniónica wavelet, puede ser aplicada en muchas y diversas áreas, tiene la ventaja de proveer mucha información para el análisis de señales. Esta desarrollada en un modelo multiresolución que puede ser ampliado como se hizo en la aplicación con la fase cuaterniónica. También podemos usar las componentes cuaterniónicas para extraer datos específicos de la imagen, aun esto dependerá de la aplicación. El alcance de emplear un kernel cuaterniónico para efectuar la transformación, esta aun inexplorado, la información y las ventajas que implica el uso de estos es tema de investigación a futuro. El día en que se vea un procesador con la TQW implementada en hardware esta lejano, pero se reconoce el potencial de una herramienta con estas características.

La decisión de usar la TQW en cierto problema depende del problema en si. No es posible aplicar la TQW para detectar una discontinuidad en una señal unidimensional, ya que solo está definida para señales de dos o mas dimensiones. Es posible que aun para problemas de análisis de imágenes, sea una herramienta demasiado grande y compleja, con resultados muy similares a los de otros métodos, solo con resultados verdaderamente excepcionales vale la pena aplicar la transformada, por la cantidad de recursos computaciones que consume. La complejidad algorítmica de la TQW es su mayor restricción. Esta restricción puede ser minimizada haciendo un análisis en la generación de las componentes wavelets, en ocasiones no necesitamos el uso de los tres detalles y si no se está interesado en la reconstrucción, se podría no calcularlos, acelerando el proceso. También se puede aplicar solo en puntos específicos de la imagen, para así aplicar la TQW solo en lo relevante. Reduciendo el numero de multiplicaciones y adiciones a efectuar. Si queremos analizar la imagen de una manzana situada en la parte superior izquierda de una imagen, realizar la transformación de toda la imagen es un desperdicio de recursos, se puede tomar solo la transformada de esa parte y mejorar el tiempo de computo. Además se puede aplicar una rotación al kernel para capturar la mayor energía o extraer características del punto de interés. Eliminando el uso del modelo de aproximaciones y detalles.

Comentarios sobre los resultados

Lo obtenido no es siempre lo deseado. Al inicio de este trabajo se tenía la esperanza de obtener resultados con la TQW superiores a los de otros métodos. Se esperaban resultados nunca antes vistos con la transformada compleja o real wavelet, al final los resultados fueron similares, mejores en ciertos casos, pero no superiores. Afortunadamente el objetivo de demostrar que la TQW es una herramienta que puede ser utilizada para resolver problemas se cumplió. Se programó y está al alcance de quien llegue a necesitar de una herramienta poderosa, que como cualquier otra tiene sus desventajas. En este momento el diseño de un problema para tratar de demostrar resultados superiores es difícil y tal vez un poco artificial. La superioridad en la solución de problemas se observará solo cuando se enfrente un problema con características que aun no son claras. La investigación de estas características puede ser tema para futuras investigaciones. Una vez encontradas dichas características, tal vez, el uso de la TQW sea más cotidiano.

Apéndices

.1 Complementos Matemáticos

Teorema de Fubini

Teorema 26 (Fubini) Si $\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x_1, x_2) dx_1 \right) dx_2 < +\infty$ entonces

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2) dx_1 dx_2 &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x_1, x_2) dx_1 \right) dx_2 \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x_1, x_2) dx_2 \right) dx_1. \end{aligned}$$

Teorema de Admisibilidad

Teorema 27 (Grossman, Morlet) Sea $\psi \in L^2(\mathbb{R})$ una función real tal que

$$C_\psi = \int_0^{+\infty} \frac{|\psi(\omega)|^2}{\omega} d\omega < +\infty \quad (1)$$

cualquier $f \in L^2(\mathbb{R})$ satisface

$$f(t) = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} Wf(u, s) \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) du \frac{ds}{s^2}, \quad (2)$$

y

$$\int_{-\infty}^{+\infty} |f(t)|^2 dt = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} |Wf(u, s)|^2 du \frac{ds}{s^2}. \quad (3)$$

Prueba. La integral derecha $b(t)$ de (2) puede ser reescrita como una suma de convoluciones. Insertando $Wf(u, s) = f \star \psi_s(u)$ con $\psi_s(t) = s^{-\frac{1}{2}} \psi\left(\frac{t}{s}\right)$ produce

$$\begin{aligned} b(t) &= \frac{1}{C_\psi} \int_0^{+\infty} Wf(\cdot, s) \star \psi_s(t) \frac{ds}{s^2} \\ &= \frac{1}{C_\psi} \int_0^{+\infty} f \star \psi_s \star \psi_s(t) \frac{ds}{s^2}. \end{aligned}$$

el operado ‘.’ indica la variable sobre la cual la convolucion es calculada. Probaremos que $b = f$ mostrando que sus transformadas de Fourier son iguales. La transformada de Fourier de b es

$$B(\omega) = \int_0^{+\infty} F(\omega) \sqrt{s} \psi^*(s\omega) \sqrt{s} \Psi(s\omega) \frac{ds}{s^2} = \frac{F(\omega)}{C_\psi} \int_0^{+\infty} |\Psi(s\omega)|^2 \frac{ds}{s}.$$

como ψ es real sabemos que $|\Psi(-\omega)|^2 = |\Psi(\omega)|^2$ y el ccambio de variable $\xi = s\omega$ prueba que

$$B(\omega) = \frac{1}{C_\psi} \int_0^{+\infty} \frac{|\Psi(\xi)|}{\xi} d\xi = F(\omega). \quad (4)$$

■

Teorema de Incertidumbre de Heisenberg

Teorema 28 (Incertidumbre Heisenberg) *La varianza temporal y de frecuencia de una función f satisface*

$$\sigma_t^2 \sigma_\omega^2 \geq \frac{1}{4}. \quad (5)$$

esta desigualdad es una igualdad sí y sólo si existe un $(u, \xi, a, b) \in \mathbb{R}^2 \times \mathbb{C}^2$ tal que

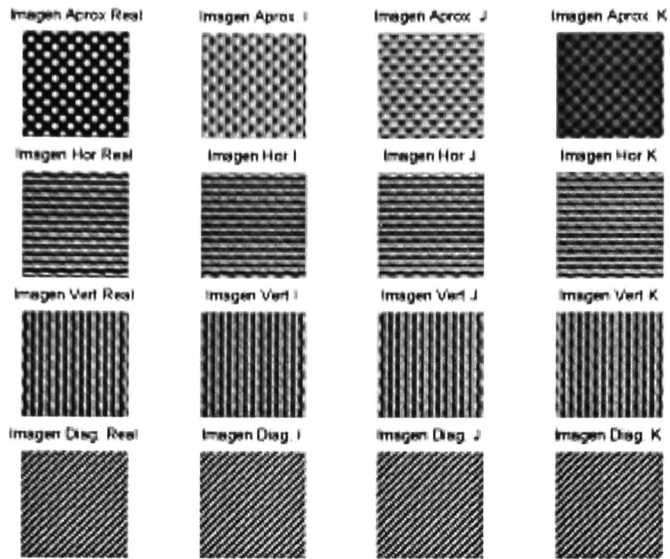
$$f(t) = a \exp[i\xi t - b(t.u)^2]. \quad (6)$$

.2 Imágenes de la Transformada Cuaterniónica Wavelet

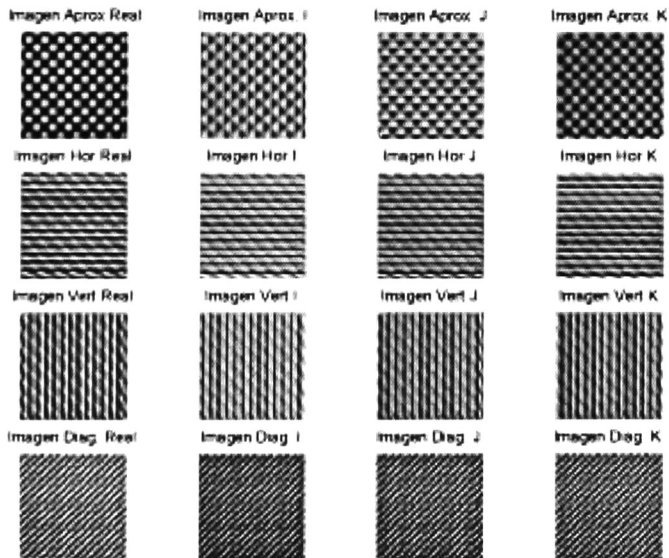
Por motivos de presentación, se presenta un número reducido de imágenes de las transformaciones. Los resultados mas importantes encuentran en el grueso de la tesis. Sin embargo, parece importante la presentación de algunas pirámides completas. Las imágenes están escaladas para su presentación. Niveles mas altos en la pirámide representan imágenes mas pequeñas.

TQW de la figura (4.2.a)

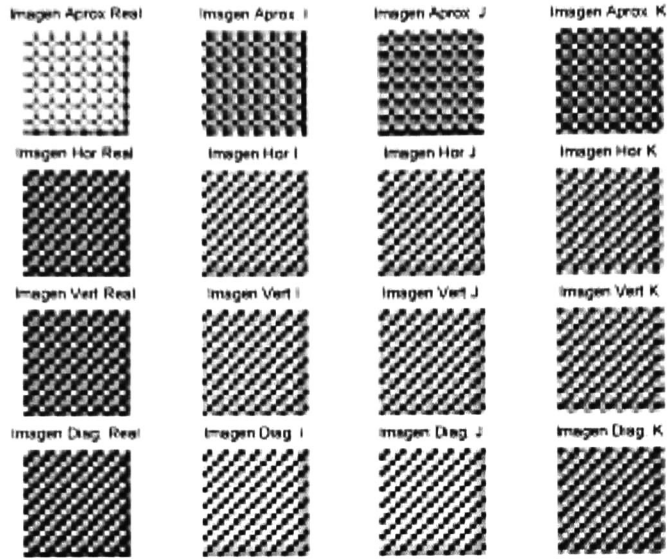
La pirámide de la TQW tiene 4 niveles.



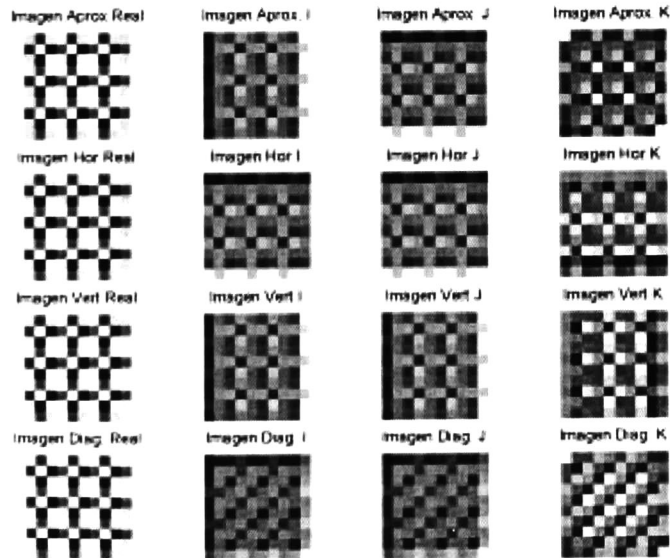
Nivel 1: tamaño de cada imagen 40×40



Nivel 2: tamaño de cada imagen 20×20

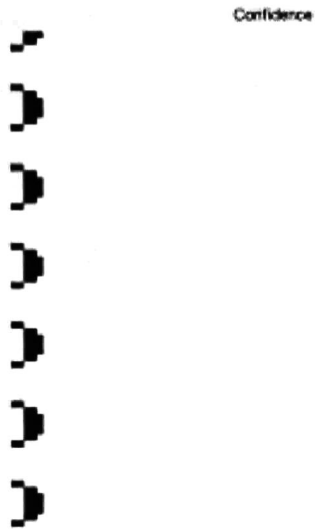


Nivel 3: tamaño de cada imagen 10×10

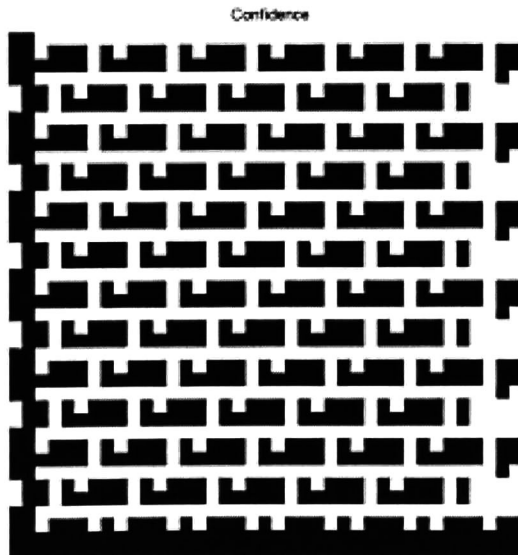


Nivel 4: tamaño de cada imagen 5×5

Imágenes de la mascara de confianza para cada nivel



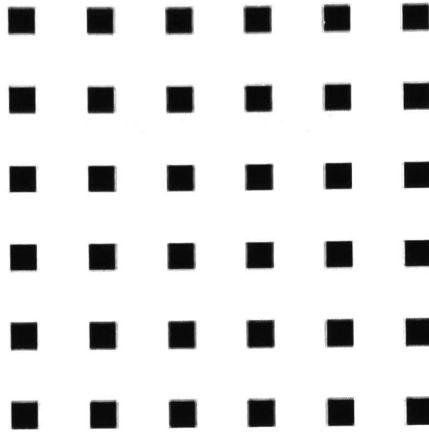
Imágen de la Confianza Nivel 1.



Imágen de la Confianza Nivel 2.

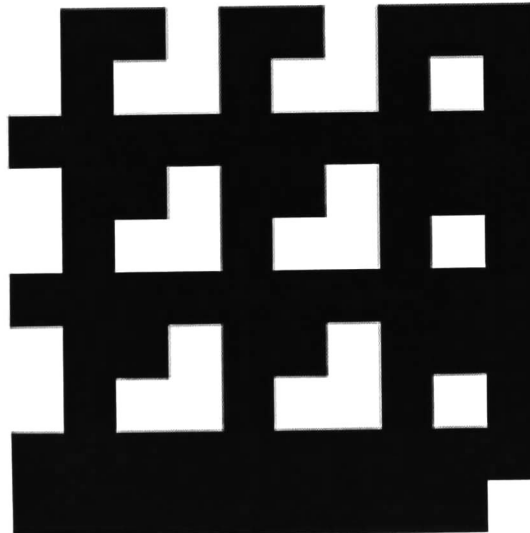
v. Basso

Confidence



Imágen de la Confianza Nivel 3.

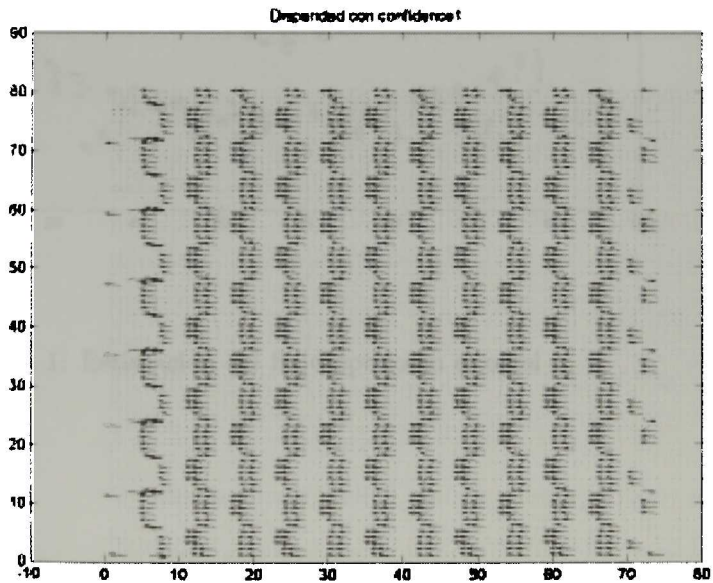
Confidence



Imágen de la Confianza Nivel 4.

Flujo óptico final

Despues de efectuar el proceso descrito en el capítulo 4 se obtiene la siguiente imagen



Flujo óptico para la función $\sin(ux + vy)$

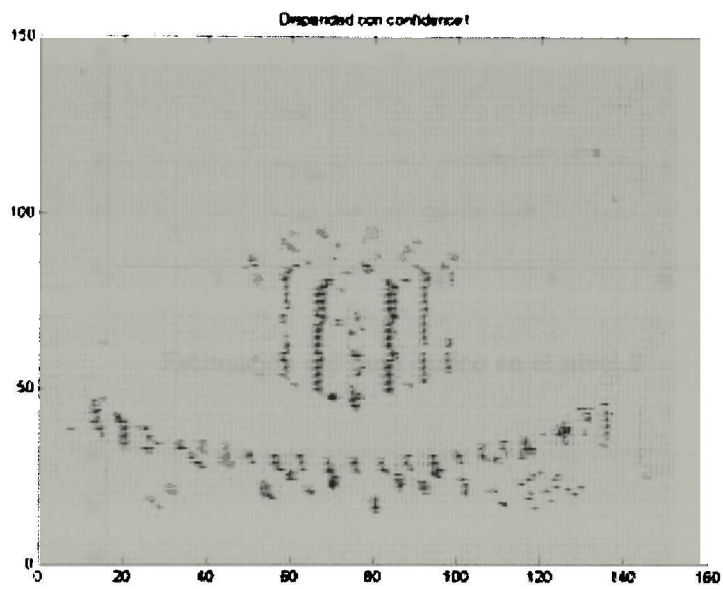
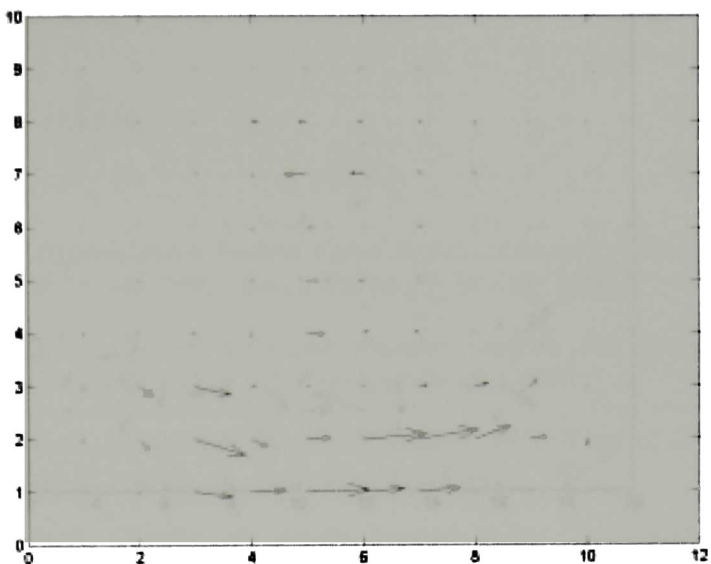
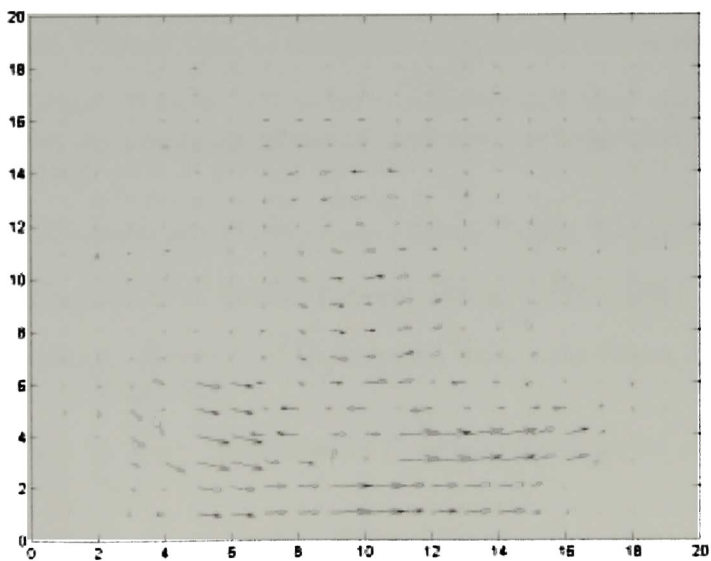


Figura 1: Estimación del flujo óptico en el nivel 1

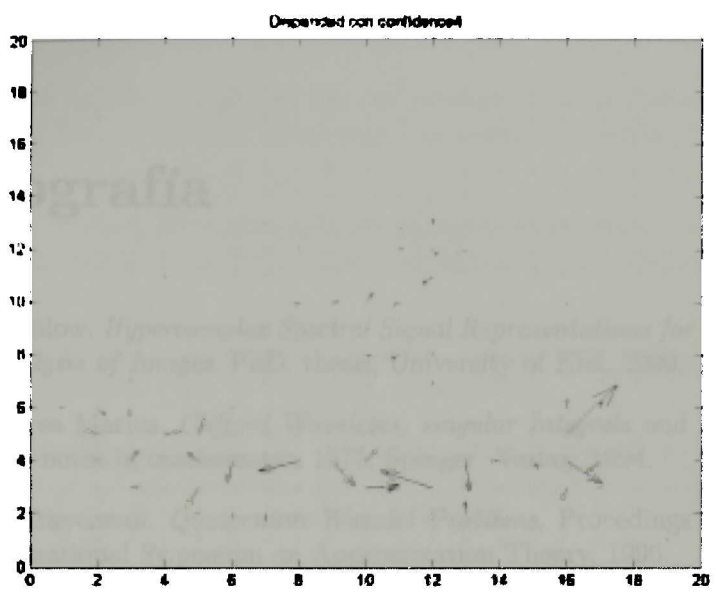
.2.1 Flujo óptico del cubo Rubik.



Estimación del flujo óptico en el nivel 2



Estimación del flujo óptico en el nivel 3



Estimación del flujo óptico en el nivel 4

Bibliografía

- [Bul99] T. Bülow. *Hypercomplex Spectral Signal Representations for the Processing and Analysis of Images*. PhD. thesis, University of Kiel, 1999.
- [Mit94] Mitrea Marius. *Clifford Waveletes, singular Integrals and Hardy spaces*. Lecture notes in mathematics 1575, Springer Verlag, 1994.
- [Tra96] L. Traversoni. *Quaternion Wavelet Problems*, Proceedings of the VIII Texas International Symposium on Approximation Theory, 1996.
- [Tra01] L. Traversoni. *Chapter 16: Image Analysis using Quaternion Wavelet*, in Geometric Algebra in Science and Engineering Book, E. Bayro Corrochano and G. Sobczyk, Springer Verlag, 2001.
- [Stra96] G. Strang, T. Nguyen. *Wavelet and Filter Banks*, Wellesley-Cambridge Press, 1996.
- [Per97] P. Lounesto. *Clifford Algebras and Spinors*. Cambridge University Press, 1997.
- [Cher95] V.M. Chernov. *Discrete orthogonal transforms with data representation in composition algebras*. In scandinavian Conference on Image Analysis, Uppsala, Sweden, 1995.
- [Hei90] W. Hein. *Structures and Clasification of Group Theory*, Springer Verlag, 1990.
- [Edd88] H.-D. Eddinghaus et al. *Zahlen*, Springer Verlag, Berlin, 1988.
- [Ham66] W.R. Hamilton. *Elements of Quaternions*. Longmans Green, London 1866. Chelsea, New York, 1969.
- [Kor68] A. Korn and M. Korn. *Mathematical HandBook for Scientist and Engineers*. McGray-Hill, 1968.
- [Mal99] Stéphane Mallat. *A wavelet Tour of Signal Processing*, Academic Press, 1999.
- [Mag97] J. Magarey and N. Kingsbury. *Motion estimation using Complex Wavelets*, PhD thesis, Cambridge University, 1997.

- [Fau95] O.D. Faugeras. Stratification of three-dimensional vision: projective, affine, and metric representation. *Journal Opt. Soc. Am. A*, 12:465-484, 1995.
- [Mag95] J. Magarey and N. Kingsbury. *Motion estimation using Complex Wavelets*, Technicakl Report TR-226, Cambridge University Engineering Departament, 1995.
- [Mal89] S. Mallat. *A Teory for multiresolution signal decomposition: the wavelet representation*. *IEEE trans. Patt. Anal. and MAch. Intell.*, 11(7):674-693, July 1989.
- [Lin97] J.-M.Lina, *Complex Daubechies Wavelets: Filters Desing and Aplications*, ISAAC conference, Univ. of Delaware, June 1997.
- [Mak95] A. Maki, L. Bretzner, and J-O. Eklundh. *Local fourier phase and disparity estimation: An Analytical study*. *Proc. Int. Conf. Computer Analysis of images and patterns CAIP*, Prag, paginas 868-874, 1995.
- [San88] T.D. Sanger. *Stereo disparity computation using Gabor filters*. *Biol. Cybernetics*, 59:405-418, 1988.
- [Wik92] J: Wiklund, C-J Westelius, H. Knutsson, *Hierachical Phase Based Disparity Estimation*. Departament of electrical engineering Linköping University, Sweden, 1992.
- [Fle90] D.J. Fleet and A.D. Jepson. *Computation of component iamge velocity from local phase information*. *IJCV*, 5:77-104, 1990.
- [Orf96] J. Orfanidis, *Signal Processing*. Prentice Hall. 1996.



Cinvestav

**Centro de Investigación y de Estudios Avanzados del
IPN**

Unidad Guadalajara

El Jurado designado por la Unidad Guadalajara del Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, aprobó la tesis: TRANSFORMADA CUATERIÓNICA DE WAVELET: TEORÍA Y APLICACIONES del(a) C. Juan Francisco ALVARADO CASAS el día 7 de Agosto de 2002 .

Dr. Valeri KORJIK
Profesor Investigador 3C
CINVESTAV GDL
Guadalajara

Dr. Jose Luis LEYVA
MONTIEL
Investigador Cinvestav 3B
CINVESTAV GDL
Guadalajara

Dr. Eduardo Jose BAYRO
CORROCHANO
Investigador Cinvestav 3B
CINVESTAV GDL
Guadalajara



CINVESTAV
BIBLIOTECA CENTRAL



SSIT000003933