

**CENTRO DE INVESTIGACIÓN CIENTÍFICA Y DE
EDUCACIÓN SUPERIOR DE ENSENADA**

**COMPRESIÓN DE IMÁGENES DIGITALES POR
MEDIO DE LA TRANSFORMADA DE ONDÍCULAS**

**TESIS
MAESTRÍA EN CIENCIAS**

JULIO CESAR ROLON GARRIDO

Ensenada, Baja California, México, junio de 1994.

RESUMEN de la tesis de **Julio César Rolón Garrido** presentada como requisito parcial para la obtención del grado de **MAESTRO EN CIENCIAS EN FÍSICA APLICADA** con opción en **ELECTRÓNICA Y TELECOMUNICACIONES**.
Ensenada, Baja California, México. Junio de 1994.

COMPRESIÓN DE IMÁGENES DIGITALES POR MEDIO DE LA TRANSFORMADA DE ONDÍCULAS

Resumen aprobado por:

Se presenta una técnica de compresión de imágenes alternativa a los métodos convencionales que usan la Transformada de Coseno Discreta.

Como principal resultado se propone un esquema de compresión de imágenes basado en la Transformada de Ondículas que aún cuando no es el mejor en términos del cuantificador y codificador, logra tasas de compresión aceptables y resalta algunos problemas abiertos cuya solución posterior contribuirá a mejorar el esquema.

La principal ventaja de la Transformada de Ondículas es su propiedad de localización espacio-frecuencia, contrario a lo que sucede con el análisis clásico de Fourier, en el cual sólo se tiene localización en frecuencia.

Relativamente, se ha realizado poco trabajo en el campo de las Ondículas, y los esfuerzos de normalización del mismo son inexistentes. Esto planteó dificultades al elegir los mejores métodos para cuantificar y codificar la imagen transformada, mismos que representan problemas abiertos, objeto de profunda investigación en la actualidad. Ha sido necesario tomar resultados normalizados obtenidos para la técnica de coseno discreta y acondicionarlos para nuestro caso con el fin de limitar el alcance del trabajo.

ABSTRACT of the thesis presented by **Julio César Rolón Garrido** in order to obtain the degree of **MASTER of SCIENCES in APPLIED PHYSICS** with the option of **ELECTRONICS AND TELECOMMUNICATIONS**. Ensenada, Baja California, México. June, 1994.

An alternative approach to traditional Discrete Cosine Transform image compression method is presented.

As the main result, a Wavelet Transform image compression scheme is proposed, which is not the best in terms of the quantizer and coder used, but it achieves acceptable compression rates and highlights some open problems whose subsequent solution will improve the scheme.

The main advantage of the Wavelet Transform is its space-frequency localization property, unlike the case of the classic Fourier analysis in which only frequency localization is achieved.

Work on wavelets field has not been greatly reported and standarization efforts are nonexistent to date. This fact was a problem when trying to choose the most appropriate methods to quantize and code the transformed image. Until today these are open problems subject to intense research efforts. It has been necessary to take standard results obtained for the Discrete Cosine Transform and adapt them in order to limit the extent of this work.

**CENTRO DE INVESTIGACION CIENTIFICA Y DE
EDUCACION SUPERIOR DE ENSENADA, B.C.**



DIVISION DE FISICA APLICADA

**DEPARTAMENTO DE ELECTRONICA Y
TELECOMUNICACIONES**

**COMPRESION DE IMAGENES DIGITALES POR MEDIO DE
LA TRANSFORMADA DE ONDICULAS**

TESIS

que para cubrir parcialmente los requisitos necesarios para obtener el
grado de **MAESTRO EN CIENCIAS** presenta:

JULIO CESAR ROLON GARRIDO

Ensenada, Baja California, México. Junio de 1994

CONTENIDO

	<u>Página</u>
I. INTRODUCCION	1
II. EL SISTEMA DE PERCEPCIÓN VISUAL HUMANO Y LA COMPRESION DE IMAGENES	4
II.1. El sistema de percepción visual humano	4
II.2. Estructura de una imagen digital	7
II.3. Redundancias de una imagen digital y su relación con la compresión	11
II.3.1. Redundancia psicovisual	13
II.3.2. Redundancia entre pixeles adyacentes	14
II.3.3. Redundancia de código	14
III. LA TRANSFORMADA DE ONDICULAS	16
III.1. Definición de ondícula y su relación con el análisis de señales	16
III.2. Aplicación de las ondículas en el análisis de resolución múltiple	22
III.3. Ondículas y bancos de filtros	28
IV. CUANTIFICACION Y CODIFICACION	32
IV.1. Cuantificación	32
IV.2. Codificación	35
IV.2.1. Codificación Huffman	36
IV.2.2. Codificación por longitud recorrida	37
V. RESULTADOS	39
V.1. Esquema de compresión utilizado	39
V.1.1. La transformada	40
V.1.2. El cuantificador	42
V.1.3. El codificador	43
V.2. Resultados	45
VI. CONCLUSIONES	52
LITERATURA CITADA	55

CONTENIDO (Continuación)

	<u>Página</u>
APÉNDICE A. PERSPECTIVAS DE LA COMPRESION DE IMÁGENES	57
A.1. Tendencias actuales.	57
A.2. Técnicas basadas en modelos.	59
A.2.1. Para objetos desconocidos.	59
A.2.2. Para objetos conocidos.	60
A.2.3. Para expresiones faciales.	61
A.3. Técnicas iterativas.	63
A.4. Normatividad.	64
ANEXO A. Tablas de codificación de longitud recorrida según la norma MPEG-2 de la ISO.	65
ANEXO B. Resumen de comandos para la operación de los programas de transformación y codificación.	69
ANEXO C. Coeficientes de Daubechies para los filtros de espejo en cuadratura.	70

LISTA DE FIGURAS

<u>Figura</u>		<u>Página</u>
1	Ojo Humano.	4
2	Muestreo de luminancia (X) y croma (O) en proporción 2:1, para el formato YUV.	10
3	(a) Ventana Gaussiana $g(x)$ y su correspondiente transformada de Fourier (a') que es la respuesta de un filtro pasa-bajas. (b) $g(x) \cos \omega_0 x$ y (c) $g(x) \cos 2\omega_0 x$, y sus respectivas transformadas de Fourier (b') y (c'). Las funciones tienen el mismo soporte pero el número de ciclos varía en (b) y (c). Las transformadas tienen el mismo ancho de banda pero diferentes posiciones en frecuencia.	17
4	Representación de las celdas de resolución en el plano frecuencia-espacio para la transformada de Gabor.	17
5	(a) Ondícula $\psi(x)$. (b) Ondícula $\psi_{s_1}(x)$ con $s_1 > 1$. (c) Ondícula $\psi_{s_2}(x)$ con $s_2 < 1$. (a'), (b') y (c') son las transformadas de Fourier de (a), (b) y (c) respectivamente. Las funciones $\psi_s(x)$ tienen un soporte que varía con la escala s , pero el número de oscilaciones es el mismo. Las transformadas tienen el mismo ancho de banda en una escala logarítmica.	19
6	Representación de las celdas de resolución en el plano frecuencia-espacio para la transformada de ondículas.	20
7	Esquema piramidal.	23
8	Esquema de descomposición por ondículas.	27
9	Etapa de análisis del esquema de descomposición por ondículas para señales bidimensionales usando filtros separables.	28
10	Zonas de influencia de $\Phi, \Psi^1, \Psi^2, \Psi^3$ en el plano frecuencial ω_x, ω_y	30
11	Cuantificador escalar típico: (a) con zona muerta, (b) sin zona muerta, (c) uniforme, (d) no uniforme.	33

LISTA DE FIGURAS (Continuación)

<u>Figura</u>		<u>Página</u>
12	(a) Esquema de compresión típico cuando se involucra una transformada. (b) Esquema de reconstrucción.	39
13	(a) Distribución de la aproximación $A_{2,i}f$ y los detalles $D_{2,i}^i f$ en una descomposición de 4 niveles. (b) En una imagen transformada.	41
14	(a) Imagen original, (b) filtrada con 2 etapas, con cuantificación en el primer nivel de resolución, (c) filtrada con 2 etapas sin primer nivel, (d) filtrada con 4 etapas, sin primer nivel, (e) filtrada con 8 etapas, sin primer nivel.	47
15	(a), (c) y (e) Imágenes originales; (b), (d) y (f) filtradas con 8 etapas, sin el primer nivel de resolución.	48
16	Imágenes con una alta entropía (a) y una baja entropía (b).	57
17	En el modelo para objetos desconocidos se hace un análisis para extraer forma y textura de los objetos de una imagen.	59
18	En el caso de objetos conocidos, se realiza una mascarilla poligonal (a) y (b), que se adapta al sujeto (c).	61
19	En el uso de modelos para expresiones faciales, un conjunto limitado de acciones posibles se almacena, como (a): ojos abiertos, boca cerrada, y (b): ojos cerrados, boca abierta.	62

LISTA DE TABLAS

Tabla
I Resultados Obtenidos

Página
49

COMPRESION DE IMAGENES DIGITALES POR MEDIO DE LA TRANSFORMADA DE ONDICULAS.

I. INTRODUCCION.

Ante el advenimiento de la era digital en la electrónica y las ventajas que introduce a esta área de trabajo el uso de los procesadores digitales monolíticos, surge el campo del procesamiento digital de imágenes. Diferentes manipulaciones se realizan sobre las imágenes, muchas de las cuales eran imposibles anteriormente. Bajo este contexto surge la necesidad de la compresión de imágenes con el fin de reducir la gran cantidad de información que se genera al digitalizarlas. Como ejemplo, una imagen típica de 512 x 512 pixeles en 256 niveles o tonos de gris necesita 262,144 octetos para su representación digital, sin considerar el color (que usando una proporción 2:1 en dos componentes de color duplicaría esa cantidad). Para aplicaciones tales como almacenamiento o transmisión esta cantidad de información haría ineficiente cualquier sistema, por lo que deben explotarse diferentes características de la imagen para lograr su compresión.

En el capítulo II se hace una breve descripción del sistema de percepción visual humano y su relación con las características mencionadas. Se resaltan los aspectos de resolución múltiple en la visión así como las redundancias que pueden ser eliminadas.

Las primeras técnicas de compresión, que generalmente no introducían pérdidas en la imagen, trabajaban en el dominio espacial, es decir, veían la imagen como un conjunto de pixeles cuyos niveles de gris diferían unos de otros en una cantidad finita. La compresión se lograba al codificar estas diferencias de tono entre un pixel y el adyacente. La compresión no era muy alta, sin embargo, la

imagen podía reconstruirse fielmente. Una de estas técnicas, que se sigue usando en la actualidad, es la conocida como DPCM (Differential Pulse Code Modulation en inglés) y se utiliza para comprimir imágenes cuyo contenido de información es crítico y debe conservarse íntegramente.

Para aumentar el grado de compresión es necesario transformar la imagen a un dominio en el cual su energía se encuentre concentrada en ciertas regiones, de manera que sean estas las que se preserven aún cuando se descartan las regiones de baja concentración de energía introduzca pérdidas. Dentro de estos métodos se agrupan todas las transformaciones espacio-frecuencia y, dentro de ellas, la transformada discreta de coseno ha sido la más explotada debido a su alta eficiencia para decorrelacionar secuencias de píxeles en el dominio espacial. Sin embargo, la principal desventaja de esta técnica ha sido que debido a las características de la transformación es necesario dividir la imagen en pequeños bloques, típicamente de 8×8 píxeles con el fin de simplificar el trabajo computacional requerido. Esto introduce fronteras o bordes artificiales a la imagen los cuales se hacen presentes al reconstruirla. Ante esto surgen las transformadas con traslape (LOTs o "Lapped Orthogonal Transforms" en inglés) que pretenden eliminar este efecto de bloquificación de la imagen eligiendo grupos traslapados de píxeles y efectuando la transformación en cada uno de ellos. Lógicamente las fronteras aún están presentes y la imagen reconstruida sufre los mismos cambios que sin el traslape.

Recientemente se ha propuesto una nueva transformación en la cual no es necesario descomponer la imagen en bloques, esta técnica agrupa los resultados de la teoría de descomposición en sub-bandas de frecuencia con la característica de resolución múltiple y las bases ortogonales para dar origen a una herramienta con propiedades ventajosas para el procesamiento de imágenes.

Esta herramienta, conocida como transformada de ondículas ("wavelets" en inglés) fue propuesta por Grossman y Morlet (1983) para el análisis de eventos sísmicos. En este trabajo no se pretende abordar la justificación de la existencia de las ondículas. En el capítulo III se mostrará que son familias de funciones idóneas para analizar eventos en imágenes, pues tienen un comportamiento similar al descrito en el capítulo II para el ojo humano. La descomposición frecuencial en bandas de una octava de ancho en un esquema de resolución múltiple es la característica de la descomposición de ondículas, tal y como sucede en el ojo humano. Si hablamos de bandas de frecuencia estaremos hablando de que es posible esta descomposición por medio de bancos de filtros de una octava de ancho de banda, y en efecto, puede demostrarse [Mallat, 1990; Daubechies, 1988] que es posible, bajo ciertas condiciones, descomponer una imagen en bandas de frecuencia por medio de un tipo de filtros especiales llamados de espejo en cuadratura o filtros de media banda (QMF en inglés) con propiedades de reconstrucción perfecta. El presente trabajo se centra en el uso de filtros de este tipo cuyos coeficientes fueron calculados por Daubechies (1988) y cuyas características fueron analizadas en el mismo trabajo.

En el capítulo IV se aborda el problema del diseño del cuantificador y codificador que se usaron para lograr la reducción real de la cantidad de información en un esquema de compresión con pérdidas. El capítulo V describe este esquema y los resultados logrados mientras que en el capítulo VI se presentan las conclusiones de este trabajo. El apéndice A presenta una breve reseña de las nuevas técnicas que se investigan actualmente, mismas que formarán toda una nueva concepción del problema de compresión de imágenes.

II. EL SISTEMA DE PERCEPCION VISUAL HUMANO Y LA COMPRESION DE IMAGENES.

II.1. El sistema de percepción visual humano.

El componente principal del sistema visual humano es el ojo, que es el elemento receptor de imágenes y a su vez transmisor de esta información al cerebro, que es quien la interpreta. El ojo humano consta de un sistema óptico muy peculiar [Encyclopaedia Britannica, 1974] (figura 1) que cuenta con una lente "ajustable" que permite proyectar sobre la retina una réplica de la imagen percibida.

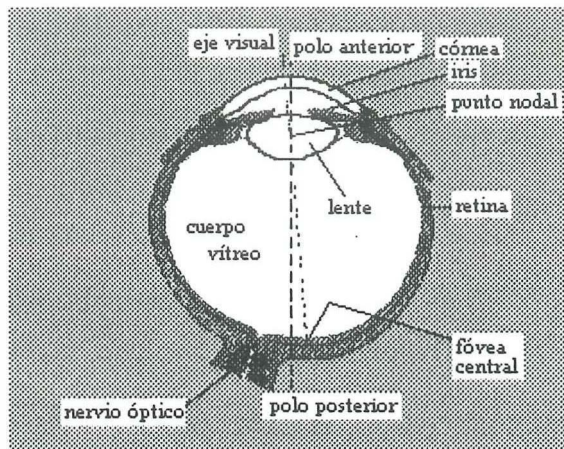


Figura 1. Ojo Humano.

La retina recibe la luz y la convierte en energía química por medio de dos clases de células fotosensibles: los conos y los bastones. Esta energía química activa las terminales nerviosas que conducen la información en forma de impulsos eléctricos al cerebro.

La retina es un arreglo complejo de tejidos en el cual las células fotosensibles están cubiertas por algunas otras capas no sensibles a la luz. Esto sucede en la mayor parte de su superficie, excepto en una región cercana al eje visual llamada fóvea central (*fovea centralis*), en la cual no existe ningún tejido que obstaculice el

paso de la luz. Es en esta región donde la agudeza visual es máxima, y es aquí donde se concentra la mayoría de los conos agrupados muy densamente. De hecho, en esta región no se encuentran bastones. Estos aparecen en grandes cantidades en la región circundante de la fovea central, llamada parafovea, y una mezcla de ambas células en la cual predominan los bastones es la que rodea a la parafovea, llamada perifovea. En total, distribuidos en la retina se encuentran alrededor de 7 millones de conos y de 75 a 150 millones de bastones en cada ojo.

La retina puede verse como un mosaico de células receptoras donde el diámetro de las mismas y su proyección sobre el punto nodal del ojo son los que determinan el poder de resolución del ojo, que encuentra el máximo para los conos de la fovea central y es de 30 segundos de arco.

Como se ha mencionado existen por lo menos 100 millones de células fotosensibles por ojo, sin embargo, solamente existen alrededor de 1 millón de nervios ópticos que se conectan al cerebro. Esto implica que deben existir agrupamientos convergentes de las terminaciones nerviosas de las células fotosensibles.

Para los conos, existe además un mecanismo de bloqueo de las trayectorias nerviosas de algunas células para enviar los mensajes de otras, es decir, mediante este mecanismo de bloqueo selectivo puede darse la impresión de que existe una conexión uno-a-uno entre estos y las terminaciones nerviosas.

Para los bastones existe un mecanismo similar, sin embargo, la cantidad de células agrupadas es mucho mayor. Esto explica el que los bastones no permitan apreciar detalles, pues la información que aportan los nervios ópticos a los cuales se conectan corresponde a grandes áreas de la retina. Las conexiones de los bastones son menos exclusivas que las de los conos, haciendo por lo tanto que no

aporten detalle a la formación de la imagen. Ahora puede comprenderse el hecho de que para visión nocturna o en condiciones de poca luz no se aprecie detalle alguno, pues en estas condiciones los grandes agrupamientos de bastones son los que aportan mayor información.

Para observar con detalle algún objeto, es necesario orientar el ojo de manera que dicho objeto se ubique en la región de la fovea central. Estos movimientos son instintivos y pueden ocurrir muchas veces por segundo si el objeto de interés es grande.

Es necesario notar también un hecho básico de la visión, y es que al voltear la mirada o abrir los ojos, la primera impresión que se recibe es un panorama general del campo de visión, definiéndose progresivamente los objetos que caen dentro de la región de la fovea. La visión general la proporcionan los bastones, y la visión aguda los conos. Puede decirse que el mecanismo de la visión es de resolución múltiple, donde varios niveles de resolución se desarrollan progresivamente hasta obtener información muy detallada de una escena.

Algunos experimentos han demostrado que esta resolución múltiple está relacionada con bandas de frecuencia de aproximadamente una octava de ancho, es decir, bandas del mismo ancho en una escala logarítmica.

El espectro visible, que se obtiene de la refracción a través de un prisma, genera diferentes colores que van del rojo al violeta (longitudes de onda de 7600 a 4000 Å). La respuesta de los conos y bastones varía de acuerdo a la longitud de onda de la luz. Se ha comprobado que los bastones tienen un máximo de sensibilidad alrededor de los 5000 Å (en la región verde-azul). Sin embargo, también se ha demostrado que aún cuando en esta región su sensibilidad es

máxima, no es posible discriminar el color de que se trata. Esto implica que la respuesta de los bastones es acromática.

Los conos, por el contrario, son sensibles a niveles mayores de iluminación. Si se hace incidir luz blanca sobre la retina, todos los receptores serán iluminados por igual y aún cuando su sensibilidad espectral sea distinta, juntos darán la sensación de luz blanca. Sin embargo, es posible generar luz blanca utilizando solamente tres colores primarios: rojo, verde y azul. Esto quiere decir que existen tres clases de conos receptores sensibles a cada una de estas longitudes de onda. Por lo tanto, será posible generar, por medio de la combinación adecuada de estos tres colores primarios, toda la gama posible de colores del espectro visible. Mas aún, existen muchos colores que pueden crearse con la combinación de solo dos de los colores primarios.

Con estos antecedentes se puede explicar que algunos objetos que bajo la luz del sol reflejan colores intensos, en la oscuridad parecen ser grises o carecer de color. Por otro lado, es comprensible entonces el hecho de que en una imagen en blanco y negro, en condiciones normales de iluminación, puedan apreciarse mejor los detalles, ya que todos los receptores disponibles (conos y bastones) estarán recibiendo la misma información, pues se carece de selección espectral.

II.2. Estructura de una imagen digital.

La retina puede verse como un mosaico de células donde cada una de ellas aporta una pequeña cantidad de información localizada espacialmente dentro del total de la imagen. Este efecto puede compararse al que ocurre en una película fotográfica, donde los "granos" o elementos sensibles a la luz contienen - una vez expuesta la película - la información de la cantidad de luz que recibieron. Si los

granos se distribuyesen uniformemente formando una matriz de dimensión $m \times n$, se tendrían $m \times n$ elementos en la imagen, que proporcionaría cada uno de ellos la información local de la cantidad de energía luminosa recibida.

En ninguno de los dos sistemas anteriores puede lograrse el alineamiento de los receptores para formar una matriz de determinadas dimensiones. Nuestro interés en lograrlo estriba en el hecho de poder manipular cantidades finitas y regularmente espaciadas de información. Las cantidades finitas de elementos de una imagen (conocidos como píxeles o pels) nos permiten conocer la cantidad de información por manipular para una determinada imagen, y el espaciamiento entre ellos nos permite determinar el poder de resolución máximo que se logrará (como en el caso de la retina mencionado en II.1). Lógicamente habrá una relación directa entre el número de píxeles y su espaciamiento para un área determinada de distribución.

La sensibilidad a la luz es el otro factor importante que completa la formación de la imagen. El sistema visual humano es sensible a grandes variaciones de luminosidad, desde el umbral de visión nocturna hasta los altos niveles de luminosidad recibidos cuando, por ejemplo, se ve directamente al sol.

Existen sistemas auxiliares que regulan la cantidad de luz que incide en los receptores, sin embargo, la intensidad luminosa toma un continuo de valores en el intervalo mencionado.

Una imagen digital es una versión limitada tanto en dimensiones físicas como en número de intensidades luminosas de una imagen real. Las limitantes son impuestas por el proceso físico de su adquisición así como por los medios físicos para su almacenamiento y representación.

Los sensores que se emplean para captar la imagen en la mayoría de los casos son verdaderas matrices de dimensión $m \times n$ cuyos elementos están uniformemente distribuidos en el espacio. La respuesta en intensidad luminosa es limitada a valores finitos y conocidos con el uso de convertidores analógico-digital, capaces de distinguir un cierto número de niveles posibles (256 niveles en el caso de un convertidor de 8 bits, aunque en aplicaciones de alta resolución, este número se incrementa a 12 e incluso a 16 bits).

Se ha visto en la sección II.1 que cualquier color puede ser generado por la combinación de los tres colores primarios, rojo, verde y azul. Lo mismo sucede en una imagen digital, donde el color de cada uno de los píxeles está dado en la misma forma.

Para la representación de imágenes en color existen diferentes formatos, entre ellos, el más evidente es el formato RGB (del inglés red, green, blue, correspondiente a rojo, verde y azul), que imita directamente el comportamiento del ojo humano.

Sin embargo, como se vió en II.1, es posible representar cualquier color con al menos dos de los tres colores primarios. Tal es el caso del formato YUV, uno de los más empleados en la actualidad, donde Y representa la luminancia o información de brillo, U es la diferencia entre la componente azul de la imagen y la luminancia (B-Y) y V es la diferencia entre la componente roja de la imagen y la luminancia (R-Y). Existen otra serie de formatos para la representación de imágenes en color, todos ellos basados en los colores primarios o en sus complementarios, llamados colores secundarios (amarillo, cyan y magenta).

La componente de luminancia contiene toda la información de detalle de la imagen, es decir, en una imagen en blanco y negro podrá apreciarse el detalle pues todos los receptores del ojo estarán recibiendo y aportando cantidades iguales de información al cerebro. Esto permite hacer el muestreo de la información de color en una proporción menor al muestreo que se hace para la información de luminancia. Así por ejemplo, para el formato YUV, las componentes U y V están muestreadas en una proporción de 2:1. Una fracción del mosaico de muestras se vería como en la figura 2.

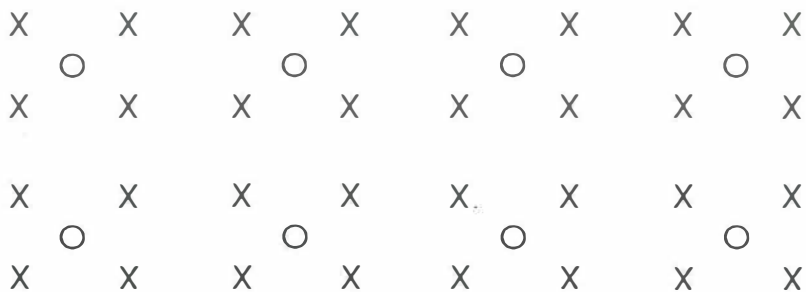


Figura 2. Muestreo de luminancia (X) y croma (O) en proporción 2:1, para el formato YUV.

En general, se dice que el formato YUV es de la forma 4:2:2. El muestrear en proporción 2:1 la información de color no introduce pérdidas o degradación perceptibles de la información, pues la señal de luminancia se preserva completamente y es ésta la que contiene la información de contornos y en general da nitidez a la imagen.

II.3. Redundancias de una imagen digital y su relación con la compresión.

En las secciones anteriores se ha explicado cómo el ojo humano percibe las imágenes y cómo esta percepción puede ser emulada por sistemas electrónicos de manera que se llegue a una imagen digital.

Como se ha visto, la representación numérica de una imagen real tiene limitantes en cuanto a que los valores de intensidad están acotados, y la cantidad de píxeles es mucho menor que la cantidad de células en el ojo. Aún con estas limitantes, las imágenes reales contienen una serie de elementos que las hacen ser redundantes en cuanto a la cantidad de información para su representación.

Un conjunto de datos se dice redundante cuando la cantidad de información que aporta el mismo puede ser representada con un conjunto menor de datos, preservando la cantidad de información.

Puede observarse que será deseable manipular conjuntos más pequeños de datos pues se logra mayor rapidez en su transmisión y se necesita menor espacio de memoria para su almacenamiento. Éste es uno de los principales objetivos de la compresión de datos, o para ser más precisos, aplicándolo a nuestro caso de estudio, de la compresión de imágenes.

Según la teoría de la información, el límite que puede lograrse en la compresión de un conjunto de datos es la cantidad de información que aporta el mismo, definida como entropía.

Se logra la compresión del conjunto de datos a cambio del requerimiento de un mecanismo para interpretar y regenerar el conjunto original a partir del conjunto comprimido. Proceso necesario después de la manipulación del mismo (almacenamiento, transmisión, etc.) . La tasa de compresión de un conjunto de

datos es una cantidad que puede medirse en relación al total de información del conjunto original. Si el conjunto original contiene n elementos, y el conjunto comprimido es de m elementos, la tasa de compresión se define como:

$$T.C. = \frac{n}{m} \quad (1)$$

y pueden distinguirse tres situaciones:

- Cuando $n < m$, se tiene una expansión de datos, caso no deseable para compresión pero que puede presentarse si no se eligen técnicas apropiadas.
- Cuando $n > m$, caso común de la compresión en el que el conjunto n contiene redundancias de algún tipo.
- Cuando $n = m$, caso límite en el cual el conjunto n o bien no tiene elementos redundantes o bien no fue comprimido ni expandido.

Una tasa de compresión de 1 (representado 1:1) indicará que hay una correspondencia unívoca entre los elementos de ambos conjuntos, es decir, no se logró compresión.

Una tasa mayor que 1, por ejemplo 25 (25:1) indicará que hay 25 elementos de n por cada uno de m .

Si expresáramos en porcentaje esta relación, hablaríamos de la fracción porcentual que significa m respecto a n .

$$F.P.C.(\%) = \frac{1}{T.C.} \times 100\% \quad (2)$$

donde: F.P.C. = Fracción Porcentual de Compresión.

Es decir, n puede representarse con sólo F.P.C.% del total de información.

Siguiendo con nuestro ejemplo tendríamos:

$$\text{F.P.C.}(\%) = \frac{1}{25} \times 100\% = 4\%$$

Esto significa que el conjunto n puede representarse por m usando sólo el 4% de la información de éste (n).

Si hablamos de redundancia en imágenes, podemos distinguir en general tres tipos que a continuación abordaremos [Gonzales, 1992], se logrará la máxima compresión al eliminarlas totalmente.

II.3.1. Redundancia psicovisual.

Si tomamos en cuenta que el ojo humano tiene un poder de resolución limitado y que generalmente observamos una imagen a cierta distancia mínima, sus contornos pequeños o regiones donde sólo haya variaciones de intensidad luminosa ligeras o nulas aparecerán a la distancia como una sola textura en el contexto de una región más amplia de la imagen. Podemos aprovechar estas limitantes del sistema visual humano para eliminar esas pequeñas variaciones y efectivamente convertirlas en texturas que harán que los procesamientos adicionales sean más fáciles. Al realizar este proceso llamado también de cuantificación, se pierde información, por lo que no se logrará una reconstrucción perfecta de la imagen pero si se eligen los parámetros adecuados, la degradación será imperceptible al ojo humano.

II.3.2. Redundancia entre píxeles adyacentes.

En las imágenes reales existe una gran correlación entre píxeles adyacentes, lo que hace que la información en la vecindad de cada píxel sea redundante. Es necesario entonces aplicar técnicas de procesamiento numérico que eliminen este efecto, es decir, logren representar la imagen como una secuencia de datos decorrelacionados.

II.3.3. Redundancia de código.

En una imagen digital a cada píxel se le asigna un código que representa la intensidad luminosa en ese punto de la imagen dentro del intervalo de valores posibles. Normalmente el código empleado es el binario natural, sin embargo, si evaluamos estadísticamente la imagen contando la cantidad de píxeles que corresponden a cada valor posible de intensidad encontraremos que siempre habrá niveles preponderantes en la imagen y otros con aparición muy escasa o incluso nula. A este procedimiento se le llama histograma y refleja en una gráfica la frecuencia de aparición de los diversos tonos de gris de una imagen. Es decir, en términos estadísticos (considerando el total de la imagen como el 100%) la probabilidad de aparición de cada uno de los tonos posibles.

Podrá verse entonces que para la mayoría de las imágenes reales habrá una concentración de la información alrededor de ciertos valores de intensidad. Si asignamos un código diferente al binario natural, que tome en cuenta esta distribución probabilística encontraremos que podemos asignar códigos más cortos a los valores de intensidad cuya aparición sea más frecuente y códigos más largos a los otros. De esta manera (si elegimos el código adecuado) estaremos efectivamente reduciendo la cantidad de información necesaria para la representación de la imagen.

Cuando el código elegido no es el óptimo, será redundante. Esto nos lleva al problema de elegir un código de longitud variable óptimo para cada imagen. El método para encontrar dicho código fué creado por Huffman en 1951 y consiste a grandes rasgos de los siguientes pasos: primero debe obtenerse el histograma de la imagen, después, de acuerdo a la probabilidad resultante para cada nivel de gris posible asignar un código único e irrepetible cuya longitud vaya en aumento conforme la probabilidad decrece en magnitud. Por último, el código empleado debe acompañar a la imagen codificada para su posterior recuperación.

La aplicación de este método no es un procedimiento trivial, por el contrario, implica un gran esfuerzo computacional. Para simplificar la tarea de eliminar la redundancia de código existen tablas que estadísticamente han sido generadas por medio del análisis de miles de imágenes comúnmente observadas por el humano, y contienen los códigos óptimos para la mayoría de los casos, así como códigos especiales para situaciones no contempladas. Estas tablas son entonces aproximaciones al óptimo y son utilizadas para la mayoría de las aplicaciones de compresión de imágenes.

III. LA TRANSFORMADA DE ONDÍCULAS.

III.1. Definición de ondícula y su relación con el análisis de señales.

La transformada de ondículas ("wavelets" en inglés) es una herramienta para la representación de señales que ha cobrado importancia en los años recientes ya que presenta ventajas respecto al análisis de Fourier, pues posee la característica de localización en el tiempo, por lo que es útil para analizar fenómenos transitorios.

Las funciones base del análisis de Fourier son de tipo senoidal, al descomponer una señal usando esta base encontraremos que no es posible localizar en el tiempo las características transitorias de la señal, pues las funciones base son periódicas.

Para tratar de resolver este problema Gabor (1946) recurrió al uso de una función ventana de tipo Gaussiano, de esta manera se contó con una localización de eventos dentro de la misma.

Esta transformada, definida como:

$$Gf(\omega, u) = \int_{-\infty}^{+\infty} e^{-j\omega x} g(x-u) f(x) dx \quad (3)$$

mide localmente, en el punto u , la amplitud de la componente senoidal de frecuencia ω (ver figura 3).

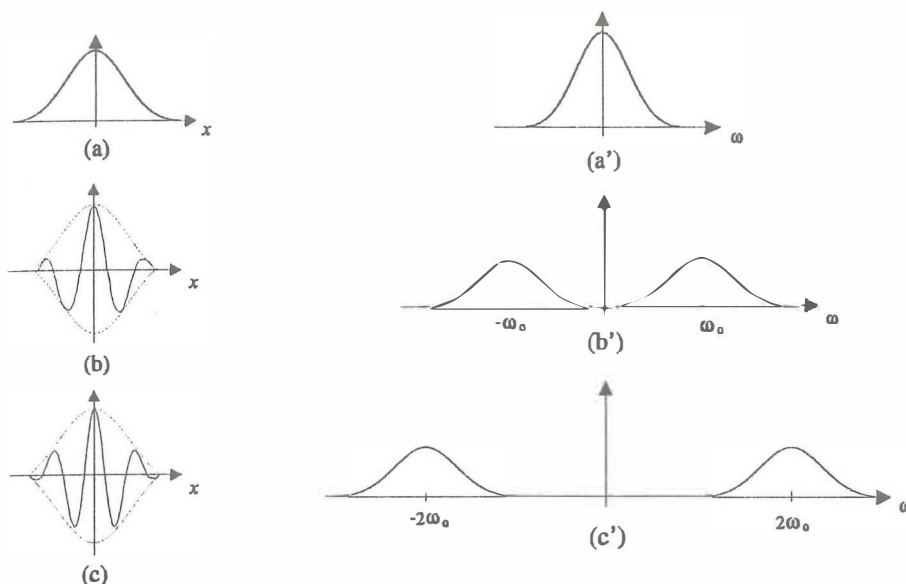


Figura 3. (a) Ventana Gaussiana $g(x)$ y su correspondiente transformada de Fourier (a') que es la respuesta de un filtro pasa-bajas. (b) $g(x)\cos\omega_0 x$ y (c) $g(x)\cos 2\omega_0 x$, y sus respectivas transformadas de Fourier (b') y (c'). Las funciones tienen el mismo soporte pero el número de ciclos varía en (b) y (c). Las transformadas tienen el mismo ancho de banda pero diferentes posiciones en frecuencia.

Consideremos σ_u como la desviación estándar de $g(x)$ y σ_ω la desviación estándar de su transformada. En el plano fase-espacio cada uno de los coeficientes $Gf(\omega, u)$ representará a $f(x)$ en una celda de resolución cuyo tamaño está dado por: $[u_0 - \sigma_u, u_0 + \sigma_u] \times [\omega_0 - \sigma_\omega, \omega_0 + \sigma_\omega]$.

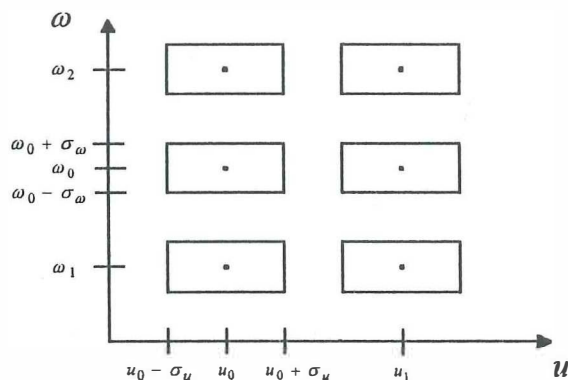


Figura 4. Representación de las celdas de resolución en el plano frecuencia-espacio para la transformada de Gabor.

En la figura 4 puede observarse el tamaño constante de las celdas, lo cual implica un límite máximo fijo a la resolución que es posible lograr para analizar una señal. Se llega a este límite cuando se usa una ventana de tipo Gaussiano, e implica un compromiso entre la resolución posible en el tiempo a cambio de la permitida en la frecuencia en un producto constante.

La transformada de ondículas está definida a partir de una función única de energía finita (implica desvanecimiento y por lo menos algunas oscilaciones) $\psi(x)$ llamada ondícula madre, que analiza una señal por medio de dilataciones y translaciones de esta ondícula en múltiples escalas, generando la familia de funciones:

$$\psi_{s,u}(x) = \sqrt{s}\psi(s(x-u)) \quad s, u \in \mathbb{R}^2 \quad (4)$$

donde s es el factor de dilatación o escalamiento y u es la translación.

La transformada de ondículas de una función está definida como:

$$Wf(s, u) = \int_{-\infty}^{+\infty} f(x)\sqrt{s}\psi(s(x-u))dx \quad (5)$$

Si la dilatación por un factor s se denota por:

$$\psi_s(x) = \sqrt{s}\psi(sx) \quad (6)$$

podemos reescribir la transformada en forma de producto interno:

$$Wf(s, u) = \langle f(x), \psi_s(x-u) \rangle \quad (7)$$

y puede considerársele como la respuesta en frecuencia de un filtro pasa banda. Esta respuesta y la característica de energía finita implican que la función madre y

por consecuencia la familia de funciones tendrán algunas oscilaciones en un pequeño intervalo de tiempo, como una pequeña onda, de ahí su nombre.

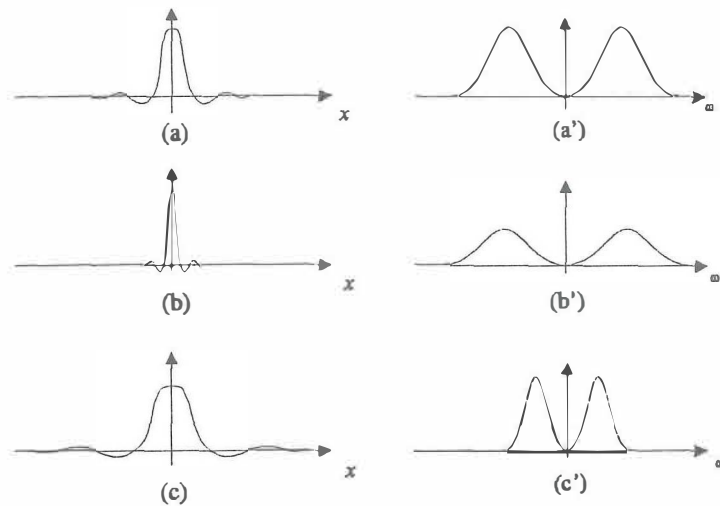


Figura 5. (a) Ondícula $\psi(x)$. (b) Ondícula $\psi_{s_1}(x)$ con $s_1 > 1$. (c) Ondícula $\psi_{s_2}(x)$ con $s_2 < 1$. (a'), (b') y (c') son las transformadas de Fourier de (a), (b) y (c) respectivamente. Las funciones $\psi_s(x)$ tienen un soporte que varía con la escala s , pero el número de oscilaciones es el mismo. Las transformadas tienen el mismo ancho de banda en una escala logarítmica.

Contrariamente a lo que sucede con la transformada de Gabor, la forma de la celda de resolución para la transformada de ondículas varía de acuerdo con la escala s .

En este caso, el centro de la banda del pasa-banda se encuentra en $s\omega_0$ y su ancho de banda es $s\sigma_\omega$. En una escala logarítmica, este ancho será el mismo para toda $s \in \mathbb{R}^+$ (distribución por octavas de frecuencia).

Primero se definirá el concepto de escala y su relación con la frecuencia [Rioul, Vetterli, 1991] para posteriormente continuar con la comparación.

Cuando una función es escalada: $f(x) \rightarrow f(sx)$ donde: $s > 0$

se tendrá una contracción si $s > 1$ y una expansión si $s < 1$. De esta manera, escalas grandes corresponden a señales "contraídas" o comprimidas, mientras que escalas pequeñas corresponden a señales "dilatadas" o expandidas en el dominio temporal.

Como se mencionó anteriormente, la función $\psi(x)$ es de tipo pasabanda, donde la escala es la que define la banda de frecuencia en la que actúa la transformación. Es preferible manejar el termino "escala" al termino "frecuencia" para referirse al dominio frecuencial. El factor s tiene la misma interpretación que la escala de los mapas, grandes escalas significan vistas globales, mientras que pequeñas escalas significan vistas detalladas.

La ondícula $\psi_s(x - u_0)$ está centrada alrededor de u_0 , con una desviación estándar de σ_u / s , de manera que el tamaño de la celda de resolución está dado por: $[u_0 - (\sigma_u / s), u_0 + (\sigma_u / s)] \times [s\omega_0 - s\sigma_\omega, s\omega_0 + s\sigma_\omega]$

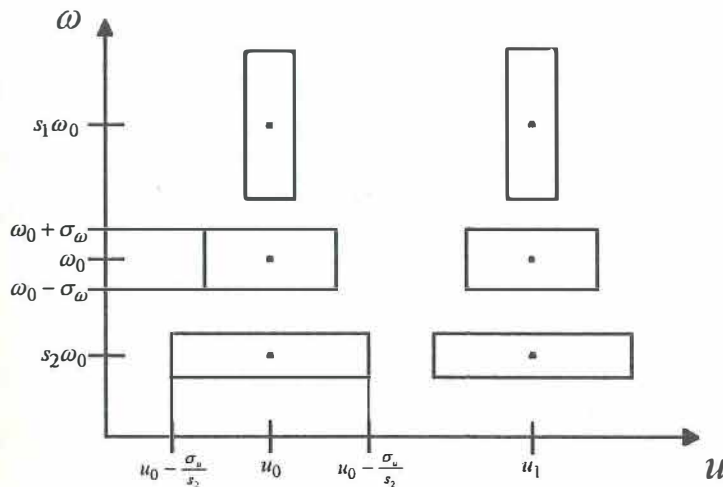


Figura 6. Representación de las celdas de resolución en el plano frecuencia-espacio para la transformada de ondículas.

En la figura 6 se puede observar que aunque el tamaño de la celda de resolución es el mismo, la resolución temporal aumenta al aumentar la escala s , mientras que la resolución frecuencial aumenta para frecuencias bajas. Entonces, fenómenos transitorios con componentes de alta frecuencia pueden ser analizados incrementando la escala s .

Para lograr representar completamente una señal, las celdas de resolución deberán cubrir completamente el plano fase-espacio. Si discretizamos la transformada $Wf(s, u)$, para lograr cubrir todo el plano, las dilataciones deben ser de tipo exponencial α^j , donde α es el factor de dilatación, el factor de muestreo debe ser proporcional a $\alpha^j, \beta / \alpha^j$. La transformada se expresa ahora como:

$$Wf(\alpha^j, n\beta / \alpha^j) = W_d f(j, n) \quad (8)$$

en forma de producto interno:

$$W_d f(j, n) = \langle f(x), \psi_{\alpha^j}(x - n\beta / \alpha^j) \rangle \quad (9)$$

Existen condiciones que α y β deben cumplir para que la transformada sea invertible, dichas condiciones fueron analizadas por Daubechies (1988) quien además estableció las condiciones para la generación de bases ortonormales de ondículas.

En particular, cuando $\alpha = 2$ y $\beta = 1$ obtendremos la familia ortogonal [Mallat, 1989c]:

$$\psi_{j,n}(x) = 2^{-j/2} \psi(x - n / 2^j) \quad j, n \in \mathbb{Z}^2 \quad (10)$$

la descomposición de una función en esta familia está dada por:

$$W_d f(j, n) = \langle f(x), \psi_{2^j}(x - n / 2^j) \rangle \quad j, n \in \mathbb{Z}^2 \quad (11)$$

Se prefiere usar familias de este tipo, con secuencias de escalamiento 2^j (llamadas dyadicas) porque los algoritmos computacionales para su cálculo son más simples.

Cualquier función podrá ser reconstruida a partir de su descomposición en una base ortonormal de ondículas por medio de:

$$f(x) = \sum_{j \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f(u), \psi_{2^j}(u - n/2^j) \rangle \psi_{2^j}(x - n/2^j) \quad (12)$$

El análisis se hace calculando el producto interno y la síntesis se hace sumando todas las proyecciones ortogonales de la señal en la base de ondículas.

III.2. Aplicación de las ondículas en el análisis de resolución múltiple.

En el análisis de resolución múltiple una señal se descompone en un conjunto de aproximaciones de la misma a diferentes resoluciones, con el fin de extraer sus características en cada uno de estos niveles. Esto equivale a descomponer el espectro de la señal en bandas de frecuencia que tendrán contenidos de información diferentes, así como niveles de energía distintos. Cuando se utiliza este análisis para la compresión, se desea agrupar la energía en bandas de frecuencia de manera que ésta se concentre en algunas de ellas (generalmente las más bajas) mientras que otras contengan tan poca información que puedan ser descartadas sin detrimento de la calidad de la imagen cuando ésta se reconstruya.

El análisis de resolución múltiple también es una alternativa para la solución de otros problemas del procesamiento digital de imágenes entre los que se pueden mencionar la extracción de características y la identificación de patrones.

La codificación por sub-bandas, las pirámides y la transformada de ondículas son esquemas similares que son considerados de resolución múltiple [Akansu et. al., 1992]. De estos, el uso de ondículas es la mejor alternativa debido a que no introduce pérdidas como en el caso de las pirámides, aún cuando en ocasiones se le considera un caso especial de la codificación por sub-bandas.

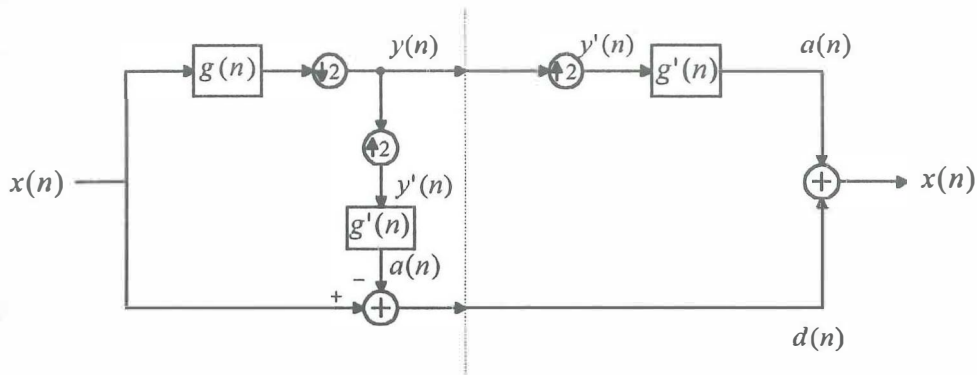


Figura 7. Esquema piramidal.

En un esquema piramidal [Burt y Adelson, 1983], la señal en la etapa de análisis es filtrada por un filtro pasabajas y posteriormente es diezmada en un factor de 2 (figura 7). $y(n)$ es la versión pasa bajas de la señal original $x(n)$ y representa una aproximación a la misma. Para obtener el detalle, $x(n)$ se reconstruye a partir de $y(n)$ y se resta al original, generando $d(n)$. Como hubo una interpolación, la señal $d(n)$ tiene el mismo tamaño de $x(n)$, mientras que $y(n)$ es de la mitad del tamaño. Sumando ambas señales tendremos un 50% más información que en el original. Este esquema puede repetirse descomponiendo sucesivamente $y(n)$ (obteniendo diferentes aproximaciones a los niveles anteriores). Como cada vez que se descompone $y(n)$ reduce su tamaño a la mitad, pareciera como si se formara una pirámide entre los distintos niveles, de ahí el nombre de este esquema. Burt y Adelson (1983) llamaron gaussiana a la pirámide formada por las distintas aproximaciones de los diferentes niveles y laplaciana a la pirámide formada por

los distintos niveles de detalle. En la etapa de síntesis, la señal se reconstruye agregando a la versión reconstruida de $x(n)$ (llamada aproximación $a(n)$) la señal de detalle $d(n)$ correspondiente a cada nivel. Lográndose una reconstrucción perfecta en ausencia de cuantificación.

Cuando la señal es bidimensional como en una imagen, existe una redundancia de alrededor de 33% en la información resultante de la suma de las señales de detalle con la de aproximación. Esto se debe a que la señal de detalle $d(n)$ no corresponde exactamente a la diferencia de información entre las aproximaciones $a(n)$ para dos niveles consecutivos de la pirámide. Si ésto fuera así, la descomposición generaría exactamente la misma cantidad de información que el original. Si descomponemos una señal en una base ortonormal de ondículas podemos evitar este defecto, pues es posible extraer la diferencia exacta entre dos niveles de resolución contiguos y así generar una descomposición en la que los diferentes niveles de resolución estén decorrelacionados. Esto es útil además porque nos permite apreciar detalles a una resolución determinada que no dependen ni están asociados a otros niveles, como sucede en la pirámide laplaciana. Podemos considerar la aproximación discreta de una función $f(x)$ a una determinada resolución como la descomposición de ésta en su proyección sobre un espacio vectorial V_{2^j} generado por una función base $\theta_{2^j}(n/2^j)$ a esa misma resolución:

$$A_{2^j} f = (f * \theta_{2^j}(n/2^j))_{n \in \mathbb{Z}} \quad (13)$$

La mejor aproximación se logrará proyectando ortogonalmente $f(x)$ en el espacio vectorial V_{2^j} , cuando $f(x)$ se descompone en la secuencia de espacios

vectoriales $(V_{2^j})_{j \in \mathbb{Z}}$ (por medio de la familia de funciones $\theta_{2^j}(x - n/2^j)$) se obtendrá una aproximación de resolución múltiple de $f(x)$.

Para lograr la proyección mencionada es necesario contar con una base ortonormal de V_{2^j} . Mallat (1989b) demostró que es posible construir dicha base por medio de dilataciones y translaciones de una función única llamada función de escalamiento $\phi(x)$, cuya transformada de Fourier tiene la forma de un filtro pasabajas y que para cualquier resolución 2^j esta definida como:

$$\phi_{2^j}(x) = \sqrt{2^j} \phi(2^j x) \quad j \in \mathbb{Z} \quad (14)$$

que genera la familia de funciones:

$$\phi_{2^j}(x - 2^{-j} n) \quad n \in \mathbb{Z} \quad (15)$$

de manera que la aproximación queda definida como:

$$A_{2^j} f = \left(\left\langle f(x), \phi_{2^j}(x - 2^{-j} n) \right\rangle \right)_{n \in \mathbb{Z}} \quad (16)$$

Consideremos la aproximación de $f(x)$ en dos resoluciones distintas 2^j y 2^{j+1} dada por la proyección ortogonal de $f(x)$ en los espacios V_{2^j} y $V_{2^{j+1}}$ respectivamente. La aproximación a la resolución 2^{j+1} debe ser mejor que la aproximación a la resolución 2^j , por lo que los espacios vectoriales deben satisfacer la relación:

$$V_{2^j} \subset V_{2^{j+1}} \quad (17)$$

Hemos definido la aproximación a una función $f(x)$ para resoluciones múltiples por medio de proyecciones ortogonales de esta función en los espacios vectoriales generados por la familia de funciones de escalamiento.

Para obtener la diferencia entre las aproximaciones logradas a las resoluciones 2^j y 2^{j+1} es necesario hacer la proyección ortogonal de $f(x)$ en el complemento ortogonal de V_{2^j} en $V_{2^{j+1}}$. Este complemento ortogonal es O_{2^j} y satisface la relación:

$$O_{2^j} \oplus V_{2^j} = V_{2^{j+1}} \quad (18)$$

Para calcular la proyección de $f(x)$ en O_{2^j} es necesaria una base ortonormal de O_{2^j} . Mallat (1989b) demostró que es posible construir dicha base por medio de dilataciones y translaciones de una función única llamada ondícula $\psi(x)$, que para cualquier resolución 2^j esta definida como:

$$\psi_{2^j}(x) = \sqrt{2^j} \psi(2^j x) \quad (19)$$

que genera la familia de funciones:

$$\psi_{2^j}(x - 2^{-j} n) \quad (20)$$

de manera que el detalle o diferencia entre las aproximaciones $A_{2^j} f$ y $A_{2^{j+1}} f$ queda definido como:

$$D_{2^j} f = \left(\langle f(x), \psi_{2^j}(x - 2^{-j} n) \rangle \right)_{n \in \mathbb{Z}} \quad (21)$$

En resumen, la función de escalamiento $\phi(x)$ se encarga de encontrar las aproximaciones para cada nivel, y la ondícula $\psi(x)$ es la encargada de encontrar la diferencia entre niveles o detalle.

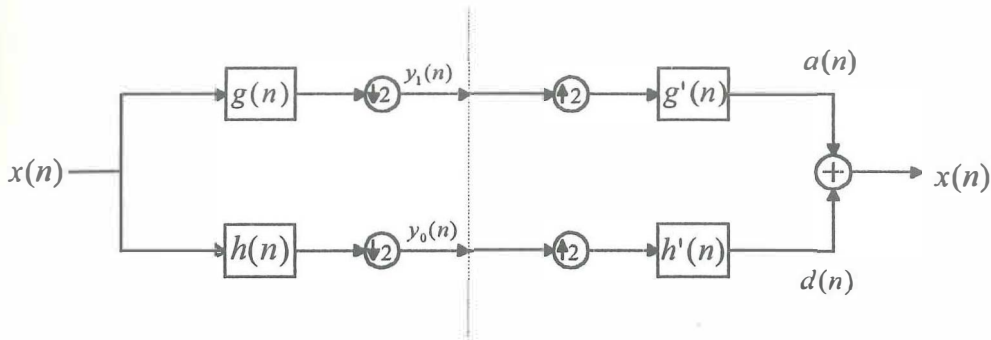


Figura 8. Esquema de descomposición por ondículas.

El esquema de descomposición se muestra en la figura 8. Como puede verse, existen ahora dos filtros, de manera que después de la descomposición, la cantidad de información es la misma, por lo que no existe redundancia de datos. Al igual que en el esquema piramidal, podemos seguir descomponiendo sucesivamente la aproximación de cada nivel, al final obtendremos una última aproximación y una serie de detalles correspondientes a cada nivel. La reconstrucción se logra recuperando la información de detalle y añadiéndola a la aproximación en cada nivel. En este caso se logra la total decorrelación de la información entre los distintos niveles y es posible apreciar detalles en cada uno de ellos que corresponden exclusivamente a él. Posteriormente se verá cómo esto es una ventaja para la compresión.

Una imagen como señal bidimensional requerirá de la extensión de este esquema a dos dimensiones. Existen dos formas de hacerlo. Utilizando filtros bidimensionales no separables y mediante una combinación de filtros monodimensionales en un esquema separable. El primer esquema es un caso complejo que no será abordado en este trabajo. El segundo esquema es el más usado por su sencillez y consiste en descomponer en una primera etapa la imagen en sentido horizontal para posteriormente descomponerla en sentido vertical llegando a un esquema como el de la figura 9 que genera cuatro sub-imágenes cada una de la cuarta parte del tamaño original.

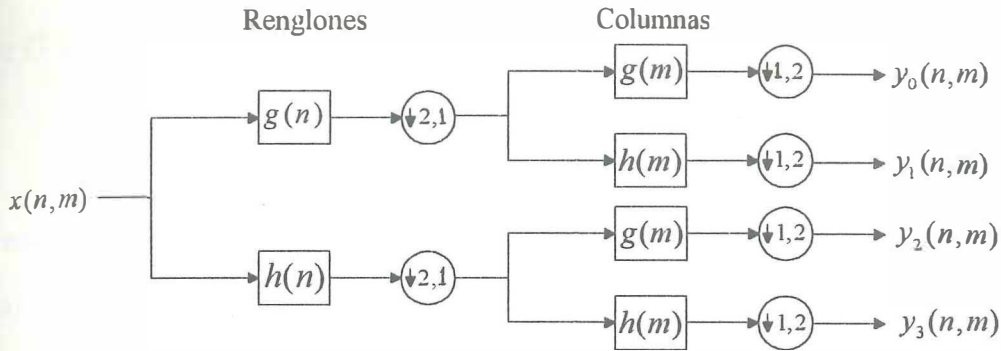


Figura 9. Etapa de análisis del esquema de descomposición por ondículas para señales bidimensionales usando filtros separables.

III.3. Ondículas y bancos de filtros.

Las ecuaciones (16) y (21) definen la aproximación $A_{2^j}f$ de $f(x)$ a la resolución 2^j y la diferencia $D_{2^j}f$ entre las resoluciones 2^j y 2^{j+1} . La aproximación $A_{2^j}f$ equivale a procesar con un filtro pasabajas la señal $f(x)$, mientras el detalle $D_{2^j}f$ se encuentra en el equivalente de filtrar la señal con un filtro pasabanda. Para dos niveles consecutivos, la función de escalamiento y la ondícula están definidas por:

$$\phi(x) = \sum_n h(n)\phi(2x - n) \quad (22)$$

$$\psi(x) = \sum_n g(n)\psi(2x - n) \quad (23)$$

en el dominio frecuencial:

$$\Phi(\omega) = \prod_{k=1}^{+\infty} H(e^{j\omega 2^{-k}}) \quad (24)$$

$$\Psi(\omega) = G(e^{j\omega/2}) \prod_{k=2}^{+\infty} H(e^{j\omega 2^{-k}}) \quad (25)$$

donde G y H son filtros que están relacionados por:

$$G(e^{-j\omega}) = e^{-j\omega} H(-e^{-j\omega}) \quad (26)$$

y forman un par de filtros espejo en cuadratura ("quadrature mirror filters" en inglés). En el dominio espacial, esta relación está dada por:

$$g(n) = (-1)^{1-n} h(1-n) \quad (27)$$

De manera que cualquier par de filtros que cumpla las condiciones anteriores servirá para descomponer una señal en una base de ondículas.

En el caso bidimensional cuando se usan filtros separables, se genera no solamente una ondícula, sino tres, dependiendo de la orientación espacial del filtro correspondiente. Esto puede observarse del esquema de la figura 9, donde se generarán las cuatro salidas siguientes:

$$\Phi(x, y) = \phi(x)\phi(y) \quad (28)$$

$$\Psi^1(x, y) = \phi(x)\psi(y) \quad (29)$$

$$\Psi^2(x, y) = \psi(x)\phi(y) \quad (30)$$

$$\Psi^3(x, y) = \psi(x)\psi(y) \quad (31)$$

que corresponden al plano ω_x, ω_y a los sectores que se muestran en la figura 10.

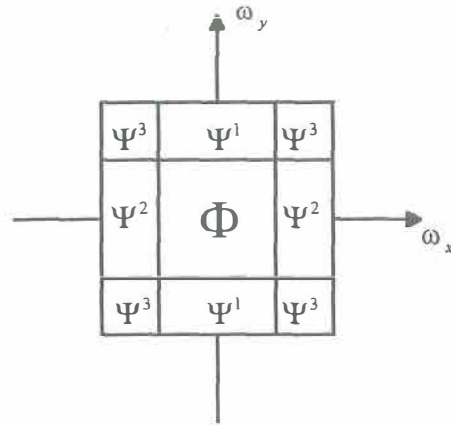


Figura 10. Zonas de influencia de Φ , Ψ^1 , Ψ^2 , Ψ^3 en el plano frecuencial ω_x, ω_y

De las ecuaciones (28) a la (31) puede deducirse [Mallat, 1989a] que la imagen será descompuesta en una aproximación dada por:

$$A_{2^j} f = \left(\left\langle f(x, y), \phi_{2^j}(x - 2^{-j}n) \phi_{2^j}(y - 2^{-j}m) \right\rangle \right)_{(n,m) \in \mathbb{Z}^2} \quad (32)$$

y tres sub-imágenes de detalle:

$$D_{2^j}^1 f = \left(\left\langle f(x, y), \Psi_{2^j}^1(x - 2^{-j}n, y - 2^{-j}m) \right\rangle \right)_{(n,m) \in \mathbb{Z}^2} \quad (33)$$

$$D_{2^j}^2 f = \left(\left\langle f(x, y), \Psi_{2^j}^2(x - 2^{-j}n, y - 2^{-j}m) \right\rangle \right)_{(n,m) \in \mathbb{Z}^2} \quad (34)$$

$$D_{2^j}^3 f = \left(\left\langle f(x, y), \Psi_{2^j}^3(x - 2^{-j}n, y - 2^{-j}m) \right\rangle \right)_{(n,m) \in \mathbb{Z}^2} \quad (35)$$

La aproximación (32) está dada por las frecuencias bajas; la sub-imagen descrita en (33) aporta las altas frecuencias verticales, es decir los contornos horizontales; la sub-imagen descrita en (34) aporta las altas frecuencias horizontales, es decir los contornos verticales; y la sub-imagen descrita en (35) aporta las altas frecuencias en ambos sentidos, es decir, aquellos contornos con alguna inclinación.

El plano frecuencial mostrado en la figura 10 puede considerarse como tan sólo la aproximación en la resolución contigua inferior, o bien, si se sigue descomponiendo la imagen, el cuadro central $\Phi(x,y)$ se descompone de igual manera.

IV. CUANTIFICACION Y CODIFICACION.

IV.1. Cuantificación.

Al transformar una imagen, sea cual sea la transformada utilizada, se logra el reagrupamiento de la energía de la misma, ya sea en forma localizada o no, de manera que es deseable que dicha energía se concentre en la menor cantidad de coeficientes posible. Una vez que se ha logrado esta concentración, es necesario descartar todos aquellos coeficientes cuyo valor esté por debajo de un umbral determinado, y reagrupar los coeficientes restantes en un conjunto finito de valores posibles, de manera que exista un conjunto limitado y establecido de valores para los coeficientes de la imagen transformada. Este proceso, llamado de cuantificación, obviamente introduce pérdidas, que harán que cuando se realice la transformada inversa el resultado sea una imagen con un cierto grado de degradación o error respecto al original.

Cuando se usan cuantificadores, es inevitable introducir errores en la imagen. Estos solamente pueden minimizarse y es precisamente esto lo que se busca en el diseño de los mismos. Existen dos grandes tipos de cuantificadores: los escalares y los vectoriales. El diseño de un cuantificador escalar requiere de un conocimiento previo de la estadística de la imagen transformada, es decir, en dónde se concentra la mayoría de los coeficientes de la misma, así como el valor y la frecuencia de aparición de los mismos dentro del total de la imagen. De esta manera es posible determinar el umbral bajo el cual se descartarán los coeficientes menores, y el número y tamaño de los diferentes intervalos de agrupamiento posibles.

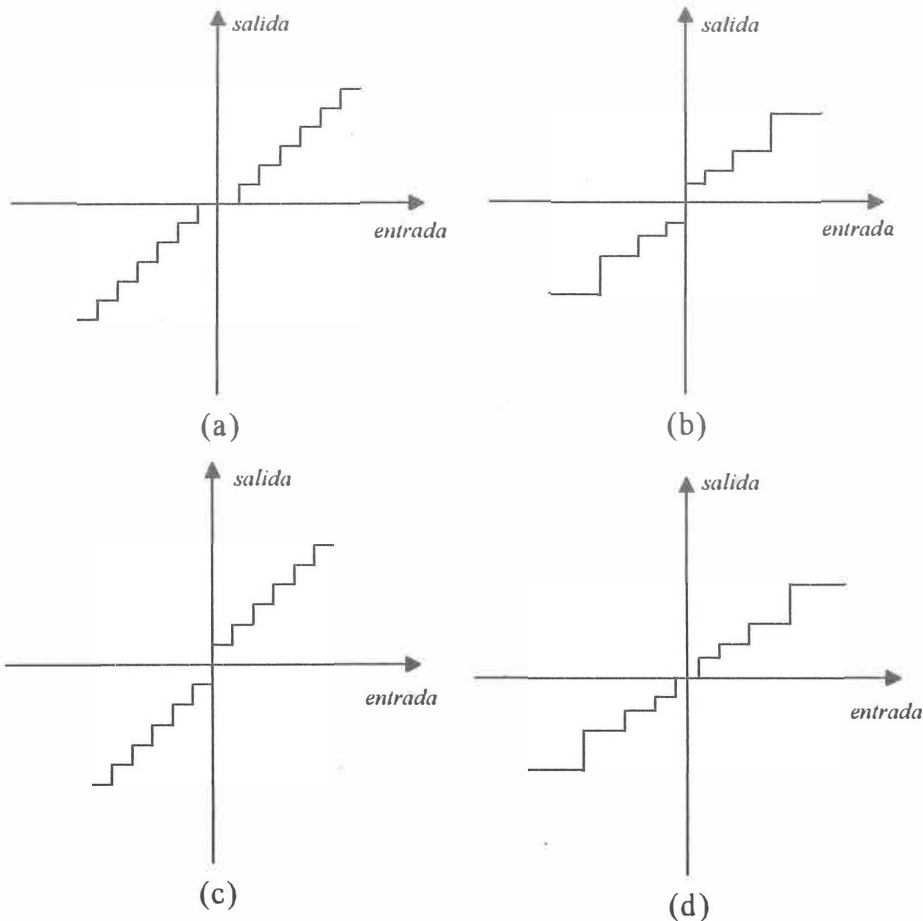


Figura 11. Cuantificador escalar típico: (a) con zona muerta, (b) sin zona muerta, (c) uniforme, (d) no uniforme.

Existen dos variantes de este tipo de cuantificador, como se muestra en la figura 11. Se dice que un cuantificador escalar tiene zona muerta cuando para un cierto intervalo de valores pequeños de entrada alrededor de cero se considera una salida cero. Cuando el tamaño de los escalones o intervalos de agrupamiento es diferente entre ellos hablamos de un cuantificador escalar no uniforme. Estas dos condiciones son particularmente deseables para una transformada como la de ondículas, donde la imagen transformada tiene grandes zonas con coeficientes cuyo valor es muy cercano a cero, así como una distribución no uniforme de los coeficientes.

Existe un compromiso entre el número y tamaño de los escalones y el grado de error que se le introduce a la imagen al cuantificarla. Por un lado, con pocos escalones el cuantificador será muy rápido, pero la distorsión será mayor y por otro, un número elevado de escalones hará que el proceso sea más lento y, aún cuando la distorsión sea considerablemente menor, existe el problema de que al haber más valores posibles de salida, los códigos que les asigne a estos el codificador serán más largos y por consiguiente el grado de compresión será menor.

Otro factor fundamental cuando se quiere diseñar un cuantificador es el comportamiento de la transformada para una serie de imágenes. Es decir, puesto que cada imagen tiene su propia estadística, el cuantificador óptimo para una imagen dada, no será necesariamente el mejor para otra diferente. Por ello, debe hacerse un estudio estadístico con un conjunto de imágenes, de manera que al diseñar el cuantificador, podamos en promedio obtener un buen comportamiento para cualquier imagen transformada. Para transformadas ampliamente desarrolladas y estudiadas como la discreta de coseno, este proceso ya se ha llevado a cabo y aún así el diseño de cuantificadores sigue siendo tema de investigación. Obviamente para la transformada de ondículas hay mucho menos trabajo desarrollado por lo que existen más problemas abiertos por resolver.

El cuantificador vectorial mejora las características del escalar porque ataca la cuantificación desde el punto de vista de vectores (no de escalares), por ello el agrupamiento de conjuntos de coeficientes y su aproximación a centroides en la vecindad de vectores posibles de salida son los que determinan qué tanto y dónde puede ser discriminada la energía.

El problema en este caso es la elección de los mejores centroides, que forman lo que se conoce como libro de códigos. Proceso que generalmente se hace por entrenamiento en base nuevamente a conjuntos de imágenes. Esta técnica es más complicada y requiere de procesos más elaborados para su realización.

Esto nos lleva a la conclusión de que para diseñar un cuantificador que pretenda ser óptimo para una cierta clase de imágenes es necesario realizar un estudio estadístico del comportamiento de dicho conjunto de imágenes para una transformada específica. Así se podrá analizar el comportamiento promedio de ellas y con esto diseñar diferentes alternativas y evaluarlas para obtener los mejores resultados tanto objetiva como subjetivamente.

IV.2. Codificación.

Cuando la imagen en el dominio de la transformada ha sido cuantificada, se ha logrado la reagrupación de la energía de la misma en un conjunto limitado de coeficientes. Sin embargo, el tamaño de estos coeficientes es el mismo en términos computacionales, es decir, tienen una longitud de palabra constante y conocida. Esto es una desventaja si pensamos que ellos forman un conjunto finito y conocido de valores y que la frecuencia de aparición de cada uno de ellos es diferente. Es obvio entonces que si encontramos una forma en la que los coeficientes cuya aparición es más frecuente puedan ocupar tamaños de palabra más cortos y aquellos cuya frecuencia de aparición sea menor ocupen tamaños de palabra más largos estaremos reduciendo efectivamente la cantidad de información necesaria para representar la imagen.

IV.2.1. Codificación Huffman.

La forma más eficiente de lograr lo anterior se conoce como codificación Huffman (1952). En este método se calcula la probabilidad de aparición de cada uno de los símbolos posibles (valores de salida del cuantificador) dentro del total de la imagen [Gonzales, 1992], se ordenan los resultados en orden de probabilidad descendente y se asigna un código más corto a aquellos valores cuya aparición es más frecuente y más largos a los que no lo son. Estos códigos son únicos e irrepetibles, de tal forma que si existen 256 valores de entrada, deberán existir 256 códigos diferentes a la salida, cuya longitud es variable. Puesto que los códigos resultantes son de longitudes diversas, se colocarán uno tras de otro sin marcas de sincronía entre ellos, agrupándolos por octetos (aunque estos nuevos octetos no representen a la imagen como tal), de manera que al reconstruir se deberá distinguir y recuperar cada valor de entrada sin error, lográndose la compresión respecto al original.

El problema aquí es nuevamente la estadística de la imagen. Sería necesario calcular el código de Huffman correspondiente a cada imagen transformada, y además anexarlo a la imagen codificada. Esto resulta impráctico si consideramos que si se ha logrado la compresión de la imagen, el anexarle código revierte un poco este resultado (es decir, expande). Por otro lado, dada una imagen, el método de Huffman proporciona la tasa de compresión máxima posible. Debido a la dificultad que representa el obtener un código Huffman para cada imagen se han buscado otras alternativas de codificación.

IV.2.2. Codificación por longitud recorrida.

En la mayoría de los casos cuando se aplica una transformada, se espera que haya una gran cantidad de ceros y valores cercanos a él como resultado de los procesos de transformación y cuantificación. Esto quiere decir que en forma típica una línea de imagen en estas condiciones tendría una secuencia de valores de salida como la siguiente:

... 1 0 0 0 0 2 0 0 0 0 0 -1 0 1 1 -1 4 0 0 ...

Hay una gran cantidad de ceros mientras que la aparición de valores grandes es muy escasa. Es posible representar de manera más compacta esta secuencia de la siguiente manera:

... 1,4,2,5,-1,1,1,0,1,0,-1,0,4, ...

donde los dígitos subrayados indican la cantidad de ceros que fueron encontrados antes de un nuevo valor diferente de cero. Es decir, la longitud recorrida antes de un nuevo código. Esta técnica (llamada "run-length" en inglés) fue desarrollada originalmente en 1950 para codificación de dos niveles (blanco y negro) y encuentra su aplicación más común en el facsímil. Sin embargo, modificaciones introducidas a la misma la han hecho útil para la codificación de imágenes.

Nuevamente es necesario recurrir a la estadística de la imagen transformada, pero ahora es necesario encontrar la probabilidad de ocurrencia de cada una de las combinaciones de longitudes recorridas y valores encontrados. En base a esta información se asignarán códigos más cortos a las combinaciones longitud recorrida-valor encontrado de mayor aparición y más largos a las de aparición menor, tal y como sucedió en la codificación Huffman. Sin embargo, existe un número infinito de posibilidades de combinar longitudes y valores, por lo que

asignar un código a cada una sería imposible. Por esta razón, el número de códigos asignados se limita en cantidad de manera que la longitud máxima asignable no sobrepase una cota establecida, y se provee de un código especial, llamado comúnmente secuencia de escape que indica que se ha encontrado un valor cuya probabilidad de aparición era muy baja y por lo tanto se encontrará a continuación tanto la longitud recorrida como el valor encontrado sin codificar.

Podrá observarse que este método ofrece ventajas para su implantación práctica respecto a la codificación Huffman, pues los códigos no necesariamente deben ser tantos si se cuenta con secuencias de escape. Por otro lado, es más versátil pues un mismo código puede acomodarse con cierto margen de error a muchas imágenes diferentes, y aún cuando no es el óptimo, es un buen compromiso dada la versatilidad lograda ya que es preferible en todo caso tener un solo código con un comportamiento promedio, al trabajo computacionalmente más intensivo de generar un código en cada caso.

De nuevo para el caso del codificador, cuando se trata de transformaciones tan ampliamente investigadas como la transformada discreta de coseno se han evaluado estadísticamente resultados de esta combinación Huffman-longitud recorrida para imágenes en el dominio de la transformada obteniéndose tablas para asignación de códigos (ver anexo A) normalizadas internacionalmente. Obviamente estas tablas no son las más recomendables cuando se trata de la transformada de ondículas, pero obtener sus similares es un problema que en la actualidad no ha sido resuelto pero que de establecerse una solución hará que los resultados de compresión sean mejores.

V. RESULTADOS.

V.1. Esquema de compresión utilizado.

En los capítulos anteriores se han abordado en forma genérica y por separado los diferentes elementos que componen un esquema de compresión de imágenes típico cuando se involucra el uso de una transformada, en este caso, la transformada de ondículas.

Dichos elementos se muestran en conjunto en la figura 12. El esquema de compresión (12a) se compone de un bloque de transformación, un cuantificador y un codificador, al final, se obtendrá una imagen comprimida, misma que será posible reconstruir mediante el uso del esquema de reconstrucción (12b) que consta de un decodificador y un bloque para la realización de la transformación inversa. Al final obtendremos una imagen reconstruida que tiene un cierto grado de degradación respecto al original. Dicho error ha sido introducido por la utilización del cuantificador.

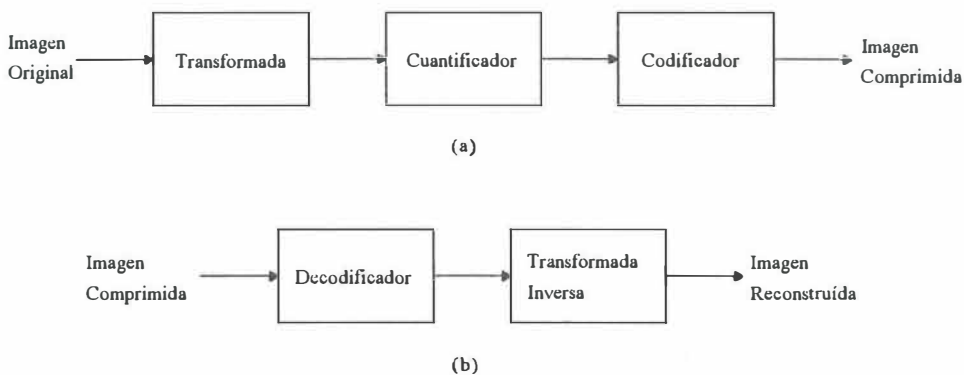


Figura 12. (a) Esquema de compresión típico cuando se involucra una transformada. (b) esquema de reconstrucción.

Se introducirán ahora los resultados del presente trabajo en cada uno de los elementos mencionados anteriormente, mismos que conforman en su conjunto la solución planteada al problema de la aplicación de la transformada de ondículas a

la compresión de imágenes. En cada uno de los tres siguientes puntos se abordarán respectivamente los bloques de la figura 12a, así como sus correspondientes de 12b.

V.1.1. La transformada.

Como se vió en el capítulo III, la transformada de ondículas puede llevarse a cabo mediante la utilización de los filtros espejo en cuadratura, la figura 10 muestra el resultado de la aplicación de estos filtros para un determinado nivel de descomposición. Las condiciones que deben cumplir los filtros ya han sido establecidas, así que cualquier ondícula puede en principio ser generadora de un banco de filtros para la transformación. Daubechies (1988) encontró una solución con la cual generó un conjunto de coeficientes que dan origen a filtros digitales de respuesta finita al impulso de 4 a 20 etapas cuya respuesta es máximamente plana. Este conjunto de coeficientes es el más utilizado en la actualidad, aunque es posible establecer un conjunto propio y diferente de coeficientes ya que aún no se ha establecido una solución genérica para el procesamiento de imágenes.

El esquema de la figura 9 fue utilizado para la transformación, utilizando cuatro niveles en la pirámide que se forma del análisis de resolución múltiple. Esto equivale a descomponer 3 veces más hacia el interior la zona Φ de la figura 10, generando una aproximación y 12 bandas de frecuencia distintas correspondientes a cada una de las $\Psi_{2^j}^i$ regiones para los cuatro niveles.

El criterio para la selección de la cantidad de niveles de descomposición fue la cantidad de distorsión que el uso de un determinado cuantificador introduciría en la imagen transformada.

Se encontró que el uso de un número mayor de niveles provoca demasiada distorsión, mientras que el uso de un número menor no provoca una mejora

sensible en la calidad de la imagen, comparada con la ganancia en la tasa de compresión que podría lograrse en la imagen cuantificada.

La distribución de la aproximación $A_{2^j}f$ (33) y los detalles $D_{2^j}^i f$ (34),(35) y (36) en la imagen transformada se muestran en la figura 13(a). En la figura 13(b) puede apreciarse como efectivamente los contornos verticales, horizontales y diagonales se resaltan en las señales de detalle (estos han sido amplificados 8 veces para que puedan ser apreciados más claramente). La aproximación $A_{2^{-4}}f$ conserva la mayoría de las texturas de la imagen original, ya que los contornos se encuentran en las imágenes de detalle.

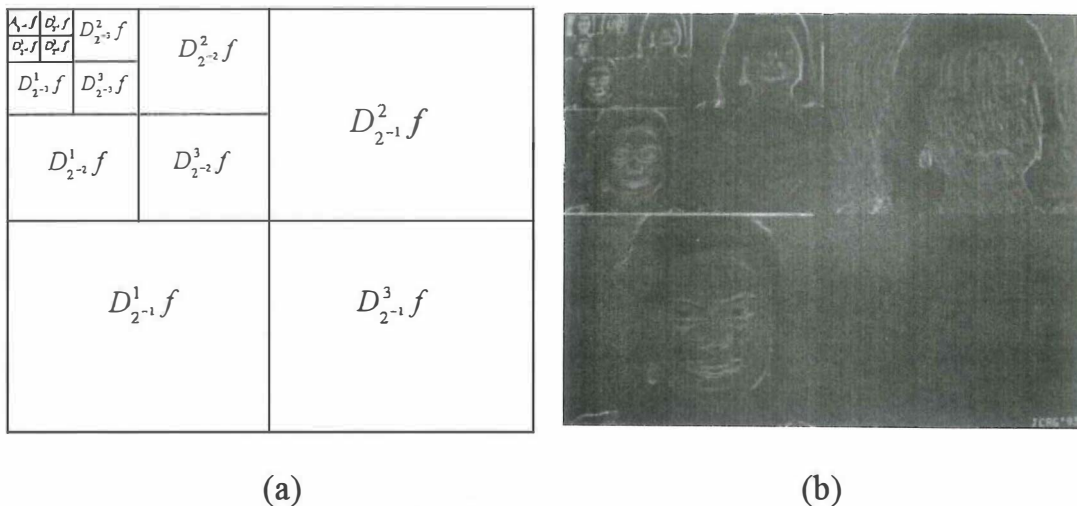


Figura 13. (a) Distribución de la aproximación $A_{2^j}f$ y los detalles $D_{2^j}^i f$ en una descomposición de 4 niveles. (b) En una imagen transformada.

Las franjas verticales en forma de paréntesis que se presentan más claramente en la región $D_{2^{-4}}^2 f$ fueron ocasionadas en la etapa de adquisición de la imagen y en realidad son sólo ruido, sin embargo, por su orientación han sido resaltadas por la transformada. Puede observarse además que este ruido está concentrado solamente en las regiones correspondientes a las frecuencias más altas de la imagen, por lo que si elegimos el cuantificador apropiado, el ruido puede ser reducido en gran medida.

Para imágenes con escaso contenido de frecuencias altas, tal como la de la figura 13(b), las regiones $D_{2^{-1}}^1 f$, $D_{2^{-1}}^2 f$, y $D_{2^{-1}}^3 f$ pueden ser eliminadas sin comprometer en gran medida su calidad, siempre y cuando la cuantificación de los demás niveles sea apropiada ya que la información de los contornos se encuentra distribuida no solamente en una banda de frecuencia (como puede observarse en 13(b)). El anexo C contiene los coeficientes de Daubechies con los que se llevó a cabo la transformación.

V.1.2. El cuantificador.

En el capítulo IV se mencionó al cuantificador escalar no uniforme con zona muerta como el más recomendable, y es este tipo el utilizado en este trabajo.

El tamaño y número de escalones, así como la forma de aplicarlo a la imagen transformada de manera de lograr la cuantificación óptima son problemas abiertos en la actualidad. Cualquier decisión para su uso es válida y son los resultados los que reflejan su eficiencia en el procesamiento.

Para aprovechar la característica de resolución múltiple, que conlleva el agrupamiento en bandas de frecuencia, el cuantificador aplicado a cada uno de los niveles de resolución es distinto, es decir, el cuantificador aplicado a los sectores de imagen $D_{2^{-1}}^1 f$, $D_{2^{-1}}^2 f$, y $D_{2^{-1}}^3 f$ difiere del aplicado a $D_{2^{-2}}^1 f$, $D_{2^{-2}}^2 f$, y $D_{2^{-2}}^3 f$. La excepción la constituye el sector de imagen correspondiente a la aproximación $A_{2^{-4}} f$ donde no se realiza cuantificación para evitar introducir distorsión en la banda más baja, correspondiente a las texturas. El tamaño de los escalones para cada cuantificador se ha definido en base a lo que cada nivel de resolución aporta a la imagen, siendo los niveles inferiores los que más información contienen, por ello su cuantificación es más fina. El hacer más fina la cuantificación hace que el número de escalones del cuantificador sea mayor. Esto plantea un compromiso

entre este número y la rapidez computacional para la ejecución de la cuantificación.

Se hicieron pruebas con diferentes diseños de cuantificadores con número y tamaño de escalón diversos. En general, uno de los aspectos más importantes fue el hecho de que en imágenes en las que no existe una gran cantidad de detalles y contornos, las zonas correspondientes al primer nivel de descomposición ($D_{2^{-1}}^1 f$, $D_{2^{-1}}^2 f$, y $D_{2^{-1}}^3 f$) aportan muy poca información al contexto general de la imagen por lo que pueden ser eliminadas sin detrimento apreciable en la reconstrucción. Sin embargo, esto no puede tomarse como regla general, pues obviamente cuando la preservación de esa información crítica es necesaria se debe cuantificar en ese sector.

Es conveniente notar que el grado de fineza en la cuantificación está directamente relacionado con la tasa de compresión que se obtendrá. Por lo tanto, éste es otro de los compromisos que debe hacerse al elegir el cuantificador. En este trabajo no se ha abordado este problema, cuya solución consiste en determinar, previo análisis de la estadística de la imagen, cómo deben distribuirse los cuantificadores y qué características deben tener estos para cumplir con una determinada tasa de compresión.

V.1.3. El codificador.

El mismo capítulo IV describe una versión modificada tanto de la codificación Huffman como de la de longitud variable que representa una técnica viable por ser intermedia, es decir, no es ni totalmente particular para una imagen, ni totalmente general para cualquier transformada. Tan es así que para el caso de las ondículas no existe particularización.

En la codificación de longitud variable que se aplica para la transformada discreta de coseno, la longitud máxima de cada secuencia es de 64 elementos, mismos que están agrupados en un bloque de 8 x 8 elementos. Este factor impone un límite en la longitud que puede ser recorrida y por ello se decidió incluir un código adicional llamado EOB ("end of block" en inglés o fin de bloque).

Puesto que en la transformada de ondículas no existen bloques, este código no tiene ningún significado práctico, sin embargo, puede usársele para indicar el final de cada una de las líneas de la imagen codificada. Estas líneas son mucho mayores de 64 elementos, lo que implica que pueden existir longitudes recorridas de más de 64 elementos. Fue necesario alterar el código correspondiente a EOB para agregar un nuevo código, llamado C64 que indica que se ha superado la cantidad máxima permisible como longitud recorrida.

Como se explicó en IV.2 existe una secuencia de escape para las secuencias no consideradas en las tablas. Como ya se ha mencionado, en el caso de la transformada discreta de coseno la longitud máxima que es posible recorrer es de 64 elementos. También se mencionó que la secuencia de escape contenía tanto la longitud recorrida como el valor encontrado. Ésta es la razón de que para el caso de la transformada de ondículas haya sido necesaria la introducción del código C64, pues el campo dedicado a la longitud recorrida en la secuencia de escape es de sólo 6 bits.

Respecto a la tabla que aparece en el anexo A, las únicas modificaciones que fueron hechas fueron el reemplazo del código EOB de 10 al 100 binario, y la introducción del código C64 con el valor binario 101.

V.2. Resultados.

Se hizo una evaluación de los diferentes filtros propuestos por Daubechies para la imagen que se muestra en la figura 14(a). Esta imagen es típica de una clase de imágenes en la que podríamos incluir todas aquellas en las que sólo aparece un sujeto sobre un fondo plano o con pocos detalles. La imagen mostrada no pertenece a las consideradas como imágenes patrón (la figura 15(e) muestra una de ellas), sin embargo, ninguna de estas imágenes patrón estuvo disponible al momento de hacer las pruebas.

La transformada de ondículas es biunívoca. Al aplicar esta transformada y posteriormente su inversa a una señal las eventuales diferencias se deben a errores de cálculo tales como el de longitud de palabra de cómputo. En este caso, el problema de longitud limitada de palabra de cómputo fue ocasionado por la necesidad de reducir a números enteros de 8 bits el resultado de la aplicación del filtrado ya que la memoria disponible era limitada.

Después del análisis realizado bajo estas condiciones se determinó que los filtros que ofrecen mejor reconstrucción son los de 2, 4 y 8 etapas. El problema al utilizar filtros de más etapas es que los errores de redondeo se acumulan, incrementándose este defecto en cada una de las recursiones del algoritmo al aplicarlo consecutivamente en los cuatro niveles de descomposición.

El problema de cuantificar escalarmente la imagen fue abordado considerando los resultados sólo de estos filtros. Se hizo la evaluación estadística de la imagen transformada y se propusieron tamaño y número de escalones diferentes en cada caso. Como se mencionó anteriormente (V.1.2.), en algunos casos, la información de ciertas zonas es tan escasa que puede ser descartada totalmente. Aún cuando no puede generalizarse esta afirmación, no existe un

método para establecer con certeza un umbral de decisión al respecto. Se evaluaron múltiples cuantificadores, de los cuales se eligió precisamente aquel que elimina la información de las zonas $D_{2^{-1}}^1 f$, $D_{2^{-1}}^2 f$, y $D_{2^{-1}}^3 f$. Para fines de comparación, las figuras 14(b) y 14(c) han sido procesadas ambas con el mismo filtro (2 etapas), mientras que en 14(b) se han cuantificado las zonas correspondientes al primer nivel de descomposición, en 14(c) se han eliminado. El efecto es prácticamente imperceptible, la diferencia principal se aprecia en el contorno del maxilar inferior, que se ve más entrecortado en 14(b). A excepción de 14(b), en las demás imágenes fue aplicado el mismo cuantificador que en 14(c). La figura 15 muestra (incisos (b), (d) y (f)) los resultados de la aplicación del filtro de 8 etapas y el mismo cuantificador que en 14(c). Los incisos (a), (c) y (e) son las imágenes originales.

El filtro de 8 etapas se utilizó para evaluar las demás imágenes porque fue el que en conjunto con el cuantificador elegido aportó los mejores resultados en la tasa de compresión lograda.



(a)



(b)



(c)

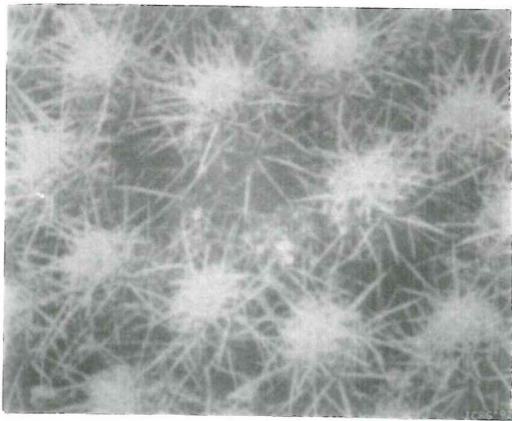


(d)

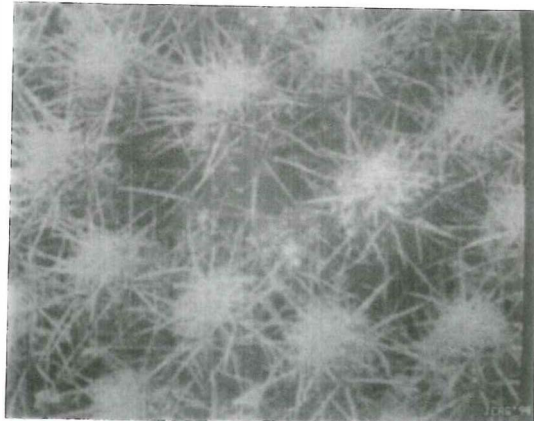


(e)

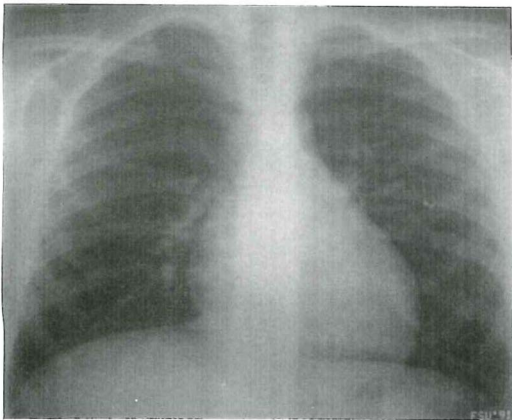
Figura 14.(a)Imagen original, (b) filtrada con 2 etapas, con cuantificación en el primer nivel de resolución,(c) filtrada con 2 etapas sin primer nivel,(d)filtrada con 4 etapas, sin primer nivel,(e)filtrada con 8 etapas, sin primer nivel.



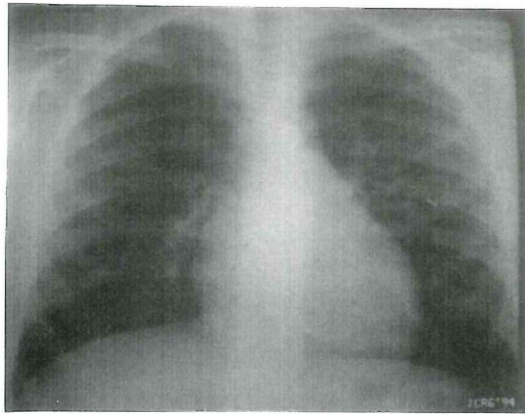
(a)



(b)



(c)



(d)



(e)



(f)

Figura 15. (a), (c) y (e) Imágenes originales; (b), (d) y (f) filtradas con 8 etapas, sin el primer nivel de resolución.

La diferencia más notoria entre 14(e) y 14(a) es la pérdida de detalles en los ojos y la disminución del brillo de la imagen. En ningún momento los contornos principales se pierden y el contexto general de la imagen se preserva.

En el caso de las imágenes presentadas en 15(b), (d) y (f) las diferencias apreciables son menores, en 15(f) pueden verse pequeños defectos en la parte izquierda del sombrero y en la línea oscura que va de la esquina superior derecha hacia el costado derecho de la cara. En 15(d) hay una pérdida de definición poco perceptible en la parte derecha de la imagen inmediatamente después del objeto blanco que ocupa la parte central (el corazón), y en 15(b) las diferencias son inapreciables. Esto ilustra el hecho de que imágenes con gran cantidad de detalles enmascaran más fácilmente los errores, pues el ojo integra la información.

Tabla I. Resultados Obtenidos.

Figura	No. de etapas del filtro	Primer nivel de resolución	Tasa de compresión obtenida	bits por pixel	PSNR(dB)
14(b)	2	cuantificado	13.31:1	0.60	29.58
14(c)	2	eliminado	15.92:1	0.50	29.46
14(d)	4	eliminado	23.83:1	0.33	23.83
14(e)	8	eliminado	25.37:1	0.31	29.34
15(b)	8	eliminado	5:1	1.59	26.70
15(d)	8	eliminado	26.18:1	0.30	25.74
15(f)	8	eliminado	10.7:1	0.74	10.7

La tabla I muestra brevemente los resultados obtenidos en la tasa de compresión para las condiciones que se indican.

El parámetro PSNR (error pico en la relación señal-ruido(reconstrucción) en inglés) es una forma objetiva de determinar el error entre la imagen original y la reconstruida.

$$PSNR(db) = 10 \log_{10} \frac{(255)^2}{MN \sum_{x=1}^{M-1} \sum_{y=1}^{N-1} (\hat{f}(x,y) - f(x,y))^2} \quad (36)$$

donde:

255 es el número máximo de niveles de gris

M es el número de píxeles en sentido vertical

N es el número de píxeles en sentido horizontal

$f(x,y)$ es la imagen original

$\hat{f}(x,y)$ es la imagen reconstruida

En teoría, mientras mayor sea este número, menor es el error entre las imágenes. Sin embargo, esto no siempre es cierto, pues la imagen puede parecer muy buena perceptualmente y no tener un PSNR muy alto. La evaluación objetiva no puede entonces determinar completamente la validez de una herramienta para la compresión, pero sí puede dar una idea de cuánto afectan los errores a la imagen reconstruida.

En las imágenes procesadas, el redondeo introducido en el cálculo de la transformada provocó una disminución del brillo de la imagen, lo que equivale a una diferencia de valor para la intensidad de cada píxel. Es decir, como existe una diferencia generalizada, el efecto neto perceptual es un cambio de brillo, pero numéricamente el resultado es un valor bajo de este parámetro.

En resumen, los resultados de la evaluación objetiva no son tan favorables si tomamos en cuenta que pueden ser mejorados sustancialmente, sin embargo, la preservación de los elementos que constituyen la imagen son la característica más favorable de esta transformada.

Los resultados obtenidos para la compresión de la imagen 14(a) no pueden generalizarse para cualquier imagen, 14(b) corrobora lo anterior, pues la tasa de compresión obtenida para ella contrasta por lo pequeña que es con la obtenida para 14(e).

VI. CONCLUSIONES.

La transformada de ondículas es una alternativa mejor que la transformada discreta de coseno para la compresión de imágenes, por sus características de preservación de contornos, de resolución múltiple (como sucede en el sistema de percepción visual humano) y una complejidad computacional relativamente menor debido a su carácter convolucional. En contraparte, su aplicación está limitada a procesamientos en tiempo diferido debido a la recursión necesaria para descomponer la imagen en los distintos niveles. Este hecho provoca serios problemas de manipulación de datos cuando la transformación se lleva a cabo en plataformas computacionales pequeñas, mismos que pueden solucionarse optimizando los algoritmos y el uso de recursos de máquina.

El trabajo que se ha realizado para la utilización de la transformada de ondículas es muy reciente y aún existen una serie de problemas abiertos entre los cuales, la selección de la ondícula madre (generadora de la familia) y la obtención de los filtros para su realización computacional son los más importantes por resolver.

Aún cuando las condiciones que deben cumplir tanto la función madre como los filtros han sido establecidas [Daubechies, 1988], existen múltiples soluciones posibles, las cuales pueden ser óptimas para una aplicación específica, pero no necesariamente para otras. Es decir, no existen soluciones genéricas.

Puesto que las normas internacionales actuales están basadas en otra transformación (discreta de coseno), para el caso de la transformada de ondículas no existen aún evaluaciones estadísticas para determinar los mejores parámetros para los cuantificadores y los codificadores de longitud variable. Esto hace muy

difícil establecer códigos y/o cuantificadores óptimos, por lo que las tasas de compresión que pueden obtenerse son susceptibles de mejora.

Los cuantificadores escalares son los más recomendables por la sencillez de su aplicación a la imagen transformada. Aún cuando los cuantificadores vectoriales parecen ofrecer ciertas ventajas, su diseño es mucho más complejo y dado que no existen resultados estadísticos del comportamiento de la transformada, su aplicación óptima sería prácticamente imposible en el estado actual.

El uso de un número y tamaño de escalón variables dependiendo de la zona de la imagen transformada por cuantificar es la mejor forma de obtener un agrupamiento eficiente por zonas de la información de la imagen. Aún cuando la elección de estos parámetros es todavía motivo de investigación, usar cuantificadores más finos en áreas de contenido relevante de información y más gruesos en áreas con menor relevancia preserva los detalles que dan nitidez a la imagen resultante, a un bajo costo en la cantidad de información total.

Ante la carencia de información estadística no se cuenta por el momento con un código de longitud variable óptimo. La condición ideal sería generar un código óptimo para cada imagen, situación por demás impráctica. Sin embargo, éste es el mejor método para lograr la máxima tasa de compresión por el momento.

Los resultados obtenidos muestran que existen una serie de parámetros que deben ser ajustados y elementos del proceso de compresión que deben ser optimizados para esta transformada.

Por otro lado, las alternativas que buscan la compresión de imágenes por medio del modelado, no son viables por el momento debido a las grandes

limitantes para la creación de formatos uniformes, inclusive para clases de imágenes. Sin embargo, es de esperarse que en el futuro representen un área importante que aporte resultados mucho mejores que los actuales, pero tomando en cuenta que será siempre sobre clases de imágenes y no en general, como en los métodos actuales. La aplicación de estas técnicas estará más orientada a las comunicaciones personales.

Las técnicas iterativas son una buena alternativa que al igual que la transformada de ondículas han sido poco estudiadas. La principal limitante es la potencia de cómputo requerida, sin embargo, este obstáculo se vencerá fácilmente en el futuro haciéndola una técnica viable sobre todo en máquinas computadoras grandes, lo que hace pensar que su aplicación no llegará a difundirse tanto como la transformada de ondículas.

LITERATURA CITADA

- Aizawa, K., H. Harashima, T. Saito 1989. "Model-Based Analysis Synthetic Image Coding for a Person's Face", *Image Communication*, vol. 1, no. 2., pp 139-152.
- Akansu, Ali N., Richard E. Haddad 1992. "Multiresolution Signal Decomposition" Academic Press, Inc. pp 376.
- Barnsley, Michael F. 1993. "Model-Based Fractal Coding". Workshop on Very Low Bitrate Video Compression. Urbana-Champaign, Ill. USA.
- Burt, P.J., E. H. Adelson 1983. "The Laplacian Pyramid as a compact image code". *IEEE Trans. on Commun.* vol. COM-31., no. 4, pp 532-540.
- Daubechies, I. 1988. "Orthonormal bases of compactly supported wavelets". *Commun. Pure Appl. Math.* vol. 41, pp 909-996.
- Encyclopaedia Britannica. vol. 7. 1974. pp. 93-107.
- Gabor, D. 1946. "Theory of Communication", *Journal of the Institute of Electrical Engineering.*, vol. 93, pp 429-457.
- Gonzales, Rafael C., Richard E. Woods 1992. "Digital Image Processing". Addison-Wesley Publishing Co. 716 pp.
- Huffman, D. A. 1952. "A Method for the construction of minimum Redundancy codes". *Proc. of the IRE*, vol. 40. no. 10. pp 1098-1101.
- Mallat, S. 1989. "A Theory for multiresolution signal Decomposition: The Wavelet Representation". *IEEE Trans. on Pattern Analysis and Machine Intelligence.* vol. 11, no. 7. pp 674-693.
- Mallat, S. 1989. "Multiresolution Approximation and Wavelet Ortonormal Bases of L^2 ". *Trans. of the American Mathematical Society.* vol. 3, no. 15, pp 69-87.
- Mallat, S. 1989. "Multifrequency Channel Decompositions of Images and Wavelet models". *IEEE Trans. on Acoustics, Speech, and Signal Processing.* vol. 37, no. 12. pp 2091-2110.

Musmann, H.G. 1993. "A Unified Point of View on Video Compression: A Model-Based Perspective". Workshop on Very Low Bitrate Video Compression. Urbana-Champaign, Ill. USA.

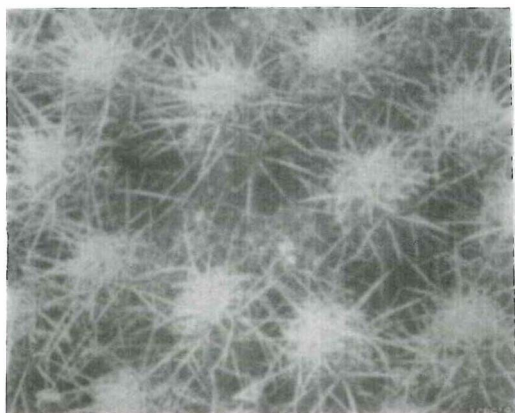
Rioul, O., M. Vetterli 1991. "Wavelets and Signal Processing". IEEE Signal Proc. Magazine. vol. 8, no. 5.

APÉNDICE A. PERSPECTIVAS DE LA COMPRESION DE IMÁGENES.

A.1. Tendencias actuales.

Cuando se procesan imágenes por medio de transformaciones espacio - frecuencia, independientemente del método de transformación que se aplique se busca reagrupar más eficientemente la energía para lograr una representación más compacta de la información. Este proceso tiene como límite la cantidad de información que contiene la imagen, conocido como entropía, y es imposible desde el punto de vista de la teoría de la información ir más allá de este límite. Por lo tanto, siempre tendremos la limitante impuesta por la entropía como el principal obstáculo cuando se desea lograr una alta tasa de compresión. Cuando una imagen contiene grandes zonas de textura similar (ver figura 16(b)), el nivel de entropía será bastante bajo, ya que la mayor parte de la imagen no aporta información. En ésta será posible lograr una alta tasa de compresión.

Por el contrario, una imagen compleja (ver figura 16(a)) tendrá un nivel de entropía bastante alto por su gran contenido de información, por lo que la tasa de compresión será menor.



(a)



(b)

Figura 16. Imágenes con una alta entropía (a) y una baja entropía (b).

En la actualidad, una de las tendencias más fuertes de investigación en el área es tratar de lograr las máximas tasas de compresión para lograr abatir el ancho de banda necesario cuando, por ejemplo, se desea transmitir las imágenes por algún medio de comunicación de ancho de banda limitado.

La otra tendencia considera que el ancho de banda del canal de comunicación no es un problema y pone como ejemplo los canales basados en fibra óptica, por lo que no se justifica la compresión a tasas tan grandes.

Si la entropía de las imágenes es la limitante de las tasas de compresión es lógico suponer que deben encontrarse alternativas a las técnicas de compresión basadas en transformaciones espacio - frecuencia en las que este factor no sea fundamental. Se han desarrollado trabajos [Aizawa, et al., 1989];[Musmann, 1993] utilizando modelos de imágenes basados en el conocimiento previo de una clase de las mismas y utilizando transformaciones iteradas [Barnsley, 1993]. La filosofía detrás de estas técnicas es que si ya no estamos hablando de una imagen digital como un conjunto de píxeles con su correspondiente energía, sino de formas u objetos comprendidos en un área determinada, es posible comparar las formas y ubicaciones de los objetos en el plano espacial ya sea a partir del conocimiento previo de clase comparándolas contra parámetros almacenados (vg. caras, manos, etc.) o a partir de una función iterada.

Estas técnicas de modelado se vuelven particularmente útiles cuando se aplican a imágenes en movimiento (video), pues es más fácil predecir o establecer vectores de movimiento de objetos completos que utilizar ventanas de búsqueda como se hace tradicionalmente.

Todas estas técnicas tienden a sintetizar artificialmente las características naturales de las imágenes como medio de reducción de información. Para algunas aplicaciones, es tolerable este "defecto" introducido a cambio de un incremento considerable en la capacidad de almacenamiento o transmisión en los medios conocidos.

Es de esperarse que en el futuro se mejore el aspecto de las imágenes resultante (se "naturalice"), y que se encuentren métodos para extender su aplicación a más clases de imágenes.

A.2. Técnicas basadas en modelos.

Estas técnicas comprenden el uso de un modelo contra el cual se analiza la imagen y cuyos parámetros representan precisamente a ésta. Existen tres vertientes de esta técnica [Musmann, 1993].

A.2.1. Para objetos desconocidos.

La imagen contiene diversos objetos que deben ser descritos. Se realiza el análisis de la imagen para obtener los parámetros de forma e intensidad luminosa (tonos de gris), mismos que posteriormente se codifican (ver figura 17).

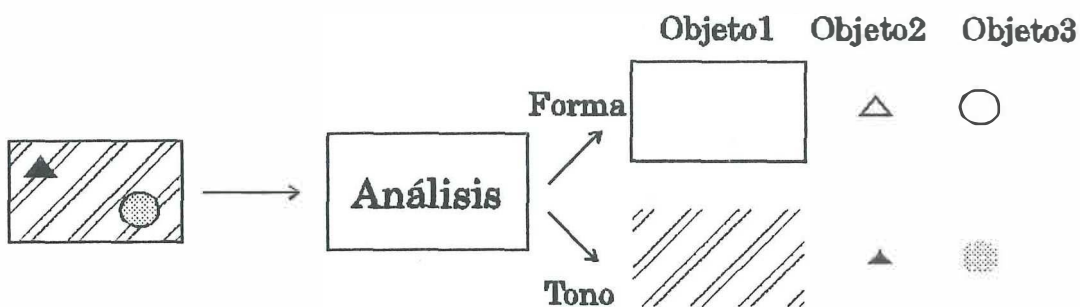


Figura 17. En el modelo para objetos desconocidos se hace un análisis para extraer forma y textura de los objetos de una imagen.

Los umbrales de detección de objetos sirven para efectuar la decisión entre diferentes objetos de tonos de gris muy similares los cuales pueden o no ser asociados a un mismo objeto.

El principal problema de este método es la determinación de umbrales de decisión, ya que para algunas imágenes habrá una infinidad de objetos pequeños muy similares entre sí (tonos muy parecidos) que podrían no ser separados sino agrupados en uno solo, lo que introduciría una gran cantidad de error. Puede verse que esta técnica está orientada a imágenes de estructura sencilla y será poco eficiente en el caso de imágenes naturales, pues el trabajo computacional de procesamiento es muy fuerte y los resultados muy probablemente no serán aceptables desde el punto de vista perceptual.

A.2.2. Para objetos conocidos.

Esta técnica se aplica a "clases" de objetos conocidos, por ejemplo, caras humanas. Por lo mismo, su aplicación está limitada a un conjunto preestablecido de imágenes y situaciones en las que el objeto principal (conocido) es prácticamente el único presente. Su aplicación se encuentra por ejemplo en la compresión de imágenes para archivos de personal.

Aquí, el modelo contiene una mascarilla poligonal con la forma del objeto conocido, de manera que al analizar la imagen, solamente se adapta la misma (alineación) al modelo y se extrae la información de textura (luminancia).

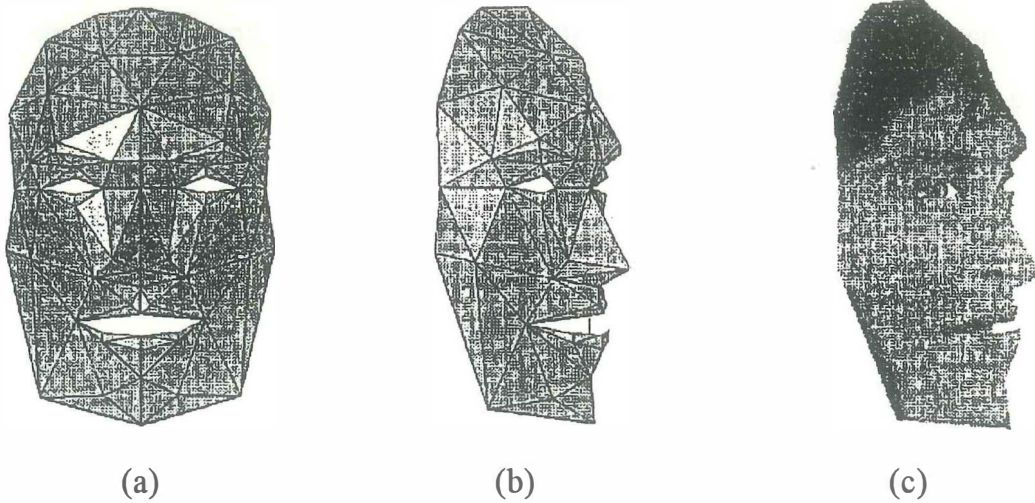


Figura 18. En el caso de objetos conocidos, se realiza una mascarilla poligonal (a) y (b), que se adapta al sujeto (c).

Los problemas a los que se enfrenta esta técnica son la adaptación a formas preestablecidas de clase que podría resultar en cierta distorsión de forma y la extracción de la textura bajo diferentes condiciones de luminosidad lo cual podría ocasionar distorsiones adicionales provocadas por el ajuste al modelo. La apariencia general podría ser bastante artificial, a cambio de una alta tasa de compresión. Siendo su principal restricción la clase de imágenes a que se destine.

A.2.3. Para expresiones faciales.

Esta es una extensión mejorada de la técnica anterior que se ha usado exclusivamente para el análisis de secuencias de imágenes (video) de individuos que efectúan acciones comunes frente a una cámara, como sería, por ejemplo, un informador en un noticiero televisivo. Es decir, la clase de imágenes se restringe aún más para poder manipular mejor el modelo.

En este caso el modelo consta de un conjunto de mascarillas poligonales contra las cuales, por medio de la detección de movimiento (posible cuando se habla de video), se comparan cada una de las imágenes de la secuencia de manera que la información necesaria para describir la secuencia es exclusivamente la "acción" que realiza el sujeto (ver figura 19). El conjunto de mascarillas se limita a las acciones posibles en este caso, como serían: ojos abiertos, boca abierta, sonrisa, ojos y boca cerrados, y otras, por lo que puede verse que puesto que el conjunto de movimientos no es muy grande, el conjunto de mascarillas no sería muy grande.

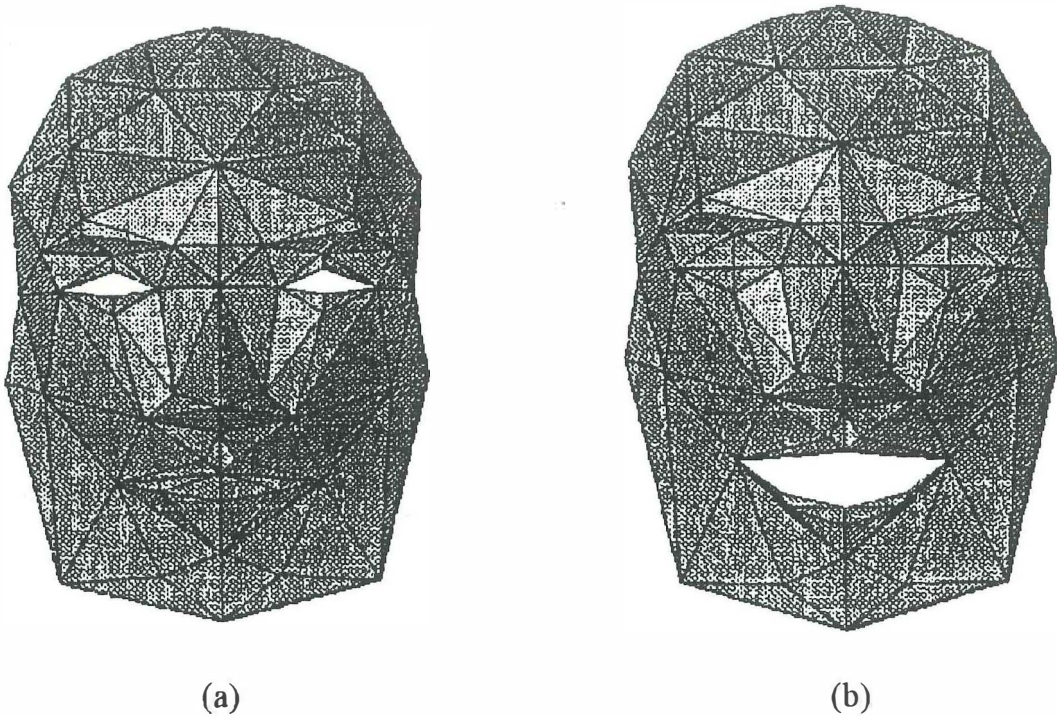


Figura 19. En el uso de modelos para expresiones faciales, un conjunto limitado de acciones posibles se almacena, como (a): ojos abiertos, boca cerrada, y (b): ojos cerrados, boca abierta.

El principal problema con este método es que las acciones están muy limitadas (al conjunto posible), aparte de los problemas mencionados en VI.2.2. Sin embargo, dentro de su margen de acción su comportamiento ha sido muy bueno logrando grandes tasas de compresión.

A.3. Técnicas iterativas.

Esta técnica se basa en el uso de transformaciones afines, formando un sistema de funciones iteradas para llevar a cabo la compresión de una imagen. La imagen comprimida estará formada por los coeficientes de la transformación para las distintas iteraciones. Una transformación afín es una función matemática compuesta por una combinación de rotación, escalamiento, inversión y translación en el espacio. Es una transformación espacio - espacio. Un sistema de funciones iteradas es un conjunto de transformaciones afines.

Una imagen natural puede considerarse como un conjunto de translaciones, rotaciones y escalamientos de muy pequeñas secciones de la misma que pueden ser descritas por una transformación afín. Como ejemplo clásico de una forma natural típica se tiene una hoja de helecho, la cual se compone de múltiples repeticiones trasladadas, rotadas, invertidas y escaladas de una pequeña sección. Mediante transformaciones afines aplicadas en forma iterativa podrá lograrse encontrar conjuntos de coeficientes que representen completamente la imagen.

Sin embargo, para imágenes naturales que no poseen la característica de auto- similitud de una hoja de helecho, el proceso se vuelve muy complejo y su complejidad computacional muy alta.

Las tasas de compresión que se logran son bastante altas, pero la plataforma computacional debe ser necesariamente muy potente.

A.4. Normatividad.

Actualmente, al contrario de lo que sucedió en su momento con la transformación de coseno discreta, se realizan esfuerzos internacionales para normalizar las técnicas y agruparlas de manera de poder unificar resultados.

El problema esta aún abierto y se considera que se tendrán las primeras propuestas de trabajo a finales de 1994, y se espera contar con una norma internacional terminada para el fin de siglo. Los trabajos que iniciaron en julio de 1992 bajo la futura norma MPEG-4 son auspiciados por la Unión Internacional de Telecomunicaciones (UIT) a través del TSS (Telecommunications Standardization Sector, antes CCITT).

**ANEXO A. TABLAS DE CODIFICACION DE LONGITUD
RECORRIDA SEGUN LA NORMA MPEG-2 DE LA ISO.**

Longitud recorrida	Nivel	Código de longitud variable
EOB	-	100*
C64	-	101 **
0	1	1s si 1er. Coef. ***
0	1	11s si no es 1er. Coef.
0	2	0100 s
0	3	0010 1s
0	4	0000 110s
0	5	0010 0110 s
0	6	0010 0001 s
0	7	0000 0010 10s
0	8	0000 0001 1101 s
0	9	0000 0001 1000 s
0	10	0000 0001 0011 s
0	11	0000 0001 0000 s
0	12	0000 0000 1101 0s
0	13	0000 0000 1100 1s
0	14	0000 0000 1100 0s
0	15	0000 0000 1011 1s
0	16	0000 0000 0111 11s
0	17	0000 0000 0111 10s
0	18	0000 0000 0111 01s
0	19	0000 0000 0111 00s
0	20	0000 0000 0110 11s
0	21	0000 0000 0110 10s
0	22	0000 0000 0110 01s
0	23	0000 0000 0110 00s
0	24	0000 0000 0101 11s
0	25	0000 0000 0101 10s
0	26	0000 0000 0101 01s
0	27	0000 0000 0101 00s
0	28	0000 0000 0100 11s
0	29	0000 0000 0100 10s
0	30	0000 0000 0100 01s
0	31	0000 0000 0100 00s
0	32	0000 0000 0011 000s
0	33	0000 0000 0010 111s

Longitud recorrida	Nivel	Código de longitud variable
0	34	0000 0000 0010 110s
0	35	0000 0000 0010 101s
0	36	0000 0000 0010 100s
0	37	0000 0000 0010 011s
0	38	0000 0000 0010 010s
0	39	0000 0000 0010 001s
0	40	0000 0000 0010 000s
1	1	011s
1	2	0001 10s
1	3	0010 0101 s
1	4	0000 0011 00s
1	5	0000 0001 1011 s
1	6	0000 0000 1011 0s
1	7	0000 0000 1010 1s
1	8	0000 0000 0011 111s
1	9	0000 0000 0011 110s
1	10	0000 0000 0011 101s
1	11	0000 0000 0011 100s
1	12	0000 0000 0011 011s
1	13	0000 0000 0011 010s
1	14	0000 0000 0011 001s
1	15	0000 0000 0001 0011s
1	16	0000 0000 0001 0010s
1	17	0000 0000 0001 0001s
1	18	0000 0000 0001 0000s
2	1	0101 s
2	2	0000 100s
2	3	0000 0010 11s
2	4	0000 0001 0100 s
2	5	0000 0000 1010 0s
3	1	0011 1s
3	2	0010 0100 s
3	3	0000 0001 1100 s
3	4	0000 0000 1001 1s
4	1	0011 0s
4	2	0000 0011 11s
4	3	0000 0001 0010 s

Longitud recorrida	Nivel	Código de longitud variable
5	1	0001 11s
5	2	0000 0010 01s
5	3	0000 0000 1001 0s
6	1	0001 01s
6	2	0000 0001 1110 s
6	3	0000 0000 0001 0100s
7	1	0001 00s
7	2	0000 0001 0101 s
8	1	0000 111s
8	2	0000 0001 0001 s
9	1	0000 101s
9	2	0000 0000 1000 1s
10	1	0010 0111 s
10	2	0000 0000 1000 0s
11	1	0010 0011 s
11	2	0000 0000 0001 1010s
12	1	0010 0010 s
12	2	0000 0000 0001 1001s
13	1	0010 0000 s
13	2	0000 0000 0001 1000s
14	1	0000 0011 10s
14	2	0000 0000 0001 0111s
15	1	0000 0011 01s
15	2	0000 0000 0001 0110 s
16	1	0000 0010 00s
16	2	0000 0000 0001 0101s
17	1	0000 0001 1111 s
18	1	0000 0001 1010 s
19	1	0000 0001 1001 s
20	1	0000 0001 0111 s
21	1	0000 0001 0110 s
22	1	0000 0000 1111 1s
23	1	0000 0000 1111 0s
24	1	0000 0000 1110 1s
25	1	0000 0000 1110 0s
26	1	0000 0000 1101 1s

Longitud recorrida	Nivel	Código de longitud variable
27	1	0000 0000 0001 1111s
28	1	0000 0000 0001 1110s
29	1	0000 0000 0001 1101s
30	1	0000 0000 0001 1100s
31	1	0000 0000 0001 1011s
ESCAPE		0000 01 ****

* Código modificado

** Código añadido

*** Código no utilizado

**** Estructura de la secuencia ESCAPE: 0000 01+(6 bits)+(8 bits)

6 bits son de la longitud recorrida, entre 0 y 63.

8 bits son del valor encontrado, entre 0 y 255.

ANEXO B. OPERACION DE LOS PROGRAMAS DE TRANSFORMACION Y CODIFICACION.

Programas de codificación y decodificación de longitud variable:

Para codificación: **codlv** narch1.ext narch2.ext

donde:

narch1.ext es un archivo que ha sido transformado y cuantificado.
narch2.ext es el archivo comprimido resultante de la codificación.

Para decodificación: **decodlv** narch1.ext narch2.ext

donde:

narch1.ext es un archivo comprimido resultante de la codificación.
narch2.ext es el archivo expandido al cual se le debe aplicar la transformada inversa para reconstruir la imagen.

Programa para transformación directa e inversa por ondículas: **ondaub**.

Opción.

Operación.

- | | |
|-----|--|
| -tn | Efectúa la transformación hasta el número de niveles indicado por n. |
| -in | Efectúa la transformación inversa hasta n veces. |
| -dn | Indica el tipo de filtro a usar, n va de 2 a 20, y corresponden a los filtros de Daubechies. |
| -? | Despliega en pantalla esta información. |

Ejemplos de su uso:

ondaub -t4 -d4 narch.ext

Indica que se realizará la transformada de ondículas en 4 niveles de resolución, utilizando el filtro de 4 etapas de Daubechies sobre la imagen narch.ext

ondaub -i4 -d4 narch.ext

Indica que se realizará la transformada inversa de ondículas en 4 niveles de resolución, utilizando el filtro de 4 etapas de Daubechies sobre la imagen narch.ext

Los archivos resultantes se identifican con el mismo nombre narch, pero la extensión tendrá la nomenclatura ttn si es transformada directa o iin si es transformada inversa, donde la n es el número de etapas del filtro utilizado.

**ANEXO C. COEFICIENTES DE DAUBECHIES PARA LOS
FILTROS DE ESPEJO EN CUADRATURA.**

Tamaño del filtro	Coefficientes
2	0.707106781187 0.707106781187
4	0.482962913145 0.836516303738 0.224143868042 -0.129409522551
6	0.332670552950 0.806891509311 0.459877502118 -0.135011020010 -0.085441273882 0.035226291882
8	0.230377813309 0.714846570553 0.630880767930 -0.027983769417 -0.187034811719 0.030841381836 0.032883011667 -0.010597401785
10	0.160102397974 0.603829269797 0.724308528438 0.138428145901 -0.242294887066 -0.032244869585 0.077571493840 -0.006241490213 -0.012580751999 0.003335725285

Tamaño del filtro	Coefficientes
12	0.111540743350 0.494623890398 0.751133908021 0.315250351709 -0.226264693965 -0.129766867567 0.097501605587 0.027522865530 -0.031582039318 0.000553842201 0.004777257511 -0.001077301085
14	0.077852054085 0.396539319482 0.729132090846 0.469782287405 -0.143906003929 -0.224036184994 0.071309219267 0.080612609151 -0.038029936935 -0.016574541631 0.012550998556 0.000429577973 -0.001801640704 0.000353713800
16	0.054415842243 0.312871590914 0.675630736297 0.585354683654 -0.015829105256 -0.284015542962 0.000472484574 0.128747426620 -0.017369301002 -0.044088253931 0.013981027917 0.008746094047 -0.004870352993 -0.000391740373 0.000675449406 -0.000117476784

Tamaño del filtro	Coeficientes
18	0.038077947364 0.243834674613 0.604823123690 0.657288078051 0.133197385825 -0.293273783279 -0.096840783223 0.148540749338 0.030725681479 -0.067632829061 0.000250947115 0.022361662124 -0.004723204758 -0.004281503682 0.001847646883 0.000230385764 -0.000251963189 0.000039347320
20	0.026670057901 0.188176800078 0.527201188932 0.688459039454 0.281172343661 -0.249846424327 -0.195946274377 0.127369340336 0.093057364604 -0.071394147166 -0.029457536822 0.033212674059 0.003606553567 -0.010733175483 0.001395351747 0.001992405295 -0.00068586695 -0.000116466855 0.000093588670 -0.000013264203