APLICACIÓN DE LA TRANSFORMADA WAVELET EN COMPRESION DE IMAGENES

FERNANDO RUIZ VERA
Ingeniería Electrónica
Universidad Distrital "Francisco José de Caldas"
jmruiz@ciencias.ciencias.unal.edu.co

Resumen: La compresión de imágenes es una herramienta esencial para aplicaciones tales como transmisión y almacenamiento de datos. Este trabajo hace uso de la técnica conocida como "Transform Coding", que aprovecha las características espaciales y frecuenciales de una imagen, utilizando la transformada Wavelet en su primera etapa. Para la descomposición wavelet se aplica el algoritmo piramidal, obteniéndose diferentes escalas de resolución para las direcciones horizontal, diagonal y vertical. En una segunda etapa se cuantizan por separado cada una de las subbandas resultantes de la transformación, para después ser codificadas aritméticamente.

Palabras Clave: Transformada Wavelet, Cuantización, Codificación Huffman, Codificación Aritmética..

I. INTRODUCCIÓN

El crecimiento de la información almacenada o transmitida en sistemas de computo ha hecho necesario representar esta información de una manera más eficiente ya sea ocupando menor espacio de almacenamiento o requiriendo menores tiempos de transmisión. Con el fin de lograr esto se aprovecha la redundancia existente entre los datos. El algoritmo aquí presentado hace uso de la técnica conocida como "Transform Coding" que es un método de compresión con pérdida de información, es decir, que la imagen recuperada del proceso de compresión no es idéntica a la original. Este método permite lograr razones de compresión entre 10:1 sin un mayor deterioro en la calidad de las imágenes, frente a los métodos de compresión sin pérdidas que recuperan fielmente la información original pero tan solo alcanzan razones de 3:1.

El proceso de compresión por "Transfrom Coding" consta de las etapas de transformación, cuantización y codificación como se muestra en la Fig. 1.

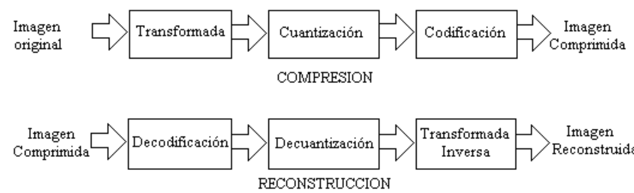

Figura 1. Proceso de compresión y reconstrucción.

El algoritmo más empleado que hace uso de esta técnica es el estándar JPEG, cuyo diagrama de bloques se muestra en la Fig. 2.

El esquema presentado en este articulo emplea en lugar de la transformada Discreta Coseno (DCT), la Transformada Discreta Wavelet (DWT), realiza una cuantización escalar y una codificación aritmética como codificador entrópico sobre cada subbanda.

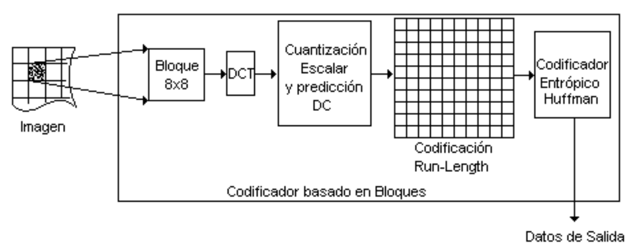

Figura 2. Diagrama de Bloques de un codificador JPEG

II. ETAPA DE TRANSFORMACIÓN - TRANSFORMADA WAVELET DISCRETA (DWT)

El objetivo de la etapa de transformación consiste en representar la información de la imagen en una base matemática distinta esperando que en esta nueva representación se manifieste la correlación o redundancia. De forma tal que se pueda despreciar gran parte de los coeficientes (datos transformados), aprovechando que la información esta almacenada en unos cuantos coeficientes.

La transformada a utilizar debe cumplir con las siguientes propiedades:

1. La transformada debe ser independiente a los datos
2. Existencia un algoritmo de rápido computo
3. La transformada debe ser capaz de remover la correlación para un gran conjunto de datos.

La transformada rápida de Fourier (FFT) podría usarse puesto que cumple con las dos primeras propiedades, sin embargo, debido a que sus bases (seno y coseno), se encuentran totalmente localizadas en frecuencia y por lo tanto totalmente dispersas en tiempo o espacio, no son capaces de revelar la correlación espacial local; razón por lo cual se busca una base que se encuentre localizada tanto en tiempo como en frecuencia.

Existen dos formas de lograr esto, una dividir el dominio espacial en regiones y usar una serie de Fourier para cada una de ellas por separado llegando a bases trigonométricas locales, como es el caso del estándar JPEG mostrado en la Fig. 2. La segunda forma es usar bases Wavelets, aprovechando que sus bases descomponen una señal en diferentes niveles de resolución.

Para realizar la transformación wavelet se debe establecer una base ortogonal, la cual es generada por una única función ψ a través de sus dilaciones y traslaciones $\psi^{a,b}(t) = |a|^{-1/2}\psi\left(\frac{t-b}{a}\right)$ (1), esta función ψ se conoce como base o madre wavelet y debe satisfacer que $\int \psi(x)dx = 0$, y $\int \psi^2(x)dx < \infty$. De esta forma una función f puede ser escrita como una superposición de wavelets, esto es:

$$f = \sum c_{m,n}(f)\psi_{m,n} \qquad (2).$$

Discretizando por razones de computo se obtiene, $a = a_0^m, b = b_0 a_0^m$, con $m,n \in Z$, y $a_0 > 1, b_0 > 0$ fijos. Siendo $\psi_{m,n}(t) = \psi^{a_0^m, nb_0 a_0^m}(t) = a_0^{-m/2}\psi(a_0^{-m}t - nb_0)$ la base wavelet ya discretizada. Como ecuación de análisis se obtiene:

$$c_{m,n}(f) = <\psi_{m,n}> = \int \psi_{m,n}(x) f(x) dx \qquad (3).$$

Para el calculo de la transformada se hace uso del algoritmo piramidal y del análisis multiresolución [4], descrito por las Ecuaciones de análisis (4), la Ecuación de Síntesis (5) y la Fig. 3.

$$\begin{aligned} c_{m,n}(f) &= \sum_k g_{2n-k} a_{m-1,k}(f) \\ a_{m,n}(f) &= \sum_k h_{2n-k} a_{m-1,k}(f) \end{aligned} \qquad (4).$$

$$a_{m-1,l}(f) = \sum_n [h_{2n-l} a_{m,n}(f) + g_{2n-l} c_{m,n}(f)] \qquad (5).$$

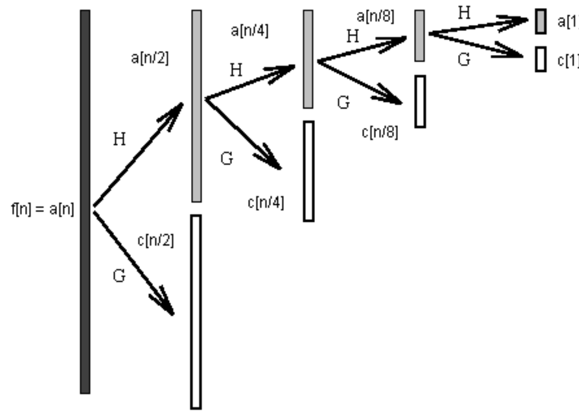

Figura 3. Transformada Discreta Wavelet como algoritmo piramidal.

La definición aquí presentada puede ser extendida a más dimensiones, y para el caso particular de imágenes existen dos formas de llevar a cabo la descomposición wavelet, una la estándar que aplica la transformada sobre cada fila y luego sobre cada columna, mientras que la no estándar aplica un paso del algoritmo piramidal sobre cada fila y luego sobre cada columna, aplicando de nuevo un paso sobre las filas y columnas resultantes hasta terminar. Ambos tipos de descomposición se describen en mayor detalle en [2].

II. Etapa de Cuantización

La Etapa de cuantización puede ser vista como una conversión Análoga/Digital, ya que con ella se restringe los valores de los coeficientes (que toman valores en un continuo), a un número limitado de posibilidades siendo esta etapa la causante de las pérdidas.

Se ha utilizado la cuantización escalar frente a la vectorial por ser más simple en su implementación aunque la segunda es más eficiente ya que codifica un grupo de coeficientes en lugar de uno único coeficiente por símbolo. En [1] se propone el uso de un cuantizador vectorial haciendo uso del algoritmo LBG desarrollado en [3].

La cuantización escalar puede ser uniforme o no, con el fin de aprovechar las propiedades estadísticas de los coeficientes wavelets, ya que ellos pueden ser fácilmente ajustados a una distribución Gaussiana Generalizada [1]. El histograma de la imagen de Lena y de sus coeficientes wavelets se muestra en la Fig. 4, en ésta se muestra como gran parte de los coeficientes son de baja magnitud, mientras que unos pocos son de alta magnitud y son estos los que deben almacenarse con mayor exactitud.

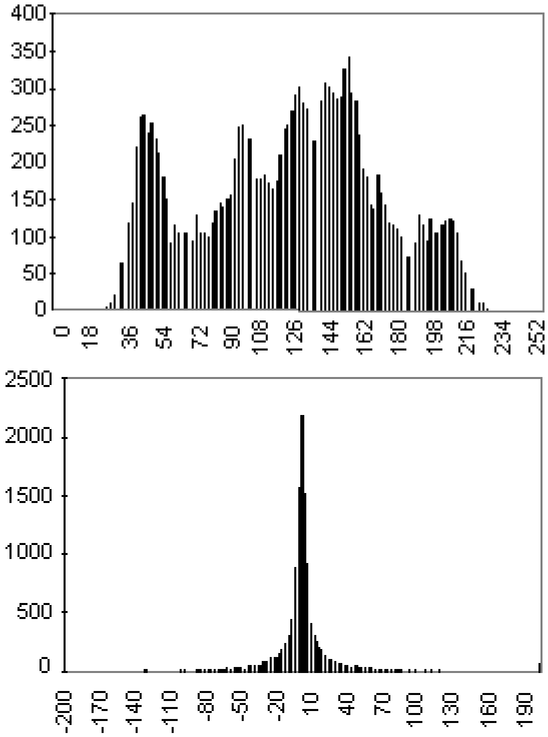

Figura 4. Histograma de la imagen de Lena (Arriba) y amplitudes de su DWT (Abajo)

En general, un cuantizador escalar puede ser definido por un conjunto de intervalos o celdas $S = \{S_i; i \in I\}$ junto a un conjunto de valores de reproducción, o niveles $C = \{y_i; i \in I\}$ ya que el cuantizador q es definido por $q(x) = y_i$ para $x \in S_i$. Lo que puede ser expresado como $q(x) = \sum_i y_i 1_{S_i}(x)$, donde $1_{S_i}(x)$ es 1 si $x \in S_i$ y 0 en otro caso.

III. Etapa de Codificación – Codificación Aritmética

El paso de codificación consiste en reemplazar, de forma reversible la cadena de símbolos provenientes del cuantizador por una cadena de bits, el objetivo es reducir el número de bits que se deben transmitir o almacenar, esto se hace posible eliminando la redundancia presente en la cadena de símbolos.

Generalmente se usa la codificación Huffman como codificador entrópico que asigna las palabras de menor longitud a los símbolos más probables y las más largas a los menos probables, sin embargo este codificador tan sólo puede codificar los símbolos en un número entero de bits, razón por la cual es superior la codificación aritmética ya que esta codifica un mensaje completo en el intervalo [0,1]. El algoritmo básico para realizar la codificación es el siguiente:

1. Dividir el intervalo [0,1] en segmentos correspondientes a los **M** símbolos, donde el segmento de cada símbolo tiene una longitud proporcional a su probabilidad.
2. Elegir el segmento del primer símbolo en la cadena del mensaje.
3. Dividir el segmento de este símbolo otra vez en **M** nuevos segmentos con longitud proporcional a las probabilidades de los símbolos.
4. De estos nuevos segmentos, elegir el correspondiente al próximo símbolo en el mensaje.
5. Continuar con los pasos 3. y 4. hasta que el mensaje este codificado.
6. Representar el valor del segmento por una función binaria.

La Fig. 5 muestra la codificación de un mensaje en especial, para ello se debe conocer la distribución de probabilidad de los símbolos, tanto en la codificación como en la decodificación, debido a esto el archivo final debe contener esta información o se debe implementar un codificador aritmético adaptativo. También se debe notar la existencia de un símbolo para fin de mensaje (!, para el ejemplo anterior). Otras consideraciones se deben tener en cuanta en la práctica para la implementación de este tipo de codificador, una descripción detallada se puede encontrar en [5].

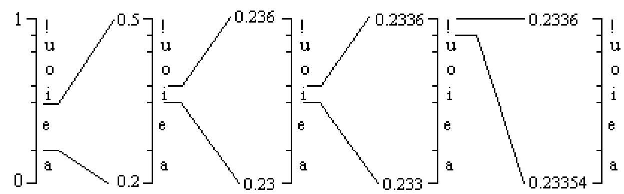

Figura 5. Codificación Aritmética del mensaje eiii.

IV. Implementación del Sistema de Compresión

La aplicación desarrollada toma una imagen en formato Bitmap con 256 niveles de gris, sobre la que se aplica el algoritmo de compresión creando un archivo final de menor tamaño, con la información necesaria para la recuperación de la imagen original. La estructura del compresor y de la creación de este archivo se muestra en la Fig. 6.

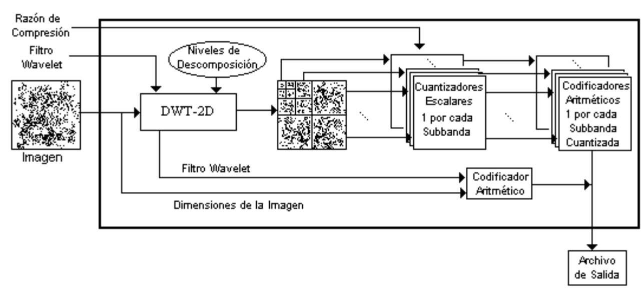

Figura 6. Estructura del Compresor Wavelet

Como se muestra en la figura anterior a la imagen original se le aplica la transformada wavelet discreta en 2 dimensiones, para lo cual la imagen es colocada en un vector de longitud = ancho * alto, colocando una fila a continuación de la anterior, a este primer bloque se le especifica el filtro wavelet a usar, las dimensiones de la imagen, y el número de niveles de resolución. Este último se encuentra fijo con un valor de 5, lo que permite comprimir imágenes desde 64x64 pixeles.

Cada subbanda resultante es cuantizada por separado, estableciendo el número de niveles y el paso de cuantización entre niveles en cada subbanda con el fin de cumplir con la razón de compresión esperada. Cada una de las subbandas cuantizadas junto a la información de niveles y paso de cuantización pasan al codificador. Con esta información se calculan las probabilidades de cada uno de los símbolos de entrada, como esta información es necesaria para la decodificación en el proceso de descompresión, también es codificada aritméticamente. Adicionalmente se agrega la información del filtro wavelet que fue usado y el ancho y alto de la imagen.

V. RESULTADOS

El esquema de compresión Wavelet aquí presentado permite alcanzar razones de compresión más altas frente al JPEG, sin un deterioro de la información como el que se sufre en este último, principalmente por causa de la descomposición subbanda provocada por la transformación gracias a la cual se logra aprovechar las características tanto en frecuencia como en espacio de la imagen. Adicionalmente la codificación aritmética permite aprovechar de mejor forma la entropía de los coeficientes ya cuantizados que con la codificación Huffman.

La Pruebas mostradas en el presente articulo fueron hechas sobre una imagen de Lena de 128x128 pixeles, usando como filtro wavelet el Daubechies 4 y descomponiendo la imagen hasta el 5° nivel de resolución.

La Fig. 7 muestra los resultados al comprimir dicha imagen, allí se observa como en el esquema wavelet crece más lentamente la medida del MSE, mientras la PSNR se mantiene baja en ambos esquemas, al aumentar la razón de compresión. En la Fig. 8 se muestra una imagen comprimida de Lena con una razón de compresión cercana a 10:1, a la izquierda se tiene el resultado de la compresión JPEG, en ella se puede observar como se revelan los bloques donde se aplica la DCT, mientras la imagen de la derecha no revela bordes ficticios.

Aunque los resultados de la compresión Wavelet son bastante buenos, el algoritmo tiene bastantes parámetros por fijar como son los niveles de resolución en que se va a descomponer la imagen, la base Wavelet a usar, el número de niveles de cuantización. Están son varias preguntas que requieren de un mayor estudio.

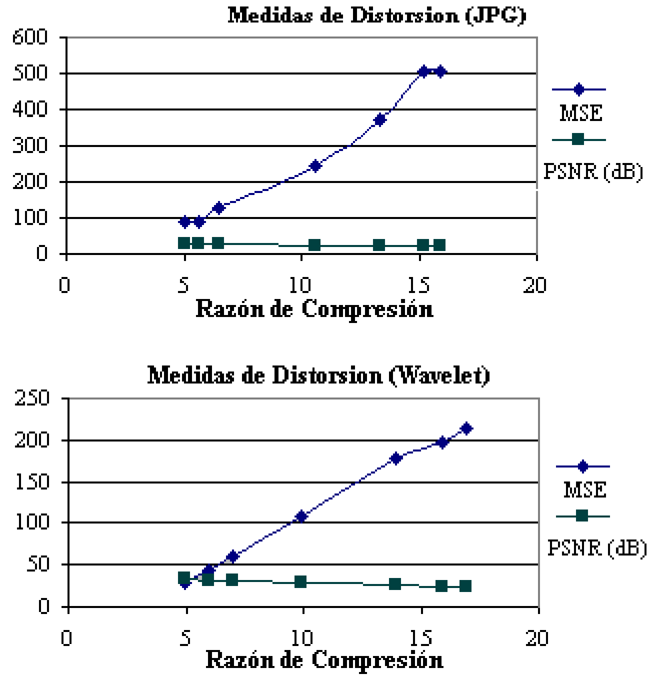

Figura 7. Medidas de distorsión, JPG (Arriba), Wavelet (Abajo)

Figura 8. Imagen Comprimida con razón 10:1. JPEG (Izquierda), Wavelet (Derecha).

VI. AGRADECIMIENTOS

El autor agradece el aporte de Haydi Perez y del profesor Jorge Ortiz de la Universidad Nacional, por su colaboración durante la elaboración de este proyecto.

VII. REFERENCIAS

[1] Antonini M, Barlaud M, Mathieu P, Daubechies I. *“Image Coding using wavelet transform”,* IEEE Trans. Image Processing. April 1992, pp. 205-220.

[2] Fournier A. *“Introduction to Wavelets, Signal Compression and Image Processing”,* Course Notes SIGGRAPH’1995: Wavelets and their Applications in Computer Graphics. pp. 5-33.

[3] Linde Y., Buzo A., y Gray R.M. "*An Algorithm for Vector Quantizer Design",* IEEE Trans on Communications, Jan. 1980, pp. 84 –95

[4] Mallat, S. "*A theory for multiresolution signal decomposition: The wavelet representation",* IEEE Trans Pattern Analysis Machine Intelligent. Vol. 11 July 1989.

[5] Witten IH, Neal RM, and Cleary JG. "*Arithmetic coding for data compression*. Comm. of the ACM", June 1987, pp. 520 – 540.

VIII. CURRICULUMS

Fernando Ruiz Vera (Student Member, IEEE) actualmente es estudiante de último semestre de Ingeniería Electrónica en la Universidad Distrital, Bogotá Colombia. Donde ocupa en este momento la presidencia del capitulo de computadores de la Rama Estudiantil del IEEE. Recibió su grado de Ingeniería de Sistemas en la Universidad Nacional de Colombia en 1999. Sus áreas de interés son principalmente el procesamiento digital de señales, en especial la compresión de imágenes y vídeo al igual que el estudio de algoritmos paralelos e informática evolutiva aplicada al procesamiento de estas.