

Uso de la Transformada de Wavelet para la compresión de imágenes

Funciones de Variable Compleja

Gonzalo Damian Acosta

Estudiante de Ingeniería en Sistemas de Computación
Universidad Nacional del Sur, Avda. Alem 1253, B8000CPB Bahía Blanca, Argentina
acosta.gonzalo@gmail.com
Marzo 2014

Resumen: Todos los días una enorme cantidad de información es almacenada, procesada y transmitida digitalmente. Debido a que un gran porcentaje de esa información son imágenes o de naturaleza pictórica, el costo de almacenamiento y comunicación requeridos para esto es inmenso. La compresión de imágenes aborda el problema intentando reducir al máximo la cantidad de datos requeridos para la representación de imágenes digitales. En esta nota de aplicación usaremos la transformada de Wavelet. En la compresión la transformada de Wavelet se utiliza para poder representar a la imagen en coeficientes que puedan ser codificados de manera más eficiente que los píxeles originales, incluso algunos pueden ser cuantificados toscamente o truncados.

Palabras clave: compresión de imagen, transformada de Wavelet, procesamiento de imágenes, DWT.

I. INTRODUCCIÓN

La función de las transformaciones en la compresión de imágenes es poder representar la misma con otros coeficientes. Como muchos de estos coeficientes calculados llevan una pequeña parte de la información visual, pueden ser cuantificados y codificados para evitar la redundancia entre coeficientes y códigos. Existen diferentes tipos de transformaciones utilizadas actualmente para la compresión de imagen, entre ellas están la transformada discreta de cosenos, y las diferentes transformadas de Wavelet: Haar, 5/3, 9/3, 9/7, 13/7, 5/11, etc. Se pueden realizar distintas técnicas de compresiones: con pérdida o sin pérdida. La compresión con pérdida puede ser usada en imágenes en donde los finos detalles pueden ser sacrificados para intentar reducir el ancho de banda utilizado o espacio en lugar de almacenamiento. Mientras que la compresión sin pérdida intenta reducir el tamaño de la imagen sin que se produzca ninguna distorsión en la imagen.

La transformación se basa en la premisa de que los píxeles que conforman la imagen muestran un grado de correlación entre vecinos. Como consecuencia, esta correlación puede ser explotada para predecir el valor del píxel a partir de sus vecinos. La transformación está definida para mapear estos datos espaciales en coeficientes (no relacionados). Así, el análisis de Wavelet se usa para dividir la información de la imagen en sub-señales de aproximación y detalles. En caso de que los detalles sean insignificantes estos pueden ser puestos a cero.

II. DESARROLLO DEL ARTÍCULO

La transformada de Fourier puede ser utilizada para traducir señales en dominio del tiempo a dominio de la frecuencia. Es decir que actúa como una especie de prisma matemático, descomponiendo las señales en tiempo en frecuencias, como un prisma lo hace con la luz. Esta transformada es útil para ver frecuencias difíciles de ver en el dominio del tiempo. Sin embargo no se adapta bien a señales breves, o que cambian repentinamente, es decir, señales no estacionarias. La razón es que muestra que señales ocurren y no cuando ocurren, y por lo tanto, no es tan útil cuando tanto el tiempo como la frecuencia son necesarios. Lo que se hizo para intentar solucionar esto es analizar la función en diferentes tamaños de ventana, de esta manera se obtiene las frecuencias en un periodo de tiempo. Sin embargo, aun así no se puede obtener el tiempo exacto para una determinada frecuencia.

La ventaja de Wavelet viene dado por el análisis multi-resolución. En vez de observar toda la señal a través de la misma ventana, diferentes partes de la señal se ven por diferentes tamaños de ventana. Para las altas frecuencias se usan ventanas más pequeñas para obtener mejor resolución temporal, mientras que para la parte

de las bajas frecuencias se usan ventanas más grandes para obtener una mayor información acerca de las frecuencias. Un aspecto a notar es que las ventanas tienen un área igual, aunque sean más anchas o largas

A. Sobre Wavelet

Sea ψ una función suficientemente regular y bien localizada. Esta función $\psi \in L^1 \cap L^2$, será llamada una wavelet si satisface la siguiente condición de admisibilidad en el dominio de frecuencia.

$$\int_{\mathbb{R}^+} \frac{|\nu(\omega)|^2}{|\omega|} d\omega = \int_{\mathbb{R}^-} \frac{|\nu(\omega)|^2}{|\omega|} d\omega < +\infty$$

Donde ν indica la transformada de Fourier de ψ . Esta condición implica, en particular, que la integral sea 0. Una condición suficiente de admisibilidad mucho más fácil de verificar es, para una wavelet ψ provista por

$$c_f(a, b) = \int_{\mathbb{R}} f(t) \overline{\psi_{a,b}(t)} dt \quad \psi, \psi \in L^1 \cap L^2, t\psi \in L^1 \text{ y } \int_{\mathbb{R}} \psi(t) dt = 0$$

Para consolidar la idea se podría decir que una wavelet oscila durante un determinado tiempo como una ola y luego se localiza debido a la amortización.

La transformada de Wavelet descompone la señal original en un conjunto de funciones básicas, estas funciones básicas son llamadas wavelets. Las wavelets son obtenidas de un prototipo wavelet $\psi(t)$, llamado Wavelet madre, mediante dilatación y corrimiento. Donde los parámetros a y b son elegidos de tal forma que $a=2^j$ y $b=k2^j$.

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad a \in \mathbb{R}^+, b \in \mathbb{R}$$

Donde a es el parámetro de dilatación y b el de corrimiento.

Para una función de energía infinita f podemos definir una transformada continua de Wavelet por la función C_f

$$C_f(a, b) = \int_{\mathbb{R}} f(t) \overline{\psi_{a,b}(t)} dt$$

La función f es luego descripta con sus coeficientes C_f de Wavelet, donde $a \in \mathbb{R}^+$ y $b \in \mathbb{R}$. Esto mide la fluctuación de f en una escala a .

B. Transformada Discreta de Wavelet (DWT)

Para poder calcular la transformada de Wavelet usando computadoras los datos deben ser discretos. La transformada discreta de Wavelet (DWT) provee la información suficiente para el análisis y síntesis de la señal, pero siendo más eficiente (para el cálculo). El análisis discreto de Wavelet se calcula utilizando el concepto de bancos de filtros, filtros con diferentes frecuencias de corte analizan la señal a diferentes escalas. Por ejemplo, en caso de que se trabaje con dos filtros, un pasa bajo y un pasa alto quedarían señales como las mostradas en la figura 1.

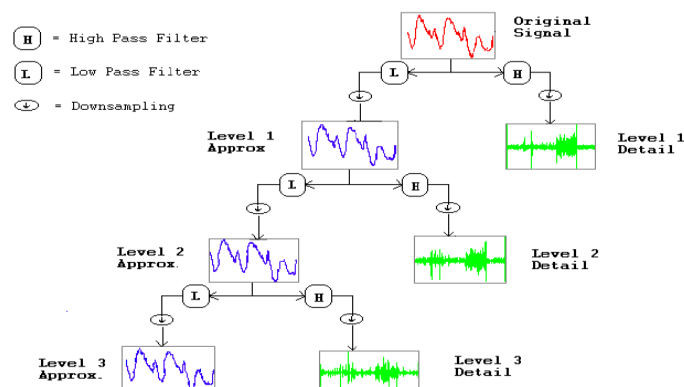


Figura 1

La DWT es obtenida a partir de todos los coeficientes de las sub-señales generadas. En combinación, los filtros, tienen el efecto de ir separando detalles cada vez más pequeños. Una propiedad importante del análisis de Wavelet es la conservación de la energía. La energía es definida como la suma de los cuadrados de los valores. Por ejemplo, la energía de una imagen es la suma de los cuadrados de los valores de los píxeles, la energía de la transformada de Wavelet de una imagen es la suma de los cuadrados de los coeficientes transformados. Sin embargo, durante la compresión, si hay una pérdida de energía ya que los umbrales cambian los valores de los coeficientes y así la versión comprimida contiene menos energía.

C. Aplicación

Las imágenes, son tratadas como señales 2D, las cuales cambian vertical y horizontalmente, por lo tanto debe ser aplicado un análisis de Wavelet 2D. Este tipo de análisis usa la misma wavelet madre pero requiere un paso extra en cada descomposición; primero se filtran los datos horizontalmente, luego, la aproximación y el detalle producido por esto, es filtrado en columnas como se muestra en la figura 2. En cada nivel cuatro sub-imágenes son obtenidas; la aproximación, los detalles horizontales, los detalles verticales y los detalles diagonales, como se muestra en la figura 3 de la imagen de Saturno. El análisis de Wavelet muestra como la imagen cambia verticalmente, horizontalmente y diagonalmente.

Los cinco pasos de la compresión y descompresión son mostrados en la figura 4 y 5. Todos los pasos de la compresión son invertibles (y por lo tanto sin pérdida), salvo por la cuantización. La cuantización se refiere a la reducción de la precisión de los valores de los números, en punto flotantes, de la transformada de Wavelet, y así usar menos cantidad de bits con que se representa cada uno. Esto lleva a un error de redondeo. La relación entre cuantización y codificación mostrada en la figura 1 es el principal aspecto de la compresión por Wavelet. Esta relación es aprovechada por diferentes algoritmos para poder comprimir aún más la imagen.

El principal beneficio de Wavelet es que produce una gran cantidad de valores que tiene ceros o magnitudes cercanas a cero. Por ejemplo, se considerará la imagen de la figura 6(a), Lena. En la figura 6(b) se muestra una transformada de Wavelet 9/7 de 7 niveles de la imagen Lena. Para esto se ha usado un umbral de 8, es decir que para todos los valores cuya magnitud eran menos de 8 se llevaron a 0. En caso de que se aplique una transformada inversa de Wavelet se obtiene la imagen 6(c) (después de redondear a valores enteros entre 0 y 255). Es difícil encontrar diferencias entre las imágenes de las figuras 6(a) y 6(c). La imagen de la figura 6(c) fue producida usando 32.498 valores no nulos en vez de los 262.144 de la original. Esto representa un ratio de compresión de 8:1.

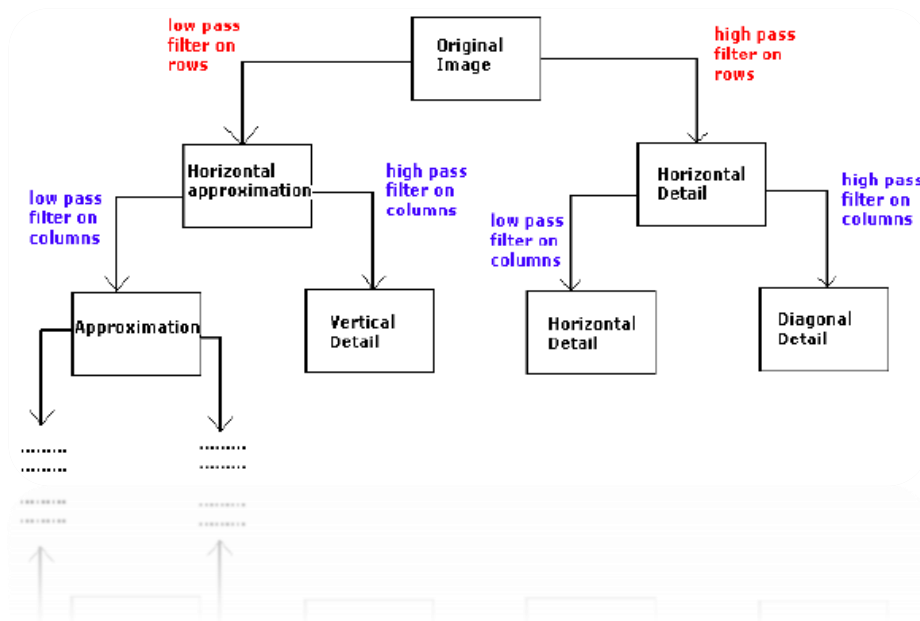


Figura 2

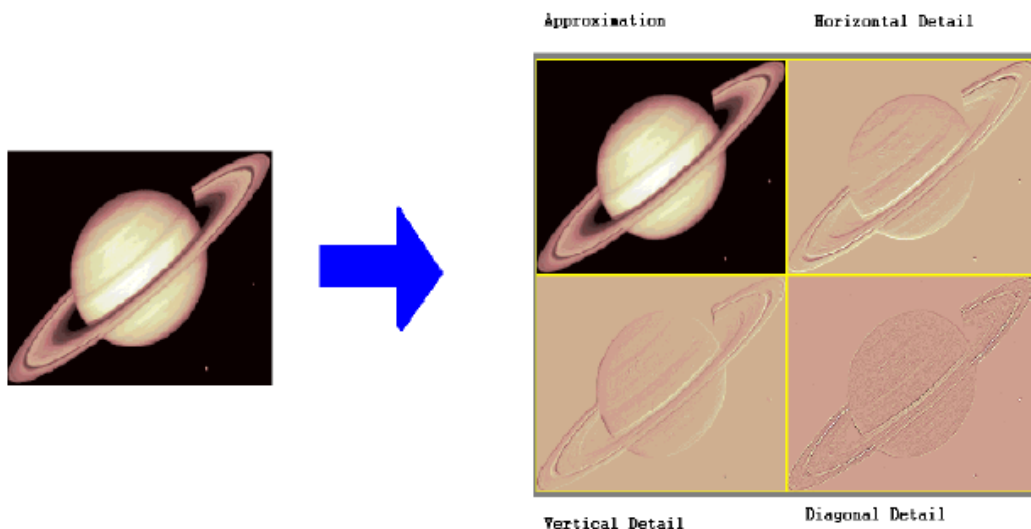


Figura 3

Las medidas de error de cuantificación más usadas son el error medio cuadrado (MSE) y la Relación Señal a Ruido de Pico (PSNR). La MSE entre dos imágenes f y g está definida por:

$$MSE = \frac{1}{N} \sum_{j,k} (f[j,k] - g[j,k])^2$$

donde la sumatoria entre j,k denota la suma de todos los pixeles de la imagen, y N es el número de pixeles totales. Para el caso del ejemplo el MSE es de 7,921. El PSNR entre las dos, en decibeles, es:

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right)$$

Por lo general se usa el PSNR debido a que es logarítmico y, al parecer, nuestro cerebro responde de esta forma a la intensidad. Para el caso del ejemplo el PSNR es de 39,14dB. Generalmente, cuando el PSNR es 40dB o mayor las imágenes son visualmente indistinguibles para el humano.

En general, el algoritmo base de compresión, consta de los dos siguientes pasos:

Primero, la posición de los valores transformados -aquellos que tienen un valor mayor que el umbral T - son determinados escaneando con la transformada. La posición de los valores significativos es codificada.

Segundo, los valores significativos de la transformada son codificados.

El algoritmo base fue uno de los primeros en utilizar Wavelets, sufre de varios defectos que luego otros algoritmos remediaron.

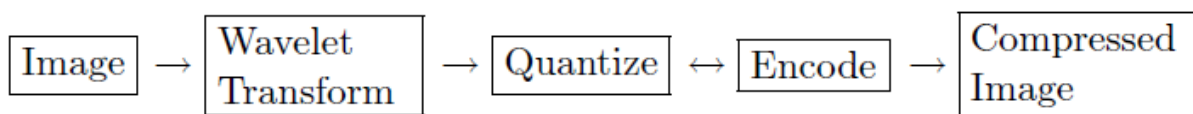


Figura 4. Compresión de imágenes

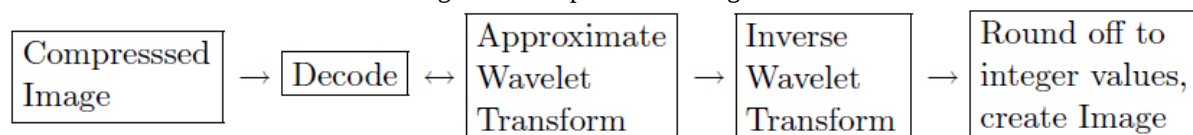


Figura 5. Descompresión de imágenes

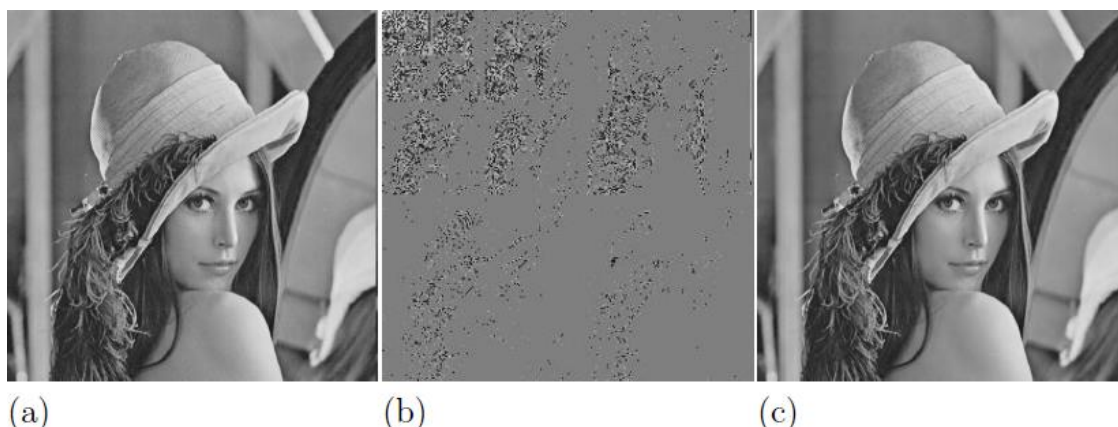


Figura 6: (a) Imagen Lena sin compresión. (b) Transformada de Wavelet de Lena. (c) Inversa de la imagen Lena transformada

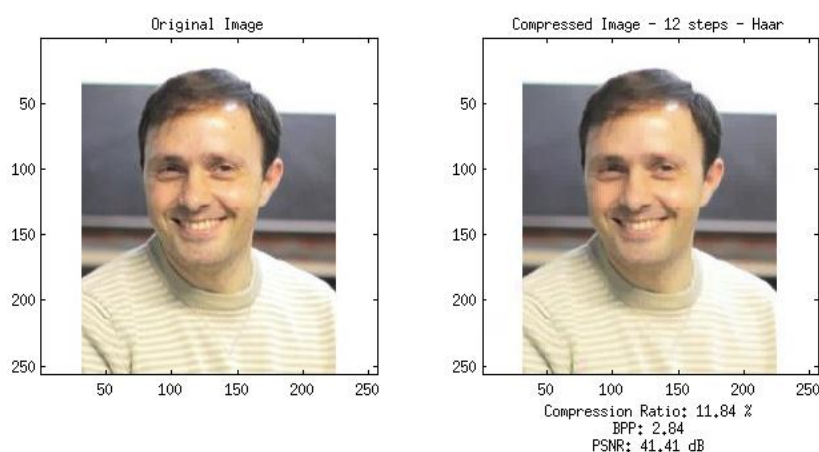


Figura 7. Imagen comprimida del Prof. Dr. Guillermo Calandrini

III. CONCLUSIÓN

El uso de la transformada de Wavelet permite optimizar el espacio de almacenamiento y ancho de banda en transferencia utilizado por las imágenes, dando la posibilidad de optimizarlo sin la necesidad de cambiar la calidad percibida por el ojo humano. Además, gracias a sus propiedades, permite obtener coeficientes más fáciles de codificar para los algoritmos utilizados actualmente. Hoy en día es utilizada en el formato más óptimo de compresión que es el JPEG 2000.

REFERENCIAS

- [1] James S.Walker, "Wavelet-based Image Compression", Sub-chapter of CRC Press book: *Transform and Data Compression*.
- [2] R. J. E Merry, "Wavelet Theory and Applications: A literature Study", DCT 2005, Eindhoven, Junio 2005.
- [3] Michel Misiti, Yves Misiti, George Oppenheim, Jean-Michel Poggi, "Wavelets and Their Applications".
- [4] Karen Lees, "Imagen Compression Using Wavelets", Mayo 2002.
- [5] Eric J. Stollnitz, Tony D. DeRose, and David H. Salesin. "Wavelets for computer graphics: A primer, part 1". *IEEE Computer Graphics and Applications*, 15(3):76-84, May 1995.
- [6] Wikipedia, *La enciclopedia libre*, [internet], disponible en <http://en.wikipedia.org/wiki>, [acceso el 28 de febrero de 2014].