

# Influencia de la eliminación de los coeficientes de la transformada coseno en la representación de imágenes digitales

E. Galarza.

Escuela Politécnica del Ejército Sede Latacunga. Latacunga, Cotopaxi, Ecuador

RESUMEN: El trabajo que se detalla a continuación corresponde al análisis que se realiza sobre la eliminación de los coeficientes reales que se obtienen a partir de la transformada coseno así como la determinación de la influencia en la calidad de las imágenes que se obtienen de esta manera. El trabajo realizado analiza la imagen como un bloque único, mas no en estructuras de 8 x 8 píxeles que es la forma en la que trabajan los sistemas de compresión de imágenes, específicamente en el formato JPEG.

Como resultado puede visualizarse las imágenes con diferentes grados de compresión, determinando, a partir de las mismas, los niveles máximos a los que se podría comprimir una imagen sin que la persona que visualiza determine una calidad que no valga la pena ser observada.

Como parte del trabajo, también se realiza un análisis del marco teórico relativo a los principios de procesamiento digital de señales, características de las imágenes, así como los principios que determinan la operación de la transformada coseno, posteriormente se incluye la metodología utilizada, los resultados obtenidos y se finaliza con las conclusiones obtenidas a partir del presente trabajo.

## 1. INTRODUCCIÓN AL PROCESAMIENTO DIGITAL DE IMÁGENES

El procesamiento digital de señales se distingue de otras áreas en las ciencias computacionales debido a que utiliza un único tipo de dato que son las señales. En la mayoría de los casos, estas señales se originan como datos que se obtienen mediante el sensado en el mundo real. Las vibraciones sísmicas, las ondas de sonido, las imágenes, son ejemplo que pueden considerarse. [Smith 2003]

Estas señales pueden describir una amplia variedad de fenómenos físicos, pueden encontrarse de diferentes formas, pero en la mayoría de los casos están contenidas en un patrón de variaciones que representa alguna forma determinada. Las señales se representan matemáticamente como funciones con una o más variables independientes en las que, en la mayoría de los casos, se utiliza al tiempo como la variable que no depende de otra variable y que se la utiliza como referencia [Oppenheim 2000].

Una imagen puede ser definida como una función bidimensional  $f(x,y)$ , en donde  $x$  y  $y$  son coordenadas espaciales y la amplitud de  $f$  en cualquier par de coordenadas se le conoce como intensidad o nivel de gris (inclusive para las imágenes a colores). Cuando  $x$ ,  $y$ , y los valores de

la amplitud de  $f$  son finitos y al mismo tiempo valores discretos, se le conoce como imagen o imagen digital. El campo del procesamiento digital de las imágenes se refiere al tratamiento de la información que representa las imágenes utilizando computadores digitales. [Gonzales - Wood 2001]

Un punto en el sistema tridimensional se define como un vector columna,  $P = (X Y Z)^t$  y la proyección de este punto sobre el plano bidimensional de las imágenes se denota por  $p=(x y)^t$ , debe considerarse que el punto genérico y el punto de la imagen se representan con respecto a su propio sistema de coordenadas, y por conveniencia, el sistema de coordenadas de la imagen se elige de manera que sea ortogonal al eje de coordenadas de  $Z$  [Farid]. Los orígenes de los dos sistemas se relacionan mediante un cambio unidimensional a lo largo del eje  $Z$  o también denominado eje óptico.

Para adquirir las imágenes, se utiliza una cámara que es un dispositivo que adquiere imágenes fijas o en movimiento. Las cámaras pueden trabajar con la luz del espectro visible o con otras frecuencias del espectro electromagnético. Una cámara, en forma general, consiste de un sistema vacío con un sistema de apertura por un lado para que ingrese la luz y de una superficie para registro que puede capturar la luz en el otro lado. La mayoría de las cámaras cuentan con lentes que se colocan al frente de la cámara para capturar la luz que ingresa y enfocar todos los elementos de la imagen en la superficie que registra a la misma [Wikipedia]. El diámetro de la apertura se controla normalmente con un mecanismo de diafragma pero otras cuentan de una apertura de tamaño fijo.

Las imágenes de la actualidad utilizan el formato de compresión JPEG, el cual utiliza sofisticadas técnicas matemáticas en la que se incluye la transformada discreta coseno para producir una escala ajustable de la compresión gráfica. Puede elegirse el grado de compresión deseado a ser aplicado a la imagen, determinando de esta manera la calidad que se obtiene [graphicsoft]. Mientras más se comprima la imagen, mayor será la degradación de la misma. Con JPEG se puede conseguir altas relaciones de compresión en valores que superan los cientos de veces con respecto al archivo original. Esto es posible lograr debido a que los algoritmos de JPEG descartan los datos que no son necesarios mientras realiza el proceso de compresión, por lo que se denomina como técnica de compresión con pérdidas.

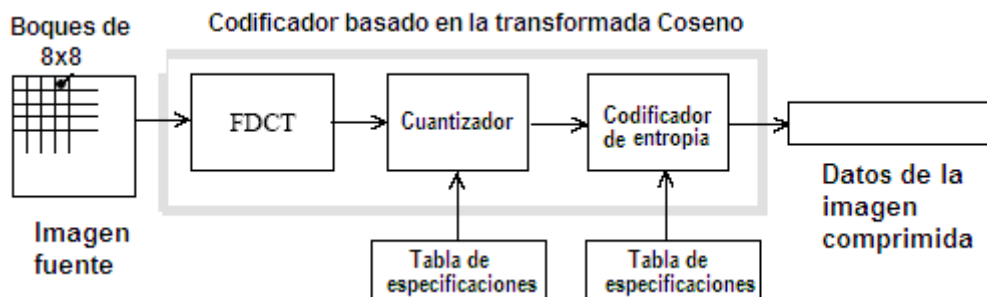


Figura 1: Sistema de codificación para imágenes en formato JPEG.

La figura 1 presenta los pasos para la codificación basada en la transformada coseno, los cuales son los pasos claves para la operación del sistema. Esta figura ilustra el caso para un componente único de la compresión de la imagen (escala de grises). La compresión se realiza aplicando el sistema para bloques de 8x8 píxeles de muestras de la imagen. La compresión de las imágenes en color puede ser obtenida mediante la compresión de múltiples imágenes representadas en la escala de gris, las cuales se comprimen de forma total o una a la vez.

Para los codificadores de modo secuencial, los cuales incluyen a los codificadores de línea base, el diagrama simplificado expresa como los sencillos componentes de la compresión trabajan en una forma adecuada que además es completa. Cada bloque de 8x8 que es ingresado, pasa por cada elemento del sistema de procesamiento y establece una salida en forma comprimida en los datos de la salida [Wallace 1991]. Para los codificadores de la transformada

discreta coseno un “buffer” de la imagen se encuentra disponible luego del codificador de entropía de tal manera que la imagen puede ser almacenada y luego puesta a disposición de acuerdo a la calidad que se desea obtener.

La representación de Fourier de las señales, desempeña un rol muy importante en el procesamiento tanto discreto como continuo. Determina un método para transportar las señales a un dominio diferente en el cual puedan ser manipuladas. Lo que hace de la transformada de Fourier extremadamente útil, es la propiedad de que la operación de la convolución en señales en el dominio del tiempo, se convierte en una multiplicación en el dominio de la frecuencia. Adicionalmente, la transformada de Fourier presenta una forma diferente de interpretar las señales y los sistemas.

La transformada de Fourier discreta en el tiempo, (DTFT) de una secuencia  $x(n)$ , que interesa para la representación de las señales digitales que han sido muestreadas a partir de un proceso real, se define la siguiente manera:

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n} \quad (1)$$

Por lo que, a partir de esta ecuación se determina que la respuesta de frecuencia de un sistema lineal invariante en el tiempo,  $H(e^{j\omega})$ , es la DTFT de la respuesta al muestreo unitario  $h(n)$ . Para que la DTFT de una secuencia exista, la sumatoria de la ecuación 1 debe converger, lo que determina que  $x(n)$  sea absolutamente sumable, es decir [Hayes 1999]

$$\sum_{n=-\infty}^{\infty} |x(n)| = S < \infty \quad (2)$$

Una desventaja de la transformada de Fourier para algunas aplicaciones es que la transformada  $X$  es un valor complejo inclusive para el caso de datos completamente reales. Una transformada relacionada, es la transformada discreta coseno (DCT), la cual no tiene el problema de la transformada de Fourier y presenta datos únicamente reales. La DCT es una transformada diferente mas no la parte real de la DFT. Esta transformada se utiliza de una manera extensiva en aplicaciones de compresión de datos de imágenes y video como en el caso de JPEG y MPEG. También se utiliza la DCT en el filtrado que utiliza formas diferentes de convolución y que se denomina convolución simétrica [Woods 2006].

La transformada coseno bidimensional que se aplica en el caso de las imágenes se define de la siguiente manera:

$$X_C(k_1, k_2) = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} 4x(n_1, n_2) \cos \frac{\pi k_1}{2N_1} (2n_1 + 1) \cos \frac{\pi k_2}{2N_2} (2n_2 + 1) \quad (3)$$

El cual es un operador separable que se obtiene a partir de la aplicación unidimensional de la DCT.

## 2. METODOLOGÍA

Para la ejecución del presente trabajo, se procedió a elaborar la fotografía en blanco y negro a partir de una imagen a color, la cual no se obtuvo mediante los programas que se utilizan normalmente en el tratamiento de fotografías, sino mediante procesamiento digital utilizando el programa MATLAB<sup>1</sup>. Con la foto base, posteriormente se procedió a realizar el análisis

---

<sup>1</sup> Propiedad de Mathworks

completo, determinando componentes de píxeles, de información horizontal y vertical así como las respuestas de frecuencia en el dominio real utilizando la transformada coseno. Finalmente se procedió a desarrollar un programa (nuevamente en MATLAB), el cual utilizaba la transformada coseno y su inversa para la recuperación de la imagen original. En el procedimiento de recuperación se reemplazaban varios valores de coeficientes con el valor “0”, lo cual se realiza en el trabajo presentado en este artículo.

La imagen utilizada para realizar el trabajo es la que se indica a continuación:



Figura 2: Imagen utilizada para el procesamiento digital (Gaby)

### 3. RESULTADOS

Los resultados obtenidos se presentan a continuación. Cada resultado expresa las características de la imagen tomada en consideración.

En la figura 3, se puede observar la representación de aproximadamente 70 píxeles de una línea horizontal seleccionados aleatoriamente, mientras que en la figura 4 se puede observar la representación de también 70 píxeles pero para el caso de una línea vertical, la cual fue también seleccionada en forma aleatoria.

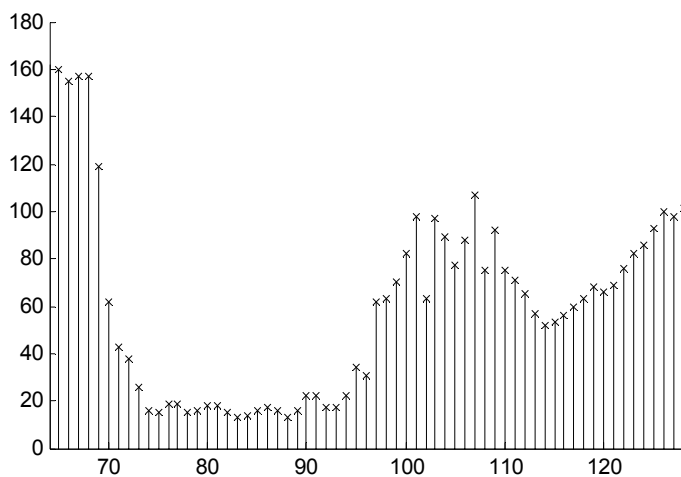


Figura 3. Representación digital de los píxeles de una línea horizontal elegida aleatoriamente

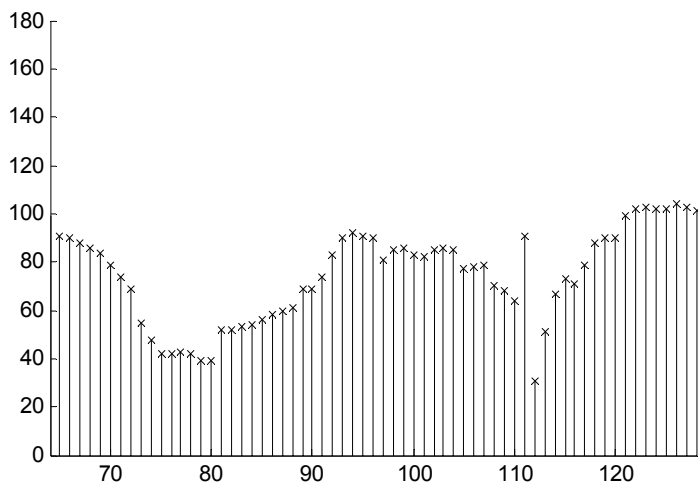


Figura 4. Representación digital de los pixeles de una línea vertical elegida aleatoriamente

Utilizando la transformada coseno, se obtienen los coeficientes de los componentes de frecuencia, y los resultados se pueden observar en la figura 5, en donde se aprecia el nivel porcentual de un grupo de cada uno de los elementos de la matriz de 256 x 256 coeficientes. En esta figura se puede observar la preponderancia del componente de la primera fila y primera columna sobre todos los 65535 coeficientes adicionales.

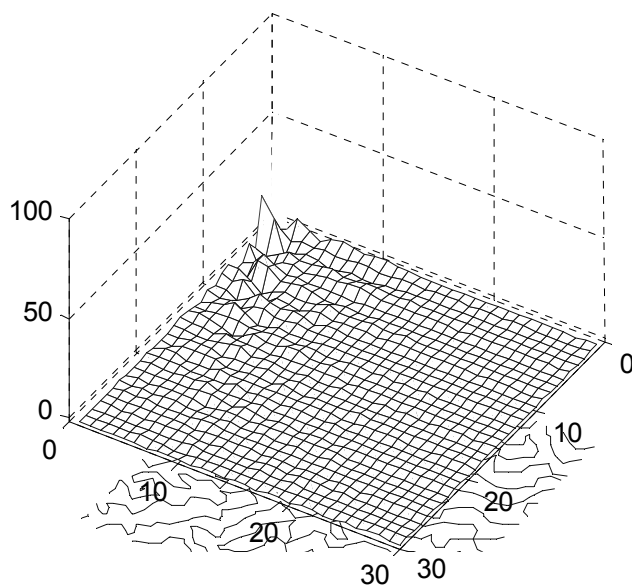


Figura 5. Componentes de frecuencia real obtenidos utilizando la transformada coseno.

Los valores obtenidos y que fueron representados en forma gráfica en la figura 5, son los que se presentan en la tabla 1, expresados en forma porcentual. Cabe indicar que únicamente se presentan 10 x 10 elementos de los 256 x 256 posibles. Puede notarse que la influencia de las bajas frecuencias es muy superior a las de las altas frecuencias (ubicación inferior derecha). La transformada puede incluir valores positivos o negativos, pero siempre reales.

Tabla 1: 10 x 10 coeficientes de la transformada coseno de la imagen "Gaby"

	1	2	3	4	5	6	7	8	9	10
1	100,0	1,0	20,3	-2,4	3,3	-3,3	-3,3	3,5	2,3	-0,6
2	-1,6	2,0	10,9	-4,1	-7,3	-0,2	-4,5	3,0	0,5	-0,7
3	6,0	-2,0	-6,1	-1,1	-7,4	6,2	5,9	-1,2	3,2	-3,7
4	1,9	0,0	-2,7	1,5	0,0	2,3	5,4	-5,5	-2,8	-0,4
5	-0,6	-0,3	-0,2	-0,1	1,5	-1,0	-0,1	-2,0	-3,7	4,2
6	1,8	0,2	-1,5	-0,1	-0,5	0,7	0,5	0,0	-0,1	2,0
7	3,6	-0,5	-3,8	0,6	0,1	-0,1	1,9	-1,1	-0,8	0,5
8	1,6	-0,8	-2,3	1,5	1,7	-0,2	0,5	-1,1	-1,9	1,2
9	0,8	-0,1	-1,5	0,1	1,5	-0,5	-1,6	0,0	0,4	1,0
10	1,9	0,4	-2,3	0,0	1,9	-1,0	-1,3	1,8	1,1	-1,0

Los resultados correspondientes se presentan en las figuras 6, 7, 8, 9, 10 y 11 pudiéndose notar la influencia de la eliminación de coeficientes en la reducción de imágenes.



Figura 6. Imagen con reducción del 0% (256x256)



Figura 7. Imagen con reducción del 43,75 (192x192)



Figura 8. Imagen con reducción del 75% (128x128)



Figura 9. Imagen con reducción del 93,75% (64x64)



Figura 10. Imagen con reducción del 98,44 (32x32) y representación en menor tamaño

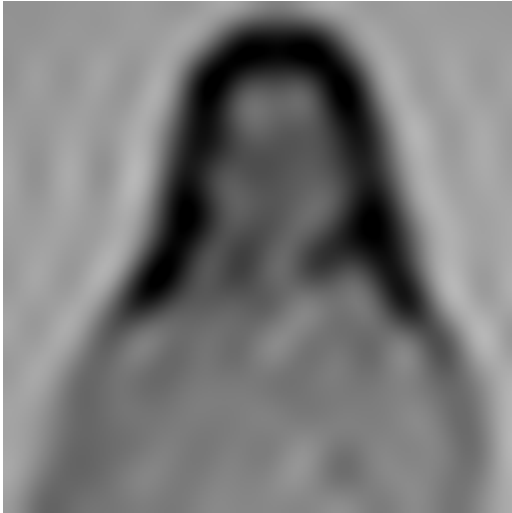


Figura 11. Imagen con reducción del 99,61 % (16x16) y representación en menor tamaño

#### 4. CONCLUSIONES

La transformada coseno bidimensional representa los componentes de frecuencia de una imagen, en una relación 1 a 1 con respecto a los niveles de intensidad de gris de una imagen, permitiendo tener ventajas sobre transformaciones como la de Fourier que requiere de una representación utilizando componentes complejos que no permiten producir un adecuado nivel de compresión de las imágenes.

Al utilizar la transformada coseno para la representación de una imagen, se puede obtener resultados visuales bastante aceptables, escogiendo únicamente los coeficientes más importantes de lo obtenido, que puede producir resultados en el orden de reducción del 75 % sin producirse mayores problemas de visualización.

#### 5. BIBLIOGRAFÍA

- Smith S. 2003. Digital Signal Processing. Newnes. 1 – 5. Burlington: Massachusetts
- Oppheim A. Willsky A. 1998. Señales y Sistemas. 2da. Ed. Pearson Education. 1 – 12. México: México
- Gonzales R. – Woods R. 2001. Digital Image Processing. Prentice Hall, New Jersey.
- Wallace G. 1991. The JPEG Still Picture Compression Standard. Multimedia Engineering Digital Equipment Corporation. 1 – 6. Maynard: Massachusetts, for publication in IEEE Transactions on Consumer Electronics.
- Monson H. 1999. Digital Signal Processing, McGraw Hill, New York.
- Woods J. 2006. Multidimensional signal, image and video processing and coding. Academic Press, Burlington: Massachusetts, 2006.

#### Enlaces:

Hany. Farid: <http://www.cs.dartmouth.edu/~farid>

<http://en.wikipedia.org/wiki/Camera>

[http://graphicssoft.about.com/gi/dynamic/offsite.htm?zi=1/XJ/Ya&sdn=graphicssoft&cdn=compute&tm=40&gps=68\\_7\\_1428\\_648&f=10&su=p284.8.150.ip\\_&tt=14&bt=1&bts=1&zu=http%3A//www.webstyleguide.com/graphics/jpegs.html](http://graphicssoft.about.com/gi/dynamic/offsite.htm?zi=1/XJ/Ya&sdn=graphicssoft&cdn=compute&tm=40&gps=68_7_1428_648&f=10&su=p284.8.150.ip_&tt=14&bt=1&bts=1&zu=http%3A//www.webstyleguide.com/graphics/jpegs.html)