

CAPÍTULO III

COMPACTACIÓN DE IMÁGENES BASADO EN LA DCT.

3.1 Introducción.

En este capítulo se presenta la transformada coseno discreta (DCT) desde el punto de vista bidimensional, donde se ven las características relevantes que le permite ser útil en la compresión de imágenes. Finalmente se presenta la aplicación de la DCT a un caso particular, el estándar JPEG (Joint Photographics Experts Group)(Grupo conjunto de Expertos Fotográficos) que ha dado como resultado el primer estándar internacional para compresión de imágenes digitales, tanto en la escalas de grises como en color.

3.2 Transformada coseno discreta (DCT).

La transformada coseno discreta es un método preferido para la compresión de imágenes, para el caso bidimensional el par transformado queda definido [1] como:

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad (3.1)$$

para $u, v = 0, 1, 2, \dots, N-1$, y

$$f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u, v) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad (3.2)$$

para $x, y = 0, 1, 2, \dots, N-1$

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & \text{para } u = 0 \\ \sqrt{\frac{2}{N}} & \text{para } u = 1, 2, \dots, N-1 \end{cases} \quad (3.3)$$

$$\alpha(v) = \begin{cases} \sqrt{\frac{1}{N}} & \text{para } v = 0 \\ \sqrt{\frac{2}{N}} & \text{para } v = 1, 2, \dots, N-1 \end{cases} \quad (3.4)$$

Como ejemplo, se aplica la DCT a la subimagen 8×8 que se presenta en la figura 3.1.

170	153	170	153	153	153	153	153
170	153	170	153	153	153	153	170
170	153	170	153	153	153	170	153
153	153	170	153	153	153	153	153
153	153	153	187	153	153	153	170
170	136	153	153	153	153	153	153
153	153	153	170	153	153	170	153
153	153	153	153	170	170	170	153

Figura 3.1 Subimagen 8×8

El resultado de este proceso se muestra en la figura 3.2

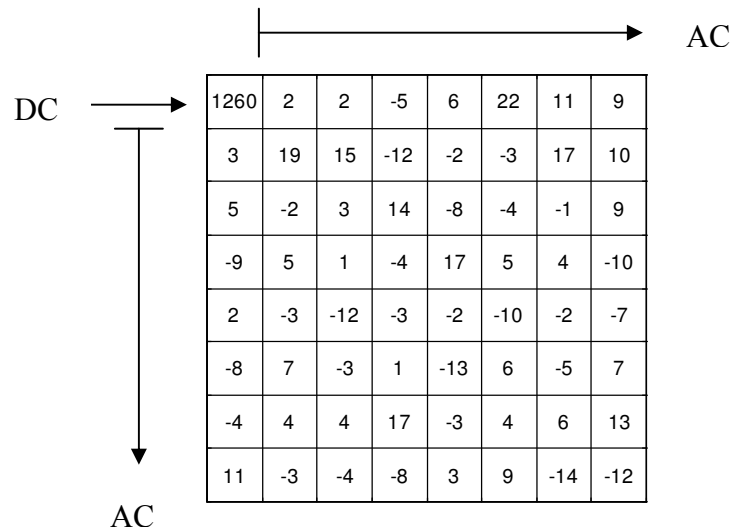


Figura 3.2. Resultado de la aplicación DCT a la subimagen de la figura 3.1

En la figura 3.2 se puede observar que el primer coeficiente, tiene el valor mas alto. En este coeficiente es concentrada la mayor parte de la energía, que está aproximadamente distribuida de manera uniforme en la subimagen de la figura 3.1. Además, este coeficiente, es el valor promedio de todos los coeficientes originales de la subimagen, llamado valor DC. Los coeficientes de la figura 3.2, excepto el primero, son los coeficientes AC resultantes.

Como puede observarse, los coeficientes AC que resultan tienen valores pequeños, por los que se pueden eliminar algunos, empezando por la parte inferior derecha, que es donde se encuentran los coeficientes de mas alta frecuencia. Estos pueden eliminarse y no afectar significativamente en la reconstrucción de la subimagen (esta eliminación depende del error que se quiera tolerar en función de la calidad de la imagen reconstruida).

En ejemplo presentado, la subimagen tiene pocos cambios de amplitud por lo que los coeficientes de altas frecuencias tienden a ser mas pequeños o bien ser cero, los detalles de la imagen son menos.

Por otro lado si los cambios son abruptos en las amplitudes de los pixels, los componentes de alta frecuencia tienden a ser mas grandes y mayores, por lo que existe un mayor detalle en la imagen.

Desde luego, la sola aplicación de la DCT, puede ser una forma de almacenar menos coeficientes(logrando compresión), con la eliminación de algunos de ellos, siempre que no exista tanta variación en la amplitud de los coeficientes de la imagen original. Si hay variaciones abruptas, la eliminación puede producir que la imagen reconstruida no sea de buena calidad.

3.2 El estándar de compresión de imágenes estáticas JPEG.

3.2.1 Introducción.

Anteriormente se vio, que mediante la aplicación de la DCT, es posible lograr alguna compresión de la imagen, bajo ciertas condiciones, que no afecten significativamente la calidad de la imagen.

Durante los últimos años un esfuerzo de estandarización conocido con el acrónimo de JPEG dio como resultado el primer estándar internacional para compresión de imágenes digitales de imágenes estáticas de tonalidad continua, tanto en la escala de grises como en color. El “Conjunto” en JPEG se refiere a la colaboración entre el Comité Consultivo de la

Telefonía y Telegrafía Internacional (CCITT, del inglés: *Consultative Committee of the International Telephone and Telegraph*) y la Organización Internacional para la Estandarización (ISO, del inglés: *International Standardization Organization*). De esta manera JPEG, es un estándar ISO y una recomendación del CCITT.

3.2.2 Pasos del proceso para la codificación basada en la DCT.

Las figuras 3.3 y 3.4 muestran los principales pasos en la codificación / decodificación basada en la DCT [4]. Las figuras ilustran el caso especial de la compresión de una imagen en la escala de grises.

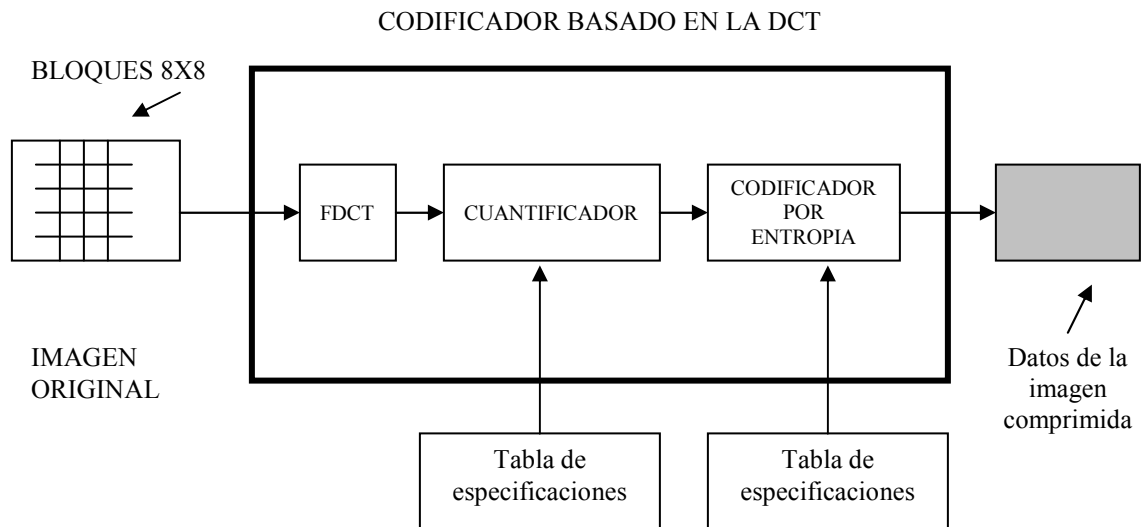


Figura 3.3 Pasos del proceso de codificación basado en la DCT.

La compresión de una imagen en escala de grises, puede pensarse como la compresión de una secuencia de bloques 8×8 que forman la imagen muestra. La compresión de imagen puede ser considerada, aproximadamente, como la compresión de múltiples imágenes en la escala de grises, las cuales son comprimidas una a la vez o son comprimidas alternativamente entrelazando los bloques 8×8 de cada una.

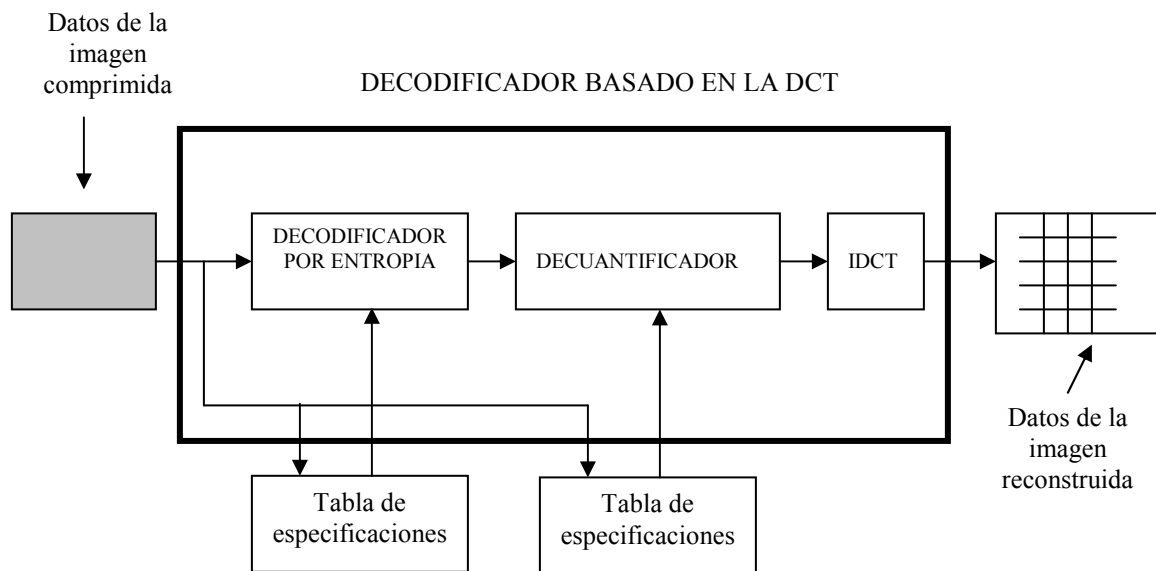


Figura 3.4 Pasos del proceso de decodificación basado en la DCT.

3.2.2.1 La FDCT e IDCT en los bloques 8×8.

En la entrada a el codificador, la imagen original se agrupa en bloque 8×8, se cambian de enteros sin signo con rango de $[0, 2^P - 1]$ a enteros con signo con rango $[-2^{P-1}, 2^{P-1} - 1]$ y se da como entrada a la DCT directa (FDCT). En la salida del decodificador, la DCT inversa devuelve los bloques 8×8 para formar la imagen reconstruida. Se utilizan las ecuaciones de la 3.1 a la 3.4 con $N = 8$.

La FDCT toma cada bloque 8×8 y lo descompone en 64 señales de una base ortogonal. Cada una de estas señales, contiene una de las 64 “frecuencias espaciales” únicas de dos dimensiones (2D) que comprenden el “espectro” de la señal de entrada. La salida de la FDCT es el conjunto de 64 amplitudes de la señal base o “coeficientes de la DCT” cuyos valores son únicamente determinados por la señal particular de entrada de 64 puntos. Al coeficiente con frecuencia cero en dos dimensiones en las dos dimensiones se le llama “coeficiente DC” y a los 63 coeficientes restantes, se les llama “coeficientes AC”

Debido a que los valores de la imagen, por lo común varían lentamente entre pixels consecutivos, el proceso de la FDCT, logra la compresión de los datos, concentrando la mayor parte de la señal en las frecuencias espaciales mas bajas. En un bloque 8x8 de una imagen típica, la mayoría de las frecuencias espaciales tienen una amplitud de cero o cercana a cero y no serán codificadas.

En el decodificador la IDCT realiza el proceso inverso. Toma los 64 coeficientes de la DCT (los cuales en ese punto ya han sido cuantificados) y reconstruye una imagen de salida de 64 punto sumando las señales base. Si la FDCT y la IDCT pudieran ser calculadas con una precisión perfecta y los coeficientes de la DCT no se cuantificaran , se podría reconstruir

exactamente la imagen original. En principio , la DCT no introduce pérdidas en la imagen original; simplemente la transforma a un dominio en la que los coeficientes pueden ser codificados mas eficientemente.

3.2.2.2 Cuantificación.

Después de salir de la FDCT, cada uno de los 64 coeficientes DCT se cuantifican uniformemente en conjunción con una tabla de cuantificación de 64 elementos, la cual debe ser especificada por la aplicación (o usuario) como entrada al codificador. Cada elemento de la tabla de cuantificación puede ser cualquier valor entero entre 1 y 255, que especifica el tamaño del paso para el cuantificador para su correspondiente coeficiente DCT. El propósito de la cuantificación es lograr una compresión mayor, representando los coeficientes DCT sin más precisión que la necesaria para lograr la calidad de la imagen deseada. Expresado de otra forma, el objetivo es este paso del procesamiento es descartar información que no es visualmente significativa. La cuantificación es fundamentalmente sin pérdidas y es la principal fuente de pérdidas en los codificadores basados en la DCT.

La cuantificación se define como una división de cada coeficiente DCT por su correspondiente tamaño de paso de cuantificación, seguido por un redondeo al entero mas cercano:

$$F^Q(u, v) = \text{Redondeo} \left(\frac{F(u, v)}{Q(u, v)} \right) \quad (3.5)$$

El valor de salida esta normalizado por el tamaño del paso de cuantificación. La decuantificación es la función inversa, que en este caso, significa simplemente, que la normalización es eliminada multiplicando por el tamaño del paso, lo que devuelve el resultado a una representación apropiada para la entrada de la IDCT:

$$F^{\mathcal{Q}}(u, v) = F^{\mathcal{Q}}(u, v) * Q(u, v) \quad (3.6)$$

Cuando el objetivo es comprimir la imagen tanto como sea posible sin artificios visibles, cada tamaño del paso idealmente debe ser escogido como el umbral porcentual para la contribución de su función base coseno correspondiente. Estos umbrales son también funciones de las características de la imagen original, de las características del visualizador y de la distancia de visión. Para aplicaciones en las que estas variables pueden estar razonablemente bien definidas, los experimentos psicovisuales se pueden realizar para determinar los mejores umbrales.

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Tabla 3.1 Tabla de la cuantificación de la luminancia propuesta por ISO

La tabla 3.1 ha sido usada experimentalmente por los miembros de JPEG y apareció en el estándar ISO como informativa, no como un requerimiento, para la cuantificación de la luminancia.

3.2.2.3 Codificación DC y secuencia zig-zag

Después de la cuantificación, el coeficiente DC se trata separadamente de los 63 coeficientes AC. El coeficiente DC es una medida del valor medio de los 64 pixels de la imagen. Debido a que normalmente hay una fuerte correlación entre los coeficientes DC de los bloques adyacentes de 8×8 , los coeficiente DC cuantificados se codifican como la diferencia entre el término DC del bloque previo en el orden de codificación, como se muestra en la figura 3.5.

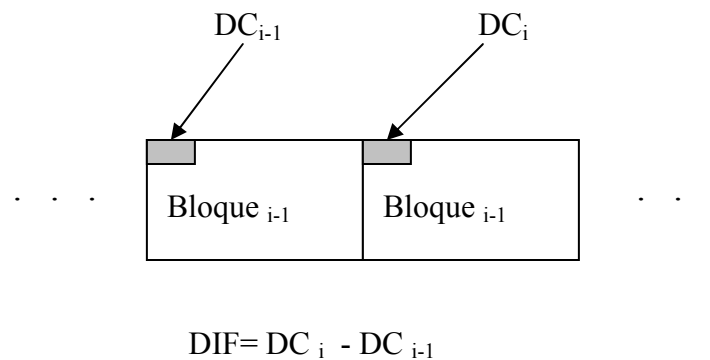


Figura 3.5 Codificación diferencial de los coeficientes DC

Este tratamiento especial vale la pena, puesto que los coeficientes DC, frecuentemente contienen una fracción significativa de la energía total de la imagen.

Finalmente, todos los coeficientes cuantificados son ordenados en una secuencia de zig-zag, como se muestra en la figura 3.6. Esta ordenación ayuda a facilitar la codificación por entropía colocando los coeficientes de baja frecuencia (que son mas probable a no ser cero) antes de los coeficientes de alta frecuencia.

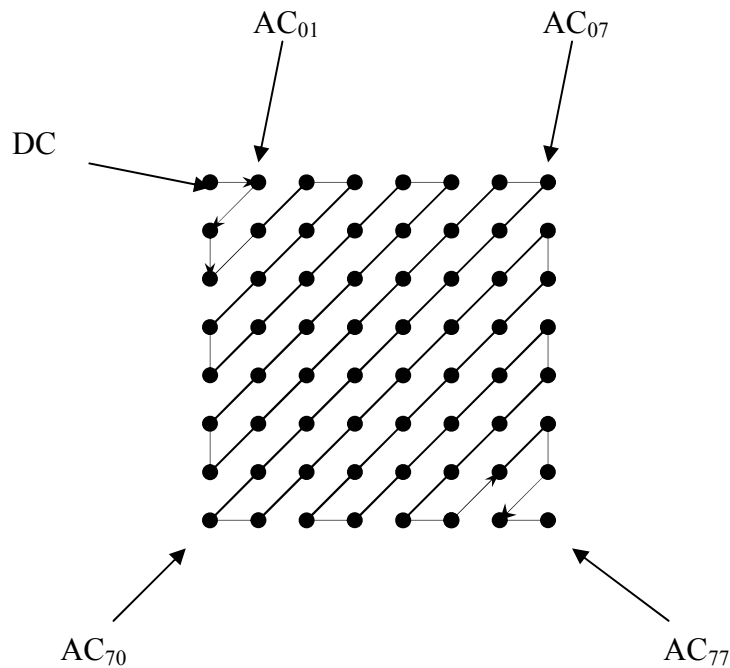


Figura 3.6 Secuencia zig-zag para la preparación de los coeficientes cuantificados

3.2.2.4 Codificación por entropía.

El paso final de la codificación basada en la DCT es la codificación por entropía. Este paso logra una compresión sin pérdida adicional, codificando los coeficientes cuantificados de la DCT mas compactamente, basándose en sus características estadísticas. La propuesta de JPEG especifica dos métodos de codificación por entropía: Codificación de Huffman y codificación aritmética. El codec (codificador-decodificador) base usa el código de Huffman pero se especifican codec con los dos métodos para todos los modos de operación. Es útil considerar la codificación por entropía como un proceso de dos pasos. El primero convierte la secuencia zig-zag de coeficientes cuantificados en una secuencia intermedia de símbolos. El segundo paso convierte los símbolos a una secuencia de datos en la que los símbolos no tienen extensas fronteras externamente identificables. La forma y definición de símbolos intermedios dependen del modo de operación y del método de codificación por entropía.

La codificación de Huffman requiere de uno o más conjuntos de tablas de código de Huffman para la aplicación. Estas deben ser la misma para comprimir y descomprimir una imagen. Las tablas de códigos de Huffman pueden ser predefinidas y usadas dentro de una aplicación como default, o específicamente calculada para una imagen.

El método de codificación aritmética propuesto por el JPEG, no requiere ninguna tabla externa, porque es capaz de adaptarse a las estadísticas de la imagen conforme la va codificando. Este tipo de codificación produce entre un 5% y un 10% mejor compresión que los códigos de Huffman para muchas de las imágenes que los miembros de JPEG han probado. Sin embargo, es más complejo que la codificación de Huffman para ciertas implementaciones.