Module 3

LOSSY IMAGE COMPRESSION SYSTEMS

Lesson
9

Discrete Cosine Transforms

Instructional Objectives

At the end of this lesson, the students should be able to

1. State the advantages of Discrete Cosine Transform (DCT) over Discrete Fourier Transform (DFT).
2. Transform a block of image into its DCT coefficients.
3. Compose the basis images of DCT.
4. Apply zonal coding over the DCT coefficients.
5. Allocate bits to the zonal coded coefficients.
6. Apply threshold coding over the DCT coefficients.
7. Encode the quantized DCT coefficients in zig-zag scanned order.
8. State the limitations of DCT.

9.0 Introduction

In lesson-8, we had pointed out some of the practical limitations of K-L Transforms (KLT). Despite its optimal performance in terms of mean-square error, it is not popular since the transformation kernel is image-dependent and fast computational algorithms and architectures are not available. Sinusoidal transforms, like the Discrete Cosine Transforms (DCT) and Discrete Fourier Transforms (DFT) use image-independent transformations. It is seen that DCT's energy compaction performance closely resembles that of KLT. Moreover, fast algorithms and architectures are available for DCT and DFT. As compared to DFT, application of DCT results in less blocking artifacts due to the even-symmetric extension properties of DCT. Also, DCT uses real computations, unlike the complex computations used in DFT. This makes DCT hardware simpler, as compared to that of DFT. These advantages have made DCT-based image compression a standard in still-image and multimedia coding standards.

In this lesson, we are going to discuss DCT-based image compression in detail. We shall first present the basic theory of DCT and show how to transform a block of image into its corresponding transformed array of DCT coefficients. The transformed array is encoded into a bit-stream using two approaches: zonal coding and threshold coding. The two approaches will be presented in this lesson. We shall show how the threshold-coded coefficients are scanned in a zig-zag order and Huffman coded. At the end of this lesson, we point out some of the performance limitations of DCT in low bit-rate situations.

9.1 Discrete Cosine Transform (DCT)

DCT is an orthogonal transformation that is very widely used in image compression and is widely accepted in the multimedia standards. DCT belongs to a family of 16 trigonometric transformations. The type-2 DCT transforms a block of image of size

N x *N* having pixel intensities *s*(*n*₁,*n*₂) into a transform array of coefficients *S*(*k*₁,*k*₂), described by the following equation:

$$S(k_1,k_2)=\sqrt{\frac{4}{N^2}}C(k_1)C(k_2)\sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1}s(n_1,n_2)\cos\left(\frac{\pi(2n_1+1)k_1}{2N}\right)\cos\left(\frac{\pi(2n_2+1)k_2}{2N}\right) \quad \text{.... (9.1)}$$

where $k_1,k_2,n_1,n_2=0,1,\cdots\cdots N-1,$ and

$$C(k)=\begin{cases}1/\sqrt{2} & \text{for } k=0\\ 1 & \text{otherwise}\end{cases}$$

The transformed array $S(k_1,k_2)$ obtained through equation (9.1) is also of the size *N* x *N*, same as that of the original image block. It should be noted here that the transform-domain indices k_1 and k_2 indicate the spatial frequencies in the directions of n_1 and n_2 respectively. $k_1=k_2=0$ corresponds to the average or the DC component and all the remaining ones are the AC components which correspond to higher spatial frequencies as k_1 and k_2 increase.

From computational considerations, it may be noted that direct application of the above equation to compute the transformed array requires $O(N^4)$ computations. Using *Fast Fourier Transform (FFT)*-like algorithm to compute the DCT, computations can be reduced to $O(2N^2\log N)$. Such fast computational approaches and use of real arithmetic has made DCT popular for image compression applications. Since all natural images exhibits spatial redundancy, not all coefficients in the transformed array have significant values. This can be demonstrated by an example. We take an 8x8 block from Lena image, whose pixel intensities are shown in Fig.9.1.

166	162	162	160	155	163	160	155
166	162	162	160	155	163	160	155
166	162	162	160	155	163	160	155
166	162	162	160	155	163	160	155
166	162	162	160	155	163	160	155
161	160	155	159	154	154	156	154
159	163	158	163	155	155	156	152
159	162	162	160	153	153	153	151

Fig 9.1 A Specimen 8 X 8 block of Lena image

We subtract 128 (i.e., the average value of intensity in 8-bit monochrome image representation) from each pixel intensity and then compute the DCT for each element of $S(k_1, k_2)$ using equation (9.1). The transformed array values are shown in Fig.9.2.

248	19	3	4	-7	9	1	-7
11	-2	3	6	-3	2	5	0
-4	2	-2	-3	0	-1	-1	0
-1	-1	1	1	2	0	-1	0
2	1	0	0	-2	0	3	0
0	0	-1	0	0	0	-1	-1
-3	0	1	0	1	0	0	0
3	0	0	0	-1	0	0	0

Fig 9.2 DCT coefficients for the 8 X 8 block

It is worth noting that most of the transformed coefficients have very small values and only a few coefficients have higher magnitudes. This shows the energy compaction capabilities of DCT.

The DCT basis images can be computed using the transformation kernel, which is the same for both forward DCT and inverse discrete cosine transformations (IDCT) and is given by

$$g(n_1,n_2,k_1,k_2)=h(n_1,n_2,k_1,k_2)=\sqrt{\frac{4}{N^2}}C(k_1)C(k_2)\cos\left(\frac{\pi(2n_1+1)k_1}{2N}\right)\cos\left(\frac{\pi(2n_2+1)k_2}{2N}\right)$$

.................................(9.2)

For each value of k_1 and k_2 ($k_1,k_2=0,1,\cdots\cdots N-1$), we obtain a basis image of size *N* x *N* by computing equation (9.2) over $n_1,n_2=0,1,\cdots\cdots N-1$. For a block size of 4 x 4 we therefore obtain 16 basis images, each of size 4 x 4 and the results are given in Fig.9.3.

Fig9_3.pgm

Selection of block-sizes in DCT is an important consideration. The images should be so subdivided that the level of redundancies between the adjacent sub-images are reduced to an acceptable level and the dimension of the sub-images should be an integer powers of 2. Increasing the block size reduces adjacent block redundancies and reduces mean square reconstruction error using truncated and quantized coefficients, but involves more computations. Most popular block sizes used in image compression are 8 x 8 and 16 x 16.

9.2 Principles of bit allocation

In Section-8.2.3, we had shown how the original image array s can be reconstructed using the transformed coefficients *S*(*k₁,k₂*) and the basis images $\mathbf{H}_{k_1,k_2}$ and the equation (8.7) is reproduced here for convenience

$$\mathbf{s}=\sum_{k_1=0}^{n-1}\sum_{k_2=0}^{n-1}S(k_1,k_2)\mathbf{H}_{k_1,k_2} \quad \text{...} \quad (9.3)$$

If all the transform coefficients are retained with full precision, it is possible to have exact reconstruction of s. However, if we decide to selectively truncate some of the transform coefficients, using a transform coefficient masking function of the form

$$\gamma(k_1,k_2)=\begin{cases}0 & \text{if } S(k_1,k_2) \text{ is truncated}\\ 1 & \text{otherwise}\end{cases} \quad \text{..} \quad (9.4)$$

for $k_1,k_2=0,1,\cdots\cdots N-1$, an approximation of s can be obtained as

$$\hat{\mathbf{s}} = \sum_{k_1=0}^{n-1}\sum_{k_2=0}^{n-1}\gamma(k_1,k_2)S(k_1,k_2)\mathbf{H}_{k_1,k_2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (9.5)$$

The reconstruction error associated with the truncated series expansion of equation (9.5) is dependent on the number of transform coefficients retained, their relative importance and the precision with which the retained coefficients are represented. Two approaches are used to select the retained transform coefficients – one is based on the basis of variance, which is referred to as *zonal coding* and the other is based on the basis of magnitude, called *threshold coding*. The overall process of truncating, quantizing and coding the transform coefficients is called the *bit allocation*.

9.2.1 Zonal Coding:

Zonal coding is based on the premise that the transformed coefficients having very high variances are the ones that carry most of the signal and should be retained, whereas the ones with less variance can be truncated. In zonal coding, it is therefore necessary to compute the variances at every position of the transformed array, based on an ensemble of representative blocks of transformed arrays or by applying global image models, such as Gauss-Markov model. *M* transform coefficients may be retained based on high values of variance. The retained coefficients will have a value of 1 in the binary zonal mask, whereas all truncated coefficients will have a value of 0. A typical zonal coding mask for an 8 x 8 block is shown in Fig.9.4.

1	1	1	1	1	0	0	0
1	1	1	1	0	0	0	0
1	1	1	0	0	0	0	0
1	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Fig 9.4 A Typical Zonal Coding mask for an 8 X 8 Block

Each block has the same zonal mask. These masks can be customized for images and in that event, the mask information needs to be encoded with the image.

Two different bit allocation policies exist to encode the retained coefficients. In one, same number of bits is assigned to each retained coefficient. Each coefficient is normalized by its standard deviation and then uniformly quantized. In the other bit allocation policy, number of bits allocated to the retained coefficients is based on the variances of those coefficients computed and more bits are to be allocated to the coefficients having high variance. In this approach, optimal Lloyd-Max quantizers are designed for every retained coefficient. A simple bit allocation algorithm as per this bit allocation scheme is presented below.

Let *B* be the number of bits available to a block for allocation. The number of bits allocated to the *i*th retained coefficient is given by

$$b_i = \frac{B}{M} + \frac{1}{2}\log_2 \sigma_i^2 - \frac{1}{2M}\sum_{i=1}^{M}\log_2 \sigma_i^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(9.6)}$$

where $\sigma_i^2 \quad i = 1,2,\cdots\cdots M$ are the variances of the *i*th retained coefficient. Summing up the bits allocated to each retained coefficients according to equation (9.6) fulfils the condition $\sum_{i=1}^{M} b_i = B$. If the retained coefficients are modeled by Gaussian density, this bit allocation approach is consistent with rate distortion theory which indicates that a Gaussian random variable of variance σ^2 cannot be represented by less than $\frac{1}{2}\log_2 \frac{\sigma^2}{D}$ bits and be represented with mean-square error less than *D*. A typical zonal bit allocation for an 8 x 8 block is shown in Fig.9.5.

10	9	8	6	4			
9	8	6	4				
8	6	4					
6	4						
4							

Fig. 9.5 *A Typical Zonal Bit Allocation for an 8x8 Block*

9.2.2 Threshold Coding:

Zonal coding is often applied over a fixed mask, which may not be optimal for all blocks and for all images. For better coding performance, the positions and the number of retained coefficients should be adaptively changed on a block to block basis, based on its transformed coefficient array. Such adaptive bit allocation is done using *threshold coding* approach, which is more often used in practice and is based on the premise that the transform coefficient of largest magnitude makes the most significant contribution to the reconstructed block quality. Only those transform coefficients, whose magnitudes exceed a threshold are significant and all the remaining ones can be discarded for image reconstruction. The transform coefficients are first quantized by the elements $T(k_1,k_2)$of a quantization matrix, having dimensions same as that of the block size. The process of quantization and thresholding can be combined in the form of a single operation to obtain the quantized transform coefficients $\dot{S}(k_1,k_2)$, as defined below:

$$\dot{S}(k_1,k_2) = NINT\left[\frac{S(k_1,k_2)}{T(k_1,k_2)}\right] \qquad (9.7)$$

where the function $NINT[.]$ perform rounding of the argument to its nearest integer. The elements of the quantization matrix are designed from psycho-visual considerations, based on the human visual system response to luminance and chrominance variations. Quantization matrices are chosen according to the source noise level and viewing distance. A typical quantization matrix used to quantize the transformed luminance coefficients is shown in Fig.9.6.

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

Fig. 9.6: *A Typical Quantization Matrix for Luminance*

When this quantization matrix is applied on the transformed coefficients obtained in Fig.9.2, the quantized transform coefficient array $\dot{S}(k_1, k_2)$ ($k_1, k_2 = 0,1, \cdots\cdots N-1$) that results is shown in Fig.9.7.

16	2	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

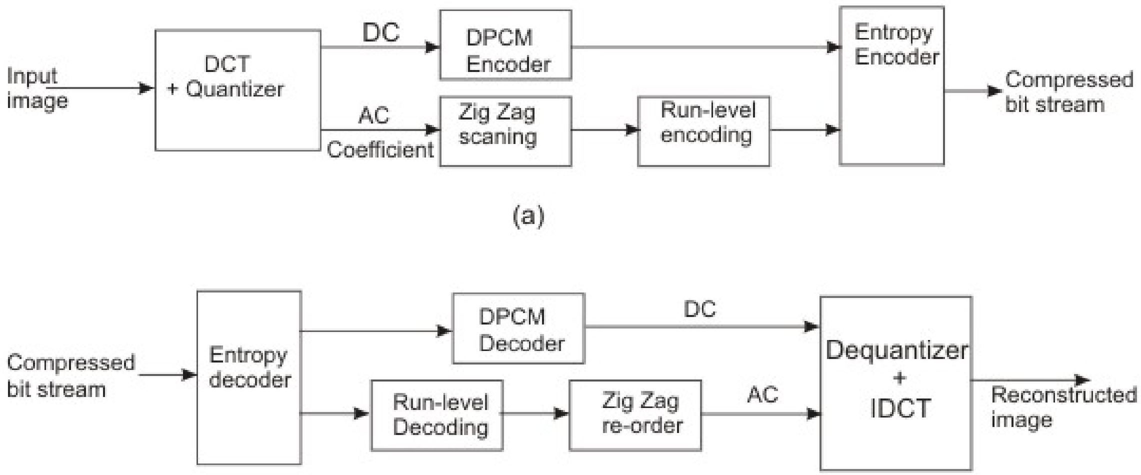
Fig 9.7 Quantized DCT coefficients for the 8 X 8 block

It may be noted that a large number of coefficients in the $\dot{S}(k_1, k_2)$ array, especially the coefficients corresponding to higher spatial frequency are zero and hence can be discarded. An efficient encoding strategy must therefore be adopted, so that the redundancies associated with large number of quantized transform coefficients having zero values can be exploited in the bit stream design. This is done by picking up the $\dot{S}(k_1, k_2)$ coefficients in a zig-zag scanned order and then encoding the (*run, level*) pairs of non-zero $\dot{S}(k_1, k_2)$ coefficients using Huffman encoding. Fig.9.8 shows the zig-zag scanned ordering of coefficients, that start with the DC coefficient ($k_1 = k_2 = 0$) and then proceeds in a zig-zag fashion to progressively pick up the higher spectral components in both n_1 and n_2 directions. Whenever any non-zero coefficient is encountered in zig-zag scanned ordering, it is encoded as a (*run, level*) pair, where *run* corresponds to the runs of 0s that precedes the non-zero coefficients in the zig-zag scanned order and *level* corresponds to the non-zero value of the quantized coefficient.

Fig.9.8 Zigzag scanning of DCT coefficients

Every (*run, level*) pair has an associated probability of occurrence and these are listed in a table, where the Huffman codes on (*run, level*) pairs are assigned. Using this scheme, variable length codes are assigned to every block of the image and the number of bits allocated to the block would depend upon the level of details present in the block. More details mean more number of non-zero coefficients in the $\dot{S}(k_1,k_2)$ array and consequently more number of bits, whereas the reverse happens for blocks having insignificant details.

Bit allocations based on the threshold coding of DCT-transformed coefficients have been adopted in the still image compression standard JPEG, prepared by the Joint Photographic Experts Group. As per this standard, the DC coefficient of a block is DPCM encoded with reference to the previous adjacent block and the threshold coding scheme, described above is applied on the AC coefficients. The block diagram of the overall encoding and decoding scheme using threshold coding is presented in Fig.9.9.



(a)

(b)

Fig 9.9 (a) Encoding and (b) decoding scheme using DCT and threshold coding

9.3 Limitations of DCT

Despite excellent energy compaction capabilities, mean-square reconstruction error performance closely matching that of KLT and availability of fast computational approaches, DCT offers a few limitations which restrict its use in very low bit rate applications. The limitations are listed below:

(i) Truncation of higher spectral coefficients results in blurring of the images, especially wherever the details are high.

(ii) Coarse quantization of some of the low spectral coefficients introduces graininess in the smooth portions of the images.

(iii) Serious blocking artifacts are introduced at the block boundaries, since each block is independently encoded, often with a different encoding strategy and the extent of quantization.

Of all the listed problems, as above, blocking artifact is the most serious and objectionable one at low bit rates. Blocking artifacts may be reduced by applying an overlapped transform, like the Lapped Orthogonal Transform (LOT) or by applying post-processing. At lower bit rates, Discrete Wavelet Transforms (DWT) (to be discussed in subsequent lessons) avoid the blocking artifacts of DCT and present better coding performance.

Questions

NOTE: The students are advised to thoroughly read this lesson first and then answer the following questions. Only after attempting all the questions, they should click to the solution button and verify their answers.

PART-A

A.1. Enlist the advantages of DCT over the DFT.

A.2. Write the expression for DCT applied on an N x N block.

A.3. How is the DCT basis images computed?

A.4. State the basic principles of applying zonal coding on DCT coefficients.

A.5. State the advantages of threshold coding over zonal coding.

A.6. State the performance limitations of DCT at low bit rates.

PART-B: Multiple Choice

In the following questions, click the best out of the four choices.
B.1 Which of the following statements is wrong

(A) An N-point DCT has N-periodicity.

(B) DCT involves real computations only.

(C) Forward and inverse DCT kernels are same.

(D) DCT exhibits good energy compaction capability.

B.2 DCT is applied on the following 2x2 pixel array:

$$\begin{bmatrix} 13 & 12 \\ 11 & 12 \end{bmatrix}$$

The DCT coefficients obtained by applying equation (9.1) on the above array are

(A) $\begin{bmatrix} 12 & 1 \\ 0 & 1 \end{bmatrix}$ (B) $\begin{bmatrix} 12 & 0 \\ 1 & 1 \end{bmatrix}$

`(C) $\begin{bmatrix} 24 & 0 \\ 1 & 1 \end{bmatrix}$ (D) $\begin{bmatrix} 24 & 1 \\ 0 & 1 \end{bmatrix}$

B.3 The DCT coefficients and quantization matrix for a 2x2 block are given by

$$S(k_1,k_2)=\begin{bmatrix} 39 & 18 \\ -23 & 15 \end{bmatrix} \quad T(k_1,k_2)=\begin{bmatrix} 16 & 21 \\ 22 & 36 \end{bmatrix}$$

The quantized array $\dot{S}(k_1,k_2)$ is given by

(A) $\begin{bmatrix} 3 & 1 \\ -2 & 1 \end{bmatrix}$ (B) $\begin{bmatrix} 2 & 0 \\ -1 & 0 \end{bmatrix}$

(C) $\begin{bmatrix} 2.44 & 0.86 \\ -1.05 & 0.42 \end{bmatrix}$ (D) $\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$

B.4 The DCT basis image $\mathbf{H}_{1,1}$ for a 2x2 block size is

(A) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ (B) $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$

(C) $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$ (D) $\frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

B.5 The 9th coefficient of a quantized 4x4 DCT array in zig-zag scanned order is:

(A) $\dot{S}(1,0)$

(B) $\dot{S}(1,2)$

(C) $\dot{S}(2,1)$

(D) $\dot{S}(3,0)$

B.6 Blocking artifacts in DCT is due to the fact that

(A) pixels within the block exhibit spatial redundancy.

(B) each block is quantized independently.

(C) quantization is followed by post-processing.

(D) none of the above.

B.7 In a zonal coding, 48 bits are to be allocated to four retained quantized DCT coefficients having variances:

$\sigma_1^2 = 256 \quad \sigma_2^2 = 64 \quad \sigma_3^2 = 64 \quad \sigma_4^2 = 16$

The bits allocated to the coefficients in decreasing order of their variances are:

(A) 13,12,12,11

(B) 11,12,12,13

(C) 16, 12, 12, 8

(D) 8, 12, 12, 16

B.8 Design of the elements of quantization matrix is based on

(A) Number of bits available for allocation.

(B) Human visual system response to spatial frequencies.

(C) Average intensity level over an ensemble of blocks.

(D) Variances of each coefficient over an ensemble of blocks.

B.9 An example $\dot{S}(k_1, k_2)$ array is as shown below:

35	5	0	0	0	0	0	0
-1	0	-1	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0
0	-1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

The zig-zag scanned (*run*, *level*) pairs for the above array are

(A) (1,35), (1,5),(1,-1), (5,-1),(15,-1),(18,1), EOB

(B) (0,35), (0,5),(0,-1), (4,-1),(14,-1),(17,1), EOB.

(C) (0,35), (0,5),(0,-1), (4,-1),(18,-1),(11,1), EOB

(D) (0,35), (0,5),(6,-1), (1,-1),(18,-1),(11,1), EOB

PART-C: Computer Assignments

C-1.
(a) Write a computer program to perform DCT and IDCT over an 8x8 array.
(b) Subdivide a monochrome image into non-overlapping blocks of size 8x8 and apply DCT on each block to obtain the transform coefficients.
(c) Apply IDCT on each block of transformed coefficients and obtain the reconstructed image.
(d) Check that the reconstructed image is exactly the same as that of the original image.

Note: Represent the DCT kernel and the transformed coefficient values in double-precision floating point and do not truncate your results, before converting the reconstructed image into an unsigned character array.

C-2.
(a) On the 8x8 transformed coefficients obtained in the above assignment, apply the quantization matrix shown in fig.9.6 and obtain the quantized DCT coefficients.
(b) Obtain inverse quantization by multiplying the quantized DCT coefficients with the corresponding elements of the quantized matrix.
(c) Apply IDCT on the coefficients as above and obtain the reconstructed image.
(d) Check that the reconstructed image is not the same as the original image and calculate the PSNR of the reconstructed image.

(e) In part-(a), multiply the elements of the quantization matrix by (i) 2, (ii) 4 and (iii) 8. In each case, repeat part-(a) to part-(d) and compute the PSNR of the reconstructed image.
(f) Check that the PSNR decreases as the multiplication factor increases.

C-3.
(a) Write a computer program to extract the (run, level) pairs from a zig-zag scanned 8x8 array.
(b) From the quantized coefficients arrays obtained in problem C-2(a), extract the (run, level) pairs. Consult the Huffman coding table from the JPEG standard and form the Huffman coded bit stream.
(c) Calculate the number of bits generated from the Huffman coder and the compression ratio achieved.
(d) Vary the multiplication factors associated with the quantization matrix and obtain a plot of PSNR versus compression ratio.

SOLUTIONS

A.1

A.2

A.3

A.4

A.5

A.6

B.1 (A) B.2 (C) B.3 (D) B.4 (D) B.5 (C)
B.6 (B) B.7 (A) B.8 (B) B.9 (B).

C.1

C.2