# Chapter 4

# TRANSFORM CODING (TC)

- TC: another efficient coding scheme based on utilization of interpixel correlation.
- TC: a core technique recommended by JPEG.
- TC: efficient in coding prediction error in motion compensated predictive coding, adopted by H. 261, H.263, and MPEG 1/2, MPEG 4.

# 4.1 Introduction

- Recall block diagram of source encoders (Figure 2.3).
- Transformation component decides which format of input source is quantized and encoded.
  - In DPCM

    $\Rightarrow$ the difference signal

    (smaller variance)
  - In TC

    $\Rightarrow$ the transformed version of a signal

    (less correlated)

## 4.1.1 Hotelling Transform

- Consider an $N$-dimensional vector $\vec{Z}_S$ .

- The ensemble of such vectors, $\{\vec{z}_s\}\ s \in I$ ,

  *I:* set of all indexes

  Can be modeled by a random vector $\vec{z}$ each of component $z_i$: a random variable.

$$\vec{z} = (z_1, z_2, \cdots, z_N)^T, \qquad (4.1)$$

- The mean vector of the population, $m_{\vec{z}}$:

$$m_{\vec{z}} = E[\vec{z}] = (m_1, m_2, \cdots, m_N)^T, \qquad (4.2)$$

$$m_i = E[z_i] \qquad (4.3)$$

- The covariance matrix of the population,

$$C_{\vec{z}} = E[(\vec{z} - m_{\vec{z}})(\vec{z} - m_{\vec{z}})^T]. \qquad (4.4)$$

The element $c_{i,j} = Cov(z_i, z_j)$

The element $c_{i,i}$: variance of $z_i$.

- $\Rightarrow C_{\vec{z}}$: real and symmetric.
- According to the theory of linear algebra, always possible to find a set of $N$ orthonormal eigenvectors of the matrix $C_{\vec{z}}$, to convert the real symmetric matrix $C_{\vec{z}}$ into a full ranked diagonal matrix.
- $N$ orthonromal eigenvectors: $\vec{e}_i$
- their corresponding eigenvalues: $l_i$
- Form a matrix $\Phi$ (orthogonal)

$$\Phi = (\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_N)^T \quad (4.5)$$

- Now, the Hotelling transform [hotelling 1933], or eigenvector transform [tasco 1971, wintz 1972]:

$$\vec{y} = \Phi(\vec{z} - m_{\vec{z}}) \quad (4.6)$$

- Two features:

1) $m_{\vec{y}} = 0.$ (4. 7)

2)

$$C_{\vec{y}} = \Phi C_{\vec{z}} \Phi^T = \begin{bmatrix} l_1 & & & & 0 \\ & l_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & l_n \end{bmatrix}$$

- The inverse Hotelling transform:

$$\vec{z} = \Phi^{-1} \vec{y} + m_{\vec{z}}$$

Or,

$$\vec{z} = \Phi^T \vec{y} + m_{\vec{z}}$$

- Note that: $C_{\vec{y}}$ is a diagonal matrix.

  ⇒the correlation previously existing between the different components of the random vector $\vec{z}$ has been removed in the transformed domain.

➢ The analogous transform for continuous data was devised by Karhunen and Loeve [karhunen 1947, loeve 1948].

➢ Alternatively, <u>the Hotelling transform can be viewed as the discrete version of the Karhunen-Loeve transform (KLT).</u>

### 4.1.2 Statistical Interpretation

- The elements in the main diagonal of $C_{\vec{y}}$:
  - <u>eigenvalues</u> of $C_{\vec{y}}$

- <u>variances</u> of the components of vector $\vec{y}$, denoted by $s_{y,1}^{2}, s_{y,2}^{2}, \cdots, s_{y,N}^{2}$.

- Arrange eigenvalues (variances) in a nonincreasing order: $l_1 \geq l_2 \geq \cdots \geq l_N$.
- Choose an integer $L$, and $L < N$.
- Using the corresponding $L$ eigenvectors, $\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_L$ to form a matrix $\overline{\Phi}$ of $L \times N$.
- Transform changes to:

$$\vec{y} = \overline{\Phi}(\vec{z} - m_{\vec{z}}). \tag{4. 8}$$

Inverse transform:

$$\vec{z}' = \overline{\Phi}^T \vec{y} + m_{\vec{z}}.$$

- It was shown [wintz 1972] that MSE between the $\vec{Z}$ and the $\vec{Z}'$ is given by

$$MSE_r = \sum_{i=L+1}^{N} s_{y,i}^2$$

(without considering the quantization error and transmission error)

- Quantizing and encoding only *L* components of vector $\vec{y}$ in the transform domain lead to higher coding efficiency. This is the basic idea behind transform coding.

- The linear unitary transform can provide the following two functions.
  - Decorrelate input data; i.e., transform coefficients are less correlated
  - Have some transform coefficients more significant than others:
    - ✓some can be discarded
    - ✓some can be coarsely quantized
    - ✓some can be finely quantized.

### 4.1.3 Geometrical Interpretation

- A binary image of a car in Figure 4.1 (a).
- Each pixel in the shaded object region: a 2-D vector with its two components being coordinates $z_1$ and $z_2$, respectively.
- The Hotelling transform
- Notice the two features.

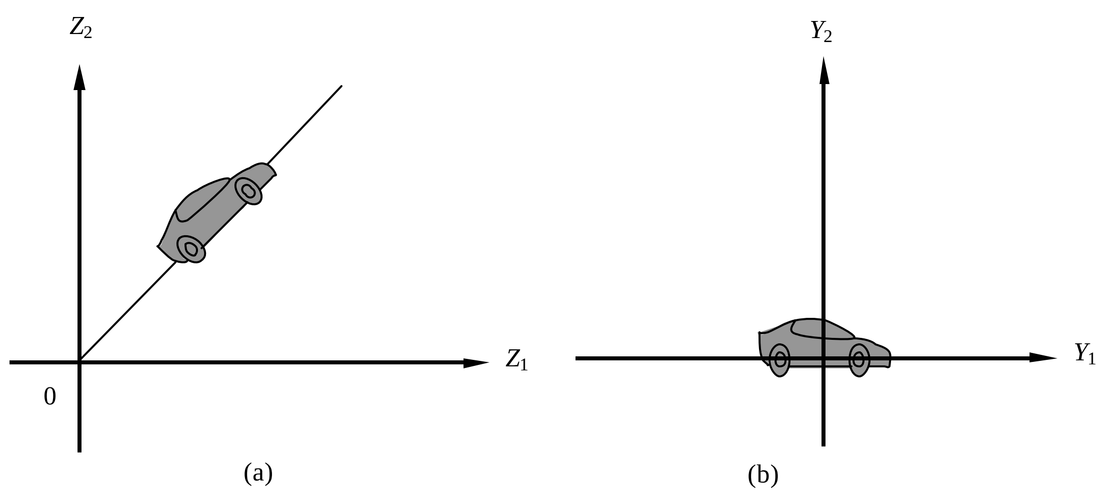



Figure 4.1 (a) A binary object in the $z_1$-$z_2$ coordinate system (b) After the Hotelling transform, the object is aligned with its principal axes.

# 4.1.4 Procedures of Transform Coding

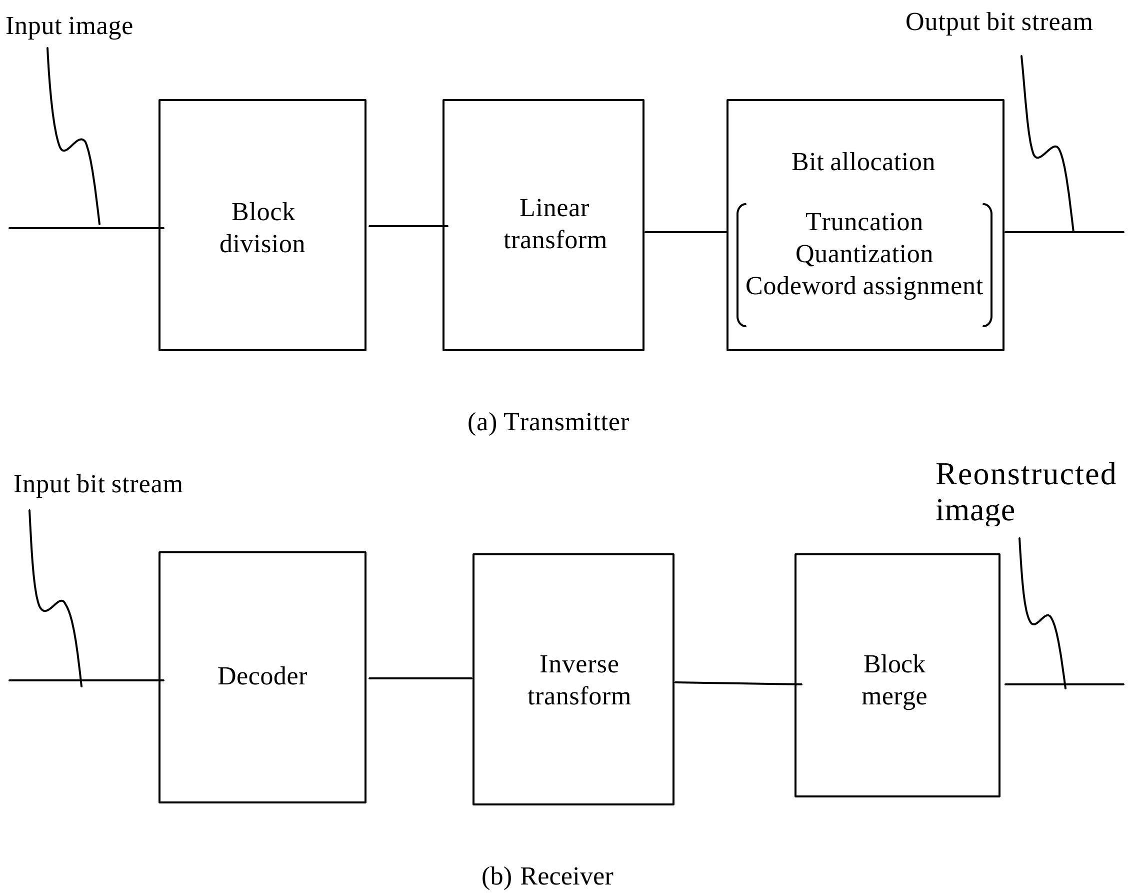



Figure 4. 2 Block diagram of transform coding

# 4.2 Linear Transforms

## 4.2.1 2-D Image Transformation Kernel

- A digital image: a 2-D array $g(x, y)$

  $(x, y)$: coordinates of a pixel in 2-D array

  $g$: gray level value of the pixel.
- A digital image $\Leftrightarrow$ 2-D array $\Leftrightarrow$ a matrix
- $T(u,v)$: 2-D transform of $g(x, y)$

  $(u,v)$: coordinates in transformed domain.
- 2-D forward and inverse transforms:

$$T(u,v)=\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} g(x,y)\, f(x,y,u,v),$$

$$g(x,y)=\sum_{u=0}^{N-1}\sum_{v=0}^{N-1} T(u,v)\, i(x,y,u,v),$$

- $f(x,y,u,v)$ and $i(x,y,u,v)$: forward and inverse ***transformation kernels***, respectively

#### 4.2.1.1 Separability

- A transformation kernel is separable

  ( transform is separable) if

$$f(x,y,u,v)=f_1(x,u)f_2(y,v)$$

$$i(x,y,u,v)=i_1(x,u)i_2(y,v)$$

- Note that a 2-D separable transform can be decomposed into two 1-D transforms.

$$T_1(x,v)=\sum_{y=0}^{N-1} g(x,y)\,f_2(y,v),$$

$$T(u,v)=\sum_{x=0}^{N-1} T_1(x,v)\,f_1(x,u),$$

#### 4.2.1.2 Symmetry

The transformation kernel is symmetric (hence, the transform is symmetric) if

$$f_1(y,v) = f_2(y,v).$$

### 4.2.1.3 Matrix Form

- If a transformation kernel is symmetric (hence, separable) then 2-D image transform can be expressed in matrix form.
- Denote an ***image matrix*** by $G$

$$G = \{g_{i,j}\} = \{g(i-1, j-1)\}.$$

- Denote ***forward transform matrix*** by $F$

$$F = \{f_{i,j}\} = \{f_1(i-1, j-1)\}.$$

- Denote ***inverse transform matrix*** by $I$

$$I = \{i_{j,k}\} = \{i_1(j-1, k-1)\}.$$

$$I = F^{-1}$$

- Note that $G$, $T$, $F$, and $I$ are of $N \times N$.
- We then have

$$T = F^T G F$$

$$G = I^T T I$$

- DFT (Complex quantities)

$$T = F^{*T} G F$$

$$G = I^{*T} T I$$

$$I = F^{-1} = F^{*T},$$

where * indicates complex conjugation.

*F* and *I* containing complex quantities

⇒ **unitary** matrices, a unitary transform.

#### 4.2.1.4 Orthogonality

- Definition: A transform is orthogonal if

$$F^T = F^{-1}$$

- An orthogonal matrix (transform) is a special case of a unitary matrix (transform): only real quantities are involved.

- All 2-D image transforms presented: separable, symmetric, and unitary.

### 4.2.2 Basis Image Interpretation

- Consider 2-D inverse transform

$$g(x,y) = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1} T(u,v)\, i(x,y,u,v)$$

(**component form**, $0 \le x, y \le N-1$)

- Consider an $N \times N$ **basis image**

$$I_{u,v} = \{ i(x,y,u,v), 0 \le x, y \le N-1 \}$$

for a specific $(u, v)$. That is,

$$I_{u,v} = \begin{bmatrix} i(0,0,u,v) & i(0,1,u,v) & \cdots & \cdots & i(0,N-1,u,v) \\ i(1,0,u,v) & i(1,1,u,v) & \cdots & \cdots & i(1,N-1,u,v) \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ i(N-1,0,u,v) & i(N-1,1,u,v) & \cdots & \cdots & i(N-1,N-1,u,v) \end{bmatrix}$$

- $0 \le u,v \le N-1 \Rightarrow N \times N$ basis images
- The inverse transform in a ***collective*** form

$$G = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} T(u,v) I_{u,v}$$

- Interpretation:
  - A series expansion of the original image $G$ into a set of $N^2$ basis images $I_{u,v}$.
  - Transform coeff. $T(u,v)$, $0 \le u,v \le N-1$, become the coefficients of the expansion.
  - **Coefficient** (weight) $T(u,v)$ : **a correlation measure between image $G$ and basis image** $I_{u,v}$ [wintz 1972].
- Comments on basis images:
  - Have nothing to do with the input image.

- Completely defined by transform itself.
- Attribute of 2-D image transforms.
- Different transforms ⇔ different sets of basis images.

• Motivation: with a proper transform, hence, a proper set of basis images, transform coefficients are more independent than the gray scales of the original input image.

- The optimum linear transform: **uncorrelated coefficients**.
- In addition, the coefficient variance varies widely.
  - ✓Insignificant coefficients ignored
  - ✓Significant coefficients are allocated more bits in encoding.
- The coding efficiency is thus enhanced.

• Figure 4.3

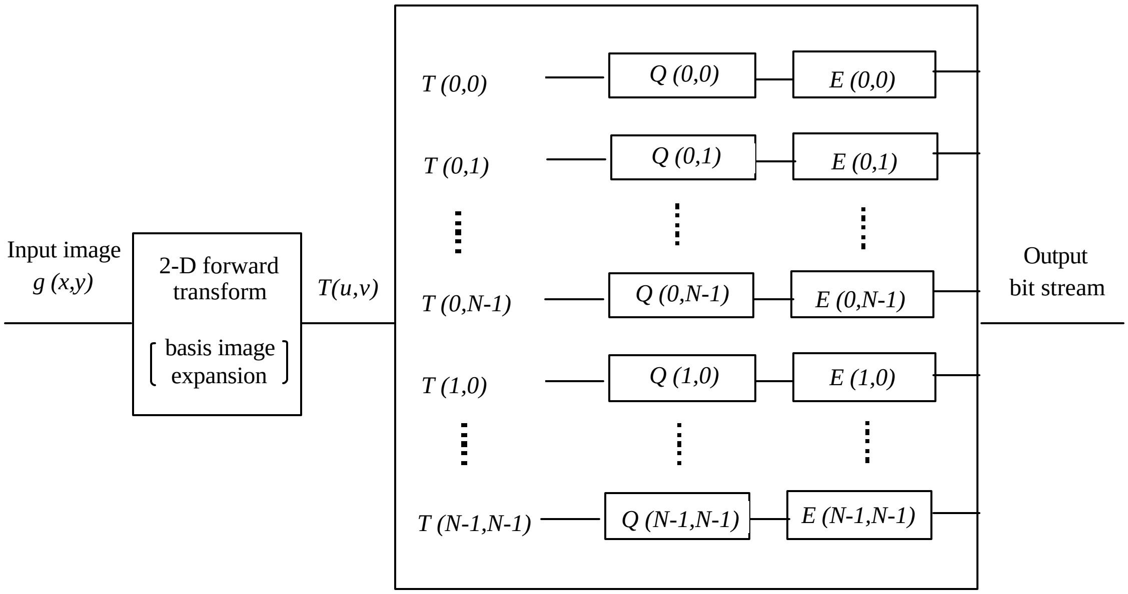
• This strategy is similar to that of subband coding.

- From this point of view, TC coding can be considered a special case of subband coding, though TC was devised much earlier than subband coding.

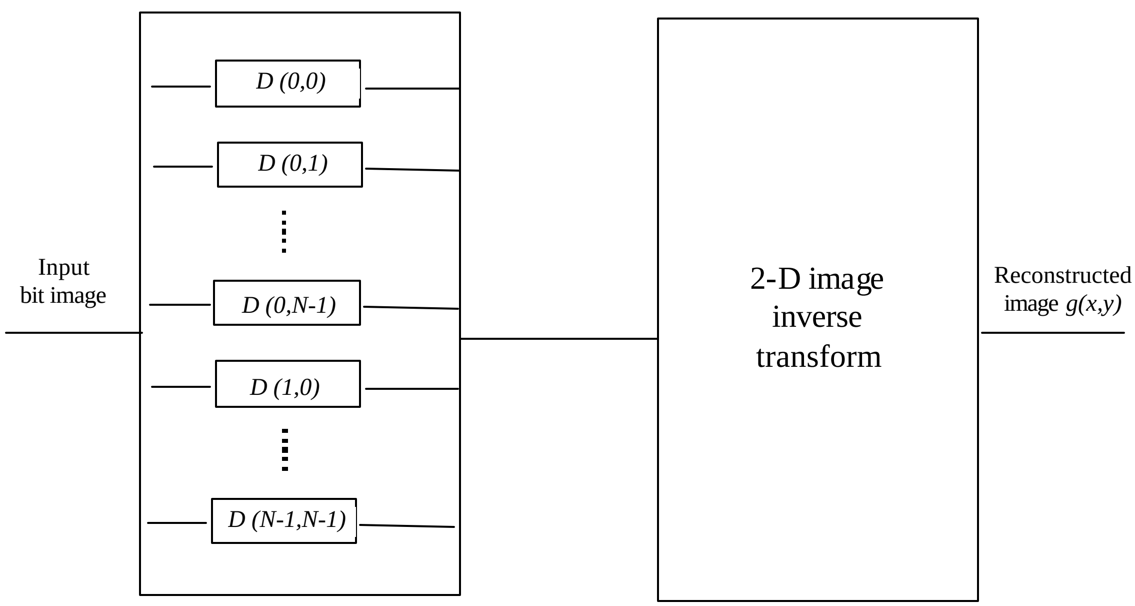
- An alternative way to define basis images.

  $I_{u,v} = \vec{b}_u \vec{b}_v^T$ : the *outer product* of

  The basis vector, $\vec{b}_u$, is the *u*th column vector of the inverse transform matrix *I* [jayant 1984].




(a) Transmitter




(b) Receiver

Figure 4.3 Basis image interpretation of TC (Q: quantizer, E: encoder, D: decoder).

### 4.2.3 Subimage Size Selection

- The larger the size the more decorrelation the TC can achieve.
- Correlation between image pixels, however, becomes insignificant when the distance between pixels becomes large, e.g., it exceeds 20 pixels [habibi 1971a].
- On the other hand, a large size causes some problems.
  - In adaptive TC, a large block cannot adapt to local statistics well.
  - As seen later, a transmission error in TC affects the whole subimage. Hence a large size implies a possibly severe effect of transmission error on quality of reconstructed images.
- In video coding, TC is used together with motion compensated (MC) coding.

  Large block size not used in motion estimation.

- Subimage sizes (N) of 4, 8 and 16 used most often.
- In particular, *N=8* is adopted by the JPEG as well as H.261, H.263, MPEG ½.

## 4.3 Transforms of Particular Interest

### 4.3.1 Discrete Fourier Transform (DFT)

### 4.3.2 Discrete Walsh Transform (DWT)

- DWT transformation kernels [walsh 1923]:

$$f(x,y,u,v)=\frac{1}{N}\prod_{i=0}^{n-1}\left[(-1)^{p_i(x)p_{n-1-i}(u)}(-1)^{p_i(y)p_{n-1-i}(v)}\right],$$

$$i(x,y,u,v)=f(x,y,u,v).$$

$$n = \log_2 N$$

- $p_i(\text{arg})$: *i*th bit in NBC of the arg,

  The <u>*o*th</u> bit: the <u>least significant</u> bit

  The (*n-1*)th bit: the most significant bit.

- For instance, consider *N=16*, then *n=4*.

  The NBC of 8: 1000.

  $p_0(8) = p_1(8) = p_2(8) = 0, \; p_3(8) = 1.$

- DWT transform matrix for N=4:

$$F = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

- The DWT implementation is simple.

- $F$ is simple
- $i(x,y,u,v) = f(x,y,u,v)$.

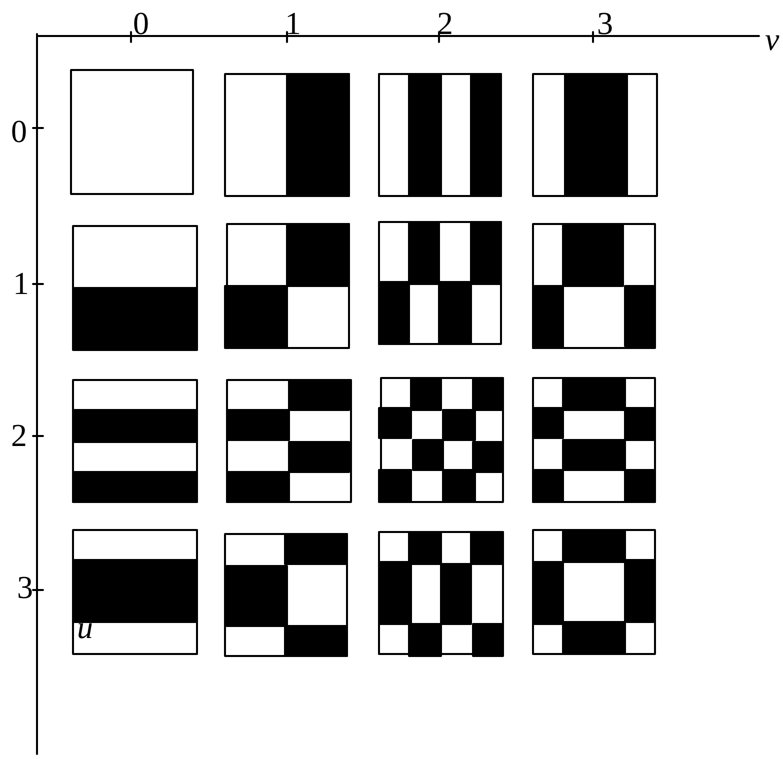



Figure 4.4 When N = 4, a set of 16 basis images of DWT

### 4.3.3 Discrete Hadamard Transform (DHT) [hadamard 1893]

- Closely related to the DWT.

$$f(x,y,u,v)=\frac{1}{N}\prod_{i=0}^{n}\left[(-1)^{p_i(x)p_i(u)}(-1)^{p_i(y)p_i(v)}\right]$$

$$i(x,y,u,v)=f(x,y,u,v),$$

$n$, $i$, and $p_i(\arg)$ are the same as in DWT.

- Term Walsh-Hadamard transform (DWHT) frequently used to represent either of the two transforms.

### 4.3.4 Discrete Cosine Transform (DCT)

#### 4.3.4.1 Background

- Most commonly used transform for image and video coding.
- Established by Ahmed, Natarajan and Rao [ahmed 1974].
- **The *basis member*** $\cos[(2x+1)up/2N]$**:** the *u*th Chebyshev polynomial $T_U(x)$ evaluated at the *x*th zero of $T_N(x)$.
- The basis vectors of 1-D DCT: a good approximation to the eigenvectors of the class of Toeplitz matrices defined as

$$\begin{bmatrix} 1 & r & r^2 & \cdots & r^{N-1} \\ r & 1 & r & \cdots & r^{N-2} \\ r^2 & r & 1 & \cdots & r^{N-3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ r^{N-1} & r^{N-2} & r^{N-3} & \cdots & 1 \end{bmatrix}$$

where $0 < r < 1$.

### 4.3.4.2 Transformation Kernel

$$f(x,y,u,v) = C(u)C(v)\cos\left(\frac{(2x+1)up}{2N}\right)\cos\left(\frac{(2y+1)vp}{2N}\right)$$

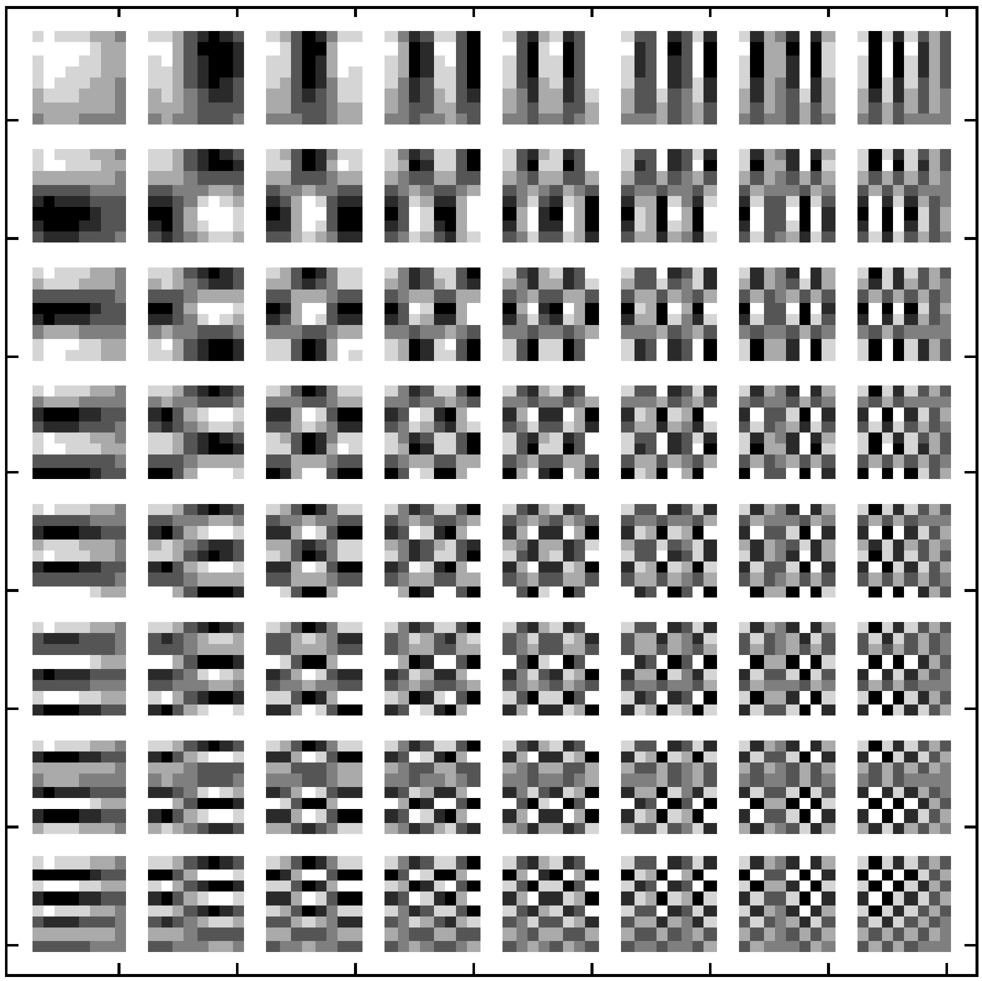
where

$$C(u) = \begin{cases} \sqrt{\frac{1}{N}} & for \quad u = 0 \\ \sqrt{\frac{2}{N}} & for \quad u = 1,2,\cdots,N-1 \end{cases}$$

$$i(x,y,u,v) = f(x,y,u,v).$$

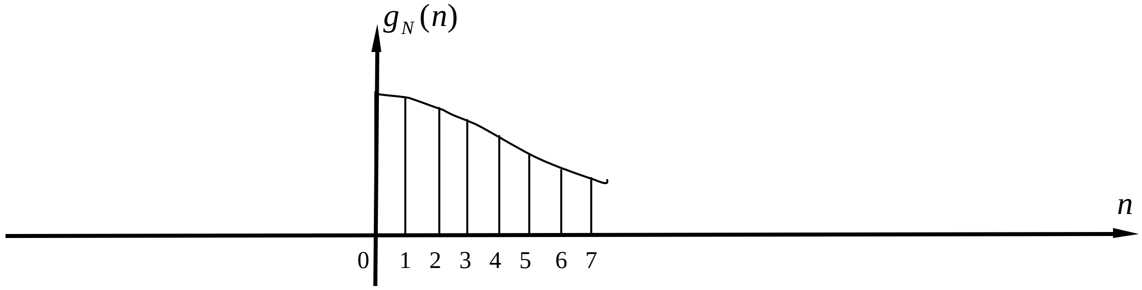
$C(v)$: defined in the same way

- The 64 basic images for N=8.

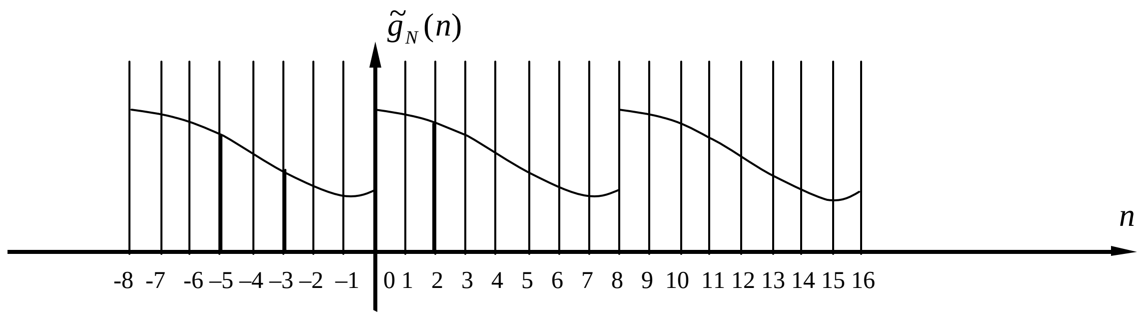


DCT basis images

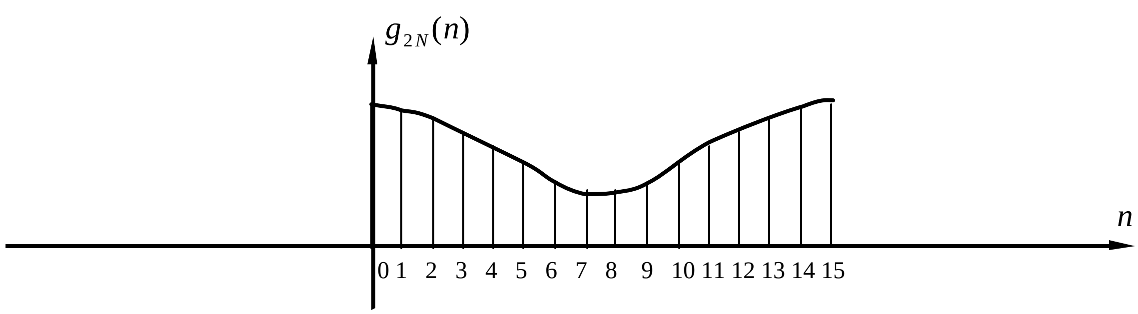
## 4.3.4.3 Relationship with DFT



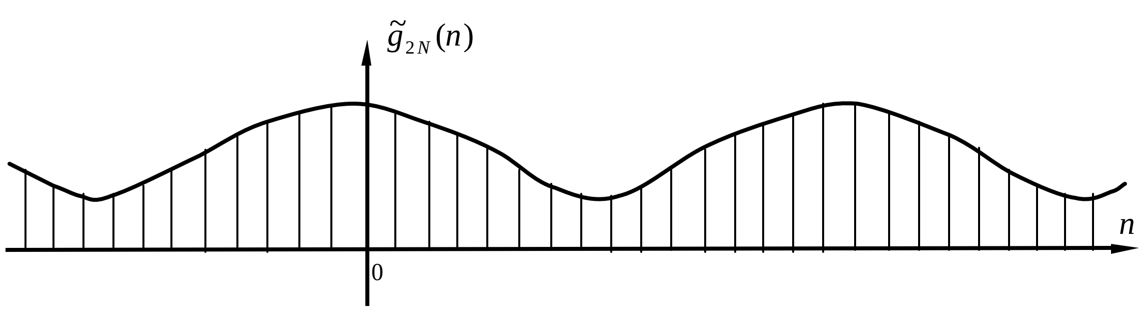

(a) Original 1-D input sequence




(b) Formation of a periodic sequence with a fundamental period of $N$ (DFT)




(c) Formation of a back-to-back $2N$ sequence




(d) Formation of a periodic sequence with a fundamental period of $2N$ (DCT)

Figure 4.5 An example to illustrate the differences and similarities between DFT and DCT.

- The DCT of an *N*-point sequence is obtained via the follows three steps.
    1) Form a *2N*-point sequence, $g_{2N}(n)$.

Then form a periodic sequence $\tilde{g}_{2N}(n)$.

Find FS coefficients of $\tilde{g}_{2N}(n)$.

    2) Only keep the *N* coefficients with indexes $0,1,\cdots,N-1$.

- These *N* FS coefficients: DCT of the given *N*-point sequence $g_N(n)$.

- Observation:
    - $\tilde{g}_N(n)$ not smooth: end-head discontinuities.

  - These discontinuities cause a high frequency distribution in the corresponding DFT.

- On the contrary, the $\tilde{g}_{2N}(n)$ does not have this type of discontinuity.

  $\Rightarrow$ DCT possesses better energy compaction in low frequencies than DFT.

- Note that the most energy of an image is contained in a small region of low frequency in the DFT domain. Vivid examples can be found in [gonzalez 1992].

- In terms of **energy compaction**, when compared with the optimal KLT, DCT is the best among DFT, DWT, DHT and discrete Harr transform.

- KLT: Not practical. Image dependent. Calculation involved complicated.

- **DCT can be implemented using the FFT**.

  DCT of an $N$-point sequence can be obtained from DFT of the $2N$-point sequence.

Even symmetry of $\tilde{g}_{2N}(n)$ makes computation required for the DCT of an *N*-point equal to that required for DFT of the *N*-point sequence.

- Because of these two merits, the DCT is the most popular image transform used in image and video coding.

## 4.3.5 Performance Comparison

### 4.3.5.1 Energy Compaction

### 4.3.5.2 Mean Square Reconstruction Error

- The performance can be compared in terms of

$$MSE_r = E\left[\left\|(\vec{z} - \vec{z}')\right\|^2\right] = \sum_{i=L+1}^{N} s_i^2,$$

where $\vec{z}'$ denotes the reconstructed vector.

- Note: quantization error and transmission error have not been included. Sometimes called: mean square approximation error.

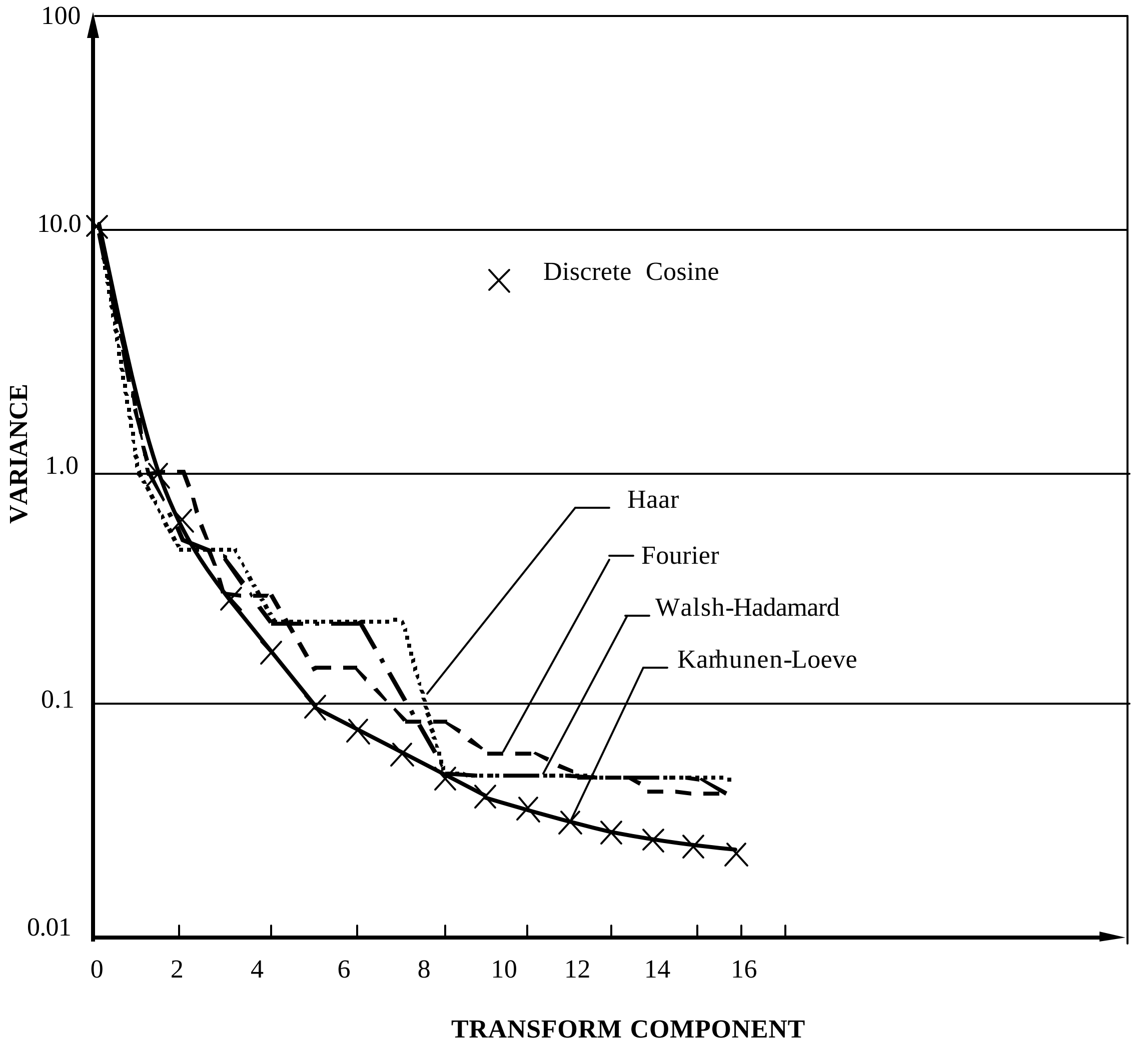



Figure 4.6 Transform coefficient variances when N=16, $\rho$=0.95 [ahmed 1974]

- A similar analysis can be carried out for the 2-D case [wintz 1972].

#### 4.3.5.3 Computational Complexity

## 4.4 Bit Allocation

- Bit allocation: truncation, quantization, and codeword assignment.
- Three types of error in TC:
  - Truncation error. That is, the majority of coefficients are truncated to zero.
  - Quantization error. (Note that truncation can also be considered a special type of quantization).
  - Transmission error
- Two different ways in truncation
  - Zonal coding
  - Threshold coding

### 4.4.1 Zonal Coding

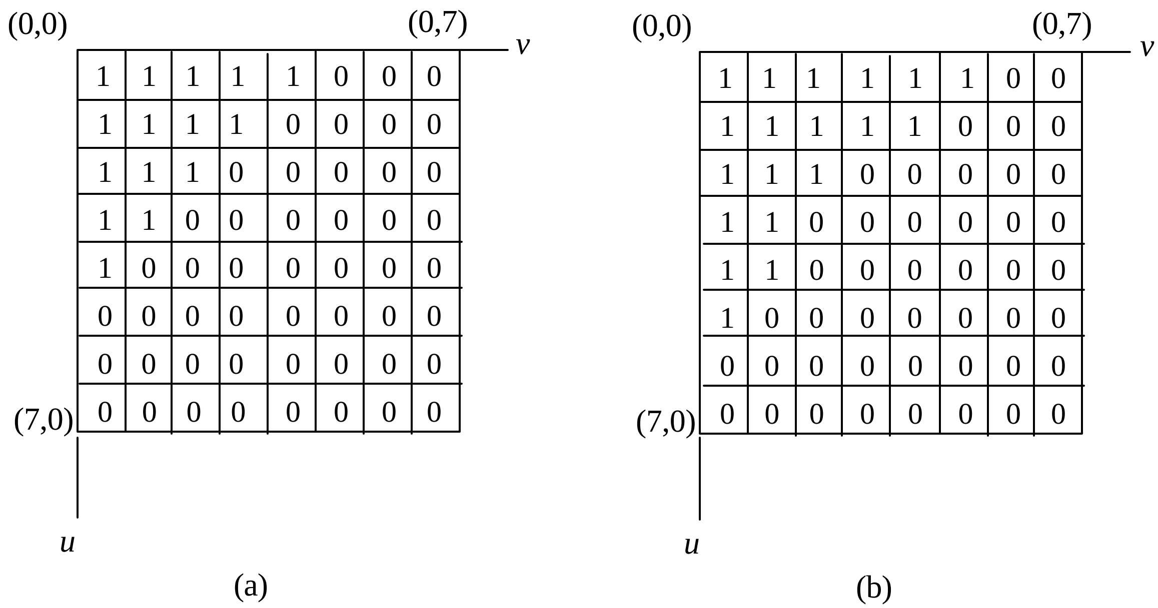



Figure 4.7 Two illustrations of zonal coding

- <u>No overhead side information</u>
- Coding efficiency, however, may not be high.
  - Some coefficients outside the zone might be large in magnitude
  - Some coefficients inside the zone may be small in quantity.

- For further improvement, an adaptive scheme has to be used.

### 4.4.2 Threshold Coding

- No a predefined zone.
- Instead, each transform coefficient compared with a threshold.
  - Smaller than threshold, set to zero.
  - Larger than threshold, retained for quantization and encoding.
- Adaptive
- Address of retained coefficients has to be sent to the receiver as side information.
- Threshold determined after an evaluation of all coefficients. ⇒ usually a two-pass adaptive technique.

❖ Chen and Pratt: an efficient adaptive coding scheme [chen 1984].

- A one-pass adaptive scheme. Hence, fast in implementation.
- With several effective techniques, it achieved excellent results in TC.

  Specifically, it demonstrated satisfied quality of reconstructed frames at a bit rate of **0.4 bits/pixel for coding of color images**, which corresponds to real-time color television transmission over a 1.5 Mbits/sec (Mbps) channel.
- This scheme was adopted by JPEG.

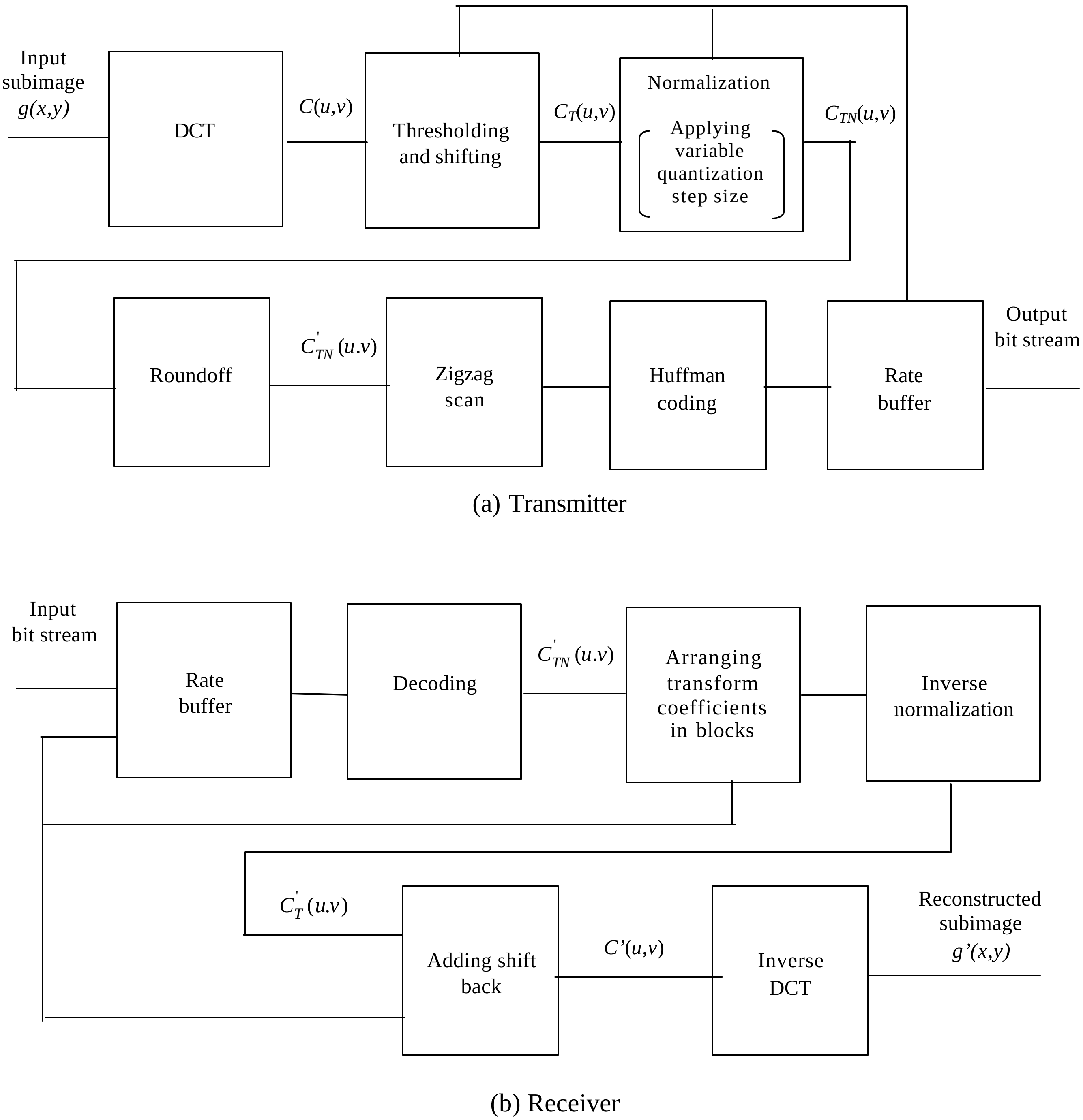



Figure 4.8 Block diagram of the algorithm proposed by Chen and Pratt [chen 1984]

#### 4.4.2.1 Thresholding and Shifting

- Formula

$$C_T(u,v) = \begin{cases} C(u,v) - T & \text{if} \quad C(u,v) > T \\ 0 & \text{if} \quad C(u,v) \le T \end{cases}$$

*T*: threshold.

- Input-output characteristic

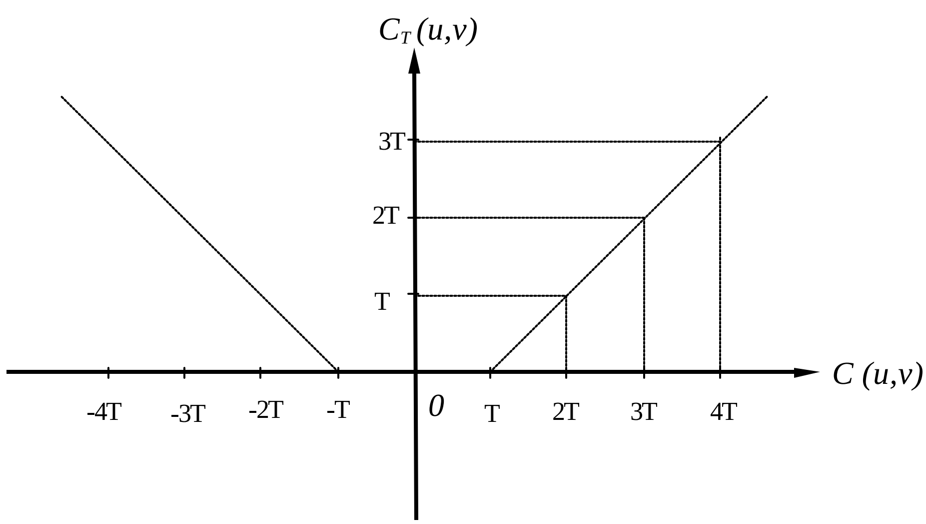



Figure 4.9 Input-output characteristic of thresholding and shifting

- Figure 4.10 demonstrates: more than 60% of DCT coefficients normally fall below a threshold value as low as 5; meaning that with a properly selected T value possible to set most of DCT coefficients equal to zero.

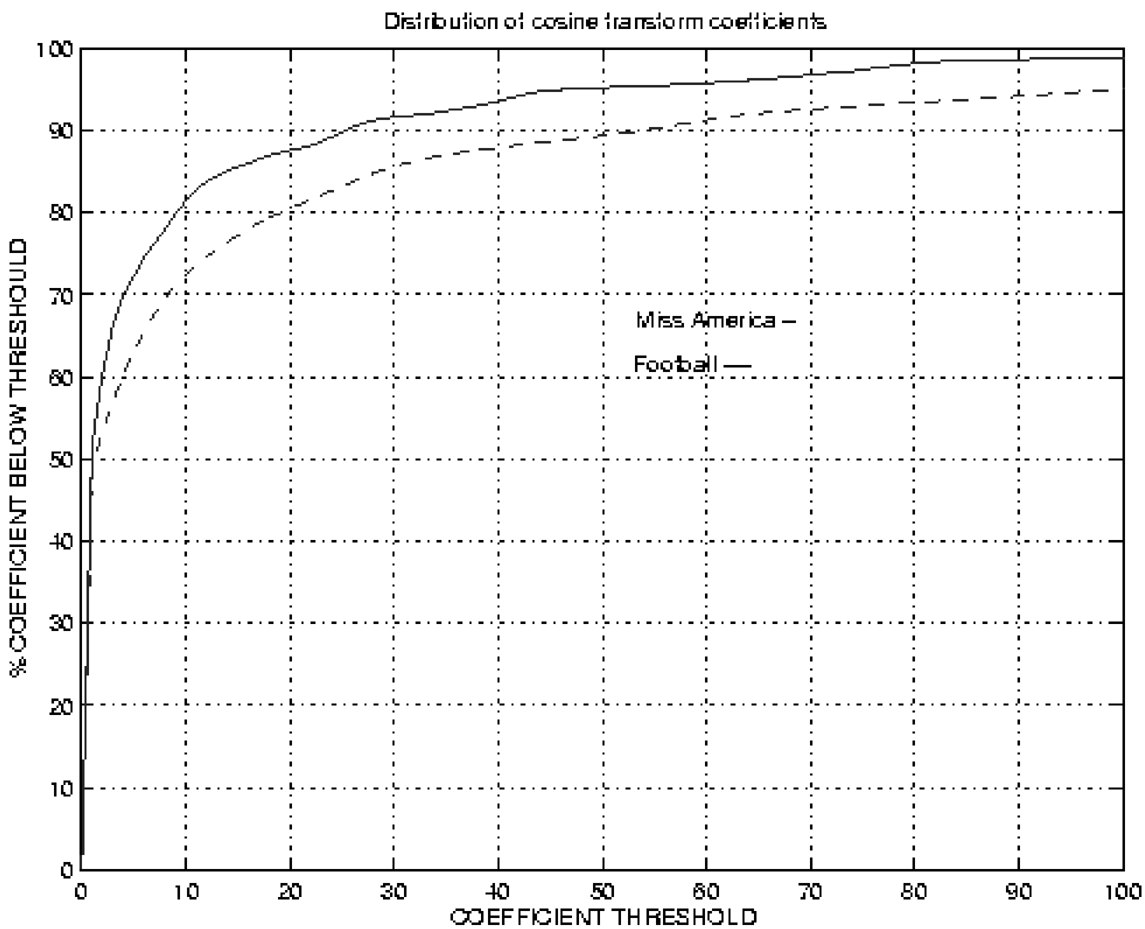



Figure 4.10 Distribution of DCT coefficients.

### 4.4.2.2 Normalization and Roundoff

- Normalization is implemented as follows:

$$C_{TN}(u,v) = \frac{C_T(u,v)}{\Gamma_{u,v}},$$

$C_T(u,v)$: T-subtracted DCT coeff.

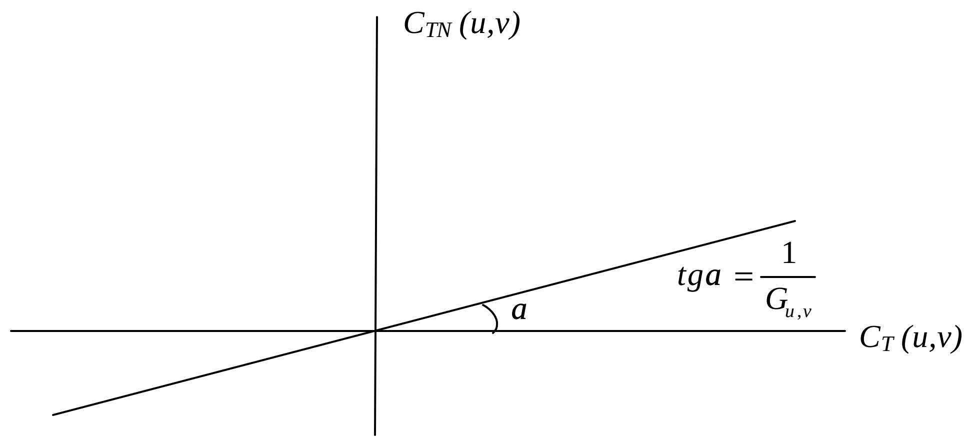
$\Gamma_{u,v}$: **Normalization factor** controlled by the rate buffer.

- The roundoff process converts floating point to integer to the nearest integer.

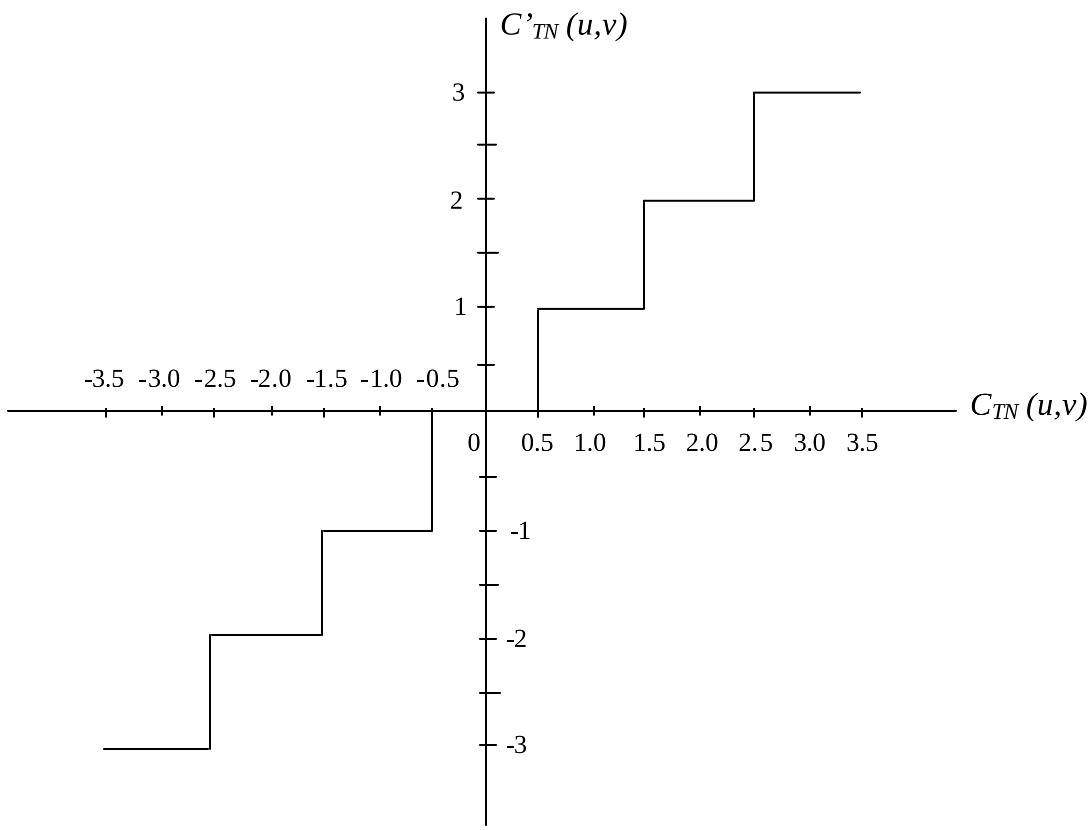
$$R[C_{TN}(u,v)] = C'_{TN}(u,v) = \begin{cases} \lfloor C_{TN}(u,v) + 0.5 \rfloor & if \quad C_{TN}(u,v) \geq 0 \\ \lceil C_{TN}(u,v) - 0.5 \rceil & if \quad C_{TN}(u,v) < 0 \end{cases}$$

Operator $\lfloor x \rfloor$: the largest integer smaller than or equal to $x$

Operator $\lceil x \rceil$: the smallest integer larger than or equal to $x$.




(a)




(b)

Figure 4.11 Input-output characteristic of (a) normalization (b) roundoff

- The roundoff: a uniform midtread quantizer with a unit quantization step.
- Normalization is a scaling process, which makes the resultant uniform midtread quantizer adapt to the dynamic range of the associated transform coefficient.
- Combination of normalization and roundoff: a uniform midtread quantizer with **quantization step size** equal to $\Gamma_{u,v}$.
- It is therefore possible for one quantizer design to be applied to various coefficients with different ranges.

♦ Obviously, by adjusting the $\Gamma_{u,v}$, variable bit rate and $MSE_q$ can be achieved.

♦ Selection of the $\Gamma_{u,v}$ for different transform coefficients can hence take statistical feature of images and characteristics of the HVS into consideration.

♦ In general, most image energy is contained in the DC and low frequency AC transform coefficients.

- ♦ HVS more sensitive to a relatively uniform region than to a relatively detailed region.
- ♦ HVS more sensitive to luminance component than to chrominance components.
- A **quantization table** in JPEG: A matrix consisting of all the normalization factors

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 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 |

(a) Luminance quantization table

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 17 | 18 | 24 | 47 | 99 | 99 | 99 | 99 |
| 18 | 21 | 26 | 66 | 99 | 99 | 99 | 99 |
| 24 | 26 | 56 | 99 | 99 | 99 | 99 | 99 |
| 47 | 66 | 99 | 99 | 99 | 99 | 99 | 99 |
| 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |
| 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |
| 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |
| 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |

(b) Chrominance quantization table

Figure 4.8 Quantization Tables

➢ In both tables the small normalization factors are assigned to the DC and low frequency AC coefficients.

- The large $\Gamma_{u,v}$s are associated with the high frequency transform coefficients.
- Compared with luminance quantization table, chrominance quantization table has larger quantization step sizes for the low and middle frequency coefficients and almost the same step sizes for the DC and high frequency coefficients, indicating that the chrominance components are relatively coarsely quantized, compared with the luminance component.

#### 4.4.2.3 Zigzag Scan

- Most quantized coefficients are zero.
- Run-length code (RLC), discussed in Chapter 6, is very efficient under these circumstances to encode the address information of nonzero coefficients.
- Run-length of zero coefficients is understood as the number of consecutive zeros in the zigzag scan.

- Zigzag scanning minimizes the use of RLC in the block.

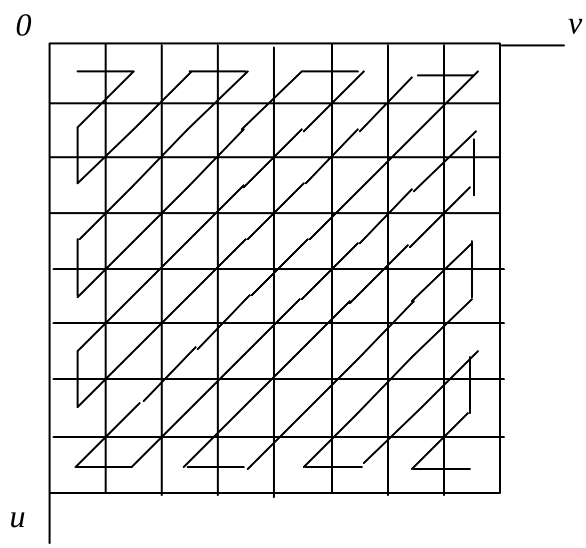



Figure 4.9 Zigzag scan of DCT coefficients within an 8×8 block

## 4.4.2.4 Huffman Coding

- Statistical studies of magnitude of nonzero DCT coefficients and run-length of zeros in zigzag scanning were conducted in [chen 1984].

- Domination of coefficients with small amplitude and short run-lengths was found.
- This justifies the application of Huffman coding to the magnitude of transform coefficients and run-lengths of zeros.

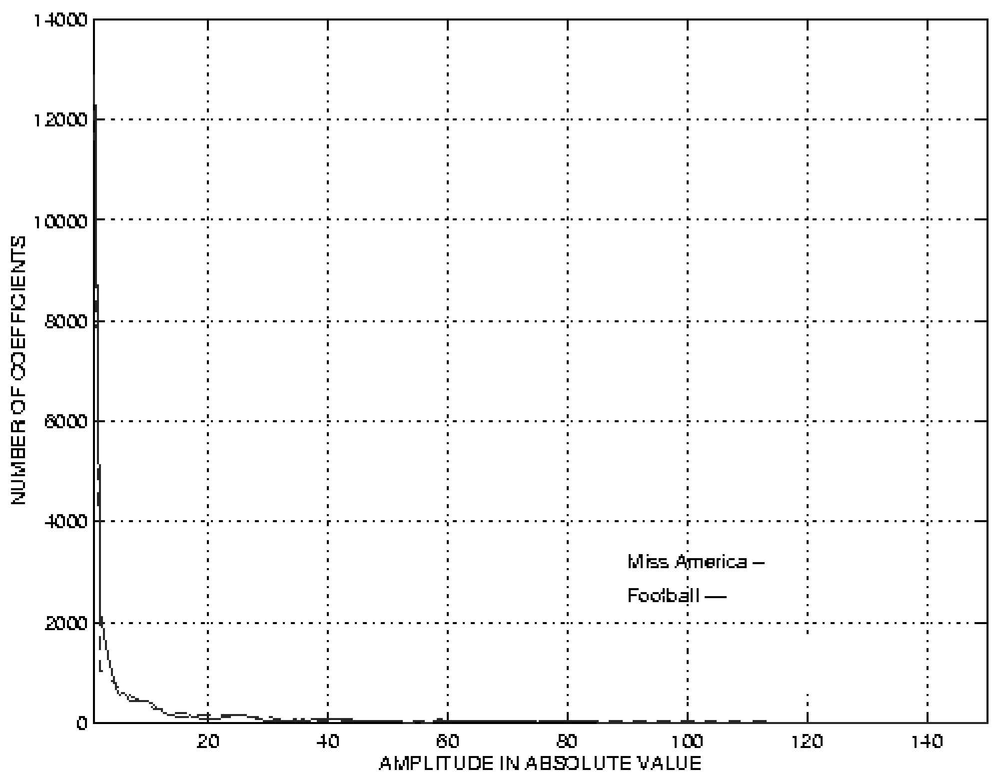



Figure 4.14 Histogram of DCT Coefficient Amplitude in Absolute Value

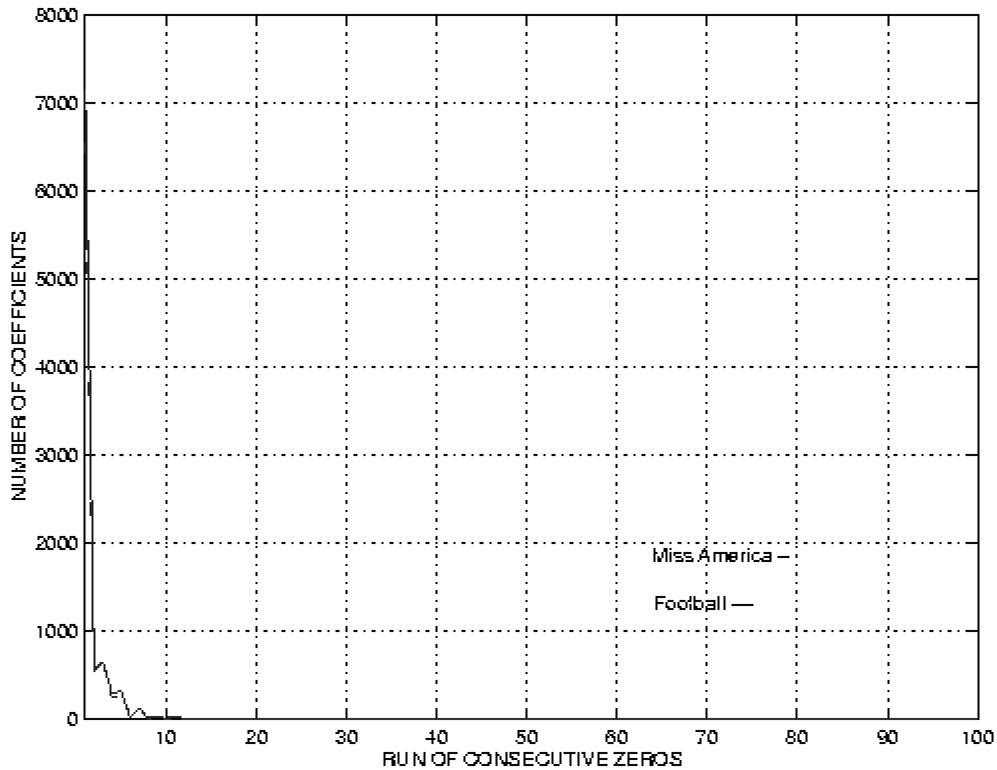



Figure 4.15 Histogram of runs of consecutive zeros

#### 4.4.2.5 Special Codewords

- Two special codewords were used in [chen 1984].
    - One: *end of block* (EOB).

      Once the last nonzero DCT coefficients in the zigzag, is coded, EOB is appended, indicating the termination of coding the block. This further saves bits used in coding.

- Another: *run-length prefix*.

  Run-length prefix is used to discriminate the run-length codewords from the amplitude codewords.

#### 4.4.2.6 Rate Buffer Feedback and Equalization

- In Figure 4.8, a rate buffer accepts a variable-rate data input from the encoding process and provides a fixed-rate data output to the channel.
- The status of the rate buffer is monitored and fed back to control the threshold and the normalization factor. In this fashion a one-pass adaptation is achieved.

# 4.5 Some Issues

## 4.5.1 Effect of Transmission Error

- In TC, each pixel in the reconstructed image relies on all transform coefficients in the subimage where the pixel is located.
- Hence, a bit reversal transmission error will be spread.
- As discussed, this is one of the reasons the selected subimage size cannot be very large.
- Depending on which coefficient is in error, the effect caused by a bit reversal error on the reconstructed image varies.
    - An error in the DC or a low frequency AC coefficient may be objectionable
    - An error in the high frequency coefficient may be less noticeable.

### 4.5.2 Reconstruction Error Sources

- As discussed, three sources contribute to the reconstruction error:
    - truncation
    - quantization
    - transmission
- Block artifacts

  Quantization and encoding of transform coefficients are also conducted blockwise.

  At the receiver, reconstructed blocks are put together to form the whole reconstructed image.
- Sometimes, even though it may not severely affect an objective assessment of the reconstructed image quality, block artifacts can be annoying to the HVS, especially when the coding rate is low.
- To alleviate the blocking effect, several techniques have been proposed.
    - One: overlapping blocks in source image.

- Another: post-filtering the reconstructed image along block boundaries.
- The selection of advanced transforms is an additional possible method [lim 1990].

- In the block overlapping method, when the blocks are finally organized to form the reconstructed image, each pixel in the overlapped regions takes an average value of all its reconstructed gray level values from multiple blocks.
- In this method, extra bits are used for those pixels involved in the overlapped regions. For this reason, the overlapped region is usually only one pixel wide.
- Due to the sharp transition along block boundaries, block artifacts are of high frequency in nature.
- Low pass filtering is hence normally used in the post-filtering method. To avoid the blurring effect caused by low pass filtering on the nonboundary mage area, low pass post-filtering is only applied to block boundaries.

- Post-filtering method has been adopted by the international coding standards since: no extra bits needed, better results.

### 4.5.3 Comparison between DPCM & TC

- Both utilize interpixel correlation, and are efficient coding techniques.
- Comparisons between these two techniques have been reported [habibi 1971b].
- In terms of computational complexity, DPCM is simpler than TC.

  Linear prediction and differencing in DPCM simpler than 2-D transform in TC.
- In terms of the memory requirement and processing delay, DPCM is superior to TC.
- The design of DPCM system, however, is sensitive to image-to-image variation, and so is its performance.

That is, an optimum DPCM design is matched to the statistics of a certain image. When the statistics change, the performance of the DPCM will be affected.

- On the contrary, TC is less sensitive to the variation in the image statistics.
- In general, the optimum DPCM coding system with a third or higher order predictor performs better than TC when the bit rate is about <u>2 to 3 bits per pixel</u> for single images.
- When the bit rate is <u>below 2 to 3</u> bits per pixel, TC is normally preferred.
- As a result, the JPEG is based on TC, whereas, in JPEG, DPCM is used for coding the DC coefficients of DCT, and information-preserving differential coding is use for lossless still image coding.

### 4.5.4 Hybrid Coding

- **Hybrid transform/waveform coding**

- In **waveform coding**, the waveform of a signal is coded. DPCM is a waveform coding technique.
- Hybrid coding combines TC and DPCM coding. That is, TC can be applied first rowwise followed by DPCM coding columnwise, or vice versa.
- In this way, the two techniques complement each other. That is, the hybrid coding technique simultaneously has TC's small sensitivity to variable image statistics and DPCM's simplicity in implementation.
- A successful hybrid coding scheme in interframe coding.
    - It uses motion compensated predictive coding. That is, the motion analyzed from successive frames is used to more accurately predict a frame.
    - The prediction error (in 2-D spatial domain) is transform coded.

- This hybrid coding scheme has been very efficient and was adopted by H.261, H.263, and MPEG 1, 2 and 4.

## 4.6 References

[ahmed 1974] N. Ahmed, T. Nararajan and K. R. Rao, "Discrete cosine transform," *IEEE Transactions on Computers*, pp. 90-93, Jan. 1974.

[andrews 1971] H. C. Andrews, "Multidimensional rotations in feature selection," *IEEE Transactions on Computers*, vol. c-20, pp. 1045-1051, Sept. 1971.

[chen 1977] W. H. Chen and C. H. Smith, "Adaptive coding of monochrome and color images," *IEEE Transactions on Communications*, vol. COM-25, pp. 1285-1292, Nov. 1977.

[chen 1984] W. H. Chen and W. K. Pratt, "Scene adaptive coder," *IEEE Transactions on Communications*, vol. COM-32, pp. 225-232, March, 1984.

[cooley 1965] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Math. of Comput*., vol. 19, pp. 297-301, 1965.

[hadamard 1893] J. Hadamard, "Resolution d'une question relative aux determinants," *Bull. Sci. Math.*, Ser. 2, vol. 17, Part I, pp. 240-246, 1893.

[habibi 1971a] A. Habibi and P. A. Wintz, "Image coding by linear transformations and block quantization," *IEEE Transactions on Communication Technology*, vol. com-19, pp. 50-60, February 1971.

[habibi 1971b] A. Habibi, "Comparison of nth-order DPCM encoder with linear transformations and block quantization techniques," *IEEE Transactions on Communication Technology*, vol. com-19, no. 6, pp. 948-956, December 1971.

[huang 1963] J.-Y. Huang and P. M. Schultheiss, "Block quantization of correlated Gaussian random variables," *IEEE Transactions on Communication Systems*, vol. cs-11, pp. 289-296, September 1963.

[jayant 1984] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*, Prentice Hall, 1984.

[lim 1990] J. S. Lim, *Two-Dimensional Signal and Image Processing*, Prentice Hall, 1990.

[pearl 1972] J. Pearl, H. C. Andrews and W. K. Pratt, "Performance measures for transform data coding," *IEEE Transactions on Communication Technology,* col. com-20, pp. 411-415, June 1972.

[reeve 1984] H. C. Reeve III and J. S. Lim, "Reduction of blocking effects in image coding," *Journal of Optical Engineering*, vol. 23, pp. 34-37, Jan./Feb. 1984.

[ramamurthi1986] B. Ramamurthi and A. Gersho, "Nonlinear space-variant postprocessing of block coded images," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 34, pp.1258-1267, Oct. 1986.

[tasto 1971] M. Tasto and P. A. Wintz, "Image coding by adaptive block quantization," *IEEE Transactions on Communication Technology*, vol. com-19, no. 6, pp. 957-972, Dec. 1971.

[walsh 1923] J. L. Walsh, "A closed set of normal orthogonal functions," *Am. J. Math.*, vol. 45, no. 1, pp. 5-24, 1923.

[wintz 1972] P. A. Wintz, "Transform picture coding," *Proceedings of The IEEE*, vol. 60, no. 7, pp.809-820, July 1972.

[zelinski 1975] R. Zelinski and P. Noll, "Adaptive block quantization of speech signals," (in German), Technical Report no. 181, Heinrich Hertz Institut, Berlin, 1975.