Fundamentals of Multimedia

Chapter 8

Lossy Compression Algorithms

Ze-Nian Li and Mark S. Drew

Outline

8.1 Introduction

- Lossless compression algorithms do not deliver *compression ratios* that are high enough.

- Hence, most multimedia compression algorithms are *lossy*.

- What is *lossy compression* ?
  - The compressed data is not the same as the original data, but a close approximation of it.
  - Yields a much higher compression ratio than that of lossless compression.

8.2 Distortion Measures

- The three most commonly used distortion measures in image compression are:

• *mean square error* (MSE) σ^2,

x_n : input data sequence

y_n : reconstructed data sequence

N : length of data sequence

$$\sigma^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n - y_n)^2$$

- *Signal to noise ratio (SNR)*, in decibel units (dB)

  σ^2_x : average square value of original data sequence

  σ^2_d : MSE

$$SNR = 10 \log_{10} \frac{\sigma_x^2}{\sigma_d^2}$$

- *Peak signal to noise ratio (PSNR)*, in decibel units (dB)

  For 8 bit image (video), $x_{peak} = 255$

$$PSNR = 10 \log_{10} \frac{x_{peak}^2}{\sigma_d^2}$$

8.3 The Rate-Distortion Theory

- Tradeoffs between Rate and Distortion (R-D).

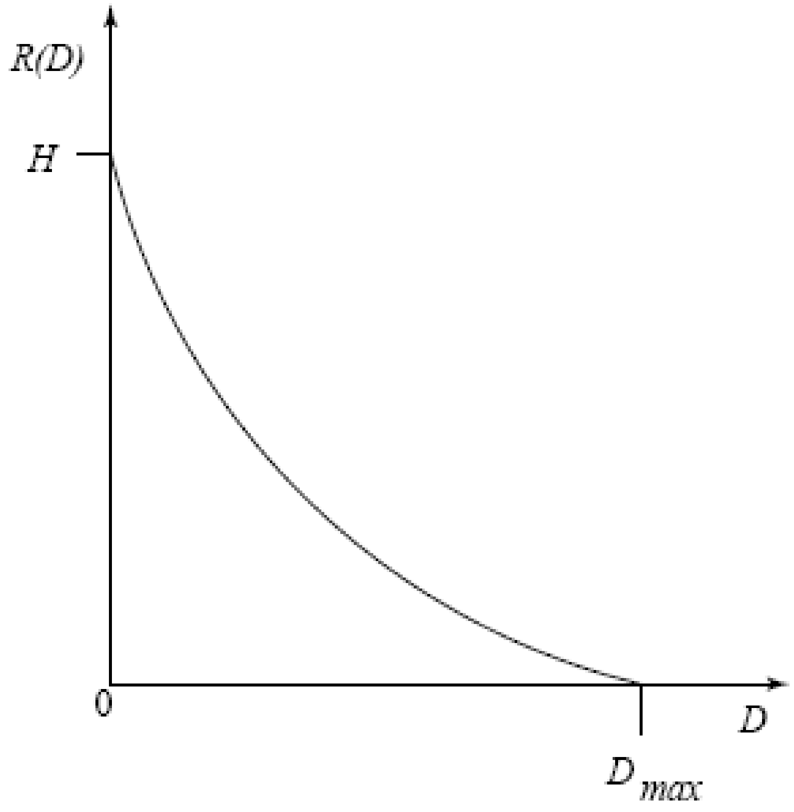


Fig. 8.1: Typical Rate Distortion Function.

8.4 Quantization

- Reduce the number of distinct output values to a much smaller set.

- Main source of the loss in lossy compression.

- Three different forms of quantization.
  - Uniform: midrise and midtread quantizers.
  - Non-uniform: companded (compress/expanded) quantizer.
  - Vector Quantization (VQ).

Uniform Scalar Quantization

- A uniform scalar quantizer partitions the domain of input values into equally spaced intervals.
  - The output or reconstruction value corresponding to each interval is taken to be the midpoint of the interval.
  - The length of each interval is referred to as the *step size*, denoted by the symbol Δ.

Uniform Scalar Quantization

- Two types of uniform scalar quantizers:
  - Midrise quantizers have even number of output levels.
  - Midtread quantizers have odd number of output levels, including zero as one of them

Uniform Scalar Quantization

- For the special case where $\Delta = 1$, we can simply compute the output values for these quantizers as:

$$Q_{midrise}(x) = \lceil x \rceil - 0.5$$

$$Q_{midtread}(x) = \lfloor x + 0.5 \rfloor$$

Uniform Scalar Quantization

- Performance of an M level quantizer.

  Let $B = \{b_0, b_1, \ldots, b_M\}$

  be the set of decision boundaries and

  $Y = \{y_1, y_2, \ldots, y_M\}$

  be the set of reconstruction or output values.

- Suppose the input is uniformly distributed in the interval $[-X_{max}, X_{max}]$. The rate of the quantizer is:

$$R = \lceil \log_2 M \rceil$$

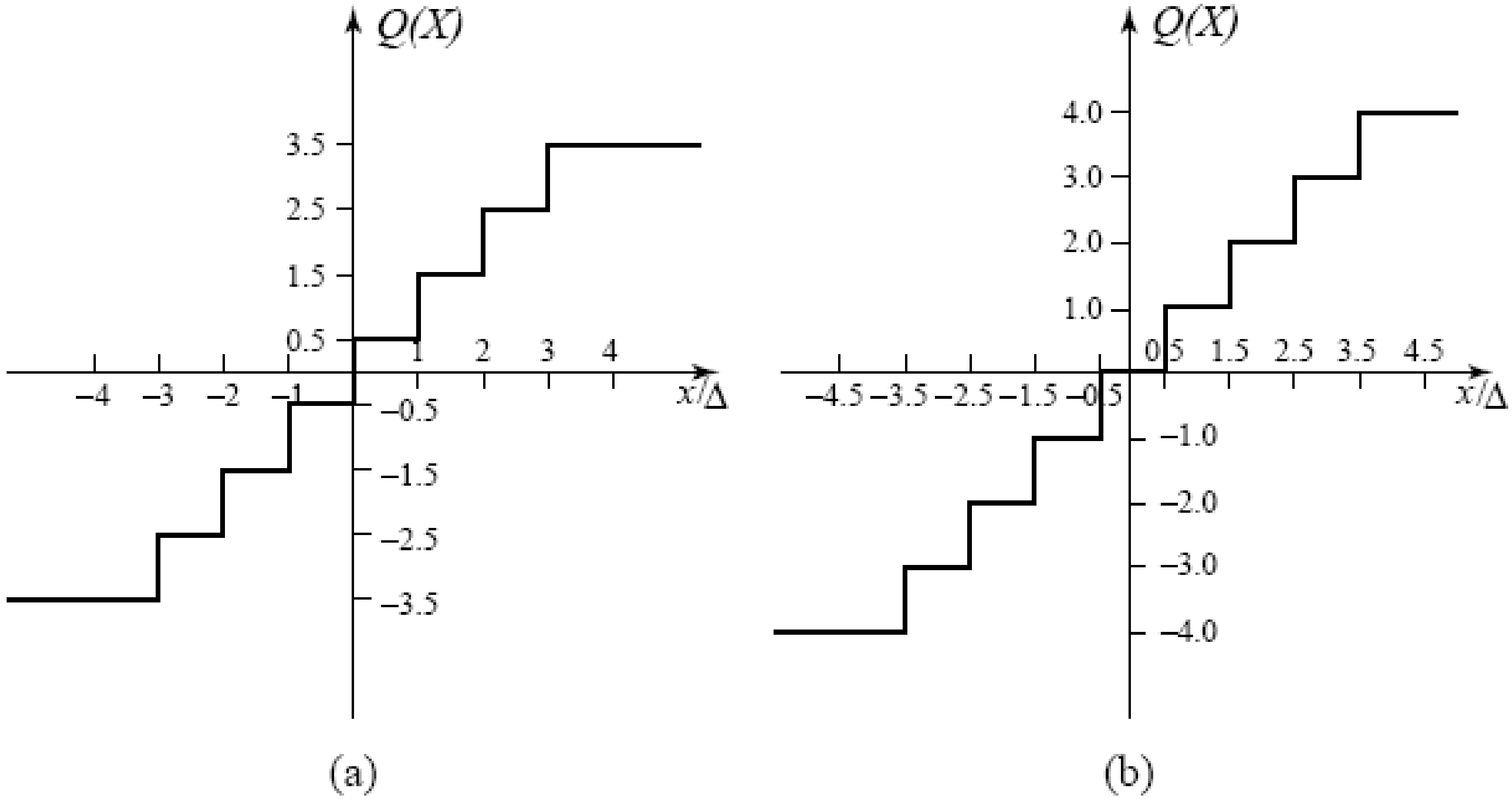


Fig. 8.2: Uniform Scalar Quantizers: (a) Midrise, (b) Midtread.

Quantization Error of Uniformly Distributed Source

- Since the reconstruction values y_i are the midpoints of each interval, the quantization error must lie within the values $[-\Delta/2, \Delta/2]$.

For a uniformly distributed source, the graph of the quantization error is shown in Fig. 8.3.

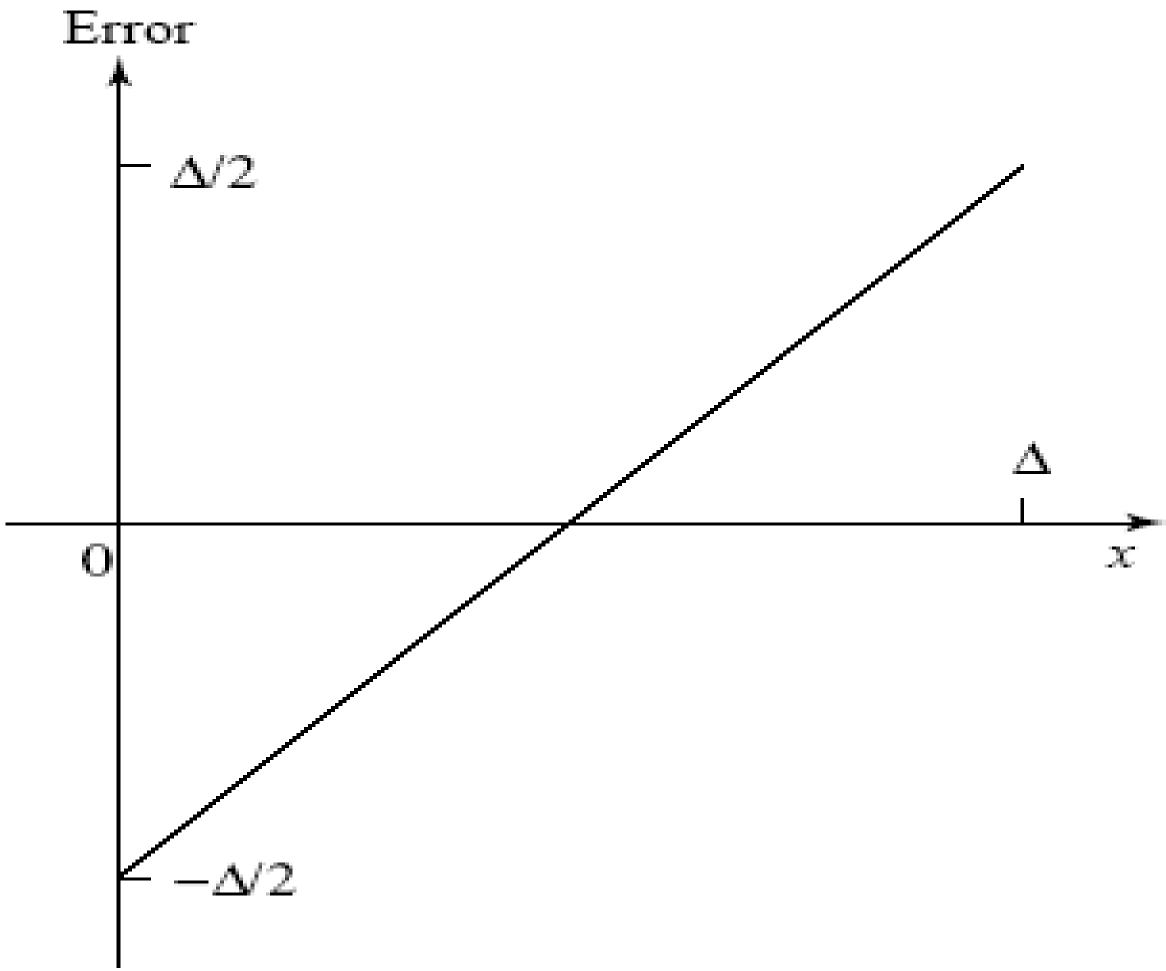


Fig. 8.3: Quantization error of a uniformly distributed source.

Vector Quantization

- According to Shannon's original work on information theory, any compression system performs better if it operates on vectors or groups of samples rather than individual symbols or samples.
- Form vectors of input samples by simply concatenating a number of consecutive samples into a single vector.
- Instead of single reconstruction values as in scalar quantization, in VQ, *code vectors* with *n* components are used.
- A collection of these code vectors form the *codebook*.

Vector Quantization

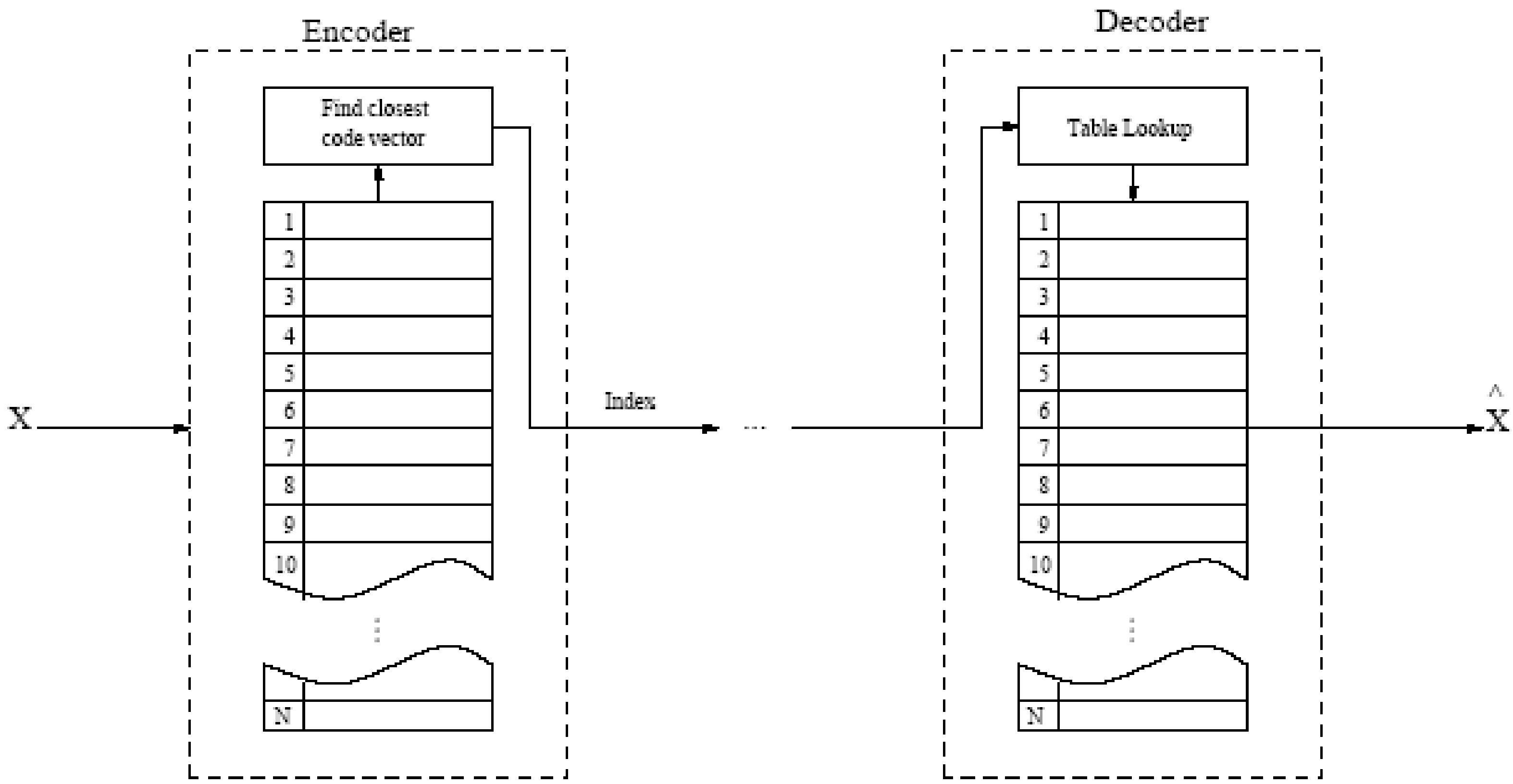


Fig. 8.5: Basic vector quantization procedure.

8.5 Transform Coding

- The rationale behind transform coding:
  If Y is the result of a linear transform T of the input vector X in such a way that the components of Y are much less correlated, then Y can be coded more efficiently than X.

- If most information is accurately described by the first few components of a transformed vector, then the remaining components can be coarsely quantized, or even set to zero, with little signal distortion.

Spatial Frequency and DCT

- Spatial frequency indicates how many times pixel values change across an image block.
- The DCT formalizes this notion with a measure of how much the image contents change in correspondence to the number of cycles of a cosine wave per block.
- The role of the DCT is to decompose the original signal into its DC and AC components; the role of the IDCT is to reconstruct (re-compose) the signal.

Definition of DCT

- f(i,j): spatial domain values
- F(u,v): (spatial) frequency domain values

  frequency values

$$F(u,v) = \frac{2\,C(u)\,C(v)}{\sqrt{MN}} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \cos\frac{(2i+1)\cdot u\pi}{2M} \cdot \cos\frac{(2j+1)\cdot v\pi}{2N} \cdot f(i,j) \tag{8.15}$$

- i, u: 1, …, M, j, v: 1, …, N

$$C(\xi) = \begin{cases} \frac{\sqrt{2}}{2} & if \quad \xi = 0, \\ 1 & otherwise. \end{cases} \tag{8.16}$$

2D Discrete Cosine Transform (2D DCT)

$$F(u,v) = \frac{C(u)\,C(v)}{4} \sum_{i=0}^{7} \sum_{j=0}^{7} \cos\frac{(2i+1)u\pi}{16} \cos\frac{(2j+1)v\pi}{16} f(i,j)$$

i, j, u, v: 0, 1, …,7

2D Inverse Discrete Cosine Transform (2D IDCT)

$$\tilde{f}(i,j) = \sum_{u=0}^{7}\sum_{v=0}^{7} \frac{C(u)\,C(v)}{4} \cos\frac{(2i+1)u\pi}{16} \cos\frac{(2j+1)v\pi}{16} F(u,v)$$

i, j, u, v: 0, 1, …,7

1D Discrete Cosine Transform (1D DCT)

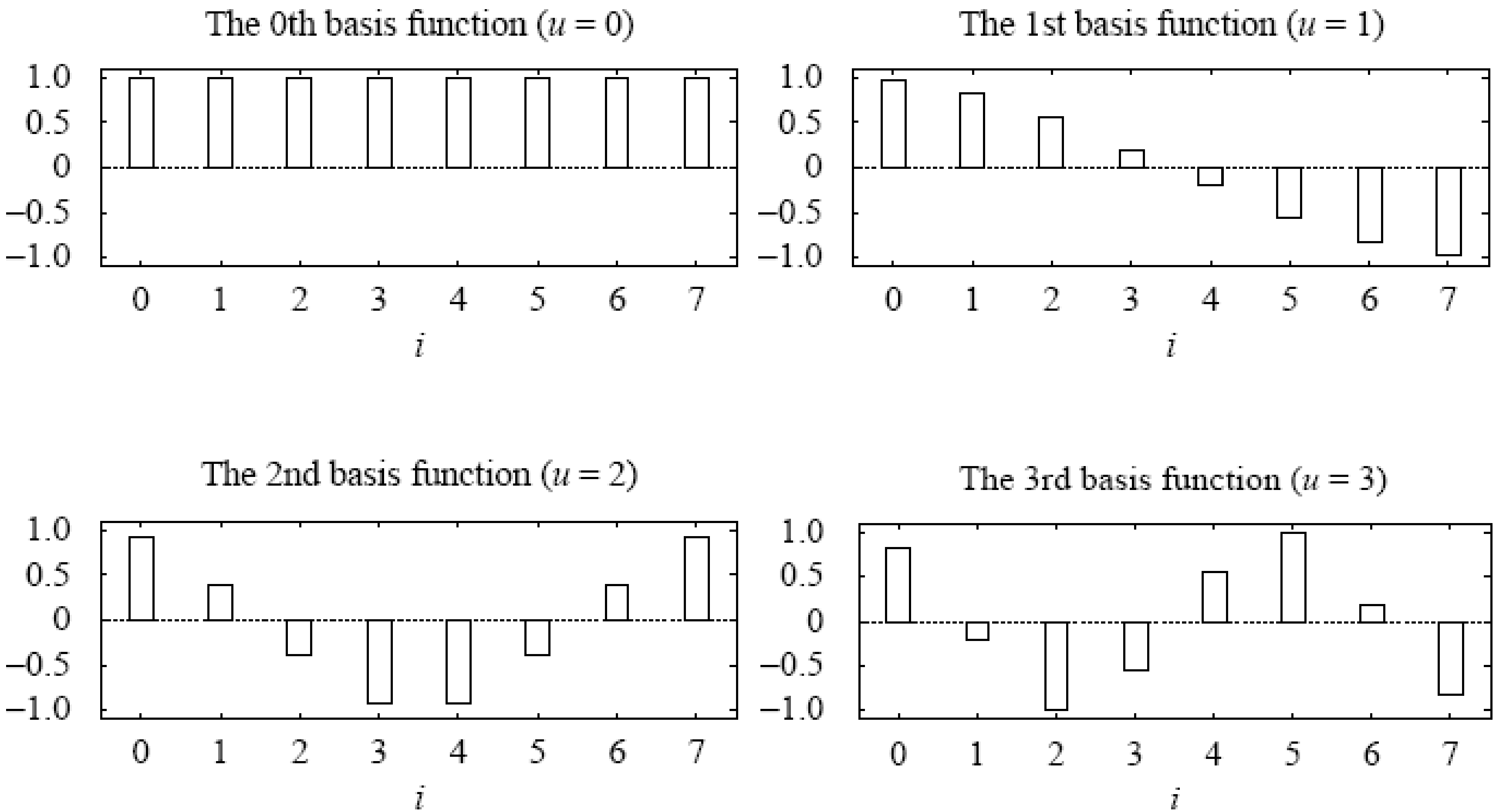
$$F(u) = \frac{C(u)}{2} \sum_{i=0}^{7} \cos \frac{(2i+1)u\pi}{16} f(i)$$

i, u: 0, 1, …,7

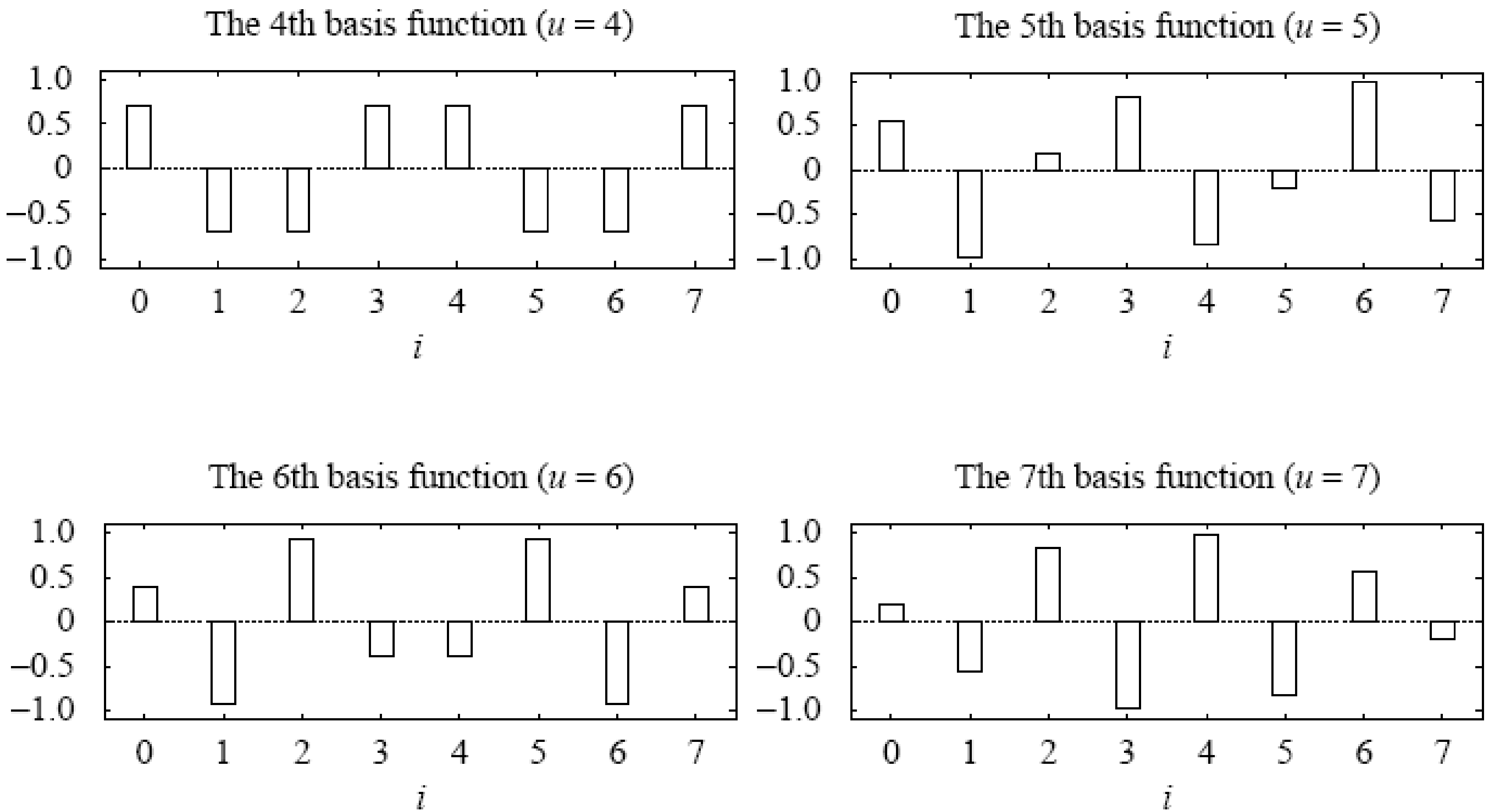
1D Inverse Discrete Cosine Transform (1D IDCT)

$$\tilde{f}(i) = \sum_{u=0}^{7} \frac{C(u)}{2} \cos \frac{(2i+1)u\pi}{16} F(u)$$

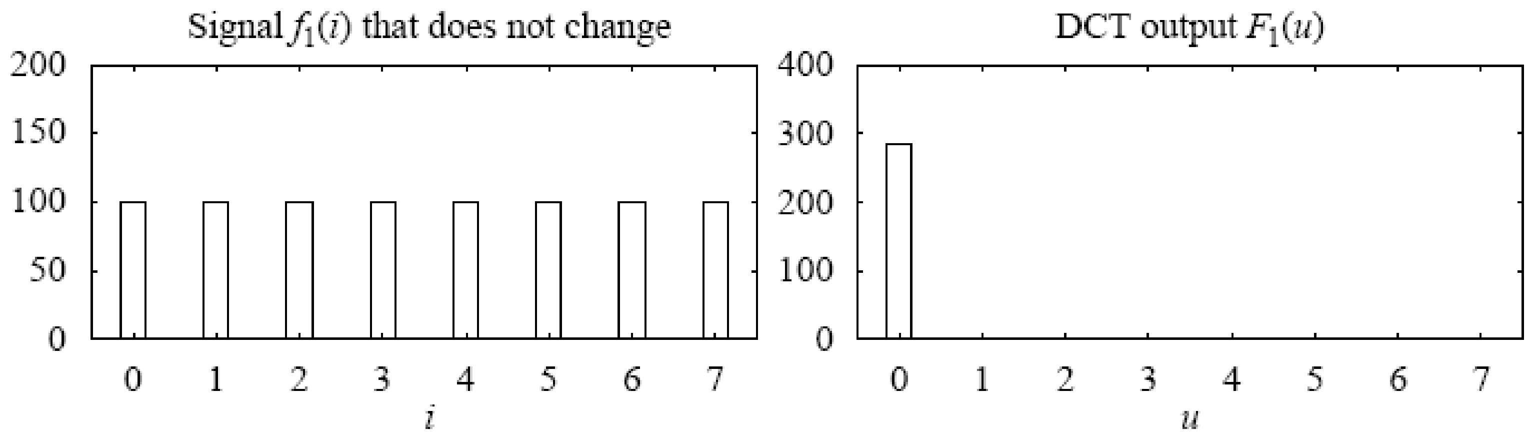
i, u: 0, 1, …,7



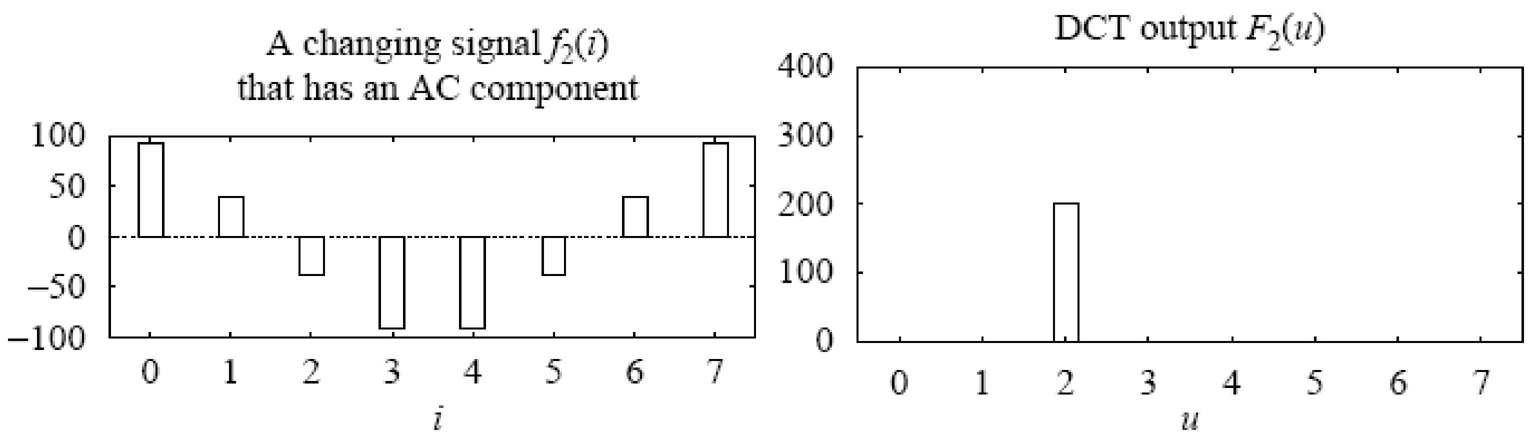
The 1D DCT basis functions.



The 1D DCT basis functions.

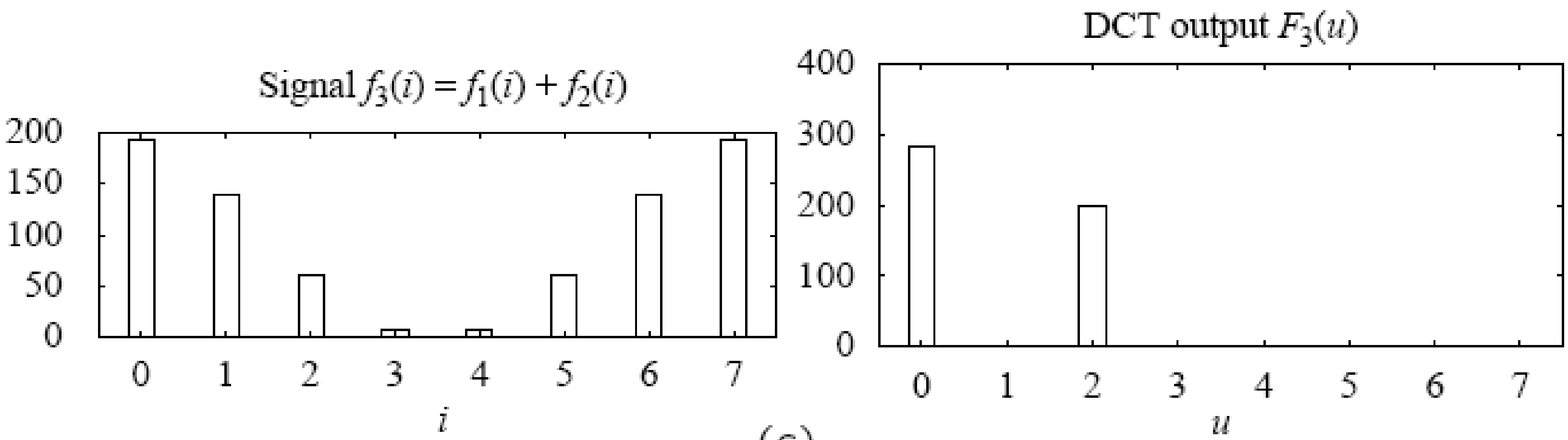


(a)

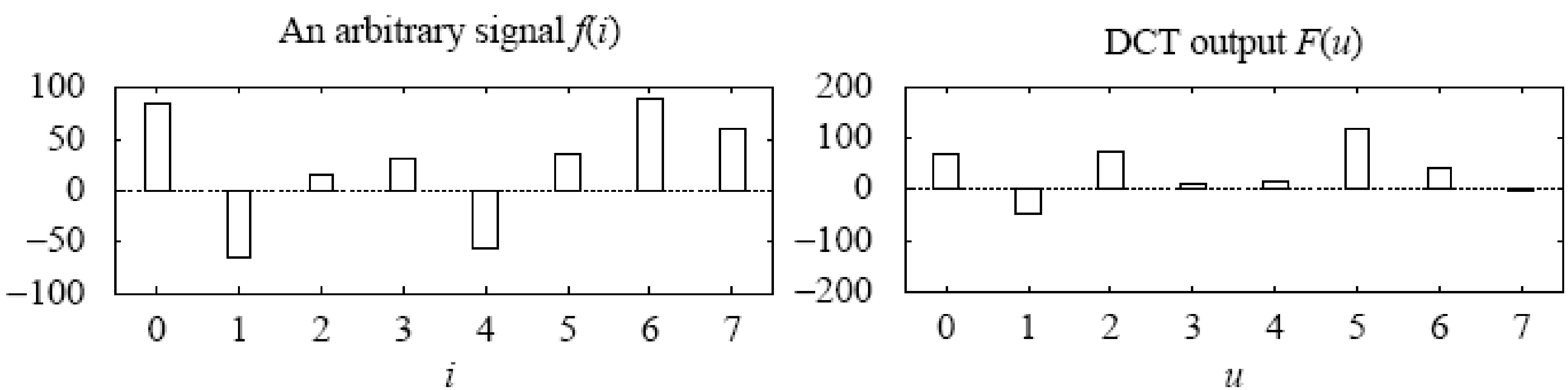


(b)

The Examples of 1D Discrete Cosine transform: (a) A DC signal $f_1(i)$, (b) An AC signal $f_2(i)$.

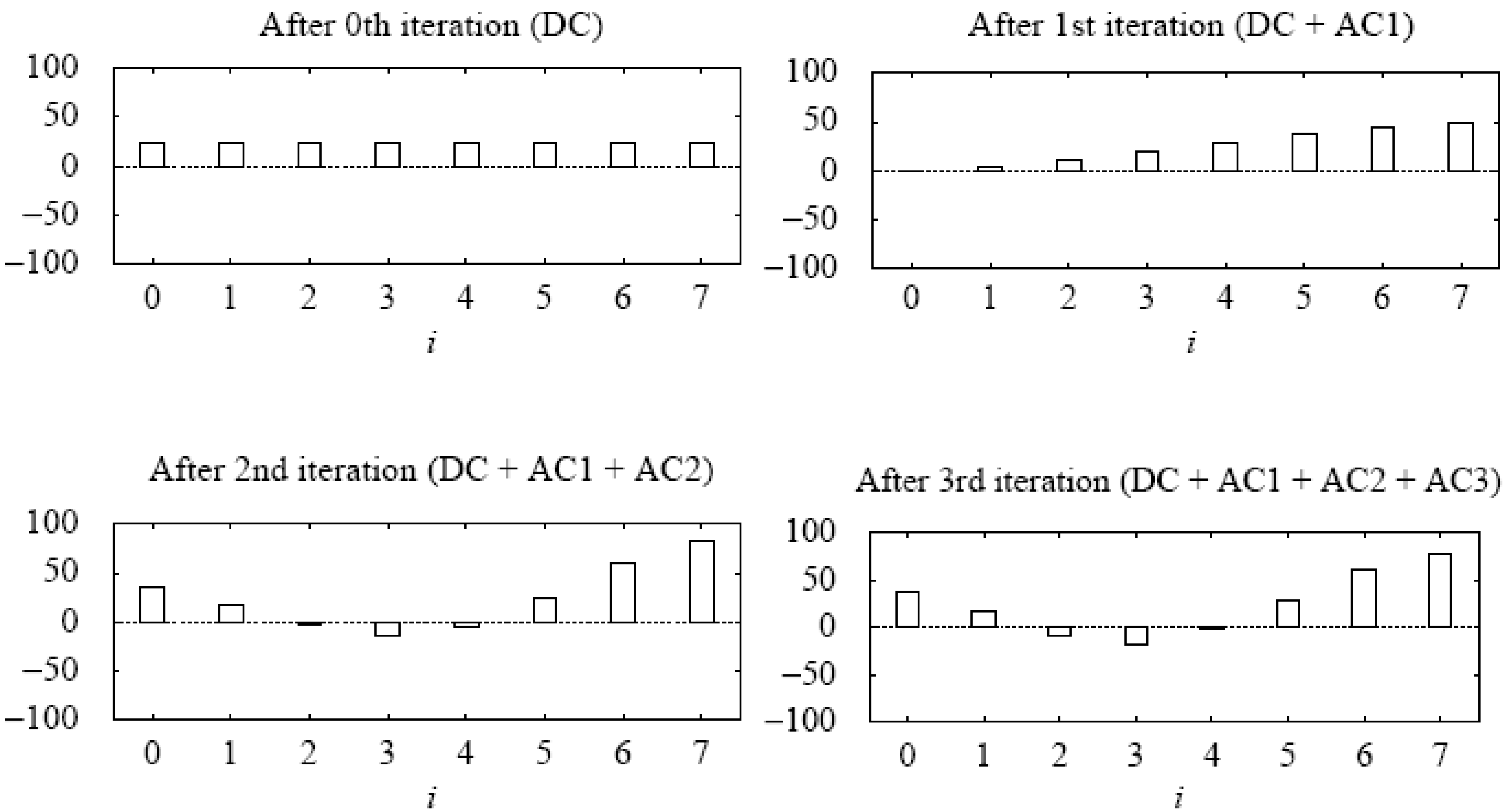


(c)

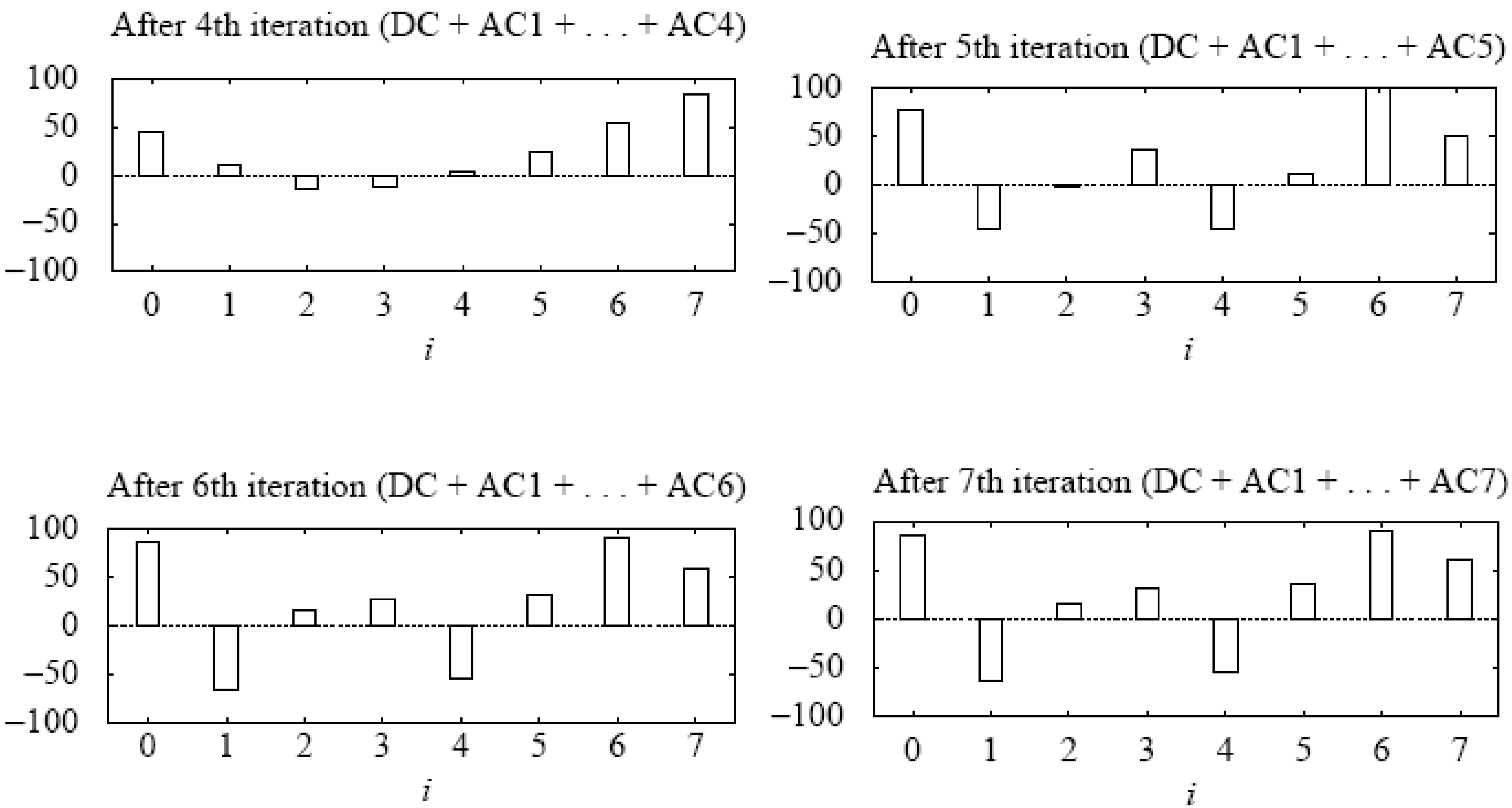


(d)

Examples of 1D Discrete Cosine Transform: (c) $f_3(i) = f_1(i) + f_2(i)$, (d) an arbitrary signal $f(i)$.



An example of 1D IDCT.



An example of 1D IDCT.

DCT is Linear Transform

- In general, a transform T (or function) is linear, iff

$$\mathcal{T}(\alpha p + \beta q) = \alpha \mathcal{T}(p) + \beta \mathcal{T}(q)$$

where α and β are constants, p and q are any functions or variables.

- This property can readily be proven for the DCT because it uses only simple arithmetic operations.

Cosine Basis Functions

- Function $B_p(i)$ and $B_q(i)$ are orthogonal, if

$$\sum_i [B_p(i) \cdot B_q(i)] = 0 \qquad if \quad p \neq q$$

- Function $B_p(i)$ and $B_q(i)$ are orthonormal,

  if they are orthogonal and

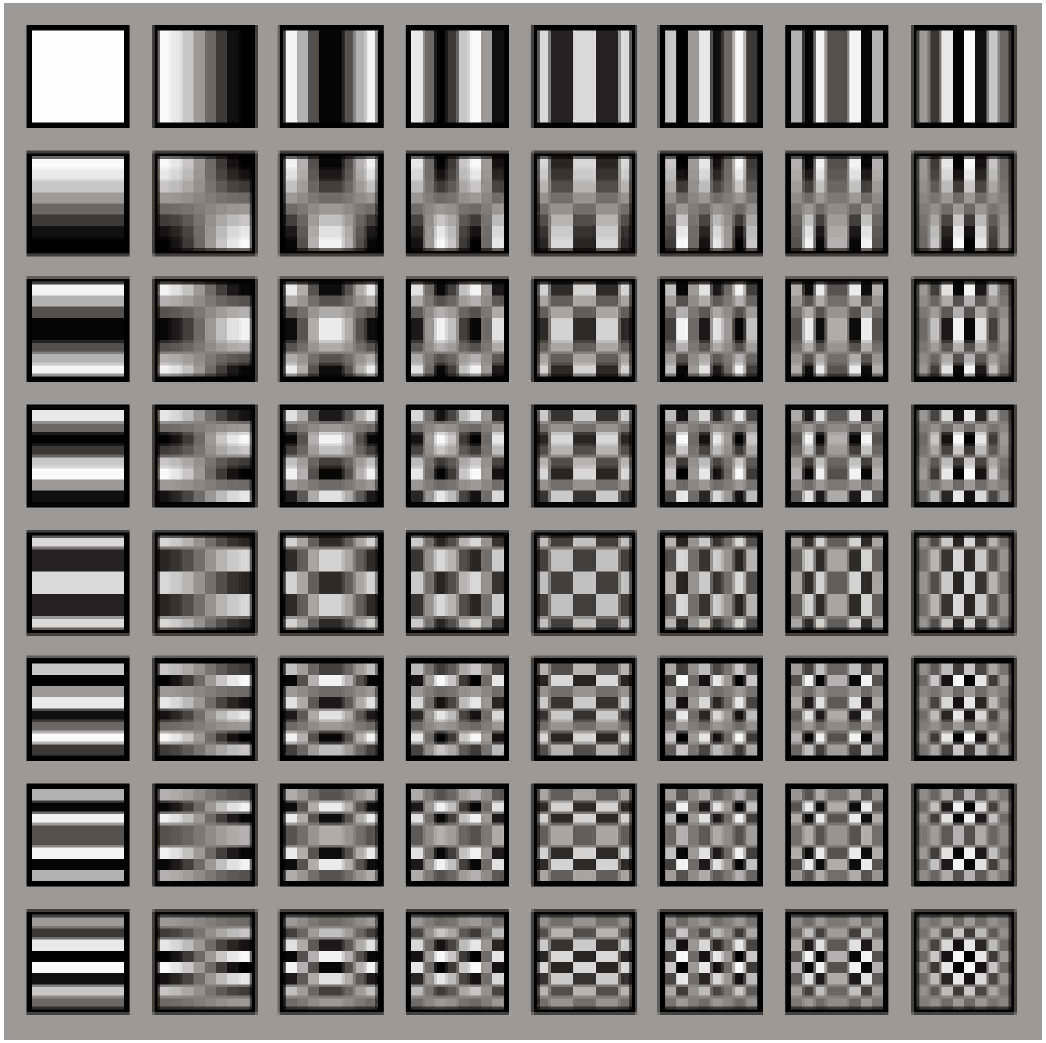
$$\sum_i [B_p(i) \cdot B_q(i)] = 1 \qquad if \quad p = q$$

Cosine Basis Functions

- It can be shown that:

$$\sum_{i=0}^{7}\left[\cos\frac{(2i+1)\cdot p\pi}{16}\cdot\cos\frac{(2i+1)\cdot q\pi}{16}\right] = 0 \quad if \quad p \neq q$$

$$\sum_{i=0}^{7}\left[\frac{C(p)}{2}\cos\frac{(2i+1)\cdot p\pi}{16}\cdot\frac{C(q)}{2}\cos\frac{(2i+1)\cdot q\pi}{16}\right] = 1 \quad if \quad p = q$$



Graphical Illustration of 8×8 2D DCT basis.

2D Separable Basis

- The 2D DCT can be separated into a sequence of two, 1D DCT steps:

$$G(i,v) = \frac{1}{2}C(v)\sum_{j=0}^{7}\cos\frac{(2j+1)v\pi}{16}f(i,j)$$

$$F(u,v) = \frac{1}{2}C(u)\sum_{i=0}^{7}\cos\frac{(2i+1)u\pi}{16}G(i,v)$$

- This simple change saves many arithmetic steps. The number of iterations required is reduced from 8×8 to $8+8$.