Network & Media Laboratory

Fundamentals of Multimedia

Authored by Ze-Nian Li Mark S. Drew

Lecturer：Lu Dongming

7、Lossy Compression Algorithms

Content

- Distortion Measures
- The Rate-Distortion Theory
- Quantization
- Transform Coding
- Wavelet-Based Coding

1. Distortion Measures

1.1 Concept of Distortion

- Distortion Measure
  - A mathematical quantity: specifies how close an approximation to its original
  - It's nature to think of the numerical difference
  - When it comes to image data, difference may not yield the intended result
  - Measures of perceptual distortion

1.2 Numerical Distortion Measures

- Many numerical distortion measures -- the most commonly used distortion measures are presented： MSE、SNR、PSNR
- Mean Square Error（MSE）：

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

  - Average pixel difference
- Signal-to-Noise Ratio（SNR):

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

  - The size of the error relative to the signal
- Peak-Signal-to-Noise Ratio (PSNR):

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

  - The size of the error relative to the peak value of the signal

1.2 Numerical Distortion Measures

- Examples of PSNR and corresponding images

original image

polluted by noise

PSNR=18.24

processed by noise filter

PSNR=39.5

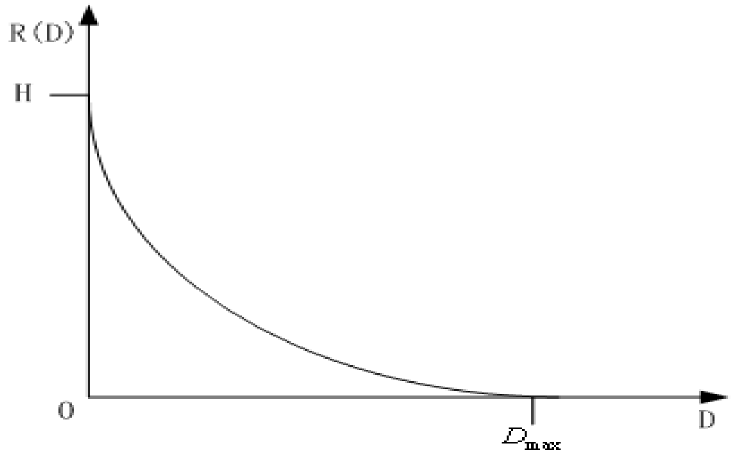
2.The Rate-Distortion Theory

2.1 Concept

- Lossy compression always involves a **tradeoff between rate and distortion**
  - **Rate** -- the average number of bits required to represent each source symbol;
  - **R(D)** note rate-distortion function;
- What is R(D)?
  - R(D) specifies **the lowest rate** at which the source data can be encoded while **keeping the distortion bounded above by D**
  - **At D＝0**, no loss, so is the entropy of the source data
  - Describe **a fundamental limit** for the performance of a coding algorithm
  - Can be used to **evaluate the performance of different algorithm**

2.2 A Typical R-D Function

- A figure of a typical rate-distortion function



- D＝0, the entropy of the source data
- R(D)＝0, nothing coded
- For a given source, it's **difficult to find a closed-form analytic description** of the rate-distortion function

3. Quantization

3.1 Functions of Quantization

- Quantization: the heart of any lossy scheme
  - Without quantization, almost no losing information
  - Reduce the number of distinct values via quantization
- Each quantizer has its unique partition of the input range and the set of output values.
  - Scalar quantizer
  - Vector quantizer

3.2 Uniform Scalar Quantization

- Uniform scalar quantizer
  - Partitions the input domain into equally spaced intervals
  - Decision boundaries: the end points of partition intervals
  - Output value: midpoint of the interval
  - Step size: the length of each interval
- Two types of uniform scalar quantizer
  - midrise： with an even number of output levels, one partition interval brackets zero；
  - midtread： odd number of output levels, zero is an output value.
- The goal of a successful uniform quantizer
  - Minimize the distortion for a given source input with a desired number of output values

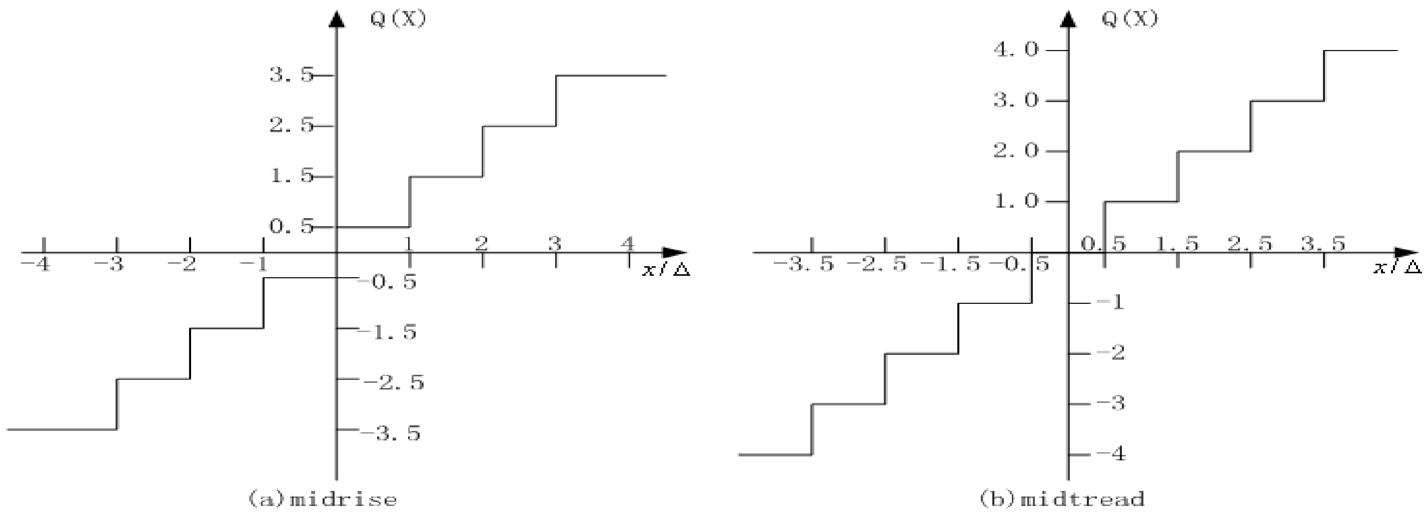
3.2 Uniform Scalar Quantization

- Given step size $\Delta=1$, output values for the two type of Quantizers be computed as:

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

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

- Two types quantizers:



3.2 Uniform Scalar Quantization

- Performance of a M level quantizer:
  - Decision Boundaries: B= { b_0, b_1, …, bM }
  - The set of output values: Y={y_1, y_2, …, yM}
  - The input is uniformly distributed: [-X_{max}, X_{max}]
  - The rate of quantizer: $R = \log_2^M$ is the number of bits required to code M things;
  - Step size is given by: $\Delta = 2X_{max}/M$
  - Granular distortion: error caused by the quantizer for bounded input
  - Overload distortion: error caused by quantizer for input values larger than X_{max} or smaller than -X_{max}

3.2 Uniform Scalar Quantization

- **Granular distortion for a midrise quantizer**
  - **Decision boundaries bi:[(i-1)Δ, iΔ], i=1..M/2, covering positive data X (another for native X values)**
  - **Output values yi: iΔ-Δ/2, i=1..M/2**
  - **The total distortion: twice the sum over the positive data:**

$$D_{gran} = 2\sum_{i=1}^{\frac{M}{2}} \int_{(i-1)\Delta}^{i\Delta} (x - \frac{2i-1}{2}\Delta)^2 \frac{1}{2X_{\max}} dx$$

  - **The error value at X is e(x)=x-Δ/2, variance of errors:**

$$\sigma_d^2 = \frac{1}{\Delta}\int_0^{\Delta} (e(x) - \bar{e})^2 dx = \frac{1}{\Delta}\int_0^{\Delta} (x - \frac{\Delta}{2} - 0)^2 dx = \frac{\Delta^2}{12}$$

3.2 Uniform Scalar Quantization

- Signal variance $\sigma_x^2 = (2X_{\max})^2/12$ ； if the quantizer is n bits， M＝2^n
- SQNR can be calculated as：

$$
\begin{aligned}
SQNR &= 10\log_{10}\left(\frac{\sigma_x^2}{\sigma_d^2}\right) \\
&= 10\log_{10}\left(\frac{(2X_{\max})^2}{12}\cdot\frac{12}{\Delta^2}\right) \\
&= 10\log_{10}\left(\frac{(2X_{\max})^2}{12}\cdot\frac{12}{\left(\frac{2X_{\max}}{M}\right)^2}\right) \\
&= 10\log_{10} M^2 = 20n\cdot\log_{10}2 \\
&= 6.02n(dB)
\end{aligned}
$$

3.3 Nonuniform Scalar Quantization

- If the input source is not uniformly distributed, a uniform quantizer may be inefficient.
- Increasing the number of decision levels within the densely distributed region can lower granular distortion
- Enlarge the region where the source is sparsely distributed can keep the total number of decision levels
- So nonuniform quantizers have nonuniforumly defined decision boundaries.
- Two common approaches for nonuniform quantization：
  - The Lloyd-Max Quantizer
  - The companded quantizer

Network & Media Laboratory

4. Transform Coding

4.1 Basic Idea

- According principles of information theory
  - Coding vectors is more efficient than coding scalars
  - Need to group consecutive samples from input into vectors
- Let $X = \{x_1, x_2, \ldots, x_k\}$ be vector of samples, there's an amount correlation among neighboring.
- If Y is the result of a linear transform T of the input vector and its components have much less correlation, then Y can be coded more efficiently than X.
  - The transform T itself does not compress any data.
  - The compression comes from the processing and quantization of the components of Y.
- DCT is a widely used transform, it can perform de-correlation of the input signal.

4.2 Discrete Cosine Transform (DCT)

- **1D Discrete Cosine Transform:**

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

- **1D Inverse Discrete Cosine Transform:**

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

$$C(u) = \begin{cases} \frac{\sqrt{2}}{2} & if \ u = 0 \\ 1 & else \end{cases}$$

- **2D transform can be used to process 2D signals such as digital images**

4.2 Discrete Cosine Transform (DCT)

DCT（2D）Definition：

- **Given a function f(i, j) over an image，the 2D DCT transforms it into a new function F(u,v), integer u and v running over the same range as i and j.**
- **The general definition of the DCT transform is：**

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

4.2 Discrete Cosine Transform (DCT)

- In the JPEG image compression standard
  - An image block is defined to have dimension M=N=8;
  - The definition of 2D DCT and its inverse IDCT are as follows:
- 2D Discrete Cosine Transform(2D DCT):

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

- 2D Inverse Discrete Cosine Transform(2D IDCT):

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

4.2 Discrete Cosine Transform (DCT)

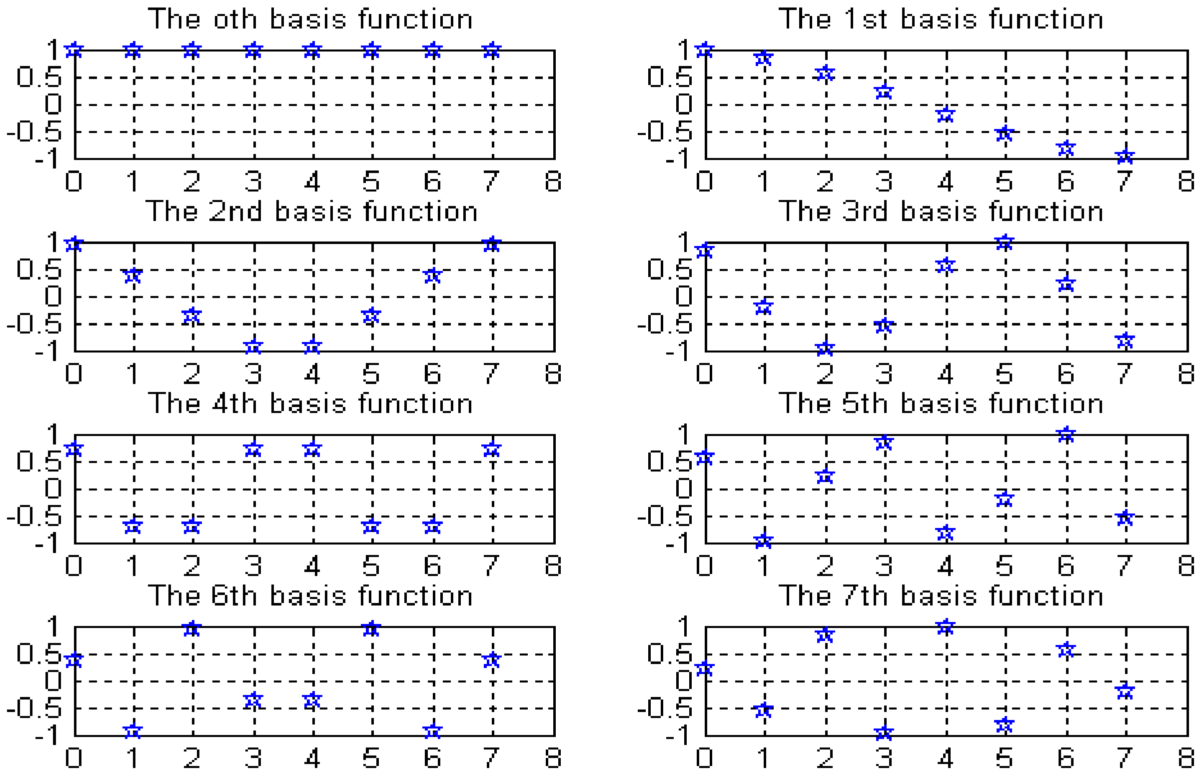
- DCT related concepts
  - Direct current (DC) and alternating current (AC)
    - Represent constant and variable magnitude respectively;
  - Fourier analysis
    - Any signal can be expressed as a sum of multiple signals that are sine or cosine wave forms.
    - An signal usually composed of one DC and several AC components;
  - Cosine Transform
    - The process used to determine the amplitude of the AC and DC components of the signal.

4.2 Discrete Cosine Transform (DCT)

- **DCT related concepts (Continue)**
  - **Discrete Cosine Transform：integer indices**
    - **U＝0, we get the DC coefficient；**
    - **U=1, 2, …, 7，we get the first up to seventh AC coefficients.**
  - **Invert Discrete Cosine transform：using DC, AC and cosine functions to reconstruct the signal**
  - **DCT and IDCT adopt the same set of cosine functions which are know as basis functions**

4.2 Discrete Cosine Transform (DCT)

- 1D DCT basis functions



4.2 Discrete Cosine Transform (DCT)

- DCT enable to process or analyze the signal in frequency domain
- Suppose f(i) represents a signal changes with time i
  - 1D DCT transforms f(i) in time domain to F(U) in frequency domain.
  - F(u) are known as frequency response, form the frequency spectrum of f(i)

4.2 Discrete Cosine Transform (DCT)

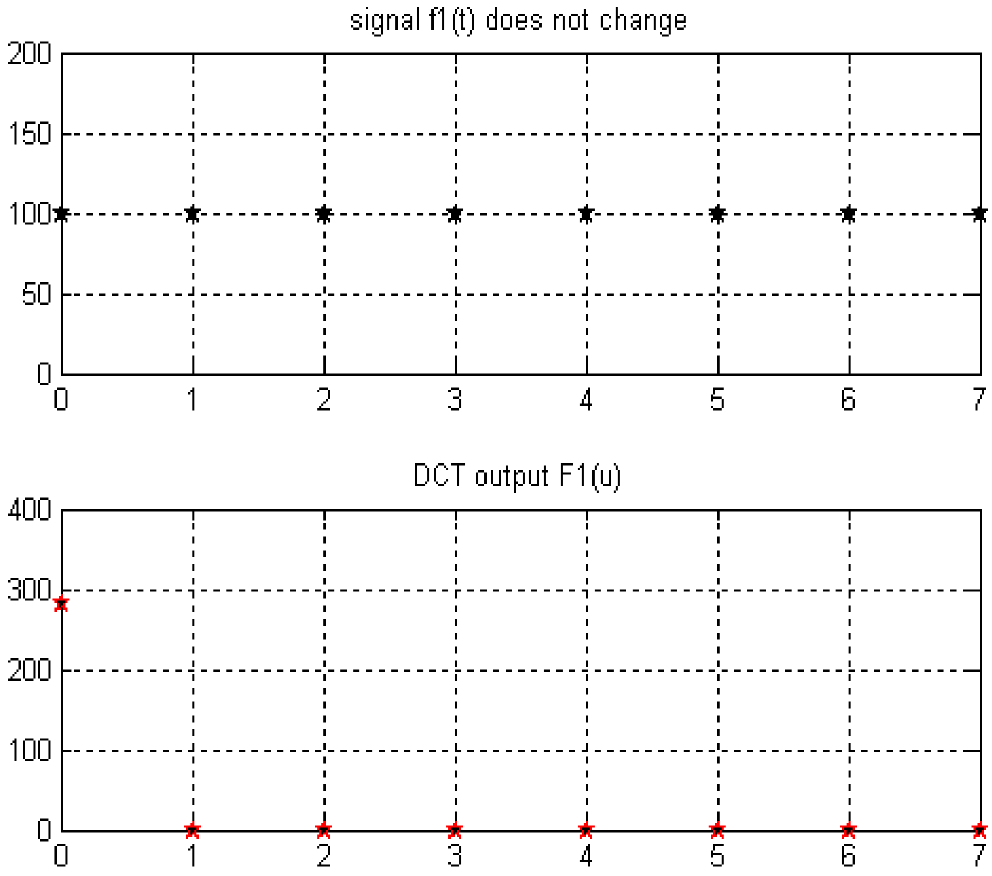
- **Example (1): $f_1(i)=100$, a signal with magnitude of 100**

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

  - **$F_1(0)=C(0)/2*（1\cdot 100+ 1\cdot 100 + 1\cdot 100 + 1\cdot 100 + 1\cdot 100 + 1\cdot 100 + 1\cdot 100 + 1\cdot 100）$ noticed that** $C(0) = \frac{\sqrt{2}}{2}$
  - **$=C(0)\cdot 400\approx 283$**
  - **$F_1(1)=$** $\frac{1}{2}(\cos\frac{\pi}{16}\cdot 100 + \cos\frac{3\pi}{16}\cdot 100 + \cos\frac{5\pi}{16}\cdot 100 + \cos\frac{7\pi}{16}\cdot 100 + \cos\frac{9\pi}{16}\cdot 100$
    $+ \cos\frac{11\pi}{16}\cdot 100 + \cos\frac{13\pi}{16}\cdot 100 + \cos\frac{15\pi}{16}\cdot 100$

    $=0$
  - **$F_1(2)= F_1(3)= F_1(4)= F_1(5)= F_1(6)= F_1(7)=0$**

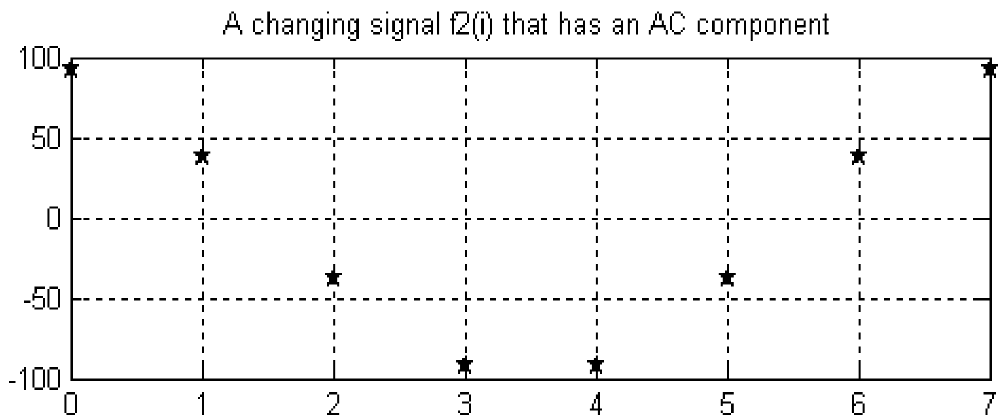
4.2 Discrete Cosine Transform (DCT)

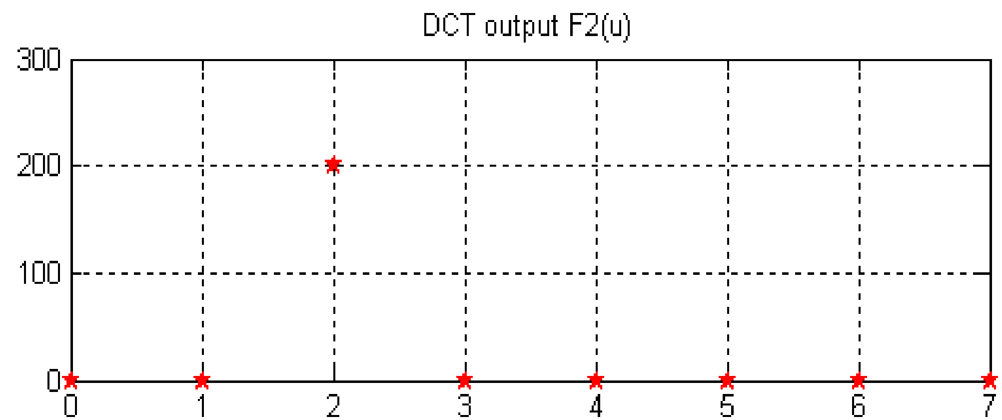


4.2 Discrete Cosine Transform (DCT)

- Example 2: a signal $f_2(i)$, has the same frequency and phase as the second cosine basis function, amplitude is100
  - F2(0)= $$\frac{\sqrt{2}}{2\cdot 2}\cdot 1\cdot(100\cos\frac{\pi}{8}+100\cos\frac{3\pi}{8}+100\cos\frac{5\pi}{8}+100\cos\frac{7\pi}{8}$$
    $$+100\cos\frac{9\pi}{8}+100\cos\frac{11\pi}{8}+100\cos\frac{13\pi}{8}+100\cos\frac{15\pi}{8})$$
    ＝0
  - F2（2)＝ $$\frac{1}{2}\cdot(\cos\frac{\pi}{8}\cdot\cos\frac{\pi}{8}+\cos\frac{3\pi}{8}\cdot\cos\frac{3\pi}{8}+\cos\frac{5\pi}{8}\cdot\cos\frac{5\pi}{8}$$
    $$+\cos\frac{7\pi}{8}\cdot\cos\frac{7\pi}{8}+\cos\frac{9\pi}{8}\cdot\cos\frac{9\pi}{8}+\cos\frac{11\pi}{8}\cdot\cos\frac{11\pi}{8}$$
    $$+\cos\frac{13\pi}{8}\cdot\cos\frac{13\pi}{8}+\cos\frac{15\pi}{8}\cdot\cos\frac{15\pi}{8})\cdot 100$$
    ＝200
- We can get other values by similar way
  - F2(1)＝F2(3)＝F2(4)＝…＝F2（7)＝0

4.2 Discrete Cosine Transform (DCT)





4.2 Discrete Cosine Transform (DCT)

Properties of DCT transform

- **DCT produces the frequency spectrum F(u) of signal f(i)**
  - **The 0th DCT coefficient F(0) is the DC component of the signal f(i);**
  - **The other seven DCT coefficients reflect the various changing components of signal f(i) at different frequencies;**
  - **If DC is a negative value, this means that the average of f(i) is less than zero;**
  - **if AC is a negative value, this means that f(i) and some basis function have the same frequency but one of them happens to be half a cycle behind.**
- **DCT is a linear transform**

$$T(\alpha p + \beta q) = \alpha T(p) + \beta T(q)$$

4.2 Discrete Cosine Transform (DCT)

- One-Dimensional IDCT
  - If F(u) contains (u=0…7):69 -49 74 11 16 117 44 -5
  - IDCT can be implemented by eight iterations:

$$Iteration\quad 0:\ \tilde{f}_i = \frac{C(0)}{2}\cos 0 \cdot F(0) \approx 24.3$$

$$Iteration\quad 1:\ \tilde{f}_i = \frac{C(0)}{2}\cos 0 \cdot F(0) + \frac{C(1)}{2}\cos\frac{(2i+1)\pi}{16}\cdot F(1)$$

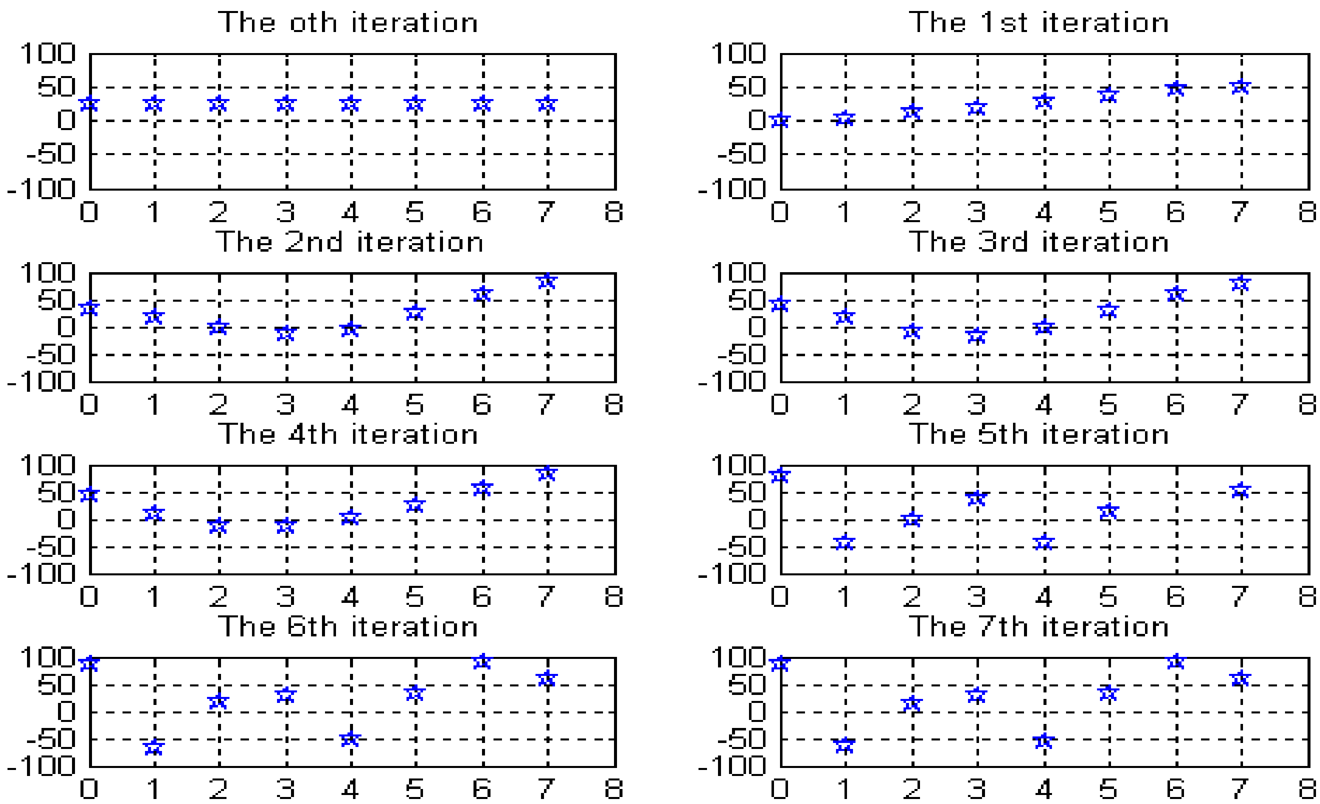
$$\approx 24.3 - 24.5 \cdot \cos\frac{(2i+1)\pi}{16}$$

$$Iteration\quad 2:\ \tilde{f}_i = \frac{C(0)}{2}\cos 0 \cdot F(0) + \frac{C(1)}{2}\cos\frac{(2i+1)\pi}{16}\cdot F(1) + \frac{C(2)}{2}\cos\frac{(2i+1)\pi}{8}\cdot F(2)$$

$$\approx 24.3 - 24.5 \cdot \cos\frac{(2i+1)\pi}{16} + 37 \cdot \cos\frac{(2i+1)\pi}{8}$$

4.2 Discrete Cosine Transform (DCT)

- An example of IDCT



4.2 Discrete Cosine Transform (DCT)

- Cosine basis functions are orthogonal

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

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

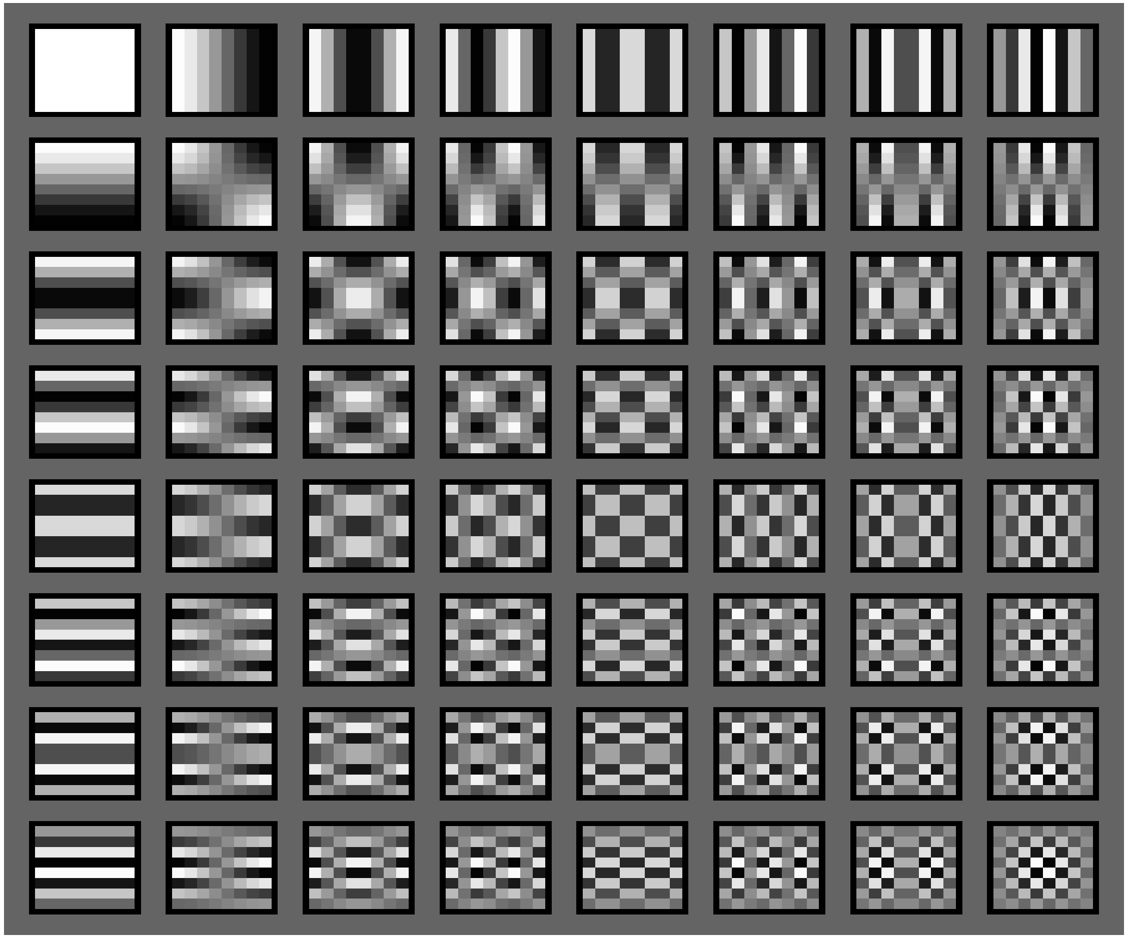
- 2D separable basis
  - Factorization 2D DCT into two 1D DCT transforms

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

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

4.2 Discrete Cosine Transform (DCT)

2D Basis Functions



4.3 Comparison of DCT and DFT

- Fourier Transform

$$F(w) = \int_{-\infty}^{+\infty} f(t) e^{-\omega t} dt$$

- Discrete Fourier Transform

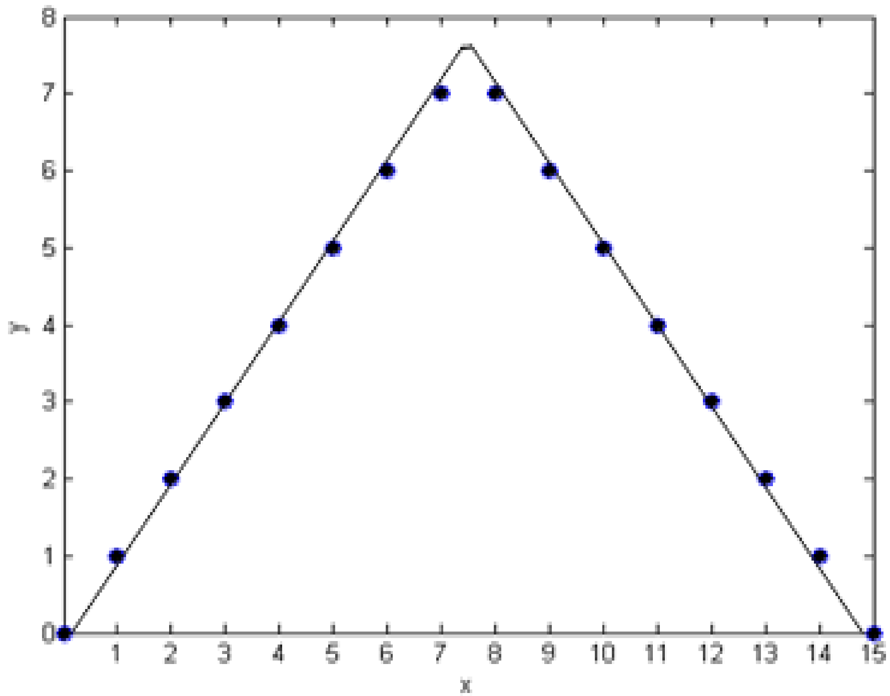
$$F_\omega = \sum_{x=0}^{7} f_x \cdot e^{-\frac{2\pi i \omega x}{8}}$$

$$F_\omega = \sum_{x=0}^{7} f_x \cdot \cos\left(\frac{2\pi wx}{8}\right) - i \sum_{x=0}^{7} \sin\left(\frac{2\pi wx}{8}\right)$$

- DCT is a close counterpart of the DFT, in signal processing, DCT is more popular.

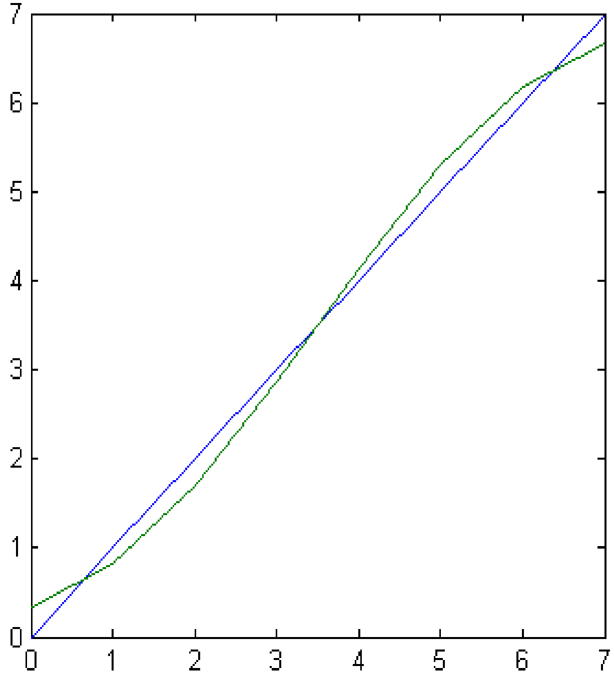
4.3 Comparison of DCT and DFT

Ramp	DCT	DFT
0	9.90	28.00
1	-6.44	-4.00
2	0.00	9.66
3	-0.67	-4.00
4	0.00	4.00
5	-0.20	-4.00
6	0.00	1.66
7	-0.51	-4.00

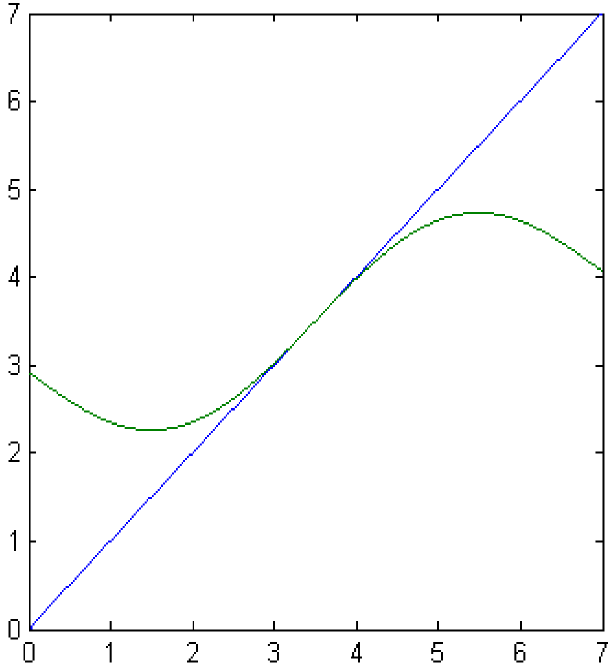


DCT and DFT coefficient of the ramp function

4.3 Comparison of DCT and DFT



(a) three-term DCT approximation

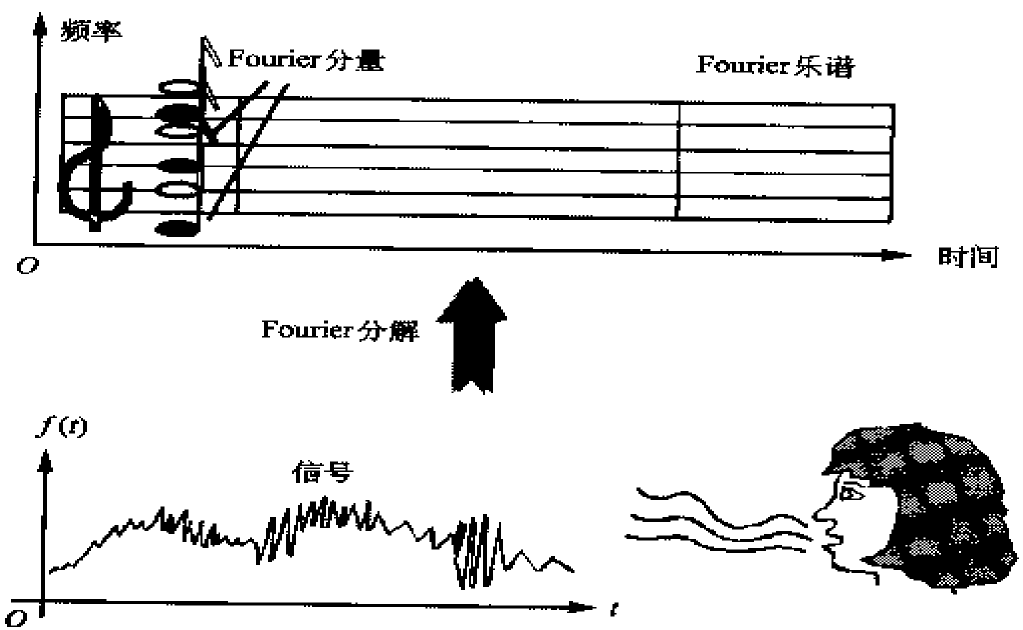


(b) three-term DFT approximation

5. Wavelet-Based Coding

5.1 Introduction

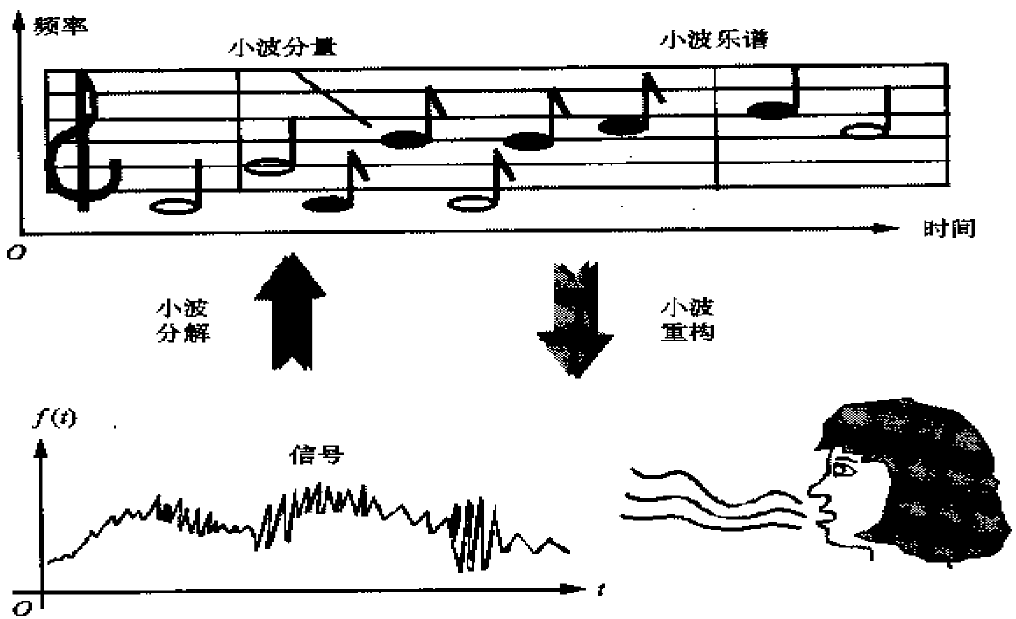
- DFT and DCT can give very fine resolution in the frequency domain, but no temporal resolution.



Fourier Decomposition of A Song Signal

5.1 Introduction

- Wavelet transform seeks to represents a signal with good resolution in both time and frequency.



Wavelet Decomposition of A Song Signal

5.2 1D Haar Transform

$$\{x_{n,i}\} = \{10, 13, 25, 26, 29, 21, 7, 15\}$$

$$x_{n-1,i} = \frac{x_{n,2i} + x_{n,2i+1}}{2}$$

$$d_{n-1,i} = \frac{x_{n,2i} - x_{n,2i+1}}{2}$$

$$\{x_{n-1,i}, d_{n-1,i}\} = \{11.5, 25.5, 25, 11, -1.5, -0.5, 4, -4\}$$

$$\{x_{n-2,i}, d_{n-2,i}, d_{n-1,i}\} = \{18.5, 18, -7, 7, -1.5, -0.5, 4, -4\}$$

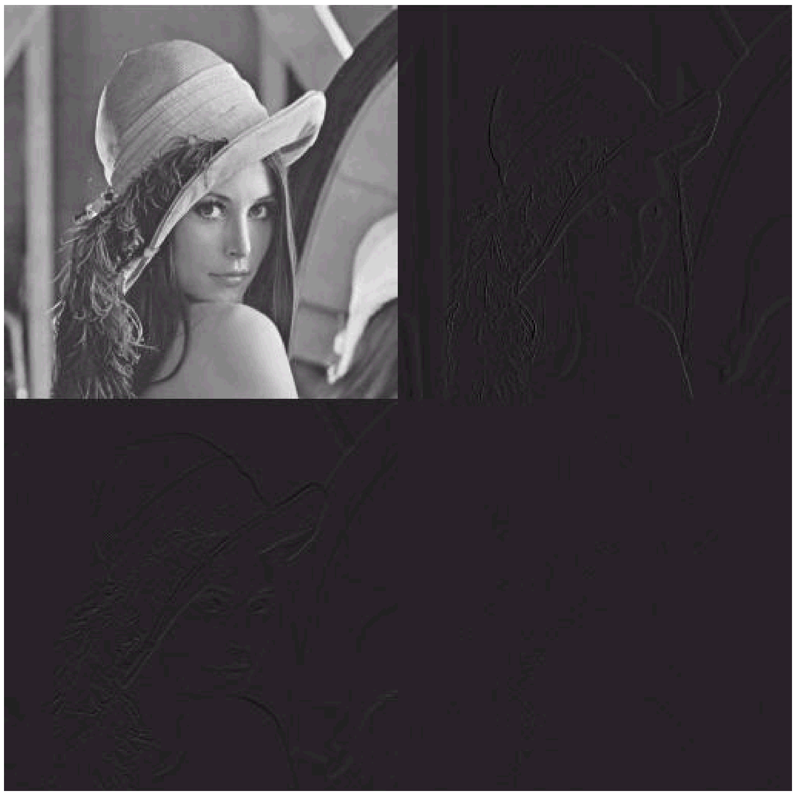
$$\{x_{n-3,i}, d_{n-3,i}, d_{n-2,i}, d_{n-1,i}\} = \{18.25, 0.25, -7, 7, -1.5, -0.5, 4, -4\}$$

5.3 2D Haar Transform

original image

wavelet horizontally transform

5.3 2D Haar Transform



wavelet horizontally and vertically transform (one level)

wavelet transformation (2 levels)

The End!

Thanks !