

- 引言
- 失真度量
- 率失真理论
- 量化
- 变换编码
- 基于小波的编码

多媒体基础——有损压缩算法 (2025年春季)

2

有损压缩算法



讲师：肖俊

软件与技术学院

1. 引言

- 无损压缩算法无法实现足够高的压缩比。因此，大多数多媒体压缩算法都是有损的。

• 什么是有损压缩？

- 压缩后的数据与原始数据不同，但非常接近。
- 比无损压缩产生更高的压缩比。

多媒体基础 - 有损压缩算法 (2025年春季)

3

2.1 失真概念

• 失真度量

- 一个数学量：指定接近程度
对其原始状态的近似
- 很自然会想到数值差异
- 当涉及图像数据时，差异可能
无法产生预期的结果
- 感知失真的度量

多媒体基础——有损压缩算法 (2025年春季)

5

2.2 数值失真度量

• 峰值信噪比示例及对应图像



原始图像

受噪声污染

峰值信噪比 = 18.24

经噪声滤波器处理

PSNR=39.5

3. 率失真理论

多媒体基础——有损压缩算法 (2025年春季)

6

多媒体基础 - 有损压缩算法 (2025年春季)

7

3.1 概念

• 有损压缩始终涉及速率和失真之间的权衡

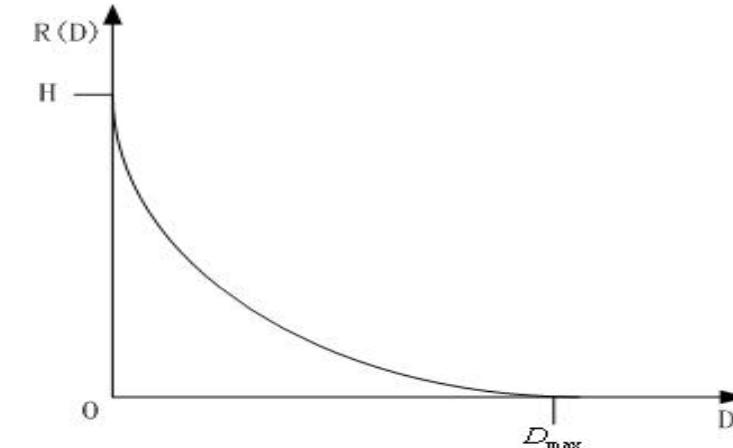
• 速率——表示每个源符号所需的平均比特数；
R(D) 记为率失真函数；

• R(D) 是什么？

• R(D) 指定了在将失真限制在 D 以内的情况下，对源数据进行编码的最低速率 - 在 D = 0 时，无损失，因此是源数据的熵 - 描述了编码算法性能的基本限制 - 可用于评估不同算法的性能

3.2 典型的率失真函数

• 典型率失真函数的示意图



• $D = 0$, 源数据的熵

• $R(D) = 0$, 未进行编码

• 对于给定的信源，很难找到率失真函数的闭式解析描述

4. 量化

4.2 均匀标量量化

• 均匀标量量化器

• 将输入域划分为等间距的区间 - 决策边界：
划分区间的端点 - 输出值：区间的中点 - 步长：每个区间的长度

• 两种类型的均匀标量量化器

• 中升型：输出电平数为偶数，有一个划分区间包含零； - 中平型：输出电平数为奇数，零是一个输出值。

• 成功的均匀量化器的目标 - 对于给定的源输入，在期望的输出值数量下使失真最小化

4.1 量化的作用

• 量化：任何有损方案的核心

- 没有量化，几乎不会丢失信息
- 通过量化减少不同值的数量
- 有损压缩中“损失”的主要来源

• 每个量化器对输入范围和输出值集合都有其独特的划分。

• 标量量化器

- 均匀
- 非均匀
- 矢量量化器

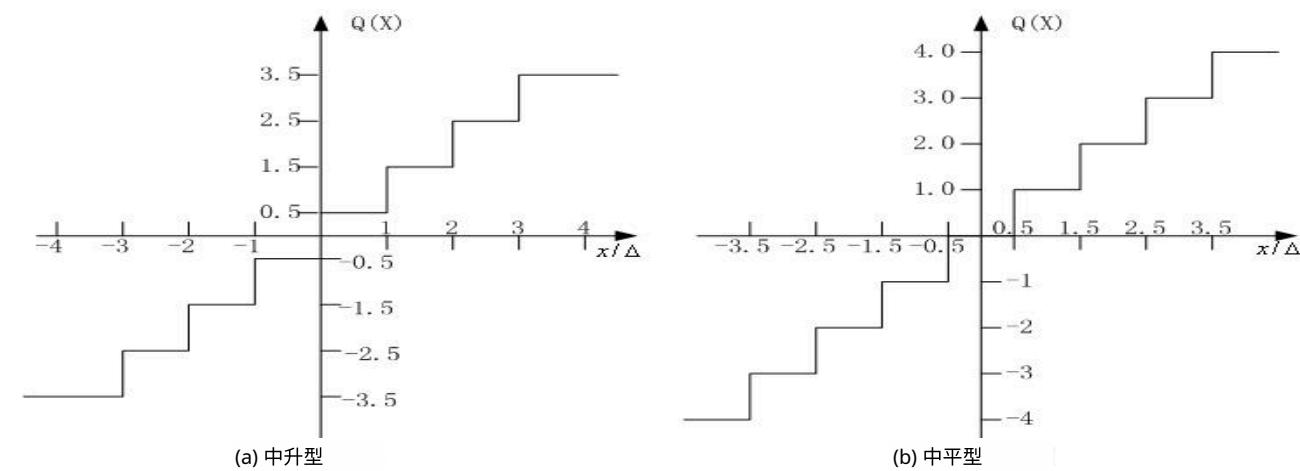
4.2 均匀标量量化

• 给定步长 $\Delta = 1$ ，两种类型量化器的输出值计算如下：

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

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

• 两种类型的量化器：



4.2 均匀标量量化

• M 级量化器的性能：

- 决策边界： $B = \{b_0, b_1, \dots, b_M\}$
- 输出值集合： $Y = \{y_1, y_2, \dots, y_m\}$
- 输入呈均匀分布： $[-X_{max}, X_{max}]$
- 量化器速率： $R = \log_2^M$ 是对 M 个事物进行编码所需的比特数；
- 步长由以下公式给出： $\Delta = 2X_{max}/M$
- 颗粒失真：量化器对有界输入造成的误差
- 过载失真：量化器对大于 X_{max} 或小于 $-X_{max}$ 的输入值造成的误差

4.2 均匀标量量化

• 中升型量化器的颗粒失真

- 决策边界 b_i : $[(i-1)\Delta, i\Delta]$, $i = 1..M/2$ ，覆盖正数据 X (另一个用于原生 X 值)
- 输出值 y_i : $i\Delta - \Delta/2$, $i = 1..M/2$
- 总失真：正数据总和的两倍

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

• 在 X 处的误差值为 $e(x) = x - \Delta/2$ ，误差方差：

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

4.2 均匀标量量化

• 信号方差 $\sigma_x^2 = (2X_{\max})^2/12$ ；如果量化器为n位， $M = 2^n$

• 信噪比 (SQNR) 可计算如下：

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

4.3 非均匀标量量化

• 如果输入源不是均匀分布的，均匀量化器可能效率不高。

• 增加密集分布区域内的决策级别数量可以降低粒度失真

• 扩大信源稀疏分布的区域可以保持决策级别的总数

• 因此，非均匀量化器具有非均匀定义的决策边界。

• 非均匀量化的两种常见方法：

- 劳埃德 - 马克斯量化器
- 压扩量化器

4.3 非均匀标量量化



图8.4：压扩量化。

• 压扩量化是非线性的。

• 如上所示，压扩器由压缩函数 G 、均匀量化器和扩展函数 G^{-1} 组成。

• 两种常用的压扩器是 μ 律和 A 律压扩器。

5.1 基本思路

• 根据信息论原理

• 对向量进行编码比对标量进行编码更高效

• 需要将输入的连续样本分组为向量

• 设 $X = \{x_1, x_2, \dots, x_k\}$ 为样本向量，相邻样本之间存在一定的相关性。

• 如果 Y 是输入向量经过线性变换 T 的结果，且其分量之间的相关性小得多，那么对 Y 进行编码比 X 更高效。

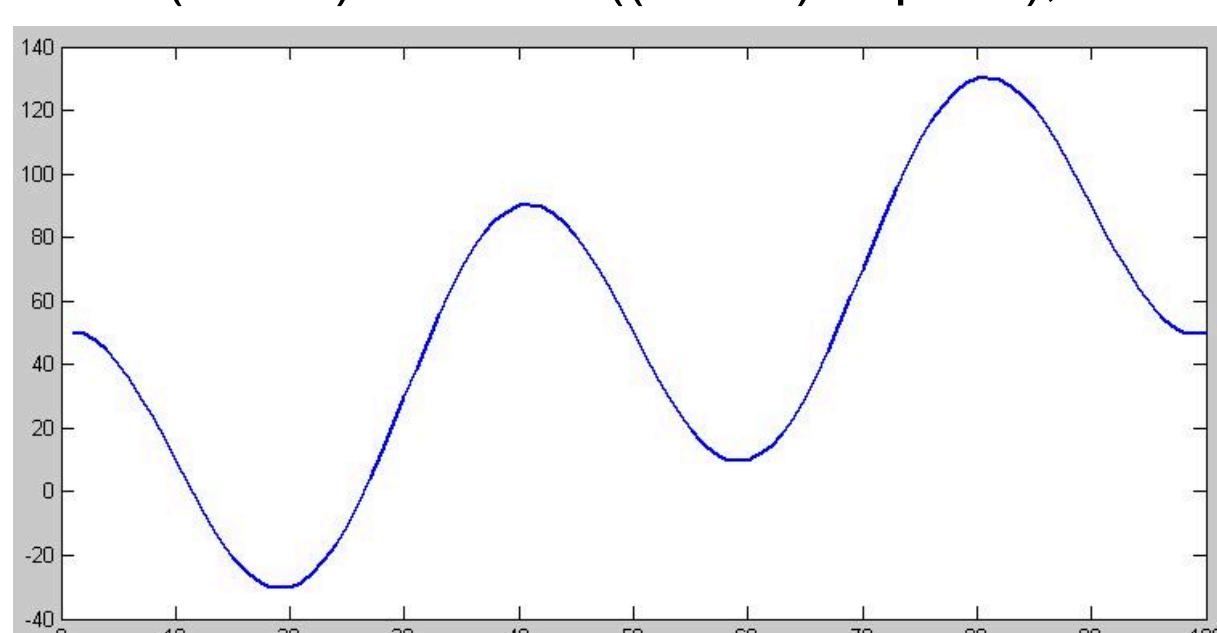
• 变换 T 本身并不压缩任何数据。

• 压缩来自对 Y 各分量的处理和量化。- 离散余弦变换 (DCT) 是一种广泛使用的变换，它可以对输入信号进行去相关处理。

如何通过离散余弦变换 (DCT) 压缩数据？

• Matlab仿真

- $x = (1:100) + 50 * \cos((1:100) * 2 * \pi / 40);$



5.1 基本思想

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

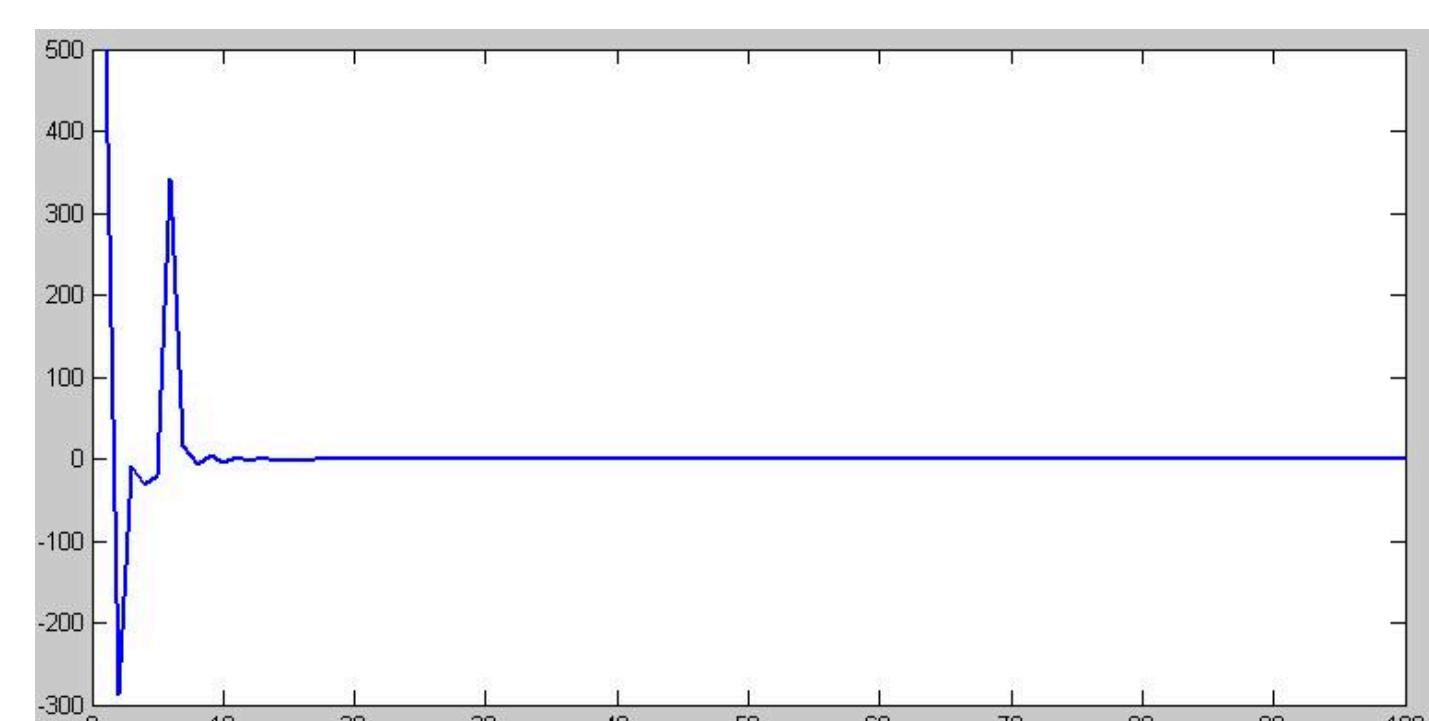
$$C(u) = \begin{cases} \frac{\sqrt{2}}{2}, & u = 0 \\ \frac{1}{2}, & u = 1, 2, 3, 4, 5, 6, 7 \\ 1, & \text{otherwise} \end{cases}$$

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

$$C(u) = \begin{cases} \frac{1}{\sqrt{N}}, & u = 0 \\ \sqrt{\frac{2}{N}}, & u = 1, 2, 3, 4, 5, 6, 7 \\ \sqrt{\frac{1}{N}}, & \text{otherwise} \end{cases}$$

如何通过离散余弦变换 (DCT) 压缩数据？

- $x_dct = dct(x);$



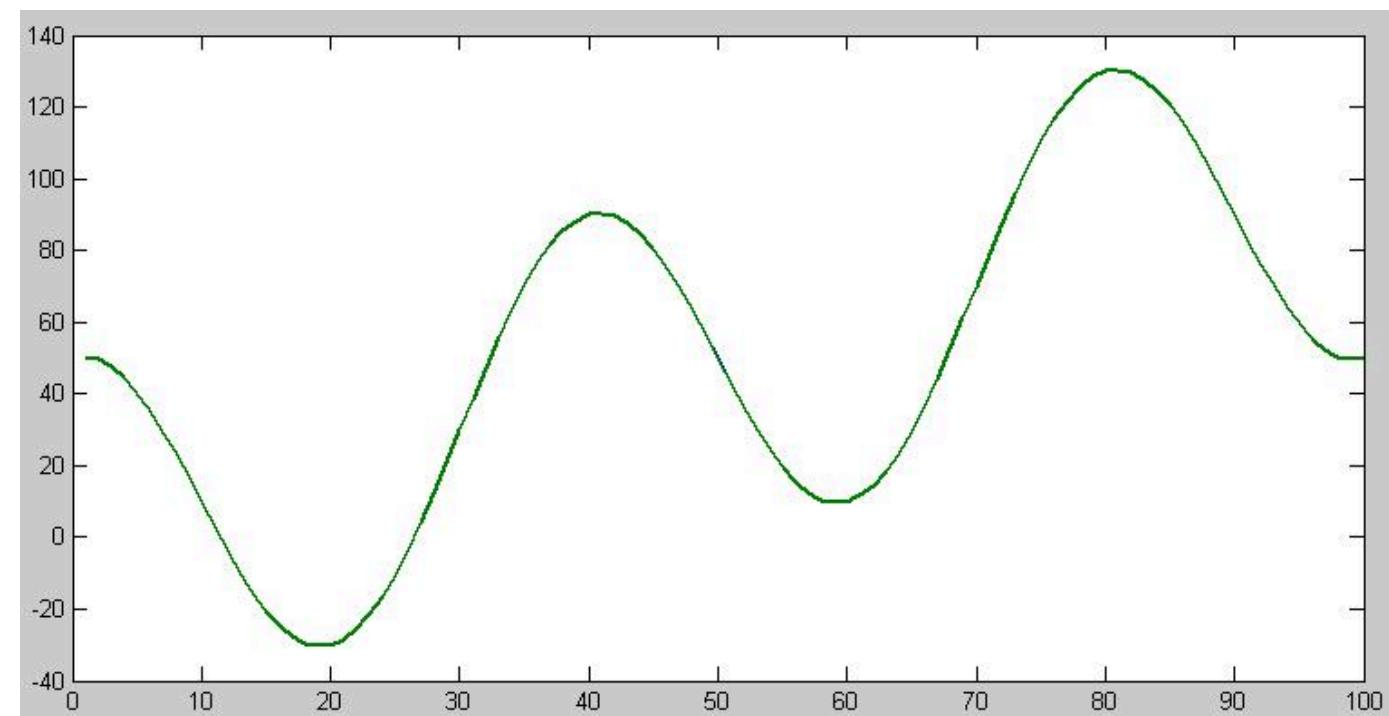
如何通过离散余弦变换 (DCT) 压缩数据?

• `x_dct`

第1列至第12列												
500.0000	-286.5678	-8.4166	-31.8304	-19.6289	341.0121	16.0468	-5.8367	4.5208	-3.5262	2.3473	-2.3565	
列 13 至第24列	1.4766	-1.6838	1.0253	-1.2617	0.7571	-0.9795	0.5831	-0.7817	0.4634	-0.6376	0.3772	-0.5294
第25列 至第36列	0.3129	-0.4461	0.2636	-0.3806	0.2249	-0.3281	0.1940	-0.2854	0.1688	-0.2502	0.1481	-0.2208
第5列 7 至第48列	0.1308	-0.1961	0.1162	-0.1750	0.1038	-0.1569	0.0931	-0.1412	0.0838	-0.1275	0.0757	-0.1155
列 9 至第60列	0.0686	-0.1049	0.0623	-0.0954	0.0567	-0.0870	0.0517	-0.0794	0.0472	-0.0726	0.0432	-0.0664
第61列 至第72列	0.0395	-0.0607	0.0361	-0.0556	0.0330	-0.0508	0.0302	-0.0464	0.0275	-0.0424	0.0251	-0.0386
第7列 } 至第84列	0.0228	-0.0350	0.0207	-0.0317	0.0187	-0.0286	0.0168	-0.0256	0.0150	-0.0228	0.0133	-0.0201
第85列 至第96列	0.0116	-0.0175	0.0100	-0.0150	0.0085	-0.0125	0.0070	-0.0102	0.0056	-0.0079	0.0042	-0.0056
第97列 至 100	0.0028	-0.0033	0.0014	-0.0011								

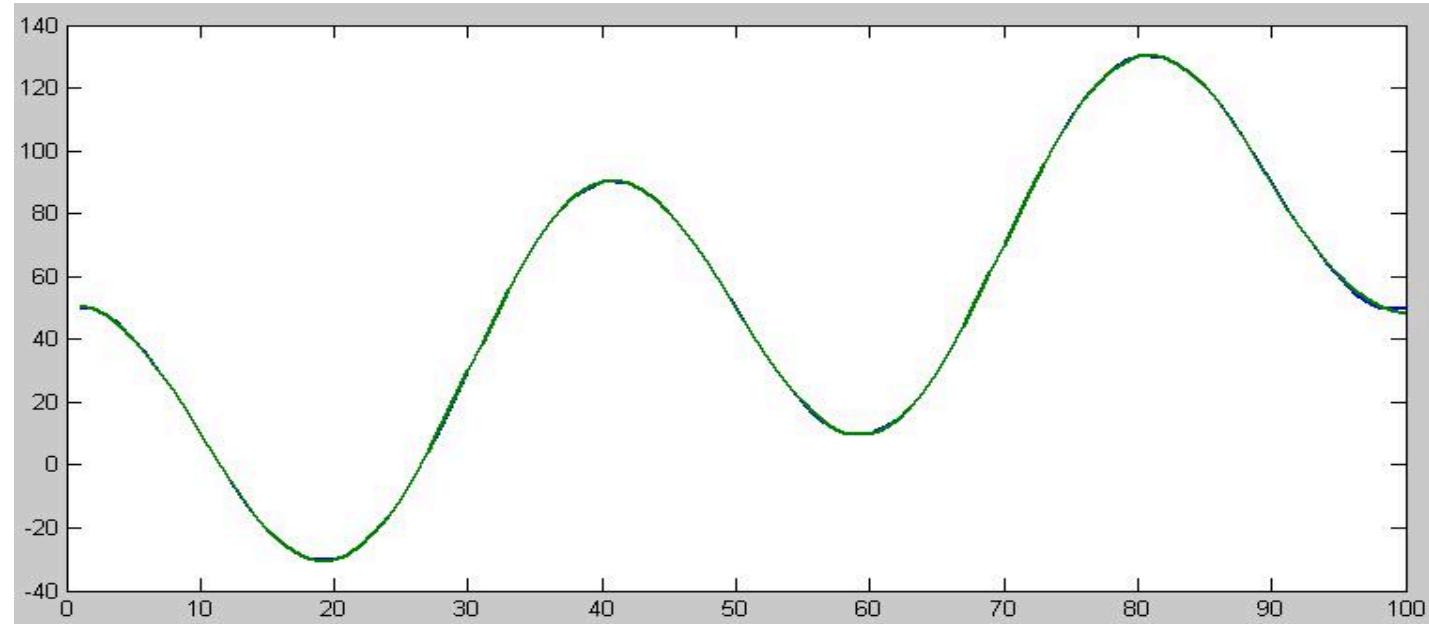
如何通过离散余弦变换 (DCT) 压缩数据?

• $y = \text{idct}(x_{\text{dct}});$



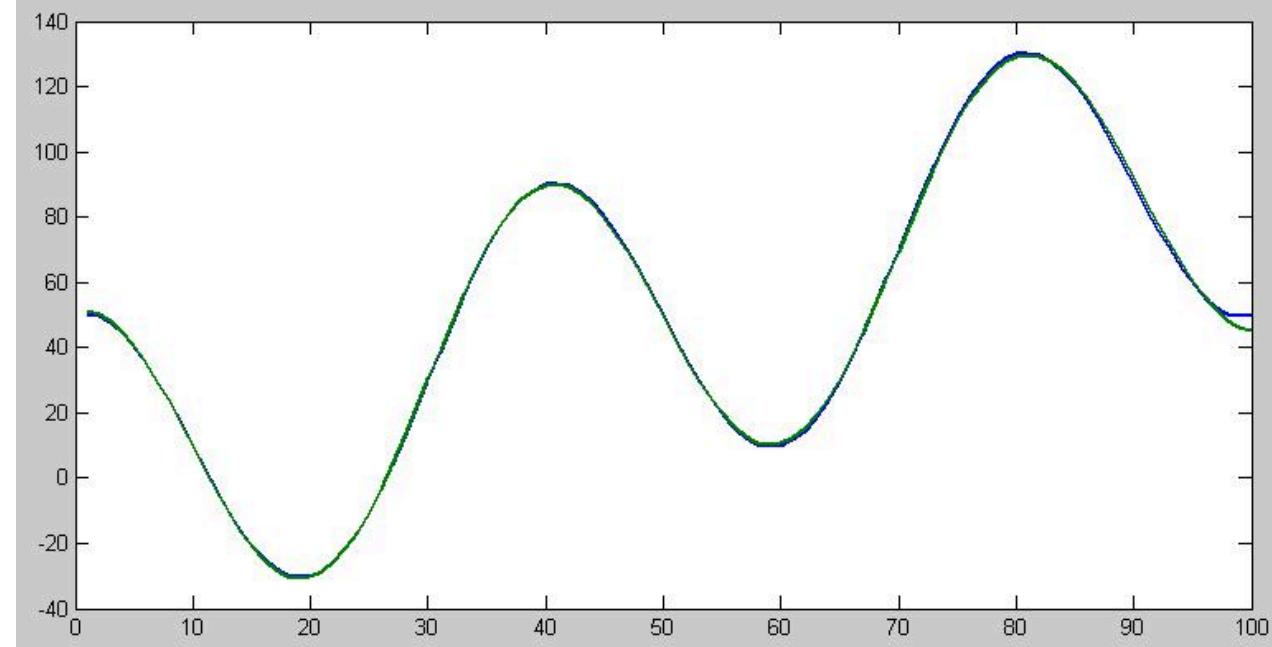
如何通过离散余弦变换 (DCT) 压缩数据?

- `x_dct(16:100) = 0;`
- `z = idct(x_dct);`



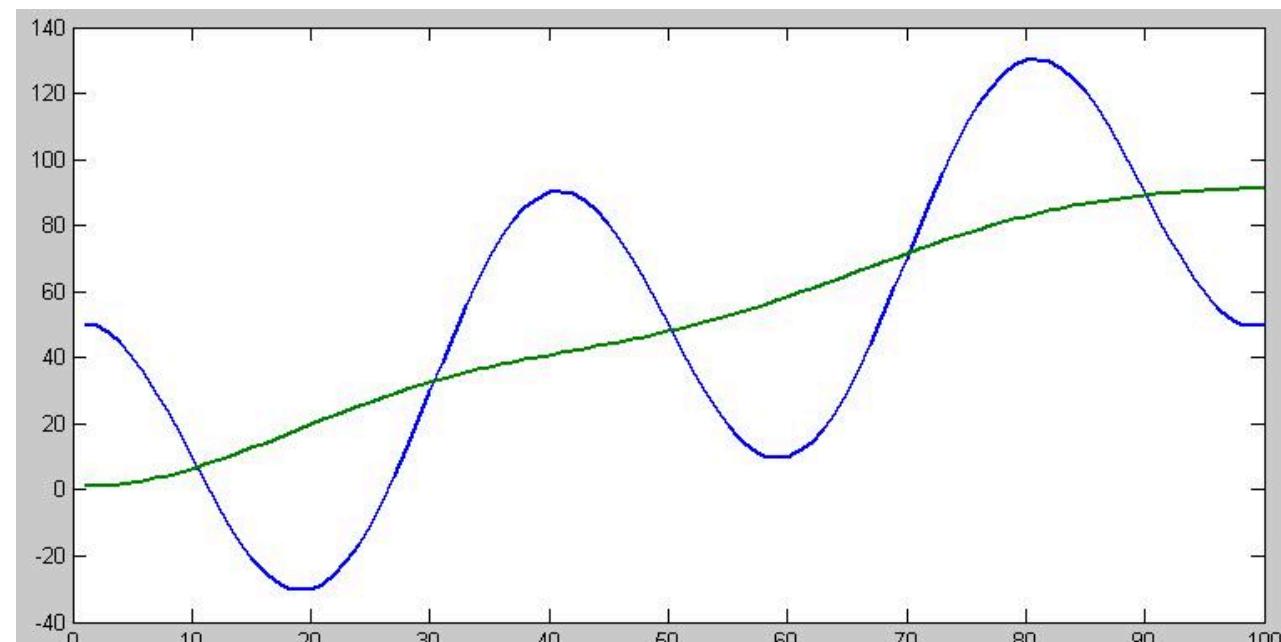
如何通过离散余弦变换 (DCT) 压缩数据?

- `x_dct(8:100) = 0;`
- `z = 离散余弦逆变换(idct)(x_dct);`



如何通过离散余弦变换 (DCT) 压缩数据?

- `x_dct(6:100) = 0;`
- `z = 离散余弦逆变换(idct)(x_dct);`



5.2 离散余弦变换 (DCT)

• 一维离散余弦变换:

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

• 一维离散余弦逆变换:

$$\tilde{f}_i = \sum_{u=0}^7 \frac{C(u)}{2} \cos \frac{(2i+1)u\pi}{16} F(u) \quad C(u) = \begin{cases} \frac{\sqrt{2}}{2} & \text{if } u = 0 \\ 1 & \text{else} \end{cases}$$

• 二维变换可用于处理二维信号，如数字图像

5.2 离散余弦变换 (DCT)

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

$$F(u) = \frac{C(u)}{2} [\cos(\frac{1u\pi}{16}), \cos(\frac{3u\pi}{16}), \cos(\frac{5u\pi}{16}), \cos(\frac{7u\pi}{16}), \cos(\frac{9u\pi}{16}), \cos(\frac{11u\pi}{16}), \cos(\frac{13u\pi}{16}), \cos(\frac{15u\pi}{16})]$$

$$\begin{bmatrix} f(0) \\ f(1) \\ f(2) \\ f(3) \\ f(4) \\ f(5) \\ f(6) \\ f(7) \end{bmatrix}$$

$$\begin{bmatrix} F(0) \\ F(1) \\ F(2) \\ F(3) \\ F(4) \\ F(5) \\ F(6) \\ F(7) \end{bmatrix} = \begin{bmatrix} \frac{1}{2\sqrt{2}} \cos(\frac{0\pi}{16}) & \frac{1}{2\sqrt{2}} \cos(\frac{0\pi}{16}) \\ \frac{1}{2} \cos(\frac{1\pi}{16}) & \frac{1}{2} \cos(\frac{3\pi}{16}) & \frac{1}{2} \cos(\frac{5\pi}{16}) & \frac{1}{2} \cos(\frac{7\pi}{16}) & \frac{1}{2} \cos(\frac{9\pi}{16}) & \frac{1}{2} \cos(\frac{11\pi}{16}) & \frac{1}{2} \cos(\frac{13\pi}{16}) & \frac{1}{2} \cos(\frac{15\pi}{16}) \\ \dots & & & & & & & \\ \frac{1}{2} \cos(\frac{7\pi}{16}) & \frac{1}{2} \cos(\frac{21\pi}{16}) & \frac{1}{2} \cos(\frac{35\pi}{16}) & \frac{1}{2} \cos(\frac{49\pi}{16}) & \frac{1}{2} \cos(\frac{63\pi}{16}) & \frac{1}{2} \cos(\frac{77\pi}{16}) & \frac{1}{2} \cos(\frac{91\pi}{16}) & \frac{1}{2} \cos(\frac{105\pi}{16}) \end{bmatrix} \begin{bmatrix} f(0) \\ f(1) \\ f(2) \\ f(3) \\ f(4) \\ f(5) \\ f(6) \\ f(7) \end{bmatrix}$$

$$\frac{C(u)}{2} [\cos(\frac{1u\pi}{16}), \cos(\frac{3u\pi}{16}), \cos(\frac{5u\pi}{16}), \cos(\frac{7u\pi}{16}), \cos(\frac{9u\pi}{16}), \cos(\frac{11u\pi}{16}), \cos(\frac{13u\pi}{16}), \cos(\frac{15u\pi}{16})]$$

基函数: B_i

$$b_i = \frac{C(i)}{2} B_i$$

示例

- 三维空间中的离散余弦变换 $y(k) = w(k) \sum_{n=1}^N x(n) \cos \frac{\pi(2n-1)(k-1)}{2N}$, $k = 1, \dots, N$

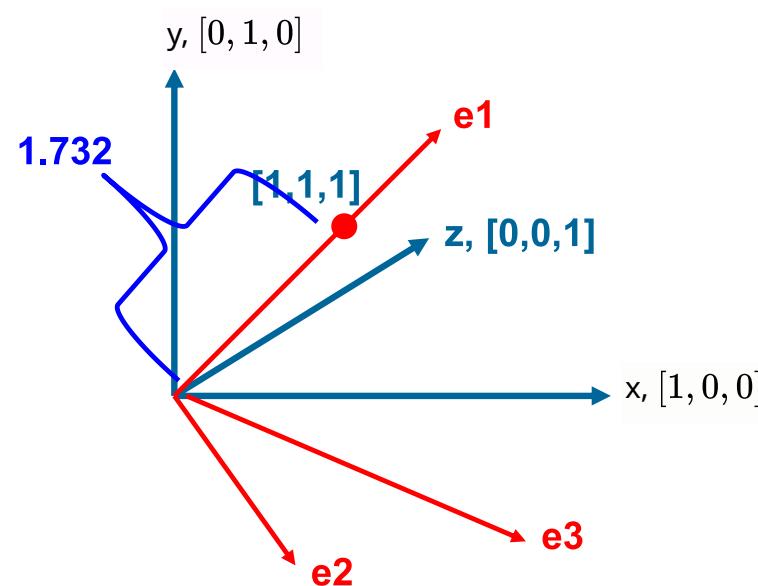
$$\mathbf{P} = [1, 1, 1]$$

$$= 1 * [1, 0, 0] + 1 * [0, 1, 0] + 1 * [0, 0, 1]$$

$$\text{Dct}(\mathbf{P}) = [1.732, 0, 0]$$

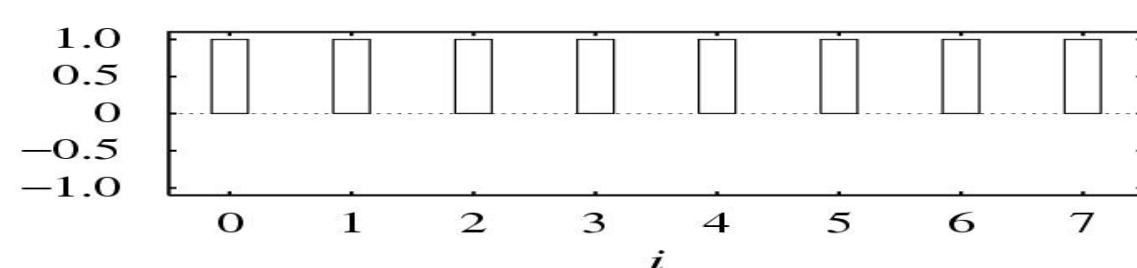
$$\begin{aligned} \mathbf{e}_1 &= [0.5774 \quad 0.5774 \quad 0.5774] \\ \mathbf{e}_2 &= [0.7071 \quad 0.0000 \quad -0.7071] \\ \mathbf{e}_3 &= [0.4082 \quad -0.8165 \quad 0.4082] \end{aligned}$$

$$\mathbf{P} = 1.732 * \mathbf{e}_1 + 0 * \mathbf{e}_2 + 0 * \mathbf{e}_3$$

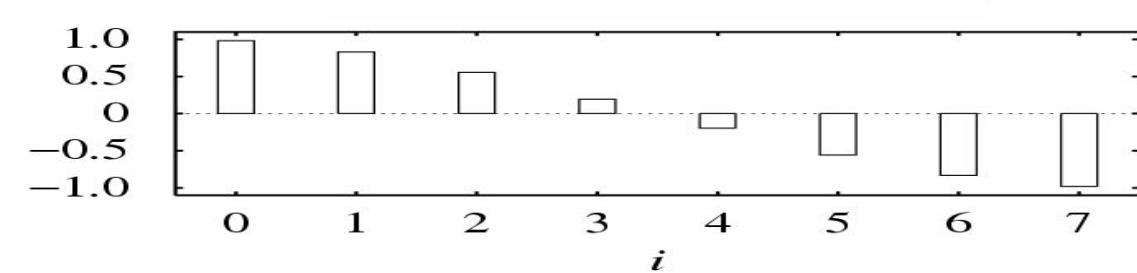


5.2 离散余弦变换 (DCT)

第零个基函数 ($u = 0$)



第一个基函数 ($u = 1$)



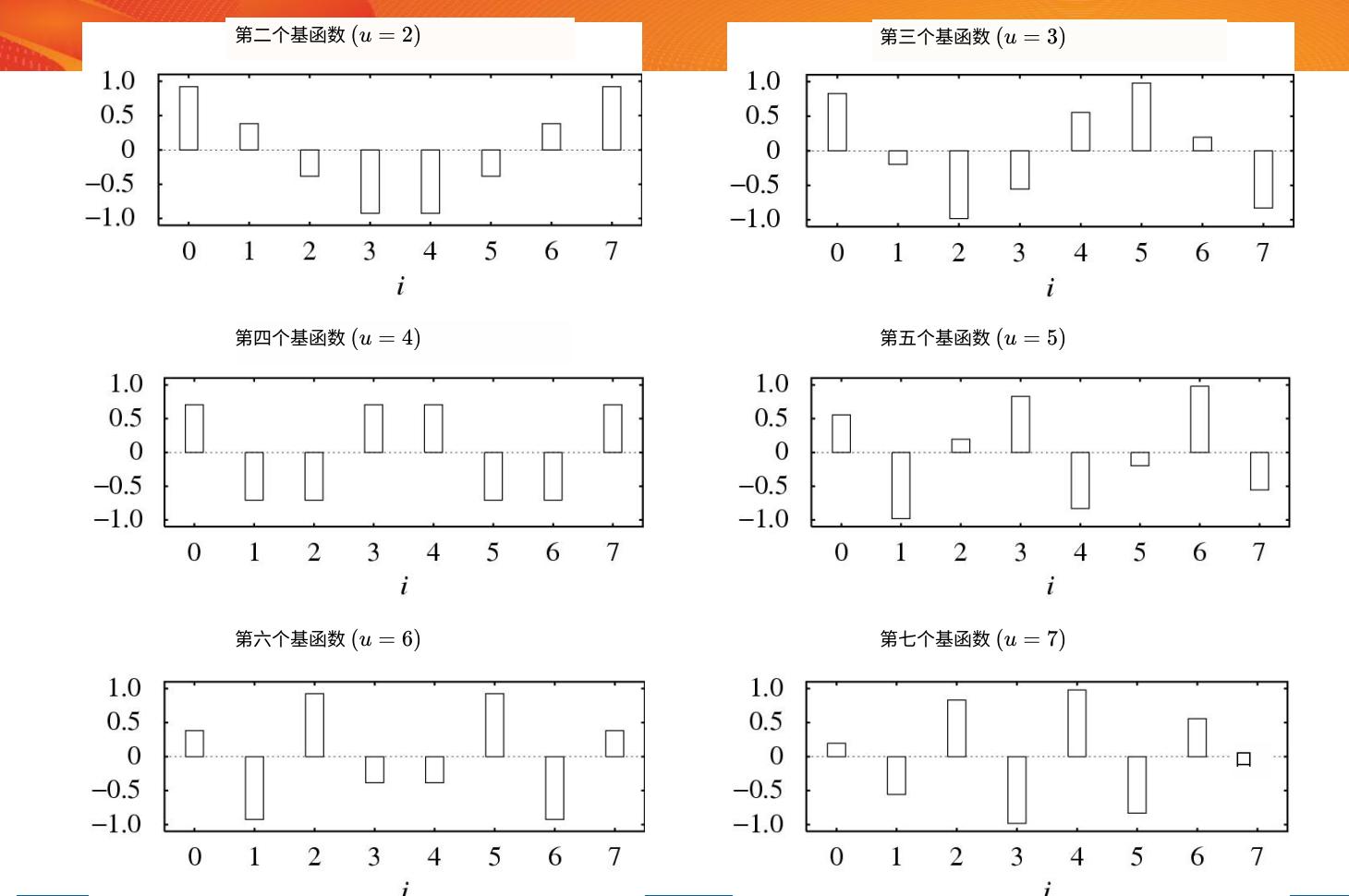
5.2 离散余弦变换 (DCT)

- 余弦基函数是正交的

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

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

数学含义：将一个向量从一个线性空间变换到另一个线性空间



5.2 离散余弦变换 (DCT)

- 示例 (1): $f_1(i) = 100$, 幅度为 100 的信号

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

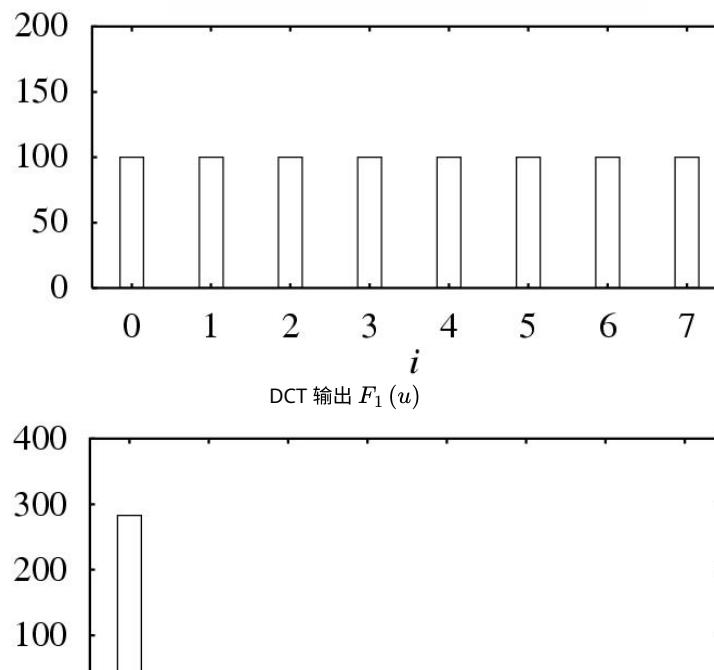
$$\begin{aligned} F_1(0) &= C(0)/2 * (1 \cdot 100 + 1 \cdot 100 + 1 \cdot 100 + 1 \cdot 100 + 1 \cdot 100 \\ &\quad + 1 \cdot 100 + 1 \cdot 100 + 1 \cdot 100) \text{ noticed that } C(0) = \frac{\sqrt{2}}{2} \\ &= C(0) \cdot 400 \approx 283 \end{aligned}$$

$$\begin{aligned} F_1(1) &= \frac{1}{2} (\cos \frac{\pi}{16} \cdot 100 + \cos \frac{3\pi}{16} \cdot 100 + \cos \frac{5\pi}{16} \cdot 100 + \cos \frac{7\pi}{16} \cdot 100 + \cos \frac{9\pi}{16} \cdot 100 \\ &\quad + \cos \frac{11\pi}{16} \cdot 100 + \cos \frac{13\pi}{16} \cdot 100 + \cos \frac{15\pi}{16} \cdot 100) \\ &= 0 \end{aligned}$$

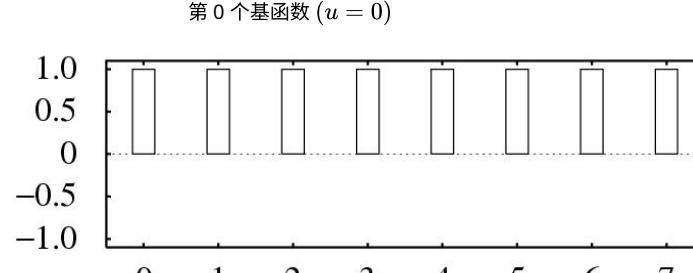
$$F_1(2) = F_1(3) = F_1(4) = F_1(5) = F_1(6) = F_1(7) = 0$$

5.2 离散余弦变换 (DCT)

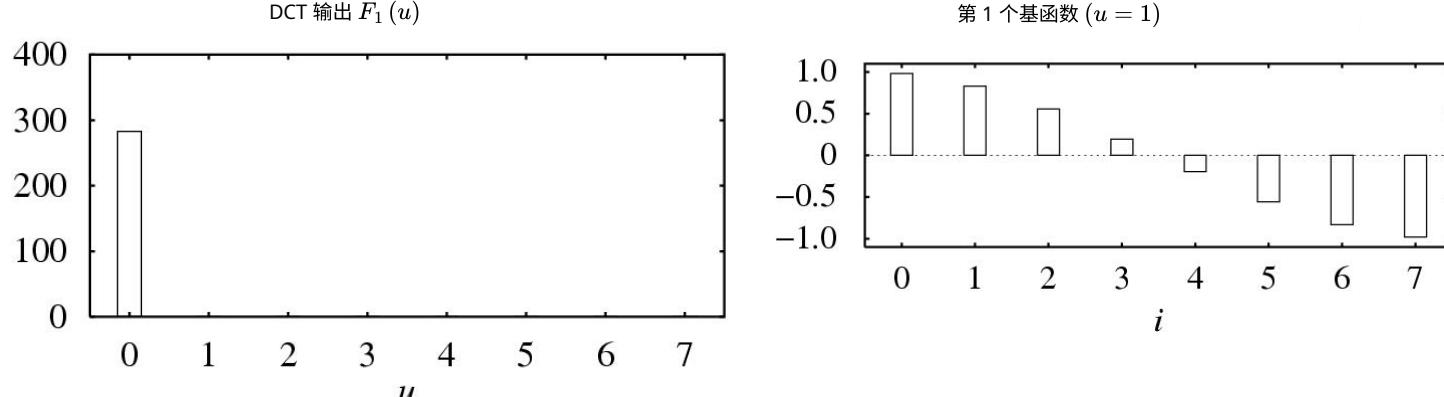
不发生变化的信号 $f_1(i)$



第 0 个基函数 ($u = 0$)



第 1 个基函数 ($u = 1$)



5.2 离散余弦变换 (DCT)

- 示例2: 信号 $f_2(i)$ 与第二个余弦基函数具有相同的频率和相位,

振幅

$$\begin{aligned} \text{is 100} \\ F_2(0) &= \frac{\sqrt{2}}{2 \cdot 2} \cdot 1 \cdot (100 \cos \frac{\pi}{8} + 100 \cos \frac{3\pi}{8} + 100 \cos \frac{5\pi}{8} + 100 \cos \frac{7\pi}{8} \\ &\quad + 100 \cos \frac{9\pi}{8} + 100 \cos \frac{11\pi}{8} + 100 \cos \frac{13\pi}{8} + 100 \cos \frac{15\pi}{8}) \\ &= 0 \end{aligned}$$

$$\begin{aligned} F_2(2) &= \frac{1}{2} \cdot (\cos \frac{\pi}{8} \cdot \cos \frac{\pi}{8} + \cos \frac{3\pi}{8} \cdot \cos \frac{3\pi}{8} + \cos \frac{5\pi}{8} \cdot \cos \frac{5\pi}{8} + \cos \frac{7\pi}{8} \cdot \cos \frac{7\pi}{8} \\ &\quad + \cos \frac{9\pi}{8} \cdot \cos \frac{9\pi}{8} + \cos \frac{11\pi}{8} \cdot \cos \frac{11\pi}{8} + \cos \frac{13\pi}{8} \cdot \cos \frac{13\pi}{8} + \cos \frac{15\pi}{8} \cdot \cos \frac{15\pi}{8}) \cdot 100 \\ &= 200 \end{aligned}$$

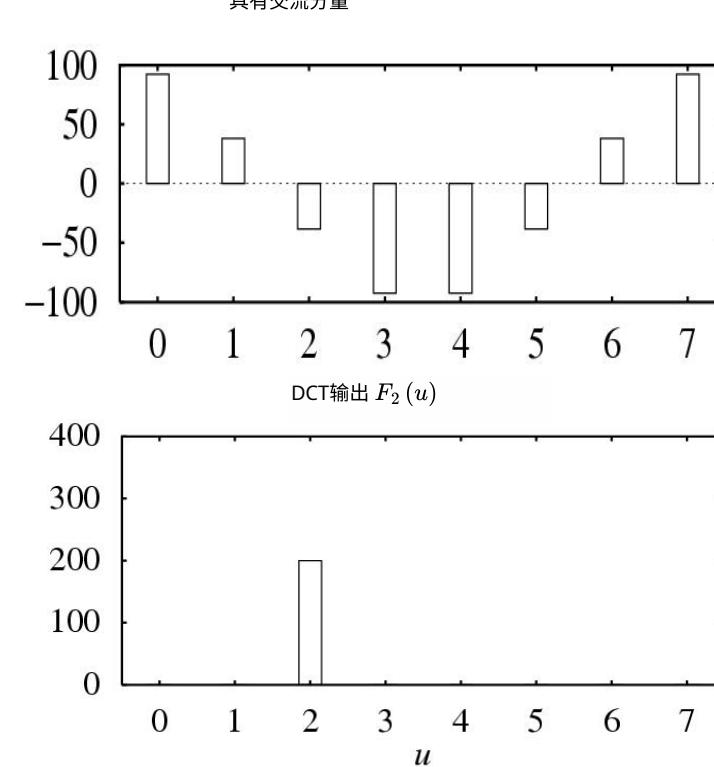
我们可以用类似的方法得到其他值

$$F_2(1) = F_2(3) = F_2(4) = \dots = F_2(7) = 0$$

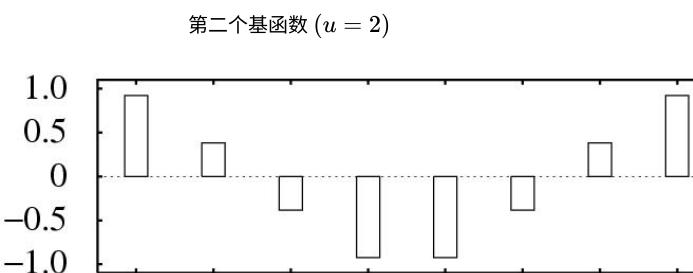
5.2 离散余弦变换 (DCT)

变化的信号 $f_2(i)$

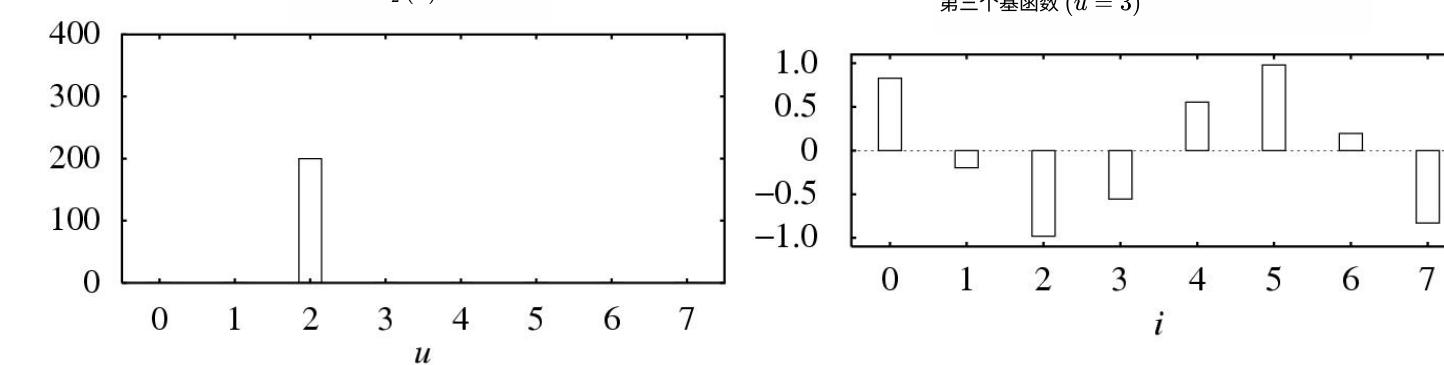
具有交流分量



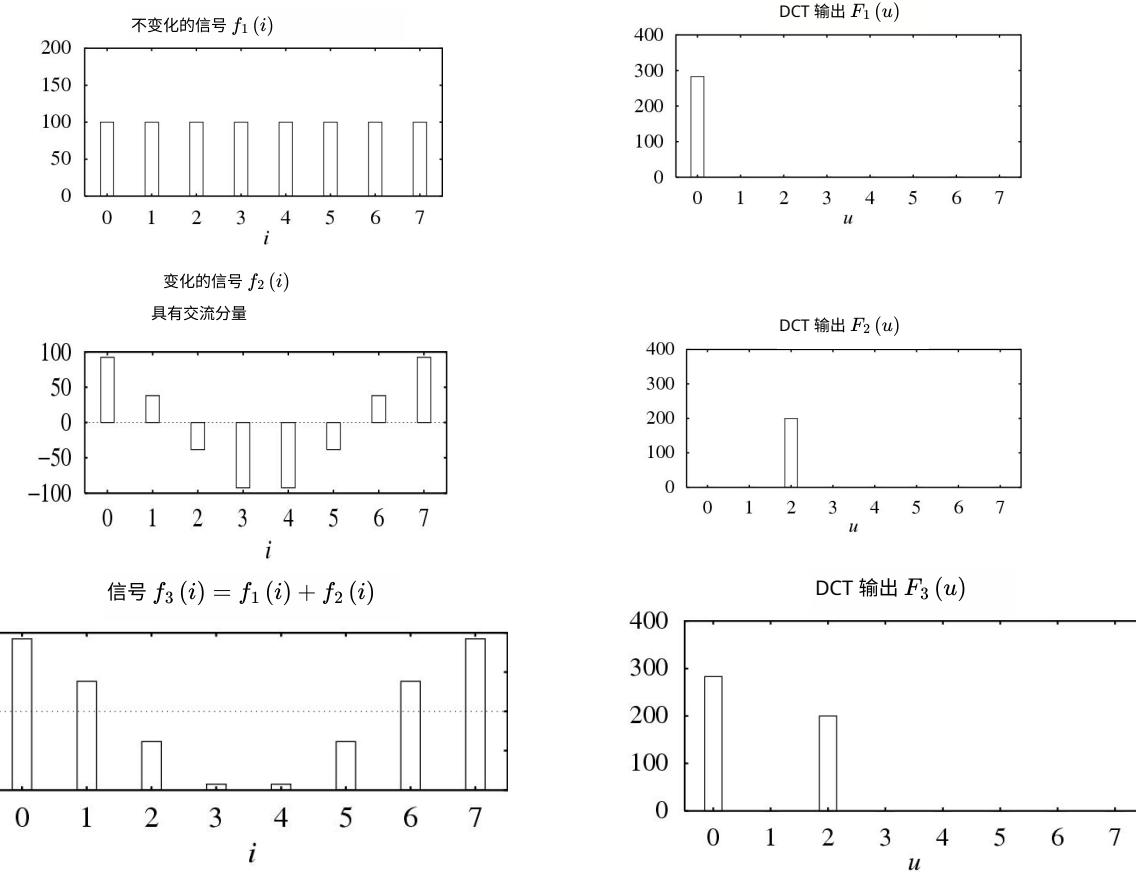
第二个基函数 ($u = 2$)



第三个基函数 ($u = 3$)



5.2 离散余弦变换 (DCT)



多媒体基础——有损压缩算法 (2025年春季)

41

5.2 离散余弦变换 (DCT)

$$\begin{aligned}
 T([f(0), f(1), \dots, f(n)]) &= [F(0), F(1), \dots, F(n)] \\
 &= [F(0), 0, \dots, 0] + [0, F(1), \dots, 0] + \dots + [0, 0, \dots, F(n)] \\
 &= T(\alpha_0 b_0) + T(\alpha_1 b_1) + \dots + T(\alpha_n b_n) \\
 &= T(\alpha_0 b_0 + \alpha_1 b_1 + \dots + \alpha_n b_n) \\
 \Rightarrow [f(0), f(1), \dots, f(n)] &\equiv \alpha_0 b_0 + \alpha_1 b_1 + \dots + \alpha_n b_n = \sum_{i=0}^n \alpha_i b_i \\
 \Rightarrow F(i) &= \alpha_i b_i b_i^T = \alpha_i \\
 \Rightarrow [f(0), f(1), \dots, f(n)] &= \sum_{i=0}^n F(i) b_i
 \end{aligned}$$

物理意义：通过基信号的线性组合来近似原始信号。

多媒体基础——有损压缩算法 (2025年春季)

5.2 离散余弦变换 (DCT)

一般来说，变换 T (或函数) 是线性的，当且仅当

$$T(\alpha p + \beta q) = \alpha T(p) + \beta T(q) \quad (8.21)$$

其中 α 和 β 是常数， p 和 q 是任意函数、变量或常数。

根据式8.17或8.19中的定义，由于离散余弦变换 (DCT) 仅使用简单的算术运算，因此可以很容易地证明该性质。

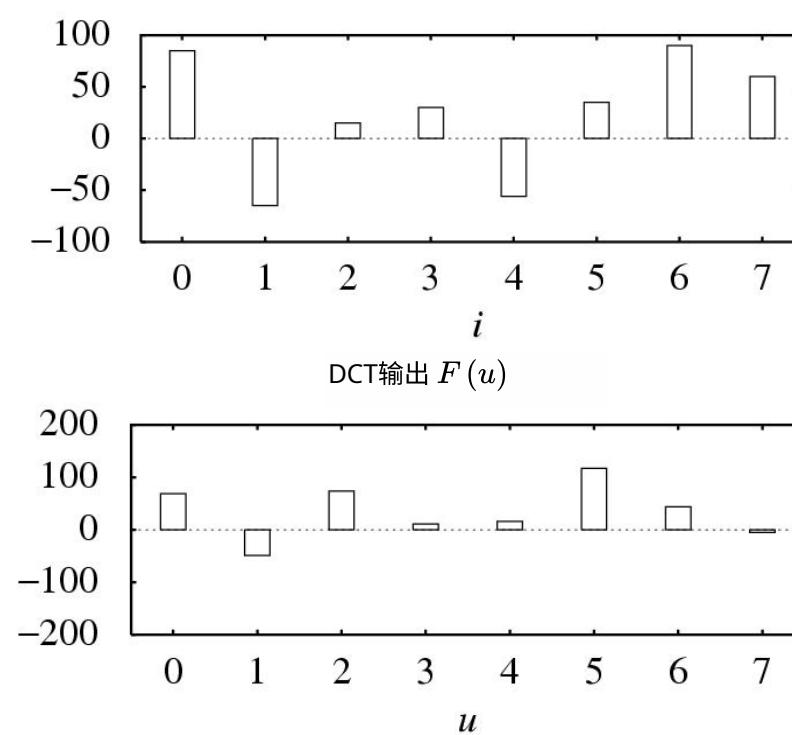
多媒体基础——有损压缩算法 (2025年春季)

42

5.2 离散余弦变换 (DCT)

85 -65 15 30 -56 35 90 60

任意信号 $f(i)$



多媒体基础——有损压缩算法 (2025年春季)

43

5.2 离散余弦变换 (DCT)

• 一维逆离散余弦变换 (IDCT)

• 如果 $F(u)$ 包含
 $(u = 0 \dots 7) : 69 - 4974111611744 - 5$ - 逆离散余弦变
 换 (IDCT) 可以通过八次迭代实现：
 迭代 0 : $\tilde{f}_i = \frac{C(0)}{2} \cos 0 \cdot F(0) \approx 24.3$

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

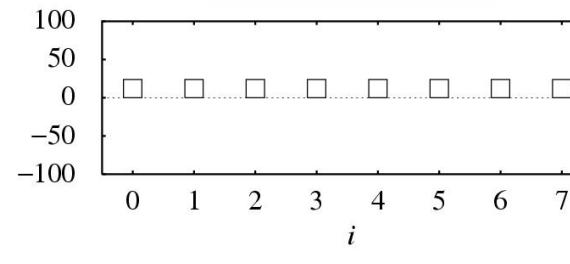
$$\approx 24.3 - 24.5 \cdot \cos \frac{(2i+1)\pi}{16}$$

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

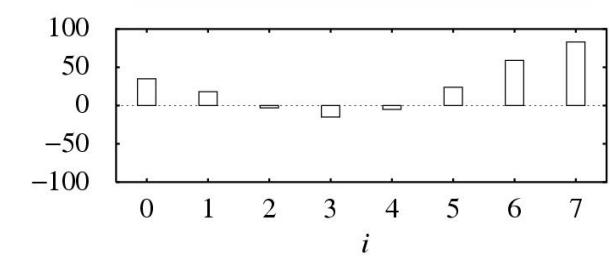
$$\approx 24.3 - 24.5 \cdot \cos \frac{(2i+1)\pi}{16} + 37 \cdot \cos \frac{(2i+1)\pi}{8}$$

5.2 离散余弦变换 (DCT)

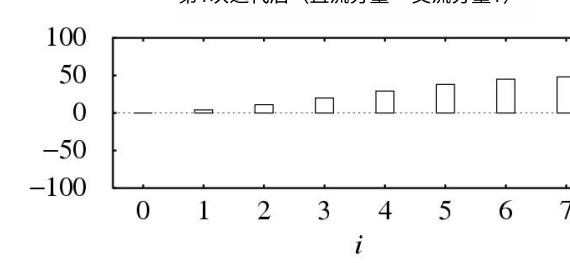
第0次迭代后 (直流分量)



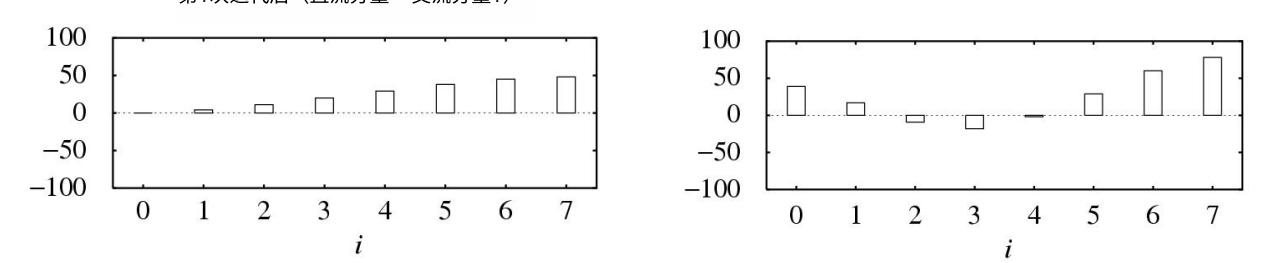
第2次迭代后 (直流分量 + 交流分量1 + 交流分量2)



第1次迭代后 (直流分量 + 交流分量1)



第3次迭代后 (直流分量 + 交流分量1 + 交流分量2 + 交流分量3)

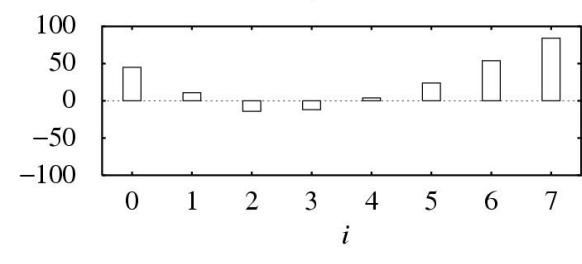


多媒体基础 - 有损压缩算法 (2025年春季)

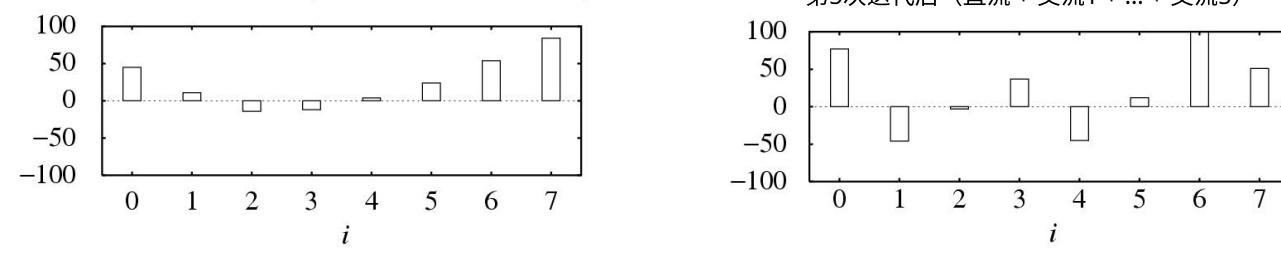
45

5.2 离散余弦变换 (DCT)

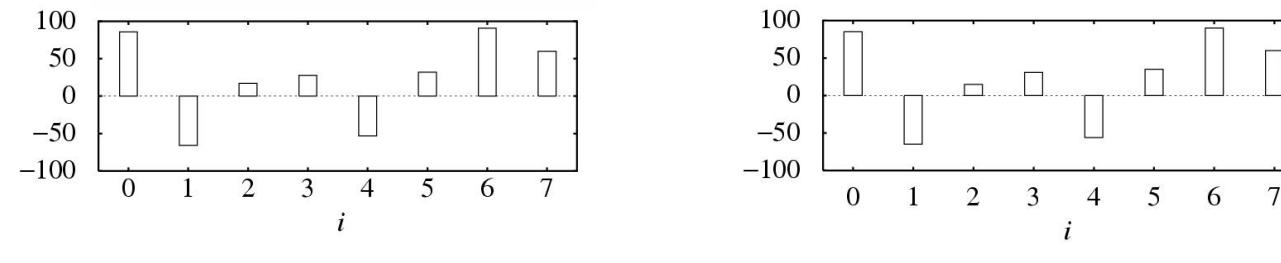
After 4th iteration (DC + AC1 + ... + AC4)



第5次迭代后 (直流 + 交流1 + ... + 交流5)



第6次迭代后 (直流 + 交流1 + ... + 交流6)



多媒体基础 - 有损压缩算法 (2025年春季)

46

5.2 离散余弦变换 (DCT)

• 与DCT相关的概念

• 直流 (DC) 和交流 (AC)

• 分别表示恒定和可变幅度；

• 余弦变换

• 用于确定信号的交流和直流分量幅度的过程。 - 离散余弦变换：整数索引 - $U = 0$ ，我们得到直流系
 数； - $U = 1, 2, \dots, 7$ ，我们得到第一个到第七个
 AC 系数。 - 逆离散余弦变换：使用直流、交流和余弦
 函数来重建信号 - DCT和IDCT采用同一组被称为基函
 数的余弦函数

5.2 离散余弦变换 (DCT)

- DCT 能够在频域中处理或分析信号
- 假设 $f(i)$ 表示随时间 i 变化的信号

- 一维 DCT 将时域中的 $f(i)$ 转换为频域中的 $F(u)$ 。
- $F(u)$ 被称为频率响应，构成 $f(i)$ 的频谱

5.2 离散余弦变换 (DCT)

离散余弦变换 (DCT) 的特性

- DCT 生成信号 $f(i)$ 的频谱 $F(u)$ - 第 0 个 DCT 系数 $F(0)$ 是信号 $f(i)$ 的直流分量； - 其他七个 DCT 系数反映了信号 $f(i)$ 在不同频率下的各种变化分量； - 如果直流分量为负值，这意味着 $f(i)$ 的平均值小于零； - 如果 AC 为负值，这意味着 $f(i)$ 和某个基函数具有相同的频率，但其中一个恰好落后半个周期。

5.2 离散余弦变换 (DCT)

离散余弦变换 (二维) 定义：

- 给定图像上的函数 $f(i, j)$ ，二维离散余弦变换将其转换为新函数 $F(u, v)$ ，整数 u 和 v 的取值范围与 i 和 j 相同。

- 离散余弦变换的一般定义为：

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

5.2 离散余弦变换 (DCT)

- 在 JPEG 图像压缩标准中

- 定义图像块的尺寸为 $M = N = 8$ ；
- 二维离散余弦变换及其逆变换（二维离散余弦逆变换）的定义如下：

- 二维离散余弦变换 (2D DCT)：

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

- 二维离散余弦逆变换 (2D IDCT)：

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

5.2 离散余弦变换 (DCT)

- 二维离散余弦变换 (2D DCT) 可分解为两个一维离散余弦变换 (1D DCT) 步骤：

一维离散余弦变换步骤：

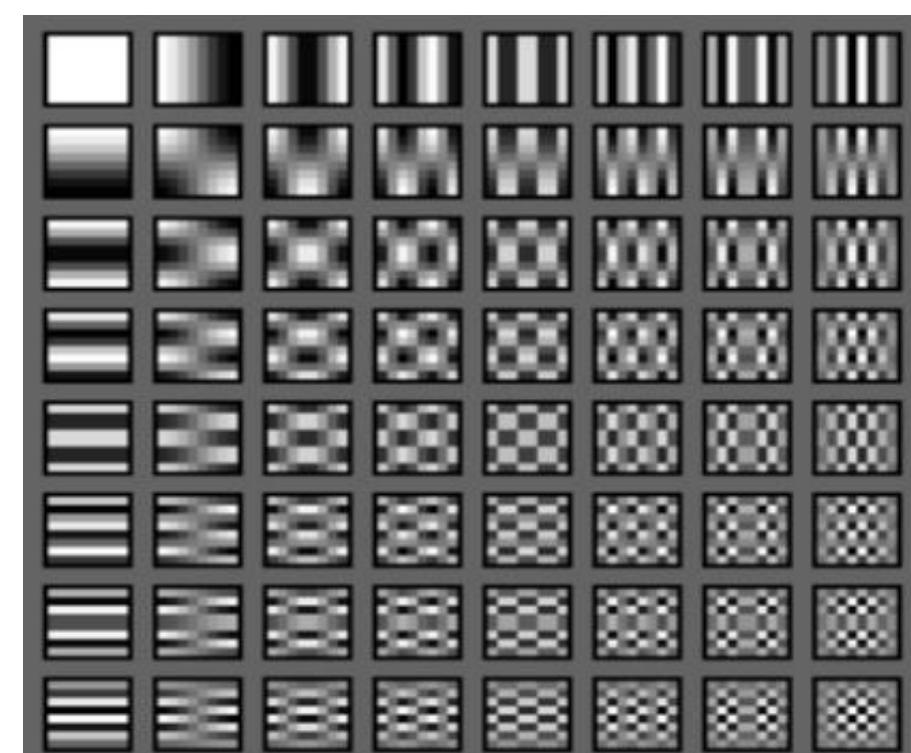
$$G(i, v) = \frac{1}{2} C(v) \sum_{j=0}^7 \cos\left(\frac{(2j+1)v\pi}{16}\right) f(i, j)$$

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

- 显然，这种简单的变换节省了许多算术步骤。所需的迭代次数从 8×8 减少到 $8 + 8$ 。

5.2 离散余弦变换 (DCT)

2D Basis Functions



5.3 离散余弦变换 (DCT) 与离散傅里叶变换 (DFT) 的比较

- DFT

- 离散余弦变换 (DCT) 与离散傅里叶变换 (DFT) 密切相关。DCT 是一种仅涉及 DFT 实部的变换。

- 对于连续信号，我们定义连续傅里叶变换 F 如下：

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

- 利用欧拉公式，我们有

$$e^{ix} = \cos(x) + i \sin(x)$$

- 由于使用数字计算机需要我们对输入信号进行离散化，我们定义一个对输入信号 $\{f_0, f_1, \dots, f_7\}$ 的 8 个样本进行操作的离散傅里叶变换 (DFT) 如下：

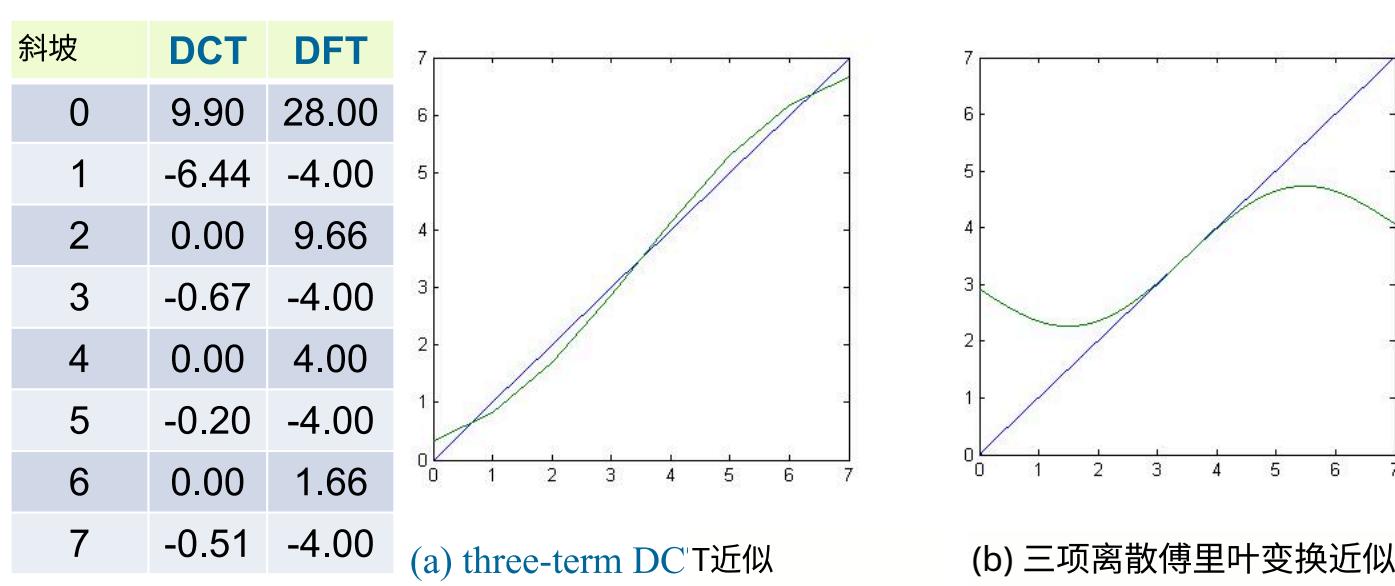
$$F_{\omega} = \sum_{x=0}^7 f_x \cdot e^{-\frac{2\pi i \omega x}{8}}$$

- 明确写出正弦和余弦项，我们得到

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

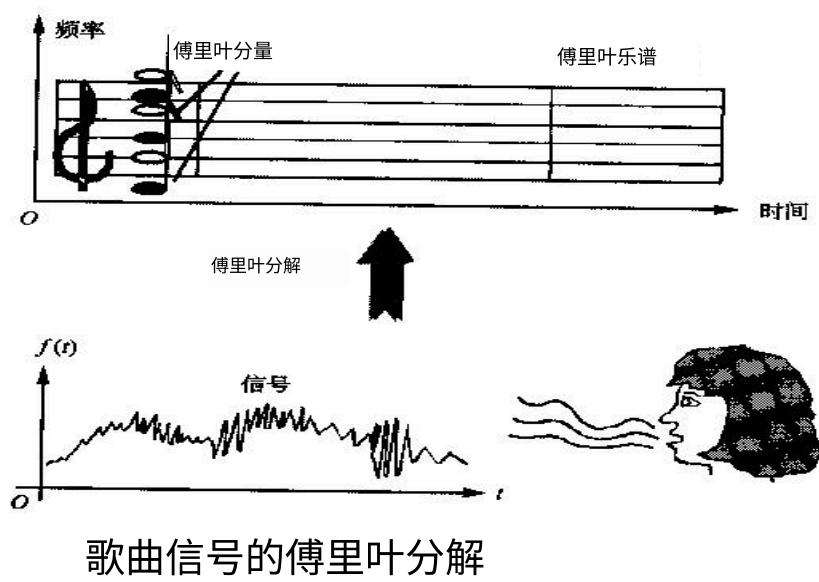
- 离散余弦变换 (DCT) 仅使用离散傅里叶变换 (DFT) 的余弦基函数的原理是，我们可以通过对原始输入信号进行对称复制来消除离散傅里叶变换 (DFT) 的虚部。

- 8 个输入样本的离散余弦变换 (DCT) 对应于由原始 8 个输入样本及其对称副本组成的 16 个样本的离散傅里叶变换 (DFT)，如图 8.10 所示。

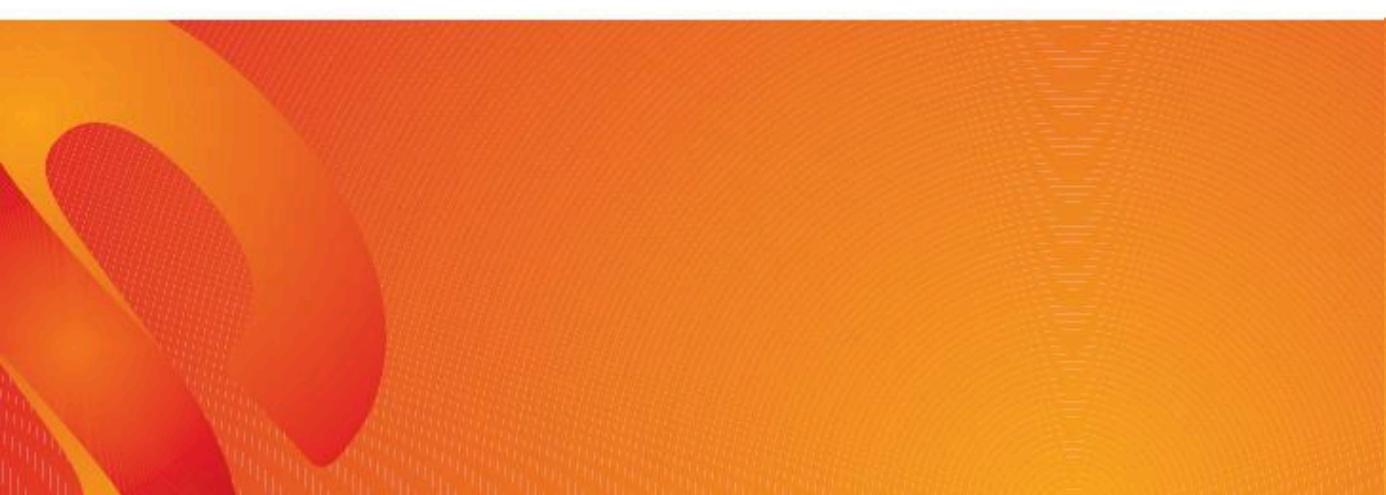


6.1 引言

- 离散傅里叶变换 (DFT) 和离散余弦变换 (DCT) 可以在频域中提供非常精细的分辨率，但没有时间分辨率。

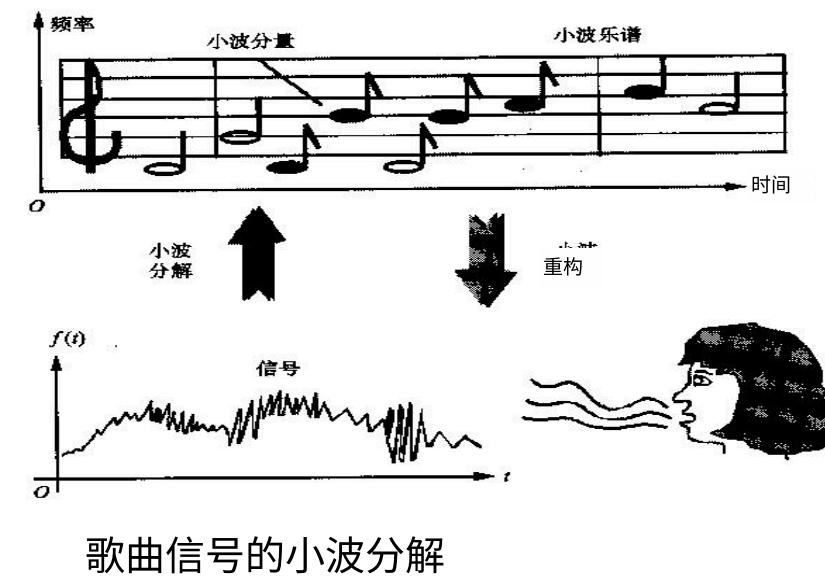


6. 基于小波的编码



5.1 引言

- 小波变换旨在时间和频率上都以高分辨率表示信号。



5.2 小波变换示例

- 假设给定以下输入序列。

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

- 考虑一种变换，用其成对平均值 $x_{n-1,i}$ 和差值 $d_{n-1,i}$ 替换原始序列，定义如下：

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

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

- 平均值和差值仅应用于输入序列中第一个元素索引为偶数的连续对。因此，每个集合 $\{x_{n-1,i}\}$ 和 $\{d_{n-1,i}\}$ 中的元素数量恰好是原始序列中元素数量的一半。

5.2 小波变换示例

- 通过连接两个序列 $\{x_{n-1,i}\}$ 和 $\{d_{n-1,i}\}$ 形成一个长度与原始序列相等的新序列。得到的序列是

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

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

- 该序列的元素数量与输入序列完全相同 - 变换并未增加数据量。

• 由于上述序列的前半部分包含原始序列的平均值，我们可以将其视为对原始信号的更粗略近似。该序列的后半部分可以视为前半部分的细节或近似误差。多媒体基础 - 有损压缩算法 (2025年春季)

5.3 一维哈尔变换

很容易验证，使用这些关系可以从变换后的序列重构原始序列

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

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

此变换为离散哈尔小波变换。

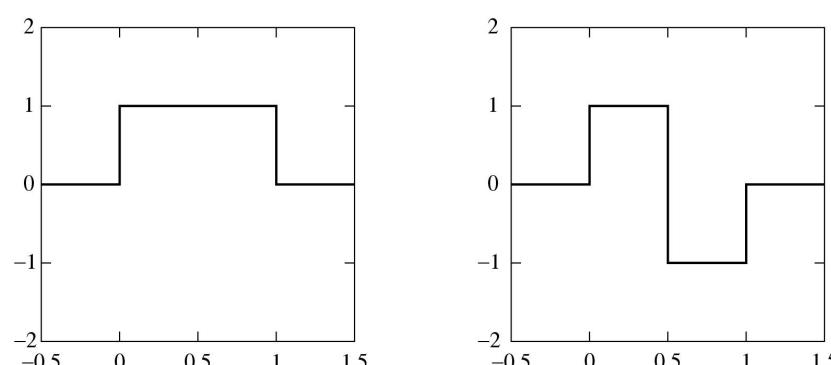


图8.12: 哈尔变换: (a) 尺度函数, (b) 小波函数。

5.4 二维哈尔小波变换

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	63	127	127	63	0	0
0	0	127	255	255	127	0	0
0	0	127	255	255	127	0	0
0	0	63	127	127	63	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

(a)

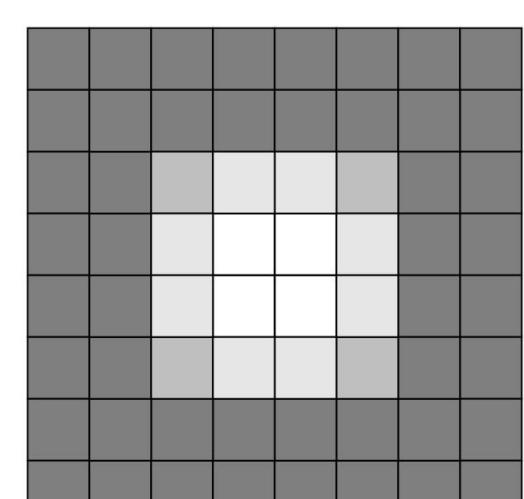


图8.13: 二维哈尔小波变换的输入图像。(a) 像素值。(b) 显示为 8×8 图像。

5.4 二维哈尔小波变换

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	95	95	0	0	-32	32	0
0	191	191	0	0	-64	64	0
0	191	191	0	0	-64	64	0
0	95	95	0	0	-32	32	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

图8.14：二维哈尔小波的中间输出
变换。

5.4 二维哈尔小波变换

0	0	0	0	0	0	0	0
0	143	143	0	0	-48	48	0
0	143	143	0	0	-48	48	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	-48	-48	0	0	16	-16	0
0	48	48	0	0	-16	16	0
0	0	0	0	0	0	0	0

图8.15：二维哈尔小波第一级的输出
变换。

5.4 二维哈尔小波变换

5.4 二维哈尔小波变换

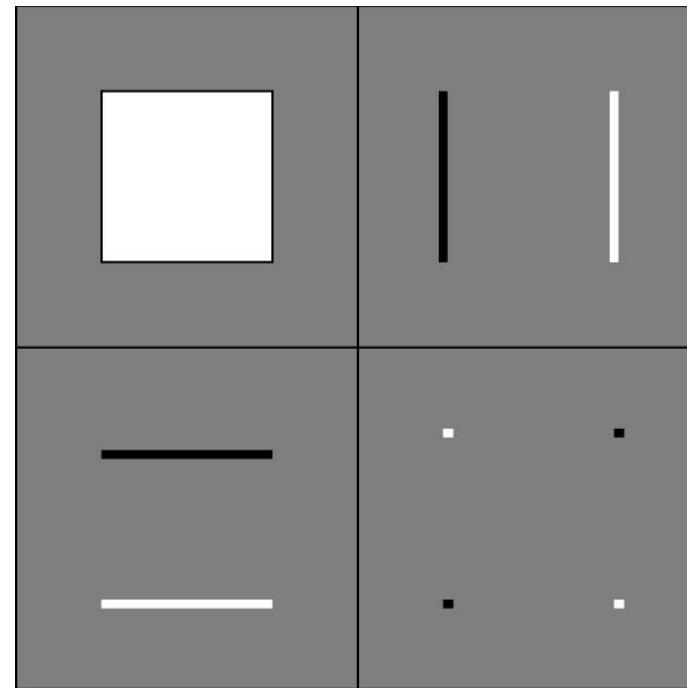


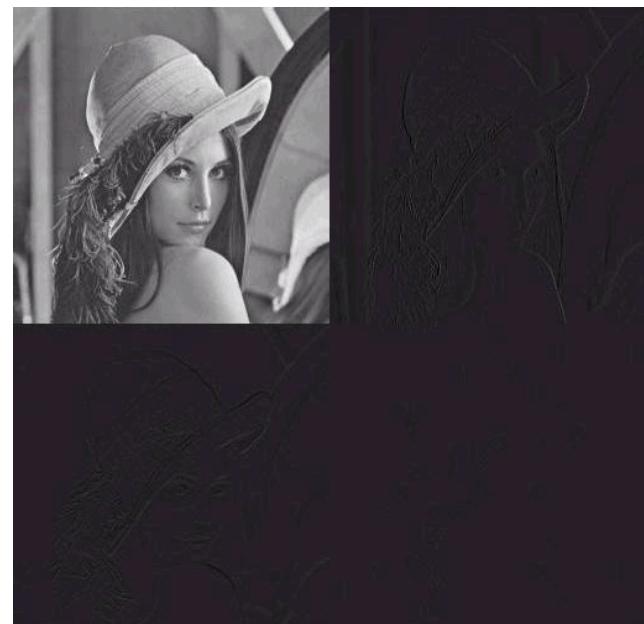
图8.16：小波的简单图形说明
变换。



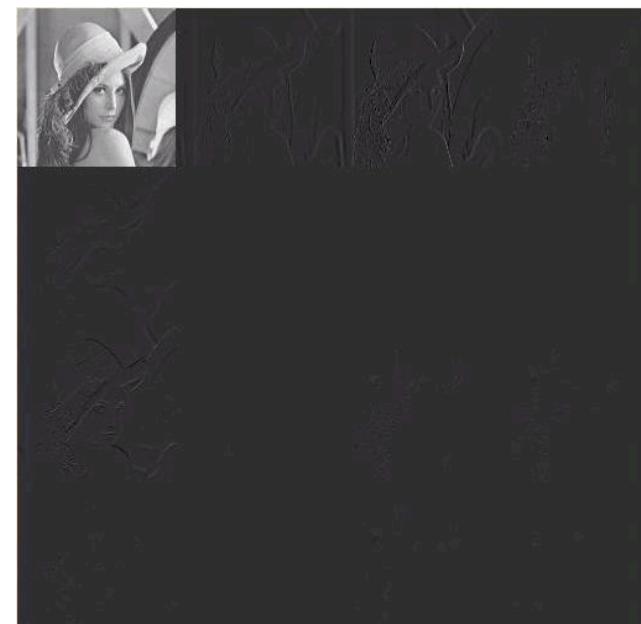
原始图像

水平小波变换

5.4 二维哈尔小波变换



小波水平和垂直变换（一级）



小波变换（二级）

结束

谢谢！

邮箱：junx@cs.zju.edu.cn