

子波域中图象数据的矢量量化*

楼生强 皇甫堪 周良柱 万建伟

(国防科技大学电子技术系 长沙 410073)

摘要 介绍了子波变换的基本理论及特点,并用它对图象进行三层分解。矢量量化优于标量量化。它已成功地应用于图象的数据压缩。文中应用矢量量化法量化图象子波域数据,取得较满意的压缩效果。

关键词 子波变换, 矢量量化, 图象编码

分类号 TN919.31

Image Coding Using Vector Quantization in the Wavelet Transform Domain

Lou Shengxiang Huangpu Kang Zhou Liangxhu Wan Jianwei

(Department of Electronic Technology, NUDT, Changsha, 410073)

Abstract In this paper, the theory and properties of wavelet transform are introduced. Vector quantization, which has been applied successfully in the image coding, is more efficient than scale quantization. We use this technique for image coding in the wavelet transform domain and get a satisfactory simulation result.

Key words wavelet transform, vector quantization, image coding

1 子波变换及其特点

近年来应用塔形分解、子波变换的图象子带编码技术取得了极大成功。它们不仅能取得较高的压缩比,而且有较好主观质量的重构图象。利用子波变换的编码技术,是具有高主观质量的重构图象、高压缩比图象数据压缩技术的重要研究方向之一。

正交二进子波变换在图象数据压缩领域中经常得到应用。二进子波基通过单个子波函数 $w(\cdot)$ 的平移及扩展得到:

$$w_{m,n}(t) = 2^{-(m/2)} w\left(\frac{t}{2^m} - n\right) \quad (m, n) \in Z$$

对任意平方可积函数 $f(t) \in L^2(R)$, 其二进子波变换 $Wf(m, n)$ 定义为

* 1996年2月9日收稿

$$Wf(m, n) = \langle f(t), w_{m,n}(t) \rangle = \int_{-\infty}^{+\infty} f(t)w_{m,n}(t)dt$$

从上式可知,二进子波变换是连续子波变换的离散化,它给出了函数 $f(t)$ 分辨率(或尺度)为 2^m 时点 n 处的近似。相反地,任意平方可积函数 $f(t) \in L^2(R)$ 能表示为二进子波基加权求和的形式

$$f(t) = \sum_{m,n} Wf(m, n)w_{m,n}(t)$$

等式成立(函数能从其二进子波变换完全重构)的条件是

$$\sum_{m=-\infty}^{m=\infty} |\hat{W}(2^m \omega)|^2 = 1$$

这里 $\hat{W}(\omega)$ 是 $w(t)$ 的傅立叶变换。实际上,上式也是 $\{w_{m,n}(t); m, n \in Z\}$ 构成 $L^2(R^2)$ 上正交基的条件。

从文献[1]可知,子波正交基能从多分辨率分析过程中得以构造。Mallat 在多分辨率过程中引入了尺度函数 $\phi(t)$, 对于正交子波变换, $\phi(t)$ 和 $w(t)$ 需满足如下条件:

$$\phi(t) = \sum_n h(n)\phi(2t - n)$$

$$w(t) = \sum_n (-1)^n g(n+1)w(2t + n)$$

其中 $H = \{h(i), i \in Z\}, G = \{g(i), i \in Z\}$ (1)

H, G 是一对滤波器,在时域上它们满足:

$$\sum_n h(n) = \sqrt{2}; \quad \sum_n g(n) = 0$$

$$g(n) = (-1)^n h(1-n)$$

相应地,在频域上它们满足:

$$(1) \quad |H(\omega)|^2 + |H(\omega + \pi)|^2 = 1 \quad (2a)$$

$$(2) \quad |G(\omega)|^2 + |G(\omega + \pi)|^2 = 1 \quad (2b)$$

$$(3) \quad H(\omega)G^*(\omega) + H(\omega + \pi)G^*(\omega + \pi) = 0 \quad (2c)$$

其中, $H(\omega), G(\omega)$ 为 $H = \{h, i \in Z\}, G = \{g, i \in Z\}$ 的傅里叶变换,子波函数和尺度函数分别能通过 H, G 的无穷迭代得到^[3]。根据 Mallat 算法,二进正交离散子波变换对一维序列的分解与合成可以用以下滤波器组实现:

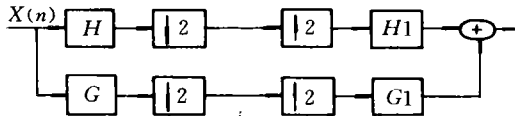


图1 一维信号的子波分解与合成, H_1, G_1 分别是 H, G 的共轭滤波器

其中, H, G 是(1)式给出的滤波器组,分别为低通和高通滤波器,而 H_1, G_1 分别为 $H,$

G 的共轭转置。Daubechies 在文[3]中从(2)式出发, 利用多项式分解, 得到滤波器组 H 、 G 的解。

子波变换能把一个非平稳信号分解为几个相对平稳的信号, 使分解后信号的统计特性, 与原始信号相比要简单一些。这对编码是非常有利的: 能用更少的数据来近似描述信号。另外, 子波变换具有较高的编码增益, 滤波器组的编码增益定义为:

$$G_{TC} = \left(\frac{1}{N} \sum_{i=1}^N \delta_i^2 \right) \left(\prod_{i=1}^N \delta_i^2 \right)^{-1/N}$$

其中, δ_i 为第 i 个滤波器输出的方差。编码增益体现了每个滤波器的输出用最优化器量化后, 引入的均方差与直接量化原始信号引起的均方误差相比减少的程度。如果变换是用于编码, 那么它的值越大, 越利于数据压缩。子波变换具有较高的编码增益, 特别是当信号的相关性较大时^[4], 对本文所选的 $512 \times 512 \times 8$ 的灰度图象 (见图 3) 进行三层子波分解后, 编码增益为 66.6。近来研究表明^[2], 人的视觉系统在处理信息时有多通道特性, 多通道的等价滤波器组不仅有相近的 Q 值 (0.6~2.0) 而且具有方向选择性, 它对水平和垂直方向的刺激最敏感。因此, 子波变换与人眼的视觉系统在一定程度上相匹配。如果根据人眼的视觉特性选择视觉敏感信息, 舍弃人眼不敏感信息, 就能取得在相同压缩比下, 主观上更高质量的重构图象。现已有在这方面作了些尝试的图象数据压缩方法, 收效不大, 只是定性地利用这一特性。

2 矢量量化 (Vector Quantization)^[5]

矢量量化优于标量量化。这是因为矢量量化能有效地利用矢量中的 4 种相关性: 线性相关、非线性相关、概率密度函数的形状以及矢量量化的维数。标量量化则只能利用数据中的线性相关、概率密度函数。矢量量化可以分为随机矢量量化和非随机矢量量化, 非随机矢量量化就是格形矢量量化。两者的区别在于: 随机矢量量化的码本是对输入数据训练后得到的, 而格形矢量量化则不考虑数据特性, 只作 R^k 空间的分隔, 每个子空间的中心作为码元, 两者各有长处。前者的压缩比高, 运算量大, 稳健性差, 而后者则相反。

矢量量化把信号每 K 个连续取样点作为一组, 形成 K 维欧氏空间中的一个矢量, 然后对矢量进行量化。它可以认为是一个 K 维欧氏空间 X 到 K 维欧氏空间的一个子集 Y 的映射 Q :

$$Q: X \rightarrow Y$$

其中 Y 是由 $\{Y_0, Y_1, \dots, Y_{M-1}\}$ 生成的子空间。称 $\{Y_0, Y_1, \dots, Y_{M-1}\}$ 为码本, 其中的元素 Y_i 称为码元。它有两个功能: 在编码端, 根据输入的数据 X , 通过映射 Q , 得到匹配码元的序号; 在解码端, 根据序号得到重构的数据矢量。

在定义失真测度 (距离或误差测度) $d(x, \bar{x})$ 后, Q 映射就是要使映射误差极小化, 失真测度是数据进行矢量量化的基础, 多数随机矢量量化的码书设计算法及编码时码元搜寻算法都要建立于这一基础之上。最简单的失真测度是均方差测度:

$$d(x, \bar{x}) = \|x - \bar{x}\|^2 = \sum_{i=1}^{K-1} (x_i - \bar{x}_i)^2$$

其中 $x = \{x_0, \dots, x_{K-1}\}$ 为数据矢量, $\bar{x} = \{\bar{x}_0, \dots, \bar{x}_{K-1}\}$ 为量化后的数据矢量, 其它失真测

度及 VQ 的特点见文献[5]。码本训练算法是矢量量化的关键技术之一。最早的码本训练算法是 LBG 法, 具体过程如下:

步骤 1 设 ϵ 为误差门限, M 为码本的码元个数, m 为迭代次数, 初始值 $m=0$, 初始码本为: $A_0 = \{y_0, \dots, y_{M-1}\}$ 。码本的训练数据序列为: $X = (x_0, x_1, \dots, x_{N-1})$, 长度为 N , 定义码本与数据序列的误差为

$$D(x, Q(x)) = \frac{1}{N} \sum_{i=1}^{N-1} d(x_i, \hat{x}_i)$$

步骤 2 根据给定的码本 A_m , 根据失真测度计算输入数据与码元的误差, 并用此作为数据分割标准。设: $P(A_m) = (S_0, S_1, \dots, S_{M-1})$ 为对应于 A_m 的分割。如果对 $l=0, \dots, M-1$, 都有 $d(x_i, y_l) \leq d(x_i, y_j)$, 则 $x_i \in S_j$ 。

步骤 3 计算最小失真:

$$D_m = D(X, Q_m(X)) = \frac{1}{N} \sum_{j=0}^{M-1} \min_{y \in A_m} d(x, y)$$

步骤 4 如果 $(D_m - D_{m-1})/D_m \leq \epsilon$, 则 A_m 为所需码本。否则, 计算 $\hat{x}(S_i)$,

$$\hat{x}(S_i) = \frac{1}{\|S_i\|} \sum_{x_j \in S_i} x_j, \quad i = 0, 1, \dots, M-1$$

用 $\{\hat{x}(S_0), \dots, \hat{x}(S_{M-1})\}$ 代替 A_m , $m=m+1$, 到步骤 2。

LBG 法对初始码本敏感, 可能陷入局部最优值。现有基于统计方法及神经网络理论的码本训练方法, 在文献[5]中作了介绍, 例如覆盖算法, 梯度算法等。在不同的条件(信源特点, 失真测度等)下, 它们各有优缺点。

仙农的率-失真函数 $R(D)$ 是矢量量化的理论基础, 率-失真函数 $R(D)$ 定义为: 对于给定的失真, 编码算法所能达到的最小码率。它给出编码器的工作性能, 不仅适合矢量量化, 而且适合所有信源编码方法。率-失真函数指出, 利用矢量量化, 编码性能可任意地接近率-失真函数, 其方法是增加维数。

3 结果与结论

对二维信号图像作子波分解时, 需考虑行列两个方向。本文采用文献[3]中四点正交基对图像作可分离的子波分解(见图 2)。

数据图像是离散的二维有限长度数据集, 而子波变换只能把无限长度序列完全重构。在不增加图像分解和合成过程中数据总量的条件下, 为了减少重构图像的

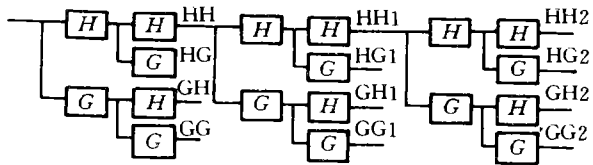


图 2 图像三层子波分解框图 (HH2, HG2, CH2, GG2, ..., HG, GH, GG) 为子带图像

边界效应, 且分解后的子带数据特性(主要是方差和数据的分布范围)要利于压缩, 对图像进行子波分解和重构之前, 需进行数据的扩展。扩展方法在分解数据利于压缩和重构图像的边界效应之间作折衷选择。文献[6]中, 讨论变换用于数据压缩时数据扩展的有关问题。本人采用周期扩展法和边界值扩展法, 对图 3 进行三层子波分解(见图 2)分解

后得到 10 个字带。结果表明,采用第一种方法几乎能完全重构图像;采用第二种方法时,虽然在重构图像中引入误差,但各高频子带数据的方差要比第一种的小,重构图像无明显边界效应。



图 3 原图像



图 4 恢复图像

分解后的 9 个细节图像分别代表原图某个频带、某个方向的信息,因此本人对不同的高频子带设计不同的码本进行矢量量化,并选用不同大小的码本得到 3 幅重构图像。结果见表 2,码书类型见表 1,图 4 是结果一的重构图像。3 幅重构图像在主观质量上与原始图像区别不大,误差主要集中于黑白对比明显的边缘,例如图像中的帽子边缘低亮区域有可分辨的灰色毛茸状的视觉误差,而在高亮区域,人眼则无法觉察。根据经验及本文的模拟结果,对高层高频子带(图 3 中的 HG、GH、GG)一般宜采用码元数较少的码本,码元的维数可以高一些,对低层高频子带则相反。从子波合成过程可知,低层子带数据点的误差对原始图像影响的范围比高层子带的要广,也就是它引入的误差分量低频占优势。这和人眼的视觉特性相符。然而现有技术水平还无法对这结论作定量分析,有待于进一步深入研究。

表 1 码本类型

| | HG | GH | GG | HG1 | GH1 | GG1 | HG2 | GH2 | GG2 |
|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 类型 1 | 128/8 | 128/8 | 128/8 | 128/4 | 128/4 | 128/4 | 256/2 | 256/2 | 256/2 |
| 类型 2 | 128/8 | 128/8 | 64/8 | 64/4 | 64/4 | 32/4 | 128/2 | 128/2 | 64/2 |
| 类型 3 | 128/8 | 128/8 | 64/8 | 16/4 | 16/4 | 8/4 | 16/2 | 16/2 | 8/2 |

* 表中数字表示码元个数及矢量维数,例 128/8 表示 128 个 8×8 码元。

表 2 压缩结果

| | 码本类型 | PSNR | SNR | 压缩比 |
|------|------|------|------|------|
| 结果 1 | 1 | 41.3 | 37.6 | 21.1 |
| 结果 2 | 2 | 38.8 | 34.1 | 24.6 |
| 结果 3 | 3 | 36.8 | 32.2 | 27.7 |

* 低频子带 HH2 采用 01 位平面编码方法,结果为:SNR=35.7,需 28657bit

VQ 法的运算量很大,特别是进行码书设计时,灵活性不够,适用面有限。数据特性发生变化时,如果不采用自适应技术,就必须重新训练码本,否则影响重构图像的质量。而如采用自适应功能,则增加编码和解码的复杂性。另外,如果码书的码元数不够,也会影响重构图像的质量。

现代图像编码技术的特点是最大可能地利用人眼的视觉特性和图像固有的特性(场景和物体的先验知识),其发展方向为模型法^[7]。具体到利用子波变换的图像编码技术而言,本人认为值得深入研究的课题有:

(1) 子波域中的视觉模型编码方法。由于子波变换与人眼多通道特性相匹配,因此子波域中选取视觉模型比在原图更有效率。然而现在人们对视觉系统的了解还不足以建立选取视觉模型的有效算法。

(2) 结合其它技术的编码方法。例如分形、神经网络理论。子波域中数据的统计特性要比原图简单一些,这有利于寻求分形的迭代函数系(IFS),而神经网络则可以用来模拟人眼对视觉信息的综合过程。此时,可以用子波变换来模拟人眼对视觉信息的分解过程。

(3) 选择使子波分解更加接近人眼分解视觉信息过程的滤波器组。这对于我们理解子波域中的图像数据以及选择子波域中高效的视觉信息是非常重要的。这是建立有效编码技术的基础。

参 考 文 献

- 1 Mallat S F. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation IEEE T—PAMI, 1989, (11): 676~693
- 2 Mallat S G. Multifrequency Channel Decomposition of Images and Wavelet Model IEEE T—ASSP, 1989, (37): 2091~2110
- 3 Daubechies I. Orthonormal Bases of Compactly Supported Wavelets Comm. Pure Applied Math. 1988, (41): 909~996
- 4 楼生强. 子波变换在图像数据压缩中的应用:[学位论文],国防科技大学,1994
- 5 胡征,杨有为. 矢量量化,西安电子科技大学出版社
- 6 Image Compression. Using Address—Vector Quantization. IEEE T—C 1990, (38): 2166
- 7 Gunnar karlsson, Extension of finite Length Signals for Sub—Band Coding. IEEE SP 1989, (17): 161~168
- 8 徐孟侠. 图像编码进展. 通信学报, 1993, (11): 40~47

(责任编辑 潘 生)