

多媒体通信中基于对象的格矢量量化

周利莉 勒中鑫

(解放军信息工程学院, 郑州 450002)

摘要 在格结构的分析与研究的基础上,详细讨论了格的重要算法,并给出了小波域图象对象的格矢量量化方法。结合小波变换系数的分布特点,对某些小波变换系数进行标量量化时,采用了一种连续逼近的方法,并利用所讨论的算法对多数小波变换系数进行了格矢量量化。实验表明,格矢量量化速度快,编码效率高,能较好地达到多媒体通信中图象压缩编码要求。

关键词 小波变换 基格 矢量量化

0 引言

近年来,随着多媒体通信的发展,宽带综合业务对图象(或视频图象)数据的压缩提出了新要求,目前处于研究之中的基于面向对象的视频图象压缩技术正是适应于这种要求而发展的。不仅数据压缩比大、重构质量高,而且传输比特率可按信息高速公路要求分配。其中一重要研究方向是基于图象对象的自适应小波变换。图象对象经过不同基的小波变换,被划分为各个不同的尺度。变换后的图象边缘更符合人的视觉系统(HVS)的要求,与原始图象相比也有利于提取纹理信息。利用这些特性可使视频图象数据获得很高的压缩比。本文主要讨论如何应用格矢量量化方法对图象对象进行有效编码。

众所周知,矢量量化的关键在于码本的设计,它不仅影响编码效率,而且也关系到量化器运算复杂度,即所谓搜索问题。LBG算法是一经典算法^[1],它在最小均方误差的意义下,获得的码本是局部最佳的。对于某一输入矢量,根据匹配准则,在码本中找到与之匹配的码矢量,必须进行“全”搜索,由于码本的不规则性,随着矢量维数的增加,所要求的运算量会呈指数增加。就此人们提出了一些解决方法^[2],并且在图象及视频编码中得到了应用。Conway和Sloane提出了格矢量量化(LatticeVQ)算法^[3,4],它利用一些规则的格建造码本,这些码本均匀地分布在矢量空间的格点上,根据这个特点,一般矢量量化方

法中所涉及的复杂码本搜索以及存储等影响编码的因素在格矢量量化中都不存在了,不仅大大加快了量化速度,也大大提高了编码效率。因此,较好地解决了多媒体通信中视频编码效率及实时处理等关键环节^[5,6]。

1 格矢量量化

先对格矢量量化所涉及的基本概念做一简介。

1.1 格

如果 a_1, a_2, \dots, a_n 是 n 维欧氏实数空间 R^n 的 n 个线性独立矢量,则称下列集合

$$L_n = \{x \in R^n \mid x = (x_1, x_2, \dots, x_n) = u_1 a_1 + u_2 a_2 + \dots + u_n a_n, u_i \in Z\}$$

为一个 n 维格。Conway和Sloane为量化器所定义的格称为根(root)格(或叫基格),包括 $S_n (n > 1)$, $A_n (n > 1)$, $D_n (n > 3)$, E_8 以及Barnes-Wall Λ_{16} 等。它们的定义分别为:

$$(1) S_n = \{x = (x_1, x_2, \dots, x_n) \mid x_i \in Z\};$$

$$(2) A_n = \{x = (x_1, x_2, \dots, x_n) \mid x_i \in Z, \sum_i x_i = 0\};$$

$$(3) D_n = \{x = (x_1, x_2, \dots, x_n) \mid x_i \in Z, \sum_i x_i \text{ 为偶数}\};$$

$\equiv 0(\text{mod } 2)\}$;

$$(4) E_8 = D_8 \cup D'_8, D'_8 = \{d_8 + \alpha \mathbf{1}\}, d_8 \in D_8,$$

$$\mathbf{1} = (1, 1, 1, 1, 1, 1, 1, 1), \alpha = 0.5;$$

(5) $\Lambda_{16} = \bigcup_{i=1}^{32} (c_i + 2d_{16})$, $d_{16} \in D_{16}$, c_i 为 Hadamard 矩阵 \tilde{H}_{16} 的行码向量以及补矩阵 $\tilde{H}_{4_{16}}$ 的行码向量。

1.2 格矢量量化

如果把上述定义的格格看成量化器,则依据均方误差最小的准则,所有位于某格点 x_i 的 Voronoi 区域的点 y 都可以用 x_i 来近似,即集合

$$V(x_i) = \{y \in R^n \mid \|y - x_i\| \leq \|y - x_j\| \text{ for all } i \neq j\}$$

中的矢量 y 可用 x_i 来表示。关于不同格的 Voronoi 区域定义见文献[7]。

格矢量量化的突出优点是能够提供一个快速编码/译码过程。一旦 VQ 码本是依据一个指定的格集建立起来的,则由于格结构的特点,最邻近(NN)搜索只会是一些特定的码矢量中进行,使得我们可以在矢量间的距离计算与比较等方面采用快速算法,简化 LBG 中的计算,大大提高量化匹配速度和编码效率。其中格 D_n 及 E_8 分别叙述如下:

1.2.1 格,对于一个输入矢量

$y = (y_1, y_2, \dots, y_n) \in R^n$, 定义映射 f 为:

$$f(y) = (f(y_1), f(y_2), \dots, f(y_n))$$

式中 $f(y_i)$ ($1 \leq i \leq n$) 依据 y_i 的取值范围而定,具体如下:

①如果 $y_i = 0$, 则 $f(y_i) = 0$, 并记 $w(y_i) = 1$;

②如果 $0 < m \leq y_i \leq m + \frac{1}{2}$, 则 $f(y_i) = m$, 并 $w(y_i) = m + 1$;

③如果 $0 < m + \frac{1}{2} < y_i < m + 1$, 则 $f(y_i) = m + 1$, 并 $w(y_i) = m$;

④如果 $-m - \frac{1}{2} \leq y_i \leq -m < 0$, 则 $f(y_i) = -m$, 并 $w(y_i) = -m - 1$;

⑤如果 $-m - 1 < y_i < -m - \frac{1}{2}$, 则 $f(y_i) = -m - 1$, 并 $w(y_i) = -m$ 。

若记 $y_i = f(y_i) + \delta(y_i)$, 可以看出 $|\delta(y_i)| \leq \frac{1}{2}$, 如果 $|\delta(y_k)| \geq |\delta(y_i)|$, 对 $1 \leq i \leq n$, 那么有

$g(y) = (f(y_1), f(y_2), \dots, w(y_k), \dots, f(y_n))$ 。很明显, $g(y)$ 和 $f(y)$ 只在坐标 k 处的值相差 1, 其余各个坐标上对应的值都相等。

利用 $g(y)$ 和 $f(y)$ 的定义, 在 D_n 中得到输入矢量的匹配矢量就可以用算法 1 描述, 即

算法 1: 给定输入矢量 $y = (y_1, y_2, \dots, y_n) \in$

R^n , 如果 $\sum_{i=1}^n f(y_i) \equiv 0(\text{mod } 2)$, 则 y 在 D_n 中的匹配矢量为 $\hat{y} = f(y) = (f(y_1), f(y_2), \dots, f(y_n))$; 否则 $\hat{y} = g(y) = (f(y_1), f(y_2), \dots, w(y_k), \dots, f(y_n))$, 其中 k 是满足条件 $|\delta(y_k)| \geq |\delta(y_i)|$ ($1 \leq i \leq n$) 的最小整数。

1.2.2 格,与算法 1 相类似

算法 2: 给定输入矢量 $y = (y_1, y_2, \dots, y_n) \in$

R^n , 分别计算 $f(y)$ 及 $g(y)$ 。如果 $\sum_{i=1}^n f(y_i) \equiv 0(\text{mod } 2)$, 则记 $\hat{y}_0 = f(y)$; 否则记 $\hat{y}_0 = g(y)$ 。对于 $y' = (y - \frac{1}{2})$, 分别计算 $f(y')$ 及 $g(y')$, 如果

$\sum_{i=1}^n f(y'_i) \equiv 0(\text{mod } 2)$, 则记 $\hat{y}_1 = f(y')$; 否则记 $\hat{y}_1 = g(y')$ 。比较 $\|y - \hat{y}_0\|$ 和 $\|y - \hat{y}_1\|$, 如果 $\|y - \hat{y}_0\| \leq \|y - \hat{y}_1\|$, 则在 E_8 中与 y 匹配的矢量为 $\hat{y} = \hat{y}_0$; 否则, $\hat{y} = \hat{y}_1$ 。

1.2.3 索引矢量

对于每一个量化器, 在搜索到输入矢量的匹配矢量后, 需要对该匹配矢量进行标记, 然后对每个标记号进行编码。格矢量量化按某一规则, 给每一个匹配矢量指定一个相应的索引矢量, 规则不同, 形成的索引也不同。本文采用下述算法 3 计算匹配格矢量的索引, 而在译码时, 由索引确定格矢量由算法 4 实现。

算法 3: 给定一个格矢量 $x = (x_1, x_2, \dots, x_n) \in L_n$, 如果限定码速率为 $(\log_2 r)$ bits/维, 则 x 的索引矢量 $k = (k_1, k_2, \dots, k_n)$, $k_i \equiv c_i (\text{mod } r)$, $c = (c_1, c_2, \dots, c_n) = xG^{-1}$ 。如果把 n 维格集看成是由 n 维欧氏实数空间 R^n 的一组基 v_1, v_2, \dots, v_n 整数张成, 则以这组基矢量作为行向量形成的矩阵 G , 称为 L_n 的生成矩阵。

算法 4: 给定索引矢量 $k = (k_1, k_2, \dots, k_n)$, 计算 $c = kG$, 根据算法 1 或算法 2 得到矢量 $\lambda = \frac{1}{r} \cdot c$ 的匹配格矢量 $\hat{\lambda}$, 则与索引矢量 k 对应的矢 $x = c - \hat{\lambda}$ 。

这里给出格 D_4 及 E_8 的生成矩阵 G_4 和 G_8 分别为:

$$G_4 = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

$$G_8 = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \end{bmatrix}$$

2 小波系数的格矢量量化

2.1 图象的小波变换

利用一个母小波的平移及伸缩对信号 $y(x)$ 进行变换得到一组系数 $\hat{y}_{m,n}$, 则 $y(x)$ 可以用 $\hat{y}_{m,n}$ 来表示, 即

$$y(x) = \sum_m \sum_n \hat{y}_{m,n} 2^{-m/2} \bar{\psi}(2^{-m}x - n),$$

$$\hat{y}_{m,n} = \int_R 2^{-m/2} \psi(2^{-m}x - n) y(x) dx.$$

把离散小波变换转化为物理可实现的滤波过程, 如图1所示:

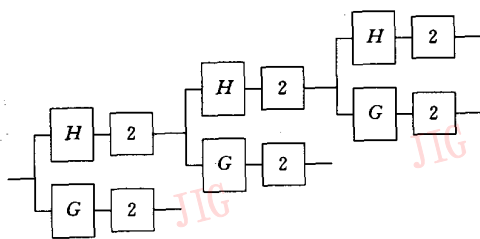


图1

图中所示的 $H(z)$ 和 $G(z)$ 为分析滤波器, 分别为低通滤波器和高通滤波器; 与此相对应是合成滤波器, 其低通滤波器为 $H'(z)$, 高通滤波器为 $G'(z)$, 它们必须满足:

$$H(z)G(-z) - H(-z)G(z) = cz^{-2m+1},$$

$$H'(z) = \frac{2}{c} z^{2m-1} G(-z),$$

$$G'(z) = -\frac{2}{c} z^{2m-1} H(-z).$$

其中 c 为任意常数。对图象做可分离的小波变换, 分别对其进行垂直方向及水平方向的小波变换, 所得小波变换系数, 其分布函数通常用 Laplacian 函数近似地表示。一般情况下, 每层所采用的小波基是相同的, 也就是说, 图1中的低通滤波器都一样, 高通滤波器也是一样的。从自适应角度来看, 垂直方向与水平方向的小波变换常常是不同的。如果在视频通信中采用自适应的小波基对信号进行小波变换, 更符合小波视频对象通信的要求。

2.2 格矢量量化的改进

上述的格矢量量化并没有限制格的大小, 其所含格点数目是无限的。实际应用时格点数目是有限的, 于是需要对格进行相应的截短及能量限制。如图2所示。其具体过程即为算法5所述。

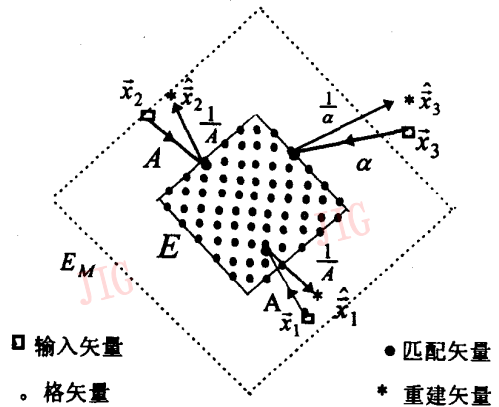


图2

算法5: (1) 根据给定的能量 E 确定格的大小, 如果量化矢量分布是独立同分布(i.i.d.)的高斯分布, 则矢量 $y = (y_1, y_2, \dots, y_n)$ 的能量为 $\sum_{i=1}^n y_i^2$; 如果量化矢量分布为 i.i.d. Laplacian 分布, 则能量为 $\sum_{i=1}^n |y_i|$ 。于是所有满足 $\sum_{i=1}^n |y_i| = E$ 或 $\sum_{i=1}^n y_i^2 = E$ 的格矢量组成码本集;

(2) 设输入矢量的最大能量为 E_{max} , $E_M < E_{max}$; 如果输入矢量 y 的能量小于或等于 E_M , 则令乘因子 $A = E_M/E$, 并根据匹配算法确定其对应的格矢量(如图2中的 x_1 及 x_2); 如果输入矢量 y 的能量输入矢量 y 的能量大于 E_M , 则令乘因子 $\alpha = E_y/E$, 并根据匹配算法确定其对应的格矢量(如图2中的 x_3);

(3) 用算法3确定索引矢量, 并对其进行熵编码^[8];

(4)结束。

需要指出的是,由于小波变换系数服从 i. i. d. Laplacian 分布,格量化区域的形状类似于超锥体^[7],所以矢量能量采用 $\sum_{i=1}^n |y_i| = E$ 的形式。由码矢量确定重建矢量利用算法 4,并且结合图 2 所示方法来实现。

2.3 小波变换系数的量化

图象对象经过 3 次小波变换后,数据结构如图 3 所示。

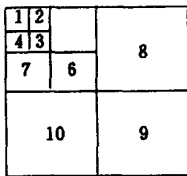


图 3

小波变换后图象的能量大部分集中在 1,对这一部分小波系数可以采用一种连续逼近的标量量化。令 L_{max} 为 1 中系数的最大值,可取 $L_0 < L_{max}$,对 1 中任意的系数 L 可用下式来近似表示^[9]:

$$L = L_0 + s_1 \cdot \frac{L_0}{2} + s_2 \cdot \frac{L_0}{4} + s_3 \cdot \frac{L_0}{8} + s_4 \cdot \frac{L_0}{16} + s_5 \cdot \frac{L_0}{32} + \dots$$

其中 s_i 等于 +1 或 -1,如图 4 所示。

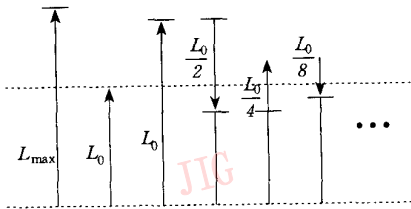


图 4

按图 3 所示的 2 ~ 10 部分,小波变换系数依方向一致的原则组成矢量,进行相应的格矢量量化。

3 实验结果

分别对 $512 \times 512 \times 8$ 的 Woman 图象及 $256 \times 256 \times 8$ 的 Lenna 图象进行 3 次双正交小波变换,通过计算变换图象各个部分的方差,可以看出其大部分能量集中在部分 1。

据此,对 1 的系数进行连续逼近的标量量化。对其余系数,首先设定一适当的门限值,然后将小波变换系数组成矢量,采用上述算法进行格矢量量化编码。图 5(a)表示 Woman 原始图象,图 5(b)为 3 次小波分解的图象,图 5(c)采用标量量化及格矢量量化后的图象,图 5(d)为重建图象。可以看出,重建图象与原始图象相比,在低频部分的失真很小,而某些边缘部分变得比较模糊了。用峰值信噪比描述图象重构质量:

$$PSNR = 10 \lg \frac{(255)^2}{\frac{1}{M \times N} \sum_{i=0}^M \sum_{j=0}^N [g(i,j) - f(i,j)]^2}$$

式中 $g(i,j)$ 表示重构图象, $f(i,j)$ 表示原始图象, $M \times N$ 表示图象大小。Woman 的平均信噪比达 34.6dB,压缩比可达 45.8:1。

表 1 Woman 变换后的方差

位置	方差	位置	方差
1	2767.05	6	2.74
2	53.06	7	5.71
3	5.85	8	8.75
4	8.38	9	1.74
5	23.91	10	2.59

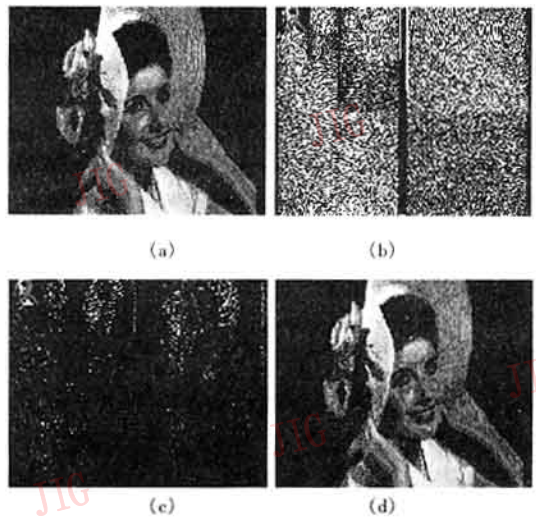


图 5

4 结论

由上述结果可知,格矢量量化是很有效的。利用小波变换系数分布特点与人视觉系统(HVS)之间的可塑性,采用格矢量量化编码以及一种连续逼近

的标量量化等综合编码方法,不仅可保持高质量重构图象所要求的重要信息,大大提高整个量化编码速度及压缩比,而且使这一方法用于多媒体通信,实现基于对象的自适应视频小波编码^[4,5]以及解决视频信号的有效传输问题成为可能。

参考文献

- 1 Linde Y, Buzo A, Gray R M. An algorithm for vector quantizer design, IEEE Trans. Commun., 1980, COM-28(1):84~95.
- 2 Makhoul J, Roucos S, Gish H. Vector quantization in speech coding. Proc. IEEE, 1995, 73:1551~1558.
- 3 Conway H C, Sloane N J. Fast quantizing and decoding algorithms for lattice quantizers and codes. IEEE Trans. Inform. Theory, 1982, IT-28(2):227~232.
- 4 Conway H C, Sloane N J. A fast encoding method for lattice codes and quantizers. IEEE Trans. Inform. Theory, 1983, IT-29(6):820~824.
- 5 Li Weiping, Cao Hugh Q, Li Shipeng, et al. A video Coding Algorithm

Using Vector-Based Techniques. IEEE Trans. CSVT, 1997, 7(1) 146~157.

- 6 Salembier P, Marques T, Pardas M, et al. Segmentation-Based Video Coding System Allowing the Manipulation of Objects. IEEE Trans. CSVT, 1997, 7(1):60~74.
- 7 Barlaud M, Sole P, Gaid T, et al. Pyramidal Lattice Vector Quantization for Multiscale Image Coding. IEEE Trans. Image Processing, 1994, 3(4):367~381.
- 8 Fischer T R. A Pyramid Vector Quantizer. IEEE Trans. Inform. Theory, 1986, IT-32(4):568~583.
- 9 Eduardo A B, da Silva, demetrios G, Sampson et al. A Successive Approximation Vector Quantizer for Wavelet Transform Image Coding. IEEE Trans. Image Processing, 1996, 5(2):299~310.



周利莉 1992年,1995年在解放军信息工程学院信息科学系分获学士、硕士学位,现在该校信息科学系攻读博士学位,主要研究方向是多媒体通信技术与计算机视觉及遥感信息处理。

An Object-Based Lattice Vector Quantization for Multimedia Communications

Zhou Lili, Le Zhongxin

(Information Engineering Institute of the PLA, Zhengzhou 450002)

Abstract Based on the analysis and research of the structure of lattice, this paper in detail discusses some important algorithms about lattice and gives lattice vector quantization algorithms of image objects in wavelet domain. In order to realize efficiently scalar quantization of some wavelet coefficients, this paper presents a successive approximation algorithm applying the distribution of the wavelet coefficients. At the same time it uses lattice vector quantization for some wavelet coefficients. It shows by examples that lattice vector quantization not only accelerates the quantization but also raises the compression efficiency, which satisfies the image compression's needs for multimedia communications.

Keywords Wavelet transform, Root lattice, Vector quantization