

一种视频图像三维子带编码方法*

王大鸣 黄慧群 皇甫堪

(国防科技大学电子技术系 长沙 410073)

摘要 介绍一种视频图像三维子带编码方法,它采用非对称滤波器进行子带划分,对各子带进行量化时,低频子带采用非对称树结构矢量量化方法,高频子带采用几何矢量量化方法。实验证明,由此构成的比特率为112kb/s视频图像编码器获得了良好的重建图像质量。

关键词 子带分解,几何矢量量化,非对称树结构矢量量化

分类号 TN911

An Algorithm of Video Coding Based on a Three-dimensional Spatio-Temporal Subband Decomposition

Wang Daming Huang Huiqun Huangfu Kan

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

Abstract This paper describes an algorithm of video coding based on a three-dimensional spatio-temporal subband decomposition, and results are shown for a 112kb/s coder that is based on an unbalanced tree-structured vector quantizer (UTSVQ) for the lowest frequency band and geometric vector quantizer (GVQ) for the higher frequency bands.

Key words subband decompose, GVQ, UTSVQ

本文介绍一种基于三维时-空子带分解的视频图像编码方法,它在进行子带分解时所采用的非对称滤波器,与自然图像的统计特性相适应,获得了较好的编码质量;它在高频子带编码时采用几何矢量量化方法 (Geometric Vector Quantization 简称GVQ),不需进行码本训练,直接根据高频子带中的边缘特征产生,算法简单、高效;在低频子带编码时采用非对称树结构矢量量化方法 (Unbalanced Tree-structured Vector Quantizer 简称UTSVQ),可根据图像特性最优地分配比特数,从而达到压缩效率与编码质量的统一。实验证明,由该算法构成的112kb/s编码器具有很好的重建图像质量。

1 非对称子带分解滤波器

在进行子带划分时,正交镜像滤波器以其线性相位和抗混叠的优点得到了广泛的应用。但存在环效应问题。考虑到人类视觉系统的心理特性和自然特性的典型特征,可采用这样的非对称滤波器进行图像子带划分:在高频子带用较宽的窗口从而便于提取特性细节,在低频子带用较窄的窗口则有利于低频带能量的保存,这与自然图像中高频信息(即边缘信息)占据极有限的空间,而低频信息占据大片空间相一致,从而可达到很好的视觉效果。表1给出一组非对称滤波器系数^[1]。与QMF相比,该滤波器同样具有线性相位,且运算量降低不少。

我们采用这组滤波器进行如图1所示的子带分解(其中HP、LP表示高低通滤波器,其下标 t, h, v 分别表示进行时间域、空间垂直方向、空间水平方向)。图像分为11个子带,每个子带代表图像的不同特性(如图2所示),其中能量集中在低频块1,高频信息主要集中在2、3,高频块8为时域分解结果,代表运动信息。由于这四个块包含图像的主要信息,故在编码过程中我们只对这四个块进行编码而抛弃其他块。对低频块1我们采用下文介绍的非对称树矢量量化(UTSVQ)的方法进行编码;对于高频块2、3、8采用几何矢量量化(GVQ)的方法,取得了很好的重建效果。

表1 子带分解滤波器系数

Tab. 1 Subband decomposition filter coefficients

12 tap	4 tap
低通滤波器 h1 (n)	高通滤波 h2 (n)
- 9.765625e-03	1.25e-01
2.9296875e-02	- 3.75e-01
3.7109375e-02	3.75e-01
- 1.89453125e-01	- 1.25e-01
- 5.07812500e-02	
6.83593750e-01	
6.83593750e-01	
- 5.07812500e-02	
- 1.89453125e-01	
3.7109375e-02	
2.9296875e-02	
- 9.765625e-03	

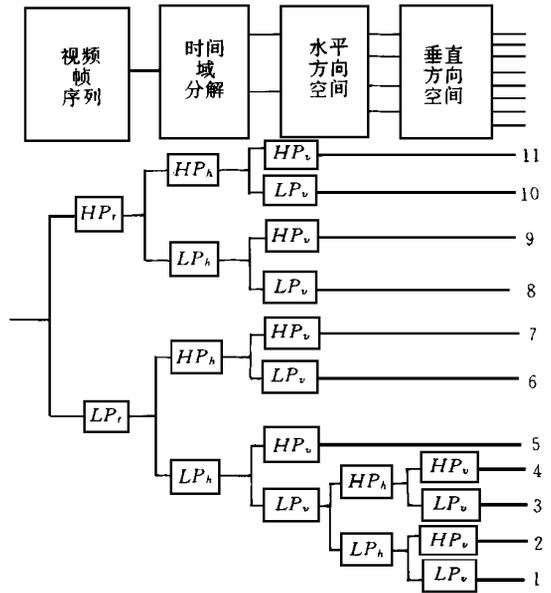


图1 子带分解示意图

Fig. 1 Subband decomposition

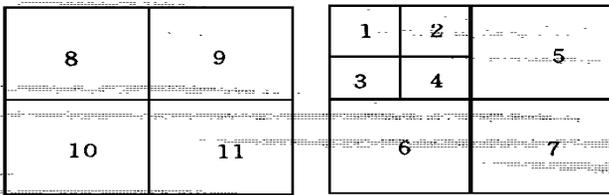


图2 标准图像 Miss America 分带示意图 (子带 1~7)

Fig. 2 Miss America subband decomposition (subband 1~7)

2 非对称树结构矢量量化 (UTSVQ) 在低频子带编码中的应用

本文采用一种非对称树结构矢量量化方法，构造非对称树的方法不同于树型结构 (TSVQ) 按层分裂结点，而是每次只分裂对全局失真影响最大的结点，这样对于图像中占大片区域的变化小的背景区域分配较短的码字、而对于图像中占小片区域的变化大的边缘分配较长的码字，从而达到最优地分配比特数。该方法称为贪婪算法^[2]，描述如下：

首先，定义品质因数 (goodness) 用来表示结点的失真度的降低：

$$\Delta i(s, t) = i(t) - P_l i(t_l) - P_r i(t_r)$$

其中 s 为用来训练码本的序列， $i(t)$ 表示结点 t 的失真度， P_l 表示结点 t 分裂到左孩子结点的概率， P_r 表示结点 t 分裂到右孩子结点的概率，则：

$$P_l = \frac{P(t_l)}{P(t)}, P_r = \frac{P(t_r)}{P(t)}$$

在分裂树的过程中，每次只对具有最大品质因数的结点进行分裂。现给定树 T ，以 \tilde{T} 代表树 T 的叶子 (终端结点)。假设 $t \in \tilde{T}$ ，将 t 分裂成 t_l, t_r 。令 D, R 分别表示对 T 树进行编码中的编码失真和速

率, D , R 分别表示对 T 树编码的失真和速率, 则 $\Delta D = D - D$, $\Delta R = R - R$ 分别表示由于 t 的分裂而导致的编码失真和速率的变化。定义 $\lambda = -\frac{\Delta D}{\Delta R}$ 作为衡量结点分裂意义的测度, 可见 λ 越大说明该次分解越有意义。再定义 $d(j) = E[d_j]$ 表示结点 j 的失真测度, $l(j)$ 表示结点 j 的深度。于是有:

$$D = \sum_{i \in T} P(i) d(i) + P(t) d(t);$$

$$R = \sum_{i \in T} P(i) l(i) + P(t) l(t)$$

$$D = \sum_{i \in T} P(i) d(i) + P(t) d(t) + P(t_r) d(t_r);$$

$$R = \sum_{i \in T} P(i) l(i) + P(t) l(t) + P(t_r) l(t_r)$$

又 $P(t) + P(t_r) = P(t); l(t) = l(t_r) = l(t) + 1$
 得 $\Delta D = P(t) d(t) + P(t_r) d(t_r) - P(t) d(t)$

$$= -P(t) [d(t) - P_{td}(t) - P_{tr}d(t_r)]$$

$$\Delta R = P(t) l(t) + P(t_r) l(t) - P(t) l(t)$$

$$= [P(t) + P(t_r)] (l(t) + 1) - P(t) l(t) = P(t)$$

(如果令 $i(\cdot) = d(\cdot)$)

$$\lambda = -\frac{\Delta D}{\Delta R} = \frac{P(t) [d(t) - P_{td}(t) - P_{tr}d(t_r)]}{P(t)}$$

$$= d(t) - P_{td}(t) - P_{tr}d(t_r) = \Delta i(s, t)$$

在构造非对称树时, 每次只分裂 λ 最大即失真度与速率比值最大的结点。在具体实现的过程 (如图3) 中, 我们选用 3×3 块作为码矢尺寸, 用大量图像训练码本, 构造出上述非对称树, 并对图像进行处理, 压缩比可达到 1.6bpp, 图像主观评价效果良好。

3 几何矢量量化 (GVQ) 在高频子带编码中应用

对于低速率图像编码, 为了获得较好的视觉效果, 应尽可能将比特位分配给能量最集中的低频子带。空间变化大以及运动特征明显的成分位于高频子带, 虽然这些子带能量相对较少, 但包含大量的图像细节。由于这些子带都具有高度的结构化稀疏特性, 可使用几何量化 (GVQ) 来对边缘信息进行提取编码。该方法不需要进行码本训练, 其码字矢量直接根据高频子带中的边缘特征产生。这种矢量量化方法在概念上和视觉模式图像编码类似, 两种方法都利用了眼对边缘强度变化的敏感性。GVQ 根据给定的误差准则对输入数据矢量进行 L 级量化^[3]。本文设计了一种基于最小均方误差准则 (MMSE) 的二级 GVQ 矢量量化方法, 下面首先介绍其码本构造方法。

子带划分后的第二子带包含垂直方向的高频信息和水平方向的低频信息, 因而图像具有大量的横条纹, 而第三子带则具有大量的垂直条纹。根据这些特征, 经大量实验, 我们选取 3×3 块作为码矢尺寸, 构造了如图4所示的9个简单几何图形做为码矢。

再进行二级 GVQ 量化方法, 描述如下:

首先将输入码矢灰度级转化成和码字矢量特征 (即具有黑 (B)、白 (W) 两色) 相匹配的值, 对于第 k 个码字有:

$$L_{1k} = \frac{1}{\# x} B_{iB} x^i, L_{2k} = \frac{1}{\# x} W_{iW} x^i$$

然后计算输入码矢与 k 码字的距离:

$$D_k = \sum_{i \in B} (x_i - L_{1k})^2 + \sum_{i \in W} (x_i - L_{2k})^2 = D_{1k} + D_{2k};$$

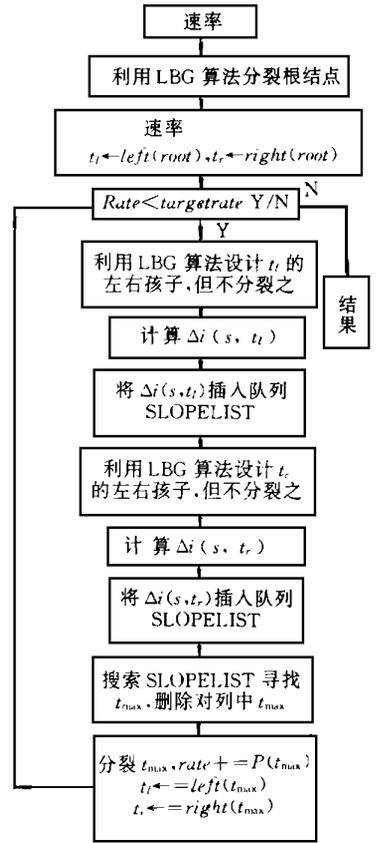


图3 贪婪算法实现框图

Fig. 3 Flow chart of Greedy algorithm

以图5所示的输入矢量为例说明其计算方法：

$$L_1 = \frac{1}{3}\{x_1 + x_5 + x_9\}; L_2 = \frac{1}{6}\{x_2 + x_3 + x_4 + x_6 + x_7 + x_8\}$$

$$D = \sum_{i \in B} (x_i - L_1)^2 + \sum_{i \in W} (x_i - L_2)^2$$

则： $D_{best} = \min\{D_k, k = 1 \dots 9\}$

对一个给定的输入矢量，通过调整每个码矢中两个区的强度信息，获得一个和输入最匹配的码字。假设输入矢量分成不重叠的两块区域 B, W, 表示候选码字矢量 k 的黑白部分，根据 MMSE 计算 L_1 , 和 L_2 , D 。对所有码字重复计算上述过程 $k = 1 \dots M$ (此处= 9)，具有最小距离 D_{best} 的码矢 k_{best} 及相应的强度值 L_{1best} , L_{2best} 为输出信息， k_{best} 可用3至8bits 表达， L_1 , L_2 值可用5bits 量化表示，这种技术类似 “gain-shape” 矢量量化。由于高频带的稀疏性，GVQ 中大量出现 NULL 块，此时， $L_1 = L_2$ 则所有 NULL 块可用统一的一种短码字表达。通过这种方法可获得很高的压缩比，一般可达到 0.01 ~ 0.3bpp。

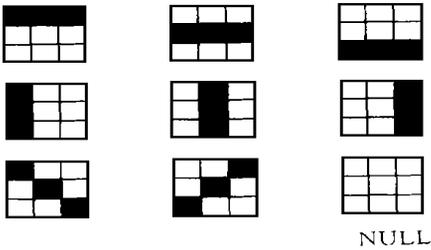


图4 GVQ 码书的9个3×3码矢

Fig. 4 GVQ Codebook of nine 3×3 Code vectors

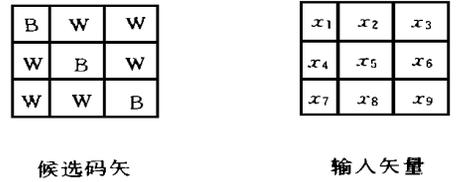


图5 量化方法示例

Fig. 5 Example of two-level GVQ block matching

4 实验结果

这里给出标准视频图像 “Miss Ameirc” 进行三维子带编码的重建图像，由图6可见，重建图像质量良好，它的编码比特分配如表2所示。子带1采用 UTSVQ 方法，压缩比为5 ~ 6倍；子带2、3、8采用 GVQ 方法，子带2、3的压缩比可达32X，子带8压缩比200X，可见GVQ 方法压缩比是很可观的。实验结果表明本文介绍的方法性能优良。

表2 Miss Americ 测试平均比特分配表

Tab. 2 Miss Americ average bit allocation

Subband	bpp	kbps
1	1.6	78
2	0.27	13
3	0.27	13
8	0.0426	8
Total Rate		112kbps

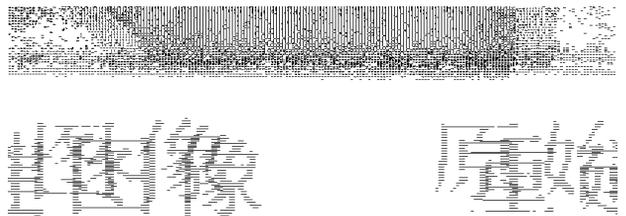


图6 Miss Americ 测试序列

Fig. 6 Miss Americ test sequence

参考文献

- 1 Egger O, Li W. Subband Coding of Images Using Asymmetrical Filter Banks. IEEE Transactions On Image Processing . 1995, 4 (4): 478 ~ 485
- 2 Riskin E, Gray R. A Greedy Tree Growing Algorithm For The Design Of Variable Rate Vector Quantizer . IEEE Trans. Signal Processing, 1991: 2500~ 2507
- 3 Christine Podilchuk I, Nikil. S. Jayant, Narriman Farvardin. Three- Dimensional Subband Coding Of Video. IEEE Transactions On Image Processing. 1995, 4 (2)