

面向边缘的分类 DCT 变换在视频编码中的应用*

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 DCT+运动估计与补偿方法是实用的视频编码方法。目前包括甚低速视频编码标准 H. 263 在内的多种视频标准均采用这种方法。考虑到视觉对于边缘信息的敏感性,我们根据边缘特性对编码块进行分类量化,使得边缘附近的量化噪声具有边缘相似性,从而最优地保留对于视觉敏感的 DCT 系数,在保证编码效率的同时较大地提高了重建图像质量。

关键词 DCT 系数,边缘,量化

分类号 TN941.1

Edge-Oriented Classified DCT for Video Coding

Wang Daming Huang Huiqun Huangfu Kan

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

Abstract The motion compensated discrete cosine transform coding is an efficient video coding technique. In recent years, it has been adopted for many standards of video coding, such as the very low video coding - H. 263. considering the sensitivity of image edge information, we classified the coding blocks based on the features of edge, made the quantization noise near an edge appear edgelike itself, then reserved the sensitive DCT coefficients optimally, and improved the image quality at high coding efficiency.

Key words DCT coefficients, edge, quantization

基于 DCT (Discrete cosine transform) 和运动估计与补偿的视频编码方法以其简单、高效的特点而广泛应用于多种视频标准中。甚低速视频编码标准 H. 263 也采用此方法。但该方法的缺点之一是当压缩比提高时,不加选择地截断了大量 DCT 高频成分,这样就没有根据 DCT 不同系数对于视觉的不同作用最优地分配比特数,从而不能获得更高的编码效率和更好的重建图像质量。

边缘信息是图像信息的重要组成部分,直接影响着图像视觉效果,重建边缘质量的下降将明显地影响整个重建图像质量,实验表明即使对于图像中边缘信息较少的情况,重建边缘的质量也极大地影响了重建图像的视觉效果。因此如何提高重建边缘的质量,即量化 DCT 系数时,如何在有限的比特数下尽可能多地保留边缘信息成为一研究课题。

人类心理学和视觉理论研究表明:人对图像边缘信息的感觉取决于边缘的尖锐性及连续性。为了使重建边缘与实际边缘的特性相一致,在量化编码时边缘的整体性必须保留,在此基础上,根据视觉的屏蔽效应,边缘两边的像素值不需精确重建,即在保留边缘的整体性的前提下,可对边缘两边的灰度进行较粗的量化,从而使得量化噪声具有“边缘相似性”,这种量化噪声只增加或减少边缘两边像素灰度差值,而不破坏边缘的整体性。因而利用人眼的屏蔽效应,允许边缘周围有较大的量化噪声,而将更多的比特数分配给边缘,以达到尽可能最优分配比特数,在保证压缩比的同时,提高重建图像质量。

本文在研究了典型时域边缘对应于 DCT 域上特性的基础上,提出了面向边缘的分类 DCT 系数量化方法,将之应用到 H. 263 的实验结果表明,该方法在保证压缩比的前提下,提高了重建图像的视觉感觉质量。

* 校预研基金资助项目
1999 年 4 月 20 日收稿
第一作者:王大鸣,男,1971 年生,博士生

1 基于 DCT 变换域的边缘分类方法

考虑到人类的视觉特点,为获得更好的重建图像主观感觉质量,有必要在通用的 DCT 编码方法中引入注意机制,以加强对图像中重要信息——边缘的注意和更合理的处理。边缘的特性包括尖锐性、连续性、位置、方向等。下文我们将分析时域边缘对应于 DCT 域上的特征,以获得时域边缘特性与 DCT 域的特征矢量(即 DCT 系数阵)之间的关系,并根据 DCT 变换域的特征矢量能获得边缘的位置和方向的特点,生成能描述典型时域边缘模板特性的 DCT 域边缘模板。这样在 DCT 量化时,根据这些 DCT 域模板,对不同边缘特性的编码块进行分类量化编码,有选择地保留反映边缘信息的高频成分,在有限比特数下有效地保留对视觉敏感的边缘信息,提高重建边缘质量,以获得更好的重建图像质量。下面我们对时域边缘与 DCT 域特征矢量的关系进行分析:

假设图像分为 4×4 的子块,子块里只含有简单边缘信息。

时域模板:

| | | | |
|-------------|-------------|-------------|-------------|
| 40 90 90 90 | 40 40 40 40 | 40 90 90 90 | 90 90 90 90 |
| 40 90 90 90 | 90 90 90 90 | 90 40 90 90 | 90 90 90 90 |
| 40 90 90 90 | 90 90 90 90 | 90 90 40 90 | 90 90 90 90 |
| 40 90 90 90 | 90 90 90 90 | 90 90 90 40 | 90 90 90 90 |

边缘提取后模板:

| | | | |
|---------|---------|---------|---------|
| 1 0 0 0 | 1 1 1 1 | 1 0 0 0 | 0 0 0 0 |
| 1 0 0 0 | 0 0 0 0 | 0 1 0 0 | 0 0 0 0 |
| 1 0 0 0 | 0 0 0 0 | 0 0 1 0 | 0 0 0 0 |
| 1 0 0 0 | 0 0 0 0 | 0 0 0 1 | 0 0 0 0 |

DCT 变换域模板:

| | | | |
|----------------------------|--------------|--------------|-----------|
| 310.0 - 65.3 - 50.0 - 27.1 | 310.0 0 0 0 | 310.0 0 0 0 | 360 0 0 0 |
| 0 0 0 0 | - 65.3 0 0 0 | 0 - 50.0 0 0 | 0 0 0 0 |
| 0 0 0 0 | - 50.0 0 0 0 | 0 0 - 50.0 0 | 0 0 0 0 |
| 0 0 0 0 | - 27.1 0 0 0 | 0 0 0 - 50.0 | 0 0 0 0 |

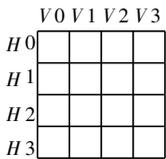


图1 DCT 系数示意

Fig. 1 DCT coefficients

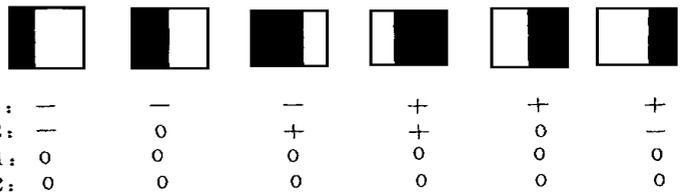


图2 垂直边缘与 V1, V2 符号关系图

Fig. 2 The relation between vertical edge and sign of V1, V2

从图 1, 图 2 得出如下规律:

- (1) 时域中不含边缘反映在 DCT 域上只有低频($V0, H0$)分量不为零。
- (2) 时域中垂直边缘反映在 DCT 域上($V0, H0$)、($V1, H0$)、($V2, H0$)、($V3, H0$) 分量不为零。
- (3) 时域中水平边缘反映在 DCT 域上($V0, H0$)、($V0, H1$)、($V0, H2$)、($V0, V3$) 分量不为零。
- (4) 时域中 45 度边缘反映在 DCT 域上($V0, H0$)、($V1, H1$)、($V2, H2$)、($V3, H3$) 分量不为零。
- (5) 由 $V1, V2$ 的符号变化, 我们不仅可以确定水平边缘的位置, 而且可以确定其极性。

由规律(1) ~ (5) 可得出如下结论: DCT 变换域上 x, y 方向与时域 x, y 方向存在对应关系, 即边缘方向可由 DCT 变换域特征矢量 $V_e = \{Vi: i = 1, 2, 3, 4\}$ 、 $H_e = \{Hi: i = 1, 2, 3, 4\}$ 得出, 此外还可得出边缘的位置和极性。一句话, DCT 变换域特征矢量能反映典型时域边缘特性, 因而可作为面向边缘分类器的特征提取空间, 并获得反映时域边缘特点的 DCT 域边缘模板, 进而在 DCT 域上对边缘进行分类。以上为针对小尺寸块情况的例子, 当块尺寸增大后, 块内边缘情况变得复杂, 该方法能否适用? 接下来我们以 8×8 尺寸块 DCT 系数作分析。

首先采用 sobel 算子对图像 Grandma(176×144) 进行边缘提取并细化。然后对边缘图像进行 DCT 变换, 并将获得的 DCT 系数进行排序, 选取绝对值最大的 16 个置为 1, 其余置为 0, 然后对这些只含 0, 1 的 8×8 块用 LBG 算法进行分类, 选取其中 8 类具有代表性的模板如图 3。

图 3 中 8 个模板的合理性分析如下: 此 8 个模板为能反映时域典型边缘的 DCT 域上的边缘模板。P0 代表不易分类的复杂边缘, P1、P4 代表以水平边缘为主的边缘块, P2、P5 代表以垂直边缘为主的边缘块, P3、P6、P7 代表以斜方向边缘为主的边缘块。由于 8×8 块尺寸较大, 内部边缘情况较为复杂, 因而仅由这 8 个模板不足以反映所有边缘块信息, 但对于一般图像压缩编码而言, 在压缩比较高时, 高频成分损失较多, 并没有根据模板 P0~P7 所反映的不同边缘信息进行分类量化, 以合理分配比特数, 因而选择这 8 个模板可作为压缩比和重建图像质量之间的一种折衷。有关 8 个模板的合理性, 将在下文中实验结果中得到进一步证明。

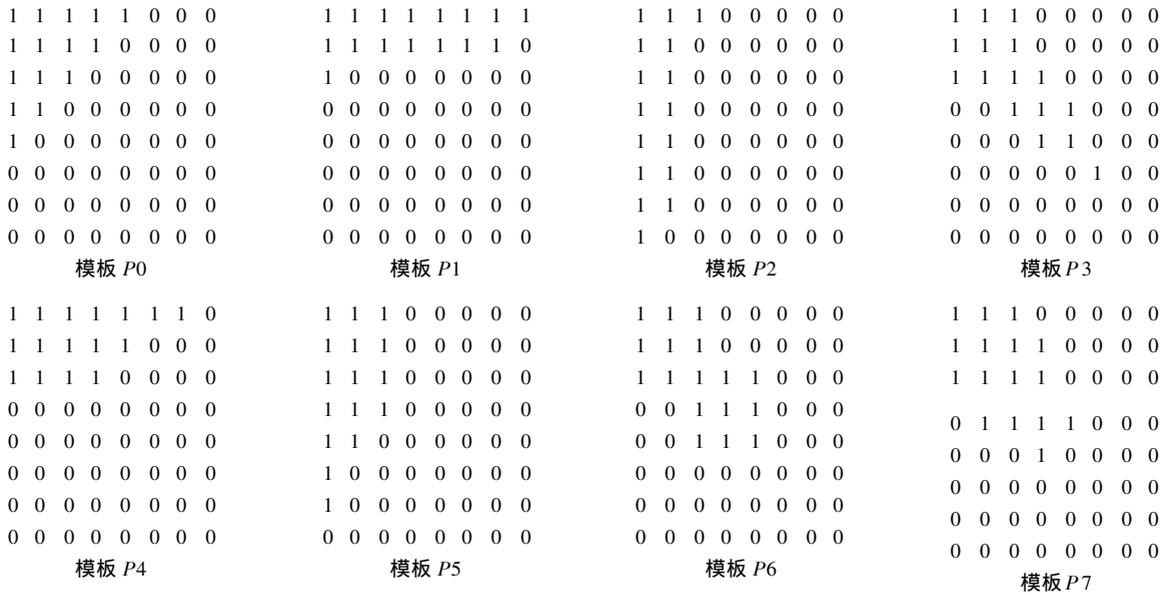


图3 DCT 系数模板

Fig. 3 DCT coefficients templet

2 面向边缘的分类 DCT 量化方法在视频编码中的应用

在视频编码中, 为降低比特率, 一般的方法是减少运动块数目, 或者在对运动预测误差 DCT 变换量化时截断更多的高频成分, 但这两种做法都会使图像质量有所降低。考虑到, 图像运动预测误差由缓变的背景和快变化的轮廓区组成, 轮廓区为图像预测误差块的最重要信息, 因而轮廓区获得好的量化编码及重建是重建视频图像质量好坏的关键。轮廓区由边缘组成, 因而有效地保存边缘的特性成为视频编码的关键。在此我们采用面向边缘的分类 DCT 变换编码的方法, 并将此方法应用于 H. 263 视频编码中。

在 H. 263 的具体实现中, 通过运动估计找出最佳运动矢量后, 不对非运动块编码, 而对运动块的预测误差进行统一 DCT 变换编码, 但不根据块内边缘特性进行分类编码。而本文提出的方法则对编码块进行分类: 将待编码块分为非运动块、含边缘的运动块、不含边缘的运动块。为了尽可能保留边缘块, 我们采用两级分类的方法, 首先将 H. 263 中区分运动块和非运动块的门限降低(这一门限在 H. 263 中选为 60, 它指运动预测误差绝对值之和与零运动矢量情况下误差绝对值之和的差), 以获得相对较多的运动块。然后对运动块进行第二级分类, 进一步分为含边缘的运动块、不含边缘的运动块和非运动块。对预测误差块的边缘提取方法采用 Sobel 算子求出梯度 $G(x, y) = G_x(x, y) + G_y(x, y)$, 及预测误差

块梯度和 $G = \sum_{x=0}^7 \sum_{y=0}^7 G(x, y)$, 并根据门限 THE 区分边缘块, 最后提高区分运动块和非运动块的门限,

再一次划分出非运动块。算法流程如图4所示。

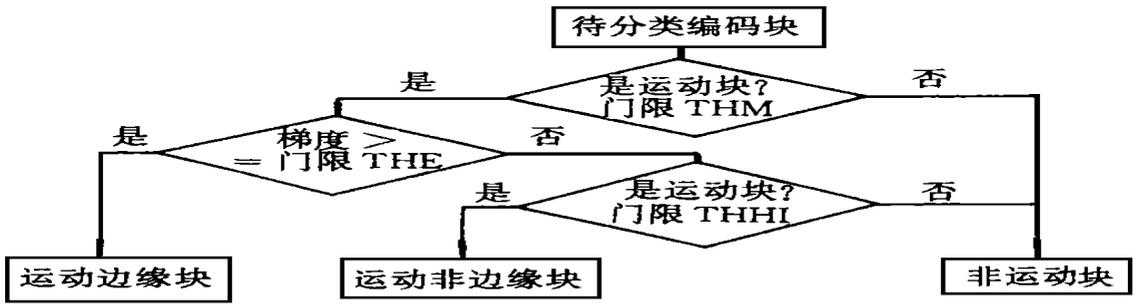


图4 两级编码块分类法

Fig. 4 The two level classification of encode block

完成编码块分类后, 不对非运动块编码, 重建图像时只需复制前一帧对应块即可, 对非边缘运动块只需保留 DCT 系数(0, 0)、(0, 1)、(1, 0)三个系数, 对于含边缘的运动块则需进一步根据上文中我们在 DCT 变换域提取 8 个边缘模板进行边缘分类, 并在此基础上采用不同的量化方法。具体方法如下:

编码器根据所选择的模板 $P_i(u, v)$ 来选择 DCT 系数 $C(u, v)$, 即 $C(u, v) = C(u, v) * P_i(u, v)$; $B_i(x, y)$ 为 $C(u, v)$ 逆 DCT 变换结果, 编码器选择模板 P_i 的标准是获得最佳的 $B_i(x, y)$, 即选择能获得最佳重建块质量的模板 i , 这样一来需作多次 IDCT, 运算量太大。考虑到 DCT 变换具有一致性, 即空间域和变换域具有相同均方误差, 又利用不同模板根据边缘特性的不同有效值分布在不同区域上这一特点, 可简化模板选取准则, 将 $C(u, v)$ 所有系数求平方和, 具有最大值的模板即为 DCT 域均方误差最小的模板, 则相应的 P_i 为所求模板。

通过上述方法对边缘块分类后, 选择相应的模板及 ZigZag 扫描方法进行 DCT 量化及变长编码。这样可根据边缘特点, 有选择地保留反映边缘的高频成分, 使得重建边缘的完整性得以保留, 提高重建边缘质量, 进而提高重建图像质量。其中 ZigZag 扫描方法如图 5 所示。

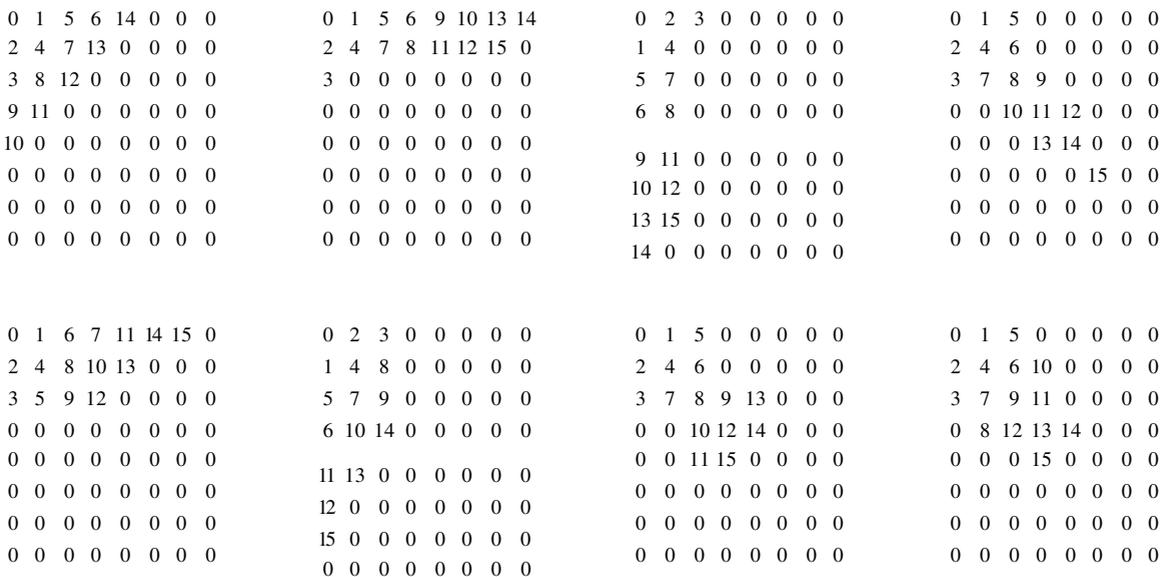


图5 不同模板对应的 ZigZag 扫描图

Fig. 5 Zigzag scan path of different templet

3 实验结果及结论

我们采用面向边缘的分类 DCT 变换方法对 H. 263 算法进行改进, 并以标准图像 Grandma(176 × 144, 30 帧/s), 进行测试, 门限 THM= 60, THHI= 80, THE= 3600。判别准则为

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \text{ dB}; MSE = \left(\frac{1}{n \cdot m} \right) \sum_{i=1}^n \sum_{j=1}^m (x_{ij} - \hat{x}_{ij})^2$$

其中 $n = 176$, $m = 144$, x_{ij}, \hat{x}_{ij} 分别表示原始和重建图像对应像素灰度值。

引入面向边缘的分类 DCT 变换方法前后重建图像如图 6 所示, 信噪比曲线如图 7 所示, 信噪比平均提高 0.3dB, 编码速率由原来 14.62kbts/s 降低为 13.83kbts/s。可见在使得编码比特率有所降低的同时, 重建图像质量得到提高。

综上: 面向边缘的分类 DCT 变换方法根据边缘块特性, 在 DCT 系数量化时有选择地保留高频成分, 提高了重建边缘质量, 进而提高了重建图像质量, 实验证明该方法行之有效。

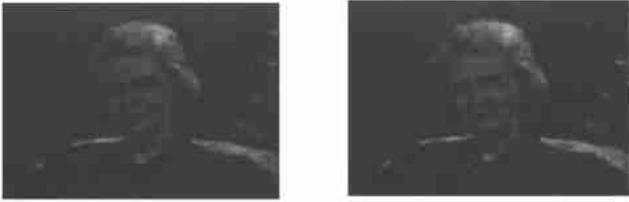


图 6 采用分类 DCT 量化方法前(左)后(右)重建图像比较

Fig. 6 Comparison of reconstruct image

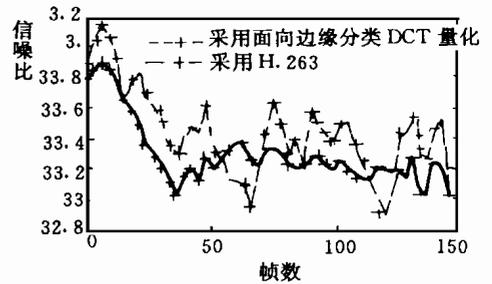


图 7 信噪比较曲线

Fig. 7 Comparison of SNR

参考文献

- 1 Chang J Y, Chang R F, Kuo W J. Edge-based motion compensated classified DCT with quadtree for image sequence coding. Signal Processing: Image Communication 11, 1998: 187 ~ 197
- 2 Bhaskar Ramamurthi, Allen Gersho. Classified Vector Quantization Of Images. IEEE Transactions On Communications, 1986, 34 (11)
- 3 Dong Sik Kim, Sang U K Lee. Image Vector Quantizer Based On a Classification in the DCT Domain. IEEE Transactions On Communications, 1991, 39(4)