

基于 DCT 的帧内视频编码器的缓存控制技术*

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 提出一种基于 DCT 的帧内视频编码器的缓存控制技术。它根据视频图像各种子块的比例及其相应的 $R-Q$ 曲线确定量化因子, 通过再编码, 调整子块 $R-Q$ 曲线的拟合参数使其自适应视频图像的场景变化。同时本文提出了一个简单有效的图像子块分类方法, 分类结果基本上与根据编码子块所需比特数分类子块的结果相一致。此法用于缓存控制的额外计算量小, 并能扩展应用于 H. 261 M PEG 等视频编码标准中。模拟结果表明, 此法在视频序列的场景发生变化时, 再编码次数少, 输出码率恒定。

关键词 缓存控制, 视频编码

分类号 TN 941. 1

A Buffer Control of DCT-Based Intrafield Video Coding

Lou Shengqiang Huang Pukan Zhou Liangzhu Wan Jianwei

(Department of Electronic Technology, NUDT, Changsha 410073)

Abstract A new buffer-control policy for DCT-based intrafield video coder is presented. It selects quantization parameter based on the number of three kinds of blocks in a picture and the rate-quantization models ($R-Q$ model) for three kinds of blocks. It adapted to the scene changes through re-encoding the picture and adjusting the $R-Q$ model. A simple and effective criterion is presented to classify the blocks, the classifying result based on it is close to the result gotten through quantizing the bits needed to code a block into three levels. This method can be applied to some other coding techniques such as H. 261 M PEG, extra computations are modest. Simulation results show that this buffer control policy can control the output bit rate steadily even at scene changes.

Key words buffer control, video coding

视频编码器的一项关键技术是缓存控制策略, 它直接影响重构视频图像质量。应用于有恒定传输码率环境中的视频编码器, 例如: 基于 PSTN 的可视电话, 其缓存策略需依据视频图像特性, 控制编码器的编码参量 (例: M PEG、JPEG 等标准中 DCT 系数的量化因子), 使其输出码率恒定。

传统的缓存控制策略利用缓存容量反馈控制编码器的系统参量, 这种方法存在如下问题: (1) 反馈具有一定的滞后性; (2) 反馈深度的选取: 若控制过强, 输出码率较低, 重构图像质量较差, 白白浪费信道容量。反之, 则易使缓存溢出, 影响视频的实时传输; (3) 控制的灵敏度: 灵敏度过低, 反馈滞后, 反之, 量化值变化过于频繁, 降低编码效率。

现已有多种缓存控制策略, 不外乎是前向控制、反向控制或两者的结合。前向控制技术, 即根据图像特性选择不同的编码器输出码率特性, 实施缓存控制, 虽提高了缓存控制效率, 但通常要增加大量计算量, 同时需一定的编码时延, 用于分析图像特性, 不利于视频编码的实时实现。现在讨论较多的是反向控制技术, 一般它们事先分析视频编码器特性, 即量化因子、输出码率、编码图像质量、图像复杂度之间的关系, 为它们建立模型, 根据模型实施缓存控制, 同时, 编码结果用于修正模型参量, 使

* 1998年7月11日收稿

第一作者: 楼生强, 男, 1969年生, 博士

模型自适应视频图像的变化。

本文提出一种新的基于 DCT 的帧内视频编码器的缓存控制策略 (编码方法采用 JPEG 标准的基本模式), 以三种子块的 $R-Q$ 曲线 (R 是编码子块所需的比特数, Q 是 JPEG 标准中 DCT 系数的量化参量) 为基础, 结合图像再编码过程, 实现编码器的缓存控制。首先, 图像子块根据一个简单有效的子块分类方法分为三类: 边缘块或强纹理块、一般纹理块、细纹理块或平滑块, 为了方便起见, 分别称为: 边界块、纹理块、平滑块。根据三种子块的 $R-Q$ 曲线及图像三种子块的数目选取量化因子, 如果场景发生变化, 根据编码结果调整 $R-Q$ 曲线, 再编码视频图像, 使其自适应图像的场景变化。

1 子块分类方法

本文将图像划分为不重叠的子块集, 尺寸为 16×16 图像子块集被分为三类: 边界块、纹理块、平滑块。对于每一子块, 首先使用 Haar 子波作三层子波分解^[1], 对于一维信号 $x(n)$, 其 Haar 子波变换为:

$$x_h(n) = (x(2n) + x(2n + 1)) / 2$$

$$x_g(n) = (x(2n) - x(2n + 1)) / 2$$

其中: $x_h(n)$, $x_g(n)$ 分别为低频子带、高频子带。子块变换结果图所示

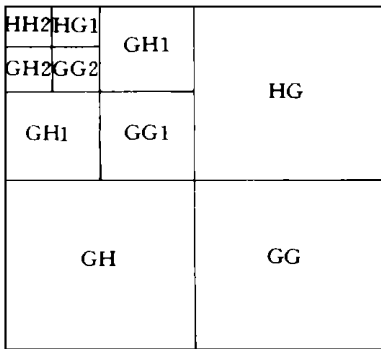


图 1 16×16 子块变换结果

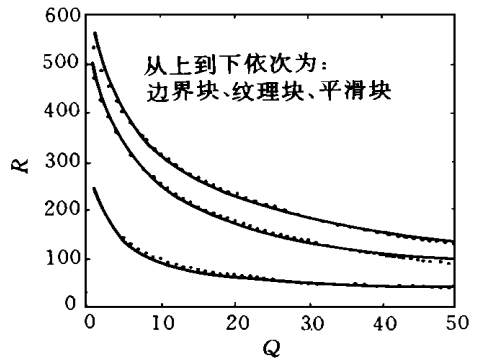


图 2 “Lenna” 图像三种子块的 $R-Q$ 曲线及拟合结果 (虚线)

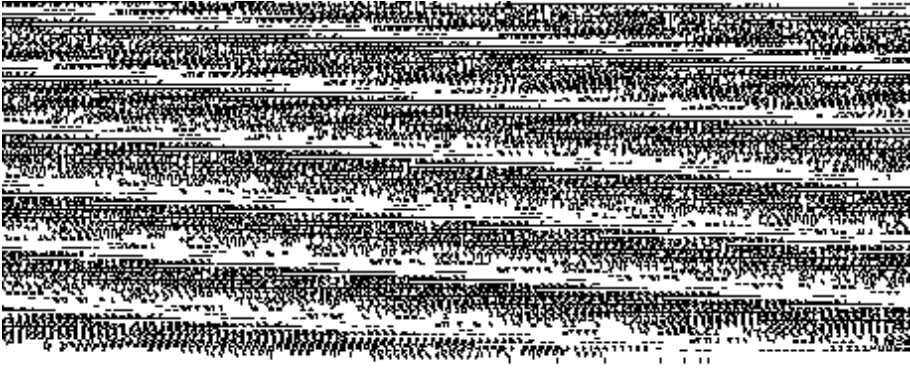
根据如下准则判别子块的类型:

- 如果 $| \max(hh2) - \min(hh2) | \ln \min(hh2) | > 0.5$ 并且 $| \min(hh2) | > 90$ 则子块为边界块。
- 如果 $| \max(hg2) - \min(hg2) | > 35$ 或者 $| \max(gh2) - \min(gh2) | > 35$ 或者 $| \max(gg2) - \min(gg2) | > 35$ 则子块为边界块
- 如果 $\text{Num}(|hg1| > 10) > 6$ 或者 $\text{Num}(|gh1| > 10) > 6$ 或者 $\text{Num}(|gg1| > 10) > 6$ 则子块为纹理块。
- 如果 $\text{Num}(|hg1| > 4) > 1$ 或者 $\text{Num}(|gh1| > 4) > 1$ 或者 $\text{Num}(|gg1| > 4) > 10$ 则子块为纹理块
- 子块为平滑块

以上 5 个判别过程的优先级逐级下降, 即: 如果某一条件成立, 则随后条件不作判别。其中 $\max(x)$, $\min(x)$ 分别表示子带 x 中系数的最大、最小值, $\text{Num}(|x| > T)$ 表示子带 x 中系数值的绝对值大于 T 的个数。

上述判别过程中使用的各门限值根据经验选取, 此准则的判别结果不一定反应子块的实际特性, 例如: 纹理块包含了渐变块, 但各类子块的 $R-Q$ 特性往往区别较大。图 3 (a) 是上述分类方法的子块分类结果, 图 3 (b) 是将编码子块所需的比特数量化为三段, 每段表示一类子块的分类结果, 图像使用 JPEG 的基本模式进行编码, $Q=35$ 。从图中可看到, 大部分子块的判别结果相同, 图 4 是另一组子块分

类结果，了块表示方法与图 3 相同，从图中可知，上述判别过程对 8 比特灰度图像而言，有较好的分类结果



(a) 根据准则分类结果

(b) 根据编码所需比特数分类结果

图 3 子块分类结果，图中‘口’表示边界块，‘△’表示纹理块，其它为平滑块，图 (b) 使用 JPEG 编码，量化因子： $Q=35$

2 $R-Q$ 模型^[3]

为了实现缓存控制，我们首先从样本图像得到三种子块的 $R-Q$ 曲线，其中 R 是编码某类子块所需的比特数， Q 是 JPEG 标准中 DCT 系数的量化因子。本文使用如下式子：

$$R = T + \frac{U}{Q^V}$$

来拟合子块的 $R-Q$ 曲线^[2]，图 4 是“Lenna”图像三种子块的 $R-Q$ 曲线及拟合结果。模拟结果表明，此式对子块 $R-Q$ 曲线有很好的拟合结果，一般地，三种曲线的位置按边界块、纹理块、平滑块顺序从上到下排列，对于曲线参数 V ，通常边界块对应的参数值较小（0.1~0.3 之间），一般纹理块此之（0.1~0.4），平滑块最大（0.2~0.9）。

3 控制策略

按如下算法选择 DCT 系数的量化参数 Q ：

- (1) 得到三种类别子块的数量；
- (2) 根据 $R-Q$ 曲线及其三种子块比例得到量化参数值；
- (3) 编码图像，如果比特率满足要求，则到第一步，编码下一帧图像
- (4) 根据编码结果，调整 $R-Q$ 曲线的拟合参数，到第 2 步。

4 结果及结论

使用上述方法，对“Miss America”序列的亮度部分进行编码（使用 JPEG 的基本模式），模拟结果表所示，同时使用上述方法对由十帧不相关图像（图 4 (a) - (j)）构成的序列进行编码，结果表所示

表 1

序列号	1	2	3	4	5	6	7	8	9	10
再编码次数	2	0	0	0	0	0	0	0	0	0
量化参数	28	28	27	27	27	27	27	27	27	28
编码比特	0.49	0.49	0.50	0.50	0.49	0.50	0.50	0.50	0.50	0.49
PSNR	40.9	41.1	41.1	41.0	40.9	41.7	41.1	41.0	41.0	41.0

表 2

序列号	1	2	3	4	5	6	7	8	9	10
再编码次数	0	1	2	0	1	0	1	0	0	1
量化参数	69	103	52	54	132	90	53	54	52	44
编码比特	0.85	0.76	0.74	0.87	0.75	0.74	0.77	0.79	0.72	0.74
PSNR	30.4	30.0	35.4	32.5	26.5	28.5	33.9	33.9	35.4	36.5



图 4 (a) 496×512

图 4 (b) 496×512



图 4 (c) 496×512

图 4 (d) 496×512



图 4 (e) 352× 512

图 4 (f) 736× 380



图 4 (g) 496× 512

图 4 (h) 736× 480

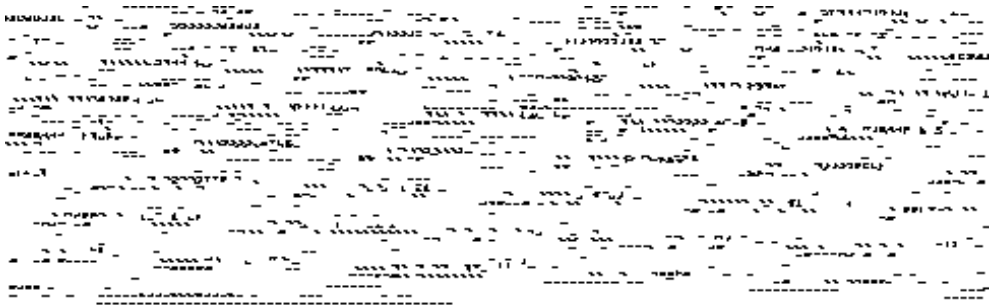


图 4 (i) 496× 512

图 4 (j) 736× 480

从表可知, 如果视频序列的场景变化不大, 一般无需再编码过程, 量化参量 Q 的变化幅度很小, 重构图像的 PSNR 也比较均匀。

从表可知, 如果视频序列场景发生变化, 表中的 Q 值和 PSNR 在一定程度上说明了图像场景的变化程度, 本方法通过再编码过程, 及时调整模型参数, 使其自适应序列的场景变化。从表中可知, 虽然 Q 值的调整范围很大, 但再编码的次数并不多。

本文提出一种基于图像子块 $R-Q$ 曲线的缓存控制策略, 这方法的优点是: (1) 以分类子块为基础的 $R-Q$ 曲线与以整副图像为基础的 $R-Q$ 曲线相比, 更加稳健。场景发生变化时, 模型曲线调整量小。(2) 用于分类子块的计算量有限, 因为判别准则以数值比较为主, 同时子块的 Haar 变换结果可直接用于子块的 DCT 计算^[5], 也就是所谓的子带 DCT。(3) 能扩展应用于 H. 261 MPEG 等视频编码标准中。

参考文献

- 1 A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. IEEE TPAMI, 1989
- 2 $\times 64\text{kb/s}$ 会议电视系统探讨. 通信学报, 1993
- 3 Rate Control of MPEG Video Coding and Recording by Rate-Quantization Modeling. IEEE T-CSVT, 1996
- 4 A Statistical Feedforward/Feedback Buffer Control for the Transmission of Digital Video Signals Compressed by DCT-Based Intrafield Coding. IEEE T-CSVT, 1996
- 5 Subband DCT: Definition, Analysis, and Application. IEEE T-CSVT, 1996