

ATM 网中可变比特率小波视频编码和码率控制¹

林 刚 刘泽民

(北京邮电大学电信工程学院 北京 100876)

摘 要 本文提出了小波视频编码系统及其码率控制的新方法。对原始图像序列进行时间、水平和垂直三维小波分解, 并采用改进的等级树集合分区的算法对小波系数进行量化。在图像组级上对该编码器进行码率控制, 使其输出码流遵循漏桶控制器的参数, 并对不同大小的漏桶缓冲器进行了讨论。

仿真实验证明了该编码系统及其码率控制方法的有效性。

关键词 ATM 网络, 视频编码, 小波变换, 码率控制

中图分类号 TN913.24

1 引 言

小波变换 (WT) 能够对信号进行多分辨率分析和时频局部化, 基于 WT 的图像压缩克服了块效应, 恢复图像符合人眼的主观特性, 在性能上优于传统的 DCT 图像压缩系统。因此, 小波编码是很有前景的压缩技术。本文提出了一种新的小波视频压缩方法, 即首先对原始图像序列进行三维小波分解, 然后采用改进的等级树集合分区的算法对小波系数进行量化。

在研究利用 ATM 网传输视频时, 大多是基于可变比特率的, 它能更有效地利用网络资源。然而, 由于交换缓冲器的容量是有限的, 视频码流的突发性可能导致拥塞的发生, 从而引起缓冲器溢出和信元丢失, 影响传输的服务质量。因此, ATM 网络需要进行流量控制。事实表明, 为可变比特率连接描述恰当的流量参数是很困难的, 因为它既要利于网络资源的分配, 又要利于码率控制。现有的网络标准更侧重于后者, 即通过一种规则 (称为漏桶控制器) 来定义流量参数。在使用该参数对视频连接进行描述时, 有两种主要途径: (1) 寻求一个最精确表征编码器特性的参数; (2) 强制编码器的输出码率符合码率控制参数。对于途径 (1), 已有一些研究工作, 如文献 [1,2], 它们都为编码器的输出码率建立各种模型。而对于途径 (2), 研究工作还不多。Reibman 和 Haskell^[3] 提出了一种码率的约束条件, 以防止漏桶缓冲器溢出。Heeke^[2] 和 Coelho^[4] 试图使编码器的输出符合某种马尔可夫链。这些码率控制都是基于 MPEG^[5] 压缩标准的。针对本文的小波视频压缩算法, 我们采用漏桶规则对编码器的输出码率进行控制, 使其符合该网络连接的流量特性参数, 如信道带宽、突发性等。

2 三维小波视频编码

2.1 时间、空间三维小波变换

视频信号中的每帧图像的相邻像素之间都存在着相关性, 低频部分能量较高。经小波变换后, 低频部分系数的幅值较大, 而高频系数幅值较小, 实现了能量集中。针对不同频带系数的不同特性进行量化, 实现压缩。同样地, 视频各帧之间也存在着相关性, 沿时间轴方向的小波变换也能消除冗余, 提高压缩效率。

将原始图像序列看作三维信号, 以 F 帧为一组进行划分, $F = 2^j, j = 1, 2, \dots$ 。对每个图像组 (GOP) 沿时间方向进行 j 级小波分解, 得到 H, LH, LLH, $\dots$ 等 $j+1$ 个时间子带。

¹ 1998-05-05 收到, 1999-01-10 定稿
邮电部专项基金资助课题

然后进行水平、垂直方向的 R_s 级二维小波变换, 共得到 $3 \times R_s + 1$ 个空间子带。图 1 是 $F = 8, R_s = 3$ 时的子带图。

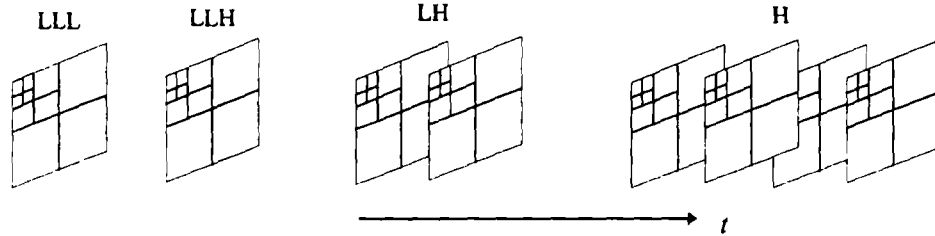


图 1 时间、空间三维小波变换子带

2.2 小波系数的量化

Shapiro 提出了嵌入式零树小波 (EZW) 算法^[6]对图像的小波变换系数进行压缩, 此方法压缩算法简单, 且所产生的码流是嵌入的。A. Said 和 W. Pearlman 又在零树编码基础上, 提出了一种等级树集合分区算法 (SPIHT)^[7], 提高了编码效率。

针对视频压缩, 本文将 SPIHT 算法从二维扩展到三维, 称为“三维等级树集合分区 (3DSPIHT)”编码。并对 SPIHT 算法作了如下的改进: 图像的小波变换系数可分为水平、垂直和对角三个空间方向, 如图 2 所示。一般地, 此三个方向的小波系数的特性是互不相同的, 各方向系数的最大值也不同, 分别设为 V_h, V_v 和 V_d , 相应的最大位次 (即表征该值所需的二进制位数) 分别为: $n_h = \lceil \log_2 v_h \rceil$, $n_v = \lceil \log_2 V_v \rceil$, $n_d = \lceil \log_2 V_d \rceil$, 则此三方向上的所有小波系数的最大位次为 $n_{\max} = \max\{n_h, n_v, n_d\}$ 。

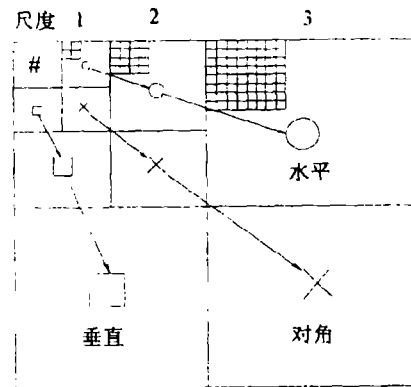


图 2 图像帧的小波变换系数的等级树结构

在进行量化时, 我们克服文献 [7] 用 $n_{\max}$ 对所有系数进行编码的缺点, 而是以初始值 n_h, n_v 和 n_d 分别对相应三个空间方向上小波系数进行编码, 这样就可减少不必要的码流。为了说明本文改进算法与 SPIHT 的区别, 将它们应用到二维静止图像编码。对 360×288 , 8 比特 / 像素的标准 CIF (Common Intermediate Format) 格式的 Claire 序列的第 1 帧进行实验, 得到 $n_h = 7, n_v = 9, n_d = 7$ 。图 3 是两种方法信噪比-压缩比曲线的比较。可见, 此改进方法可大大提高编码效率。

而且, 本压缩算法得到的码流是嵌入式的。可在任意时刻中止码流, 而不会影响图像的恢复。此特性有利于对输出码率进行精确控制, 是其它量化方法所不具有的。例如应用较为广泛的矢量量化方法, 其码率的控制通过调整量化台阶来实现的。量化台阶与码流大小之间的关系不够明确, 只能通过统计的方法来获取。而且, 不同的信号也具有不同的关系。因此, 欲对它进行精确的码率控制, 是非常困难的。

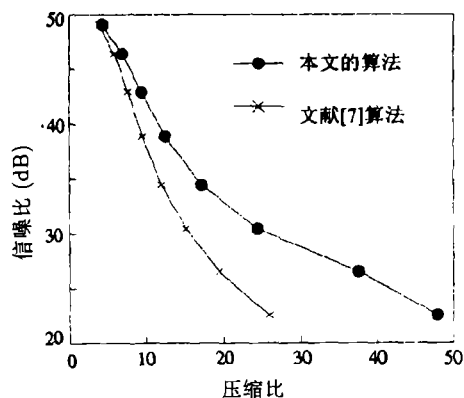


图 3 改进算法与 SPIHT 算法
信噪比-压缩比曲线比较

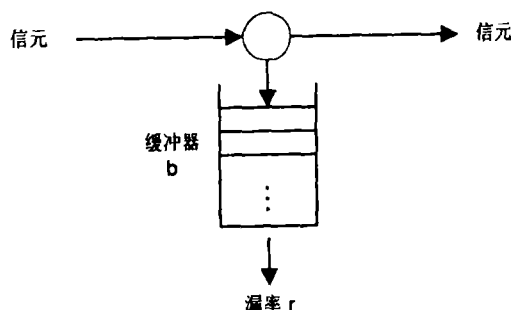


图 4 一个基本漏桶模型的示意图

3 码率控制方法

在 ATM 网中, 用户参数控制 (UPC) 的功能是根据服务质量的约定, 实现对输入流量的控制。该机制的本质就是将符合网络规则的信元与不符合网络规则的信元分开。我们使用漏桶控制方法来实现这一功能。研究表明, 它优于其它的监管机制, 如移动窗、跳动窗等。图 4 是一个基本漏桶 $LB(r, b)$ 的模型。它由一个大小为 b 的缓冲器组成, 此缓冲器中的标志 (flag) 以常数漏率 r 减少。每到来 1 个信元时, 此缓冲器就相应增加 1 个标志。如果缓冲器尚未填满, 则信元允许通过。反之, 当缓冲器充满时, 到来的信元就被丢弃。可见, 流量通过漏桶 $LB(r_i, b_i)$ 时必须满足突发性约束条件, 其中 r_i 、 b_i 为第 i 个信源所对应的漏桶的漏率及缓冲器容量。设 $N_i(s, t)$ 表示在时间段 (s, t) 之间由信源 i 输入到网络的比特数, 则漏桶必须满足以下不等式:

$$N_i(s, t) \leq r_i(t - s) + b_i \quad (1)$$

输入码率的均值限制为 r_i , 峰值速率 p_i 时的最大突发性为 $b_i p_i / (p_i - r_i)$ 。本文仅以漏桶所定义的描述参数来进行流量控制。对于许多信源来说, 该参数是否合适还是有争议的。特别地, 对于视频连接, 其有效性还依赖于视频编码的算法。

如第 2 节所述, 本文的 3DSPIHT 算法是基于图像组的。在每个图像组中, 码率分配是优化的。因此, 在进行码率控制时, 只需在图像组级上进行。其原则是: 在码流满足漏桶约束条件 ((1) 式) 的情况下, 给那些活动性强的图像组分配较多的比特数, 给那些活动性弱的图像组分配较少的比特数, 以尽可能地使恢复图像的质量平稳。实现过程如下:

(1) 测定各图像组的活动性程度。在不进行任何约束 (开环) 的情况下, 让各个图像组的量化精确到某个位次 n , 得到码率 $R_{\text{open}}(k)$, $k = 1, 2, \dots, N$ 。显然, 码率高的图像组的活动性强。

(2) 计算在该精确位次 n 下, 各图像组的平均码率 $r_{\text{ave}} = \left(\sum_{k=1}^N R_{\text{open}}(k) \right) / N$ 。

(3) 根据网络分配给该连接的码率 r , 即漏桶的漏率, 重新调整各图像组的码率。计算公式为: $R'_{\text{open}}(k) = R_{\text{open}}(k) \times r / r_{\text{ave}}$ 。这样 $R'_{\text{open}}(k)$ 既反映了各图像组的活动性强弱, 其均值又等于网络所分配的码率。

(4) 在漏桶的约束条件下, 求得各图像组的控制码率 $R(k)$: 对于活动性大的图像组 ($R'_{\text{open}}(k) \geq r$), 当漏桶缓冲器中的码流 $X(k) \approx 0$ 时, 取 $R(k) = R'_{\text{open}}(k)$, 类似于开环情形, 而当 $X(k) \approx b$ 时, 取 $R(k) = r$, 类似于不变比特率情形; 对于活动性弱的图像组 ($R'_{\text{open}}(k) < r$), 当 $X(k) \approx 0$ 时, 取 $R(k) = r$, 类似于不变比特率情形, 而当 $X(k) \approx b$ 时, 取 $R(k) = R'_{\text{open}}(k)$, 类似于开环情形。以上考虑的是漏桶缓冲器取两种极限填充情形下的码率分配。当缓冲器部分填充时, 取上述两种极限情况的线性组合。公式如下:

$$\begin{cases} R(k) = (1 - f_1(x)) \times R'_{\text{open}}(k) + f_1(x) \times r, & R'_{\text{open}}(K) \geq r; \\ R(k) = f_2(x) \times R'_{\text{open}}(k) + (1 - f_2(x)) \times r, & R'_{\text{open}}(K) < r, \end{cases} \quad (2)$$

其中 $x = X(k)/b$ 表示漏桶填充的百分比。函数 $f_1(x)$ 、 $f_2(x)$ 的取值范围为 $[0, 1]$, 且它们都是随填充程度 x 而单调递增的。可见, 控制后的码率 $R(k)$ 一定介于开环码率 $R'_{\text{open}}(K)$ 与不变比特率 r 之间。

本码率控制方法实现简单, 只需考虑漏桶中的两个参数 r 、 b 。其中 r 的大小是受网络连接带宽的限制, 由网络分配给该连接的码率而决定的。而参数 b 的选取会影响到视频质量和网络性能。 b 值越大, 各图像组的码率的变化幅度就越大, 但重建图像的质量就越平稳。由于漏桶缓冲器只是一个虚构的缓冲器, 因此 b 的大小对编码时延没有影响。

4 仿真实验和结论

本文所用的原始图像序列是 360×288 , 25 帧/s, 8bit/pel 的标准 CIF 格式的 “Miss America” 序列。选取函数 $f_1(x) = x$ 、 $f_2(x) = x$ 。图 5(a) 是在网络连接带宽 $r = 180\text{K b/s}$ 时, 该序列前 17 个图像组 (即前 136 帧) 分别在 (1) 开环; (2) 不变比特率; (3) 漏桶控制下的信噪比曲线的比较, 其中漏桶缓冲器 b 的大小选取为 58kb , 相当于 8 帧图像码流的空间。图 5(b) 是三者所对应的码率曲线。从图中看出, 开环情形时, 图像信噪比最平稳, 但

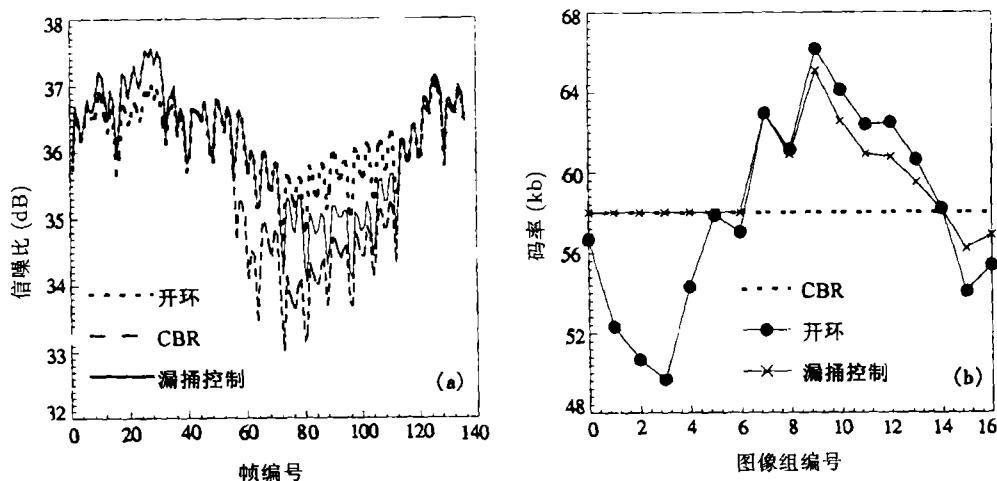


图 5 “Miss America” 序列分别在开环、不变比特率和漏桶控制下的
(a) 恢复图像的信噪比较; (b) 码率比较

码率的变化幅度最大; 在不变比特率下, 图像信噪比变化大, 恢复图像的质量最差。而本文的漏桶控制介于二者之间。

由于漏桶缓冲器参数 b 的选取非常重要, 它直接影响着网络流量的突发性大小, 因此, 我们将 b 取几种不同的值, 做了实验。图 6(a), 6(b) 中的三条曲线分别对应 $b = 29\text{kb}$ 、 $b = 87\text{kb}$ 、 $b = 145\text{kb}$ 时的信噪比和比特率曲线。可见, 随着 b 的增加, 曲线逐渐从 CBR 移向开环情形, 即图像质量逐渐趋于平稳, 但码率变化也趋于剧烈。

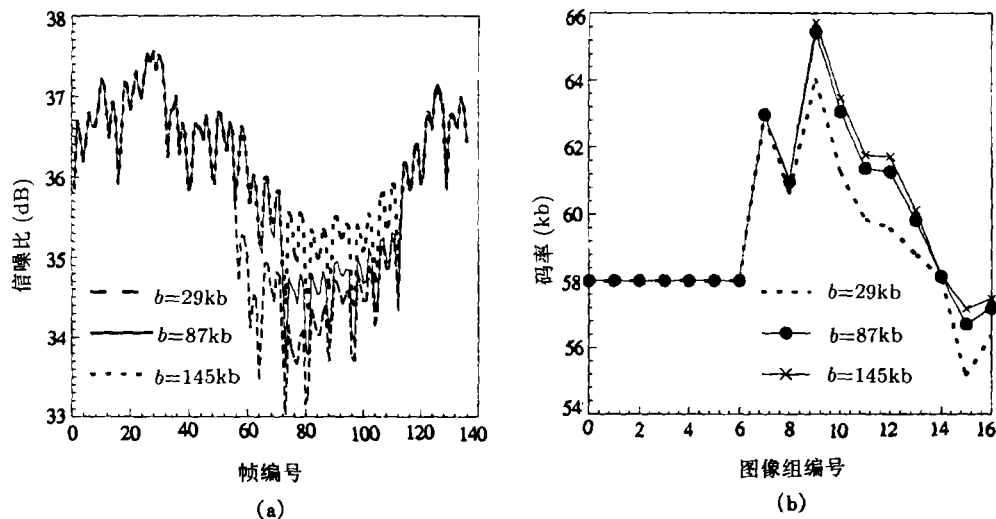


图 6 “Miss America” 序列分别在 $b = 29\text{kb}$ 、 $b = 87\text{kb}$ 、 $b = 145\text{kb}$ 时的
(a) 恢复图像的信噪比较; (b) 码率比较

对于基于 ATM 的宽带网中的多媒体通信, 视频编码尤为重要, 本文提出了一种新的基于三维小波变换的视频编码方法, 其压缩效率高于传统的基于 DCT 的压缩算法。为了避免网络拥塞的发生, 针对该编码系统, 提出了一种基于漏桶的码率控制方法。在遵循网络连接突发性约束条件下, 尽可能地兼顾各个图像组的运动性强弱, 以便使译码后的恢复图像的质量最佳。由于编码器的输出码流是嵌入的, 可根据漏桶的流量描述参数, 在任意时刻截止码流, 实现精确的码率控制。而且本码率控制是在视频图像组级上进行的, 易于实现。

参 考 文 献

- [1] Garrett M W, Willinger W. Analysis, modeling and generation of self-similar VBR video traffic. in Proc. SigComm., ACM, New Mexico, USA: Sept, 1995.
- [2] Heeke H. A traffic control algorithm for ATM networks. IEEE Trans. on Circuits and Syst., Video Technol., 1993, CAS VT-3(3): 182-189.
- [3] Reibman A R, Haskell B G. Constraints on variable bit rate video for ATM networks. IEEE Trans. on Circuits and Syst., Video Technol., 1992, CAS VT-2(6): 361-372.
- [4] Coelho R, Tohme S. Video Coding Mechanism to Predict Video Traffic in ATM Network. in Proc. IEEE GLOBECOM'93, Orlando, USA: 1993, 447-451.
- [5] LeGall D J. The MPEG video compression algorithm. Signal Processing: Image Commun., 1992, 4(2): 129-140.
- [6] Shapiro J M. Embedded image coding using zerotrees of wavelets coefficients. IEEE Trans. on Signal Processing, 1993, SP-41(12): 3445-3462.

- [7] Said A, Pearlman W A. A new, fast, and efficient image codec based on set partitioning in hierarchical trees. IEEE Trans. on Circuits and Syst., Video Technology, 1996, CAS VT-6(3): 243-249.

VBR WAVELET VIDEO CODING AND RATE CONTROL IN ATM NETWORKS

Lin Gang Liu Zemin

(*Radio Eng. Dept., Beijing University of Posts and Telecomm., Beijing 100876*)

Abstract In this paper, a new video coding system based on wavelet transform is presented, and a rate control algorithm is proposed for it. First, 3D wavelet transform is performed to the original image sequence, and an extension of set partitioning in hierarchical trees algorithm is used to quantize the wavelet coefficients. Then, the rate of the coder is controlled at GOP scale, ensuring that output conforms to the parameters of a leaky bucket controller. Several leaky buckets with different sizes are discussed too. The simulation shows the efficiency of our coding system and rate control algorithm.

Key words ATM networks, Video coding, Wavelet transform, Rate control

林 刚: 男, 1972 年生, 博士生, 研究方向为视频编码、小波分析和 ATM 技术.

刘泽民: 男, 1927 年生, 博士生导师, 研究领域为信号处理、智能技术在高速通信网中的应用.