

# 智能集成 VBR MPEG 视频流量预测模型

刘晓颖<sup>1</sup>, 戴琼海<sup>2</sup>, 刘晓冬<sup>1</sup>

(1 清华大学深圳研究生院宽带网多媒体研究中心, 广东深圳 518055; 2 清华大学自动化系宽带网数字媒体实验室, 北京 100084)

**摘 要:** 作为数字媒体网络视频通信的主要方式, VBR MPEG 视频流量的预测能力是直接关系缓冲区设计、动态带宽分配及拥塞控制等提高网络服务质量的关键因素. 因此针对 MPEG 视频流的复杂特性, 充分利用人工智能方法的优势, 提出并建立了基于模糊神经网络的智能集成 VBR MPEG 视频流量预测模型. 采用模糊预测模型提高预测精度, 利用神经网络解决预测的实时性问题. 实验结果表明, 与标准 AR 预测模型相比, 该模型预测的准确度和可靠性显著提高, 且算法简单易于推广到其他方法中使用.

**关键词:** 可变比特视频流; 视频流量; 模糊神经网络; 智能集成; 预测模型

**中图分类号:** TP39 **文献标识码:** A **文章编号:** 0372-2112 (2006) 05-0833-04

## Intelligent Integrated Traffic Prediction Modeling for VBR MPEG Video Transmission

LU Xiaoying DAI Qionghai LU Xiaodong

(1. Broadband Network & Multimedia Research Center, Graduate School at Shenzhen, Tsinghua University, Shenzhen, Guangdong 518055, China; 2. Department of Automation, Tsinghua University, Beijing 100084, China)

**Abstract** As a main video transmission mode for digital media networks, the capability to predict VBR video traffic can significantly improve the effectiveness of quality of services. Therefore, aiming at the complex characteristics of MPEG videos, a novel intelligent integrated traffic prediction model is proposed based on fuzzy and neural network. The prediction error is reduced by the fuzzy predictor and the implementation of neural network is used to lower prediction computation for real time. Simulation results show that the proposed method is able to predict the original traffic more accurately than the normal AR method and can be easily applied in other methods.

**Key words** variable bit rate of moving picture experts group (VBR MPEG); video traffic; fuzzy neural network (FNN); intelligent integrated prediction model

### 1 引言

视频已经成为网络通信的主要组成部分, 为了提高网络业务质量, 研究建立精确的视频流量预测模型对于设计更好的缓冲区、更合理的动态带宽分配以及更高效的拥塞控制方案, 从而保证网络视频传输的稳定、高性能和效率, 起着不可替代的作用<sup>[1-3]</sup>.

传统的流量管理方案采用的是基于模型参数的方法. 然而, 这种方法在 MPEG (Moving Picture Experts Group) 视频服务应用中存在几个问题: 首先, 由于 MPEG 视频流量的复杂特性, 难以建立精确的数学模型; 其次, VBR (Variable Bit Rate) 视频传输的突发性容易导致网络资源的崩溃和低利用. 为了解决上述问题, 近年来提出了动态资源分配方法. 与确定固定带宽的模型参数方法相比, 动态方

法优越性在于利用实时测量为每个用户调整带宽分配, 从而合理地利用了网络资源, 提高了网络传输效率. 流量预测作为动态方法中使用的最普遍的技术, 预测的准确性和实时性与 QoS (Quality of Services) 密切相关. 因此, 研究 VBR MPEG 视频流量的预测模型技术对于提高网络服务质量具有重要意义<sup>[4-5]</sup>.

近来 VBR MPEG 视频流量在线预测取得的进展主要有: 将递归最小二乘法和时滞神经网络的方法应用于 JPEG (Joint Photographic Expert Group) 和 MPEG 视频流量预测, 多复用 MPEG 视频流量预测的扩展递归最小二乘法, 基于最小均方的线性预测方案, 以及为了快速收敛提出的自适应小波预测方法等<sup>[3]</sup>. 由于人工智能方法在实时性要求很高且无法准确描述的问题处理上具有不可比拟的优势, 与传统的基于数学的预测方法相比, 能够更加适

收稿日期: 2004-11-25 修回日期: 2006-03-03

基金项目: 国家重点自然科学基金 (No. 60432030); 国家杰出青年科学基金 (No. 60525111)

应复杂视频网络传输的要求,采用智能方法进行视频流量的预测研究已成为当前的研究热点之一<sup>[6]</sup>.

因此,本文以人工智能方法为主体,在对 MPEG 视频特性进行分析的基础上,将模糊逻辑无需对象精确模型的优点与神经网络的高度并行性求解能力相结合,建立了智能集成 VBR MPEG 视频流量预测模型.采用模糊逻辑预测模型提高流量预测的准确性,在多步预测中引入前馈神经网络减少计算量以满足网络通信的实时性要求.采用标准视频测试序列对方法进行验证.与标准 AR(autoregressive)模型<sup>[7]</sup>相比,实验结果表明该模型预测误差显著减小,且计算方便,易于在网络的连接准许控制、参数控制及拥塞控制等算法中实现.

## 2 MPEG 视频特性

典型的 MPEG 视频主要特征在于图像类型产生的方式. MPEG 视频包含三种帧型: I 帧和 B 帧.尽管没有标准要求,但是给定视频序列的帧一般都按照确定的规则模式排列,称为 GOP(Group of Picture)模式. GOP 模式的特点主要在于两个参数:从一个 I 帧图像到下一个 I 帧图像的帧距  $N$ ,以及从一个 I 帧图像到下一个 P 帧图像或两个连续 P 帧图像的帧距  $M$ .由于视觉信息在压缩过程中,编码类型和帧周期的通信模式均取决于 GOP 模式,使得 VBR MPEG 广播视频表现出复杂的和突发的传输特性.

通常数字视频源采用的是 MPEG-2 编码方式,例如,数字标准测试视频序列中很多都是以 MPEG-2 编码器编码,然后将输出流存为磁盘文件用作分析.在研究中经常作为测试标准的视频序列<sup>[2]</sup>短片有“flower garden”、“mobile and calendar”等,长片有“star wars”等,它们都是从具有中等级别场景变换的纪录片和长片中选出的代表性场景.经过混合编码算法(运动补偿、离散余弦变换)产生的输出比特流周期地包含 I 帧和 B 帧. I 帧采用无需参考其他图像的帧内编码,因此压缩比最低. P 帧根据先前的 I 或 P 帧作为运

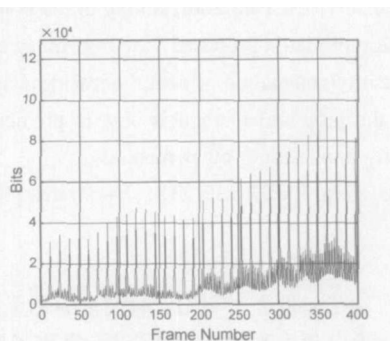


图1 完整的 MPEG 视频流量波形图

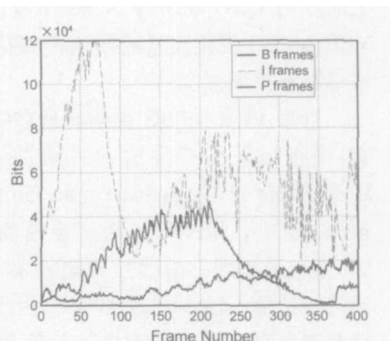


图2 分离的 MPEG 视频流量波形图

动预测补偿的参考,压缩比较高. B 帧采用双向预测编码方案,需要以前和以后的 I 或 P 帧用于插值计算以获得最大压缩比.从“star wars”中截取的视频流量轨迹如图 1 所示,可以看出,完整的编码流包含了突发的和某种程度上周期性的比特率变化.而分离的 I、P 和 B 帧则表现出较为缓和的比特率变化,如图 2 所示.因此,对分离的 I/P/B 帧可以更准确地预测流量.

## 3 基于模糊神经网络的视频流量预测模型

### 3.1 线性 AR 预测模型

自回归估计是研究预测使用的非常普遍的方法.对于接近高斯分布的视频流,AR 预测模型表述如下.

对于次数为  $n$  且具有均值  $\mu$  的 AR 过程  $X_t$ , 可以表示为:

$$X_{t+1} - \mu = \sum_{k=0}^{n-1} \alpha_k (X_{t-k} - \mu) + Z_{t+1} \quad (1)$$

其中,  $\alpha_k$  是参数,  $Z_t$  为过程残差,在理想情况下都是以零平均值正态分布.显然,可得到单步预测的应用方程如下:

$$X_{t+1} - m_u = \sum_{k=0}^{n-1} \alpha_k (X_{t-k} - m_u) \quad (2)$$

其中,  $m_u$  是组成预测单元的  $X_t$  的  $u$  个先前值的平均值,  $X_{t+1}$  表示在时刻  $t$  的前步预测帧率,  $u$  过大或过小都会增大预测误差,而且  $u$  取值过大也会增加获得  $\alpha_k$  的计算量.

由式(2),可得多步预测的方程式如下:

$$X_{t+L} - m_u = \sum_{k=0}^{n-1} (\alpha_k)^L (X_{t-k} - m_u), \quad L = 1, 2, \dots, n \quad (3)$$

其中,  $L$  为预测步数,  $X_{t+L}$  表示在时刻  $t$  的前  $L$  步预测帧率.如果单步预测过程残差是  $Z_{t+1}$ , 则多步预测过程残差为

$$\sum_{i=0}^{L-1} \sum_{k=0}^{n-1} (\alpha_k)^i Z_{t+1-i}$$

AR 预测模型所需的估计参数  $\alpha_k$  个数与其阶数相同,因此模型的阶数越高,需要计算的参数也越多,计算量也随之上升,但是提供的结果并不一定有明显的改善.

### 3.2 模糊预测模型

由于一般通信模式大多表现为非高斯或非稳定的过程,参数与估计误差之间不存在相关性,无法建立精确的数学模型,所以采用自回归方法难以获得满意的预测结果.根据 MPEG 视频特性分析结果可知,图像流呈现出突发的和周期性变化特点,因此,利用模糊数学原理能将复杂的预测问题转换成线性规划问题求解的优势,并结合时间序列技术能够克服偶然因素产生的随机性影响的优点,建立了模糊预测模型.

指数平滑(Exponential Smoothing)是根据以前的实际值和预测值的加权平均值估算未来预测值的.亦即,未来预测是由原预测值与实际值间的误差而形成的原预测的增加部分函数.设调整增量为  $\alpha$  实际帧率为  $X_t$ , 则时刻  $t$  的单步预测可表示如下:

$$X_{t+1} = \alpha X_t + (1 - \alpha) X_t \quad (4)$$

其中,  $X_{t+1}$  是  $X_t$  的未来预测帧率,  $X_{t+1}$  是  $X_t$  的预测帧率 (即  $X_t$  的原预测帧率),  $\alpha$  是平滑常数 ( $0 \leq \alpha \leq 1$ ).

采用递归方法, 由式 (4) 可得时刻  $t$  的多步预测表示如下:

$$X_{t+1L} = \alpha X_{tL-1} + (1 - \alpha) X_{tL}, \quad L = 1, 2, \dots, n \quad (5)$$

其中,  $L$  为预测步数,  $X_{t+1L}$  是  $X_t$  的前  $L$  步未来预测帧率,  $X_{tL}$  是  $X_t$  的前  $L$  步预测帧率,  $X_{tL-1}$  是  $X_t$  的前  $(L - 1)$  步预测帧率. 可以看出, 当  $L = 1$  时, 多步预测式 (5) 即等价于单步预测式 (4).

由式 (4) 和 (5) 可知,  $\alpha$  作为过去作用的权重, 是唯一可调的参数.  $\alpha$  越大, 表明越倚重近期数据所载的信息, 修正的幅度也越大;  $\alpha$  越小, 修正的幅度也越小. 简单地说, 就是预测误差越小, 则  $\alpha$  越大, 反之亦然. 因此研究如何选择  $\alpha$  是提高模型预测准确性的主要方法. 由于  $X_t$  的分布是未知的, 采用模糊控制器来决定  $\alpha$  取值是最简单有效的解决方法. 在初始阶段选择较大的  $\alpha$  则给当前样本值的权数就大, 从而减少了由于初始值选择不当而引起的偏差.

以相对预测误差作为模糊控制器的输入, 根据式 (4) 定义为:

$$pe = |X_{t+1} - X_t| / X_{t+1} \quad (6)$$

式中,  $pe \in [0, 1]$  表示相对预测误差. 相对预测误差越小, 表明预测精度越大.  $\alpha$  作为模糊控制器的输出. 将  $pe$  和  $\alpha$  在  $[0, 1]$  范围内分为 {大 (Large), 中 (Medium), 小 (Small)} 三个等级, 则  $\alpha$  的模糊规则推论如下:

- ⑧ F  $pe = \text{Large}$  THEN  $\alpha = \text{Small}$
- ⑧ F  $pe = \text{Medium}$  THEN  $\alpha = \text{Medium}$
- ⑧ F  $pe = \text{Small}$  THEN  $\alpha = \text{Large}$

$pe$  和  $\alpha$  的隶属度函数采用三角函数表示, 模糊控制器的灵敏度参数是从相对预测误差  $pe$  到  $\alpha$  的映射. 由于输入采用了模糊集, 输出利用了质心反模糊化过程, 因此该映射是一个有理函数, 灵敏度参数用  $m_{pe}$  和  $m_\alpha$  表示. 在 0.05 和 1 处的导数可给定:

$$\frac{-m_\alpha(4-m_\alpha)}{2m_{pe}(1-m_\alpha)}, \quad 0, \quad -\frac{(m_\alpha+3)(1-m_\alpha)}{2m_\alpha(1-m_{pe})} \quad (7)$$

3.3 神经网络实施

由于多步预测采用了递归的方法, 计算量将随着预测步数的增加而呈线性上升趋势, 为了限制多步预测的计算量, 既满足系统实时性的要求, 同时又能保证预测精度, 因此利用神经网络的高度并行性求解问题的能力, 在多步预测计算中引入了一个三层前馈神经网络 (Feed forward neural networks FFNN). 在 FFNN 中, 输入层和输出层为线性转移函数, 隐含层采用 Hypertangent Signoid 转移函数, 输入层和隐含层的神经元个数相等. 对 FFNN 进行训练时, 从模糊预测模型中获得的输入和输出数据集作为训练事例, 训练序列选择在生产方式中得到的具有相似均方差的值, 输入序列的个数为所需预测步数, 输出量即为所求预测值.

将神经网络与模糊预测模型结合建立的智能集成视频流量预测模型如图 3 所示.

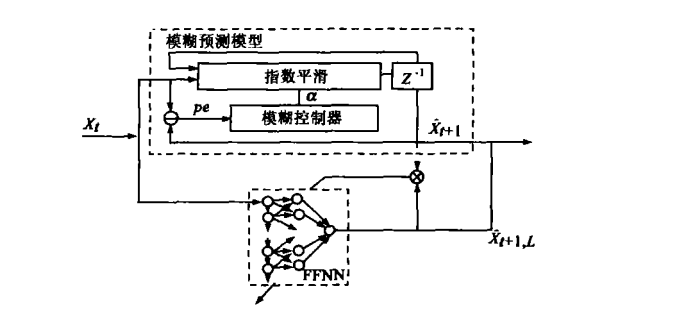


图 3 智能集成 VBR MPEG 视频流量预测模型

整个模型充分利用了模糊预测模型和 FFNN 的优点进行互补, 单步预测时采用模糊预测模型, 多步预测时转入 FFNN 模型, 既提高了预测精度, 又降低了计算量. 此外, 为了符合视频变化要求, 在模型中设计了自学习功能, 将模糊预测模型与神经网络预测结果的误差作为神经网络的训练参数, 从而使其参数不断进行调整更新, 保证了模型预测的准确性.

4 仿真比较

采用标准视频测试序列作为仿真对象对提出的方法进行验证. 如图 2 所示, 从 “star wars” 中获取 1 图像流 3333 帧作为原始轨迹, 平均值  $m = 4.401 \times 10^4$  (Bits/Fram e), 标准偏差.  $\sigma = 1.414 \times 10^4$ . 将预测得到的  $m$  和  $\sigma$  与原始轨迹的  $m$  和  $\sigma$  采用相对误差方法进行分析, 分别用  $E(m)$  和  $E(\sigma)$  表示, 各实验方法的预测误差分析比较结果如表 1 所示, 预测结果对原始轨迹的跟踪曲线 (以多步预测为例) 如图 4 所示. 其中, AR (1) ~ AR (4) 表示 AR 的一阶 ~ 四阶预测模型, FNN 表示提出的预测模型.

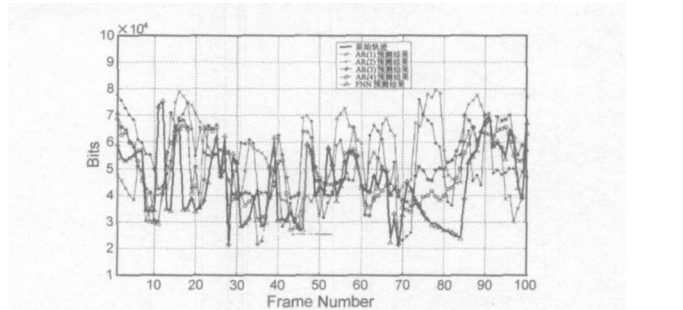


图 4 多步预测 ( $L=20$ ) 结果轨迹跟踪对比

表 1 预测误差分析比较结果				
实验方法	单步预测		多步预测 ( $L = 20$ )	
	$E(m) (\%)$	$E(\sigma) (\%)$	$E(m) (\%)$	$E(\sigma) (\%)$
AR (1)	17.79	15.21	31.09	29.45
AR (2)	9.45	10.71	19.78	20.34
AR (3)	2.78	4.47	9.58	11.07
AR (4)	1.73	4.08	7.39	10.46
FNN	0.11	0.50	1.01	2.37

对 AR 与 FNN 的仿真结果进行分析可知, AR 模型的预测误差不仅没有随着阶数的增加而明显减少, 相反因为阶数个数等于所需计算的参数个数, 导致计算更加复杂. FNN 模型的预测误差比 AR 明显减少, 不仅只需要计算一个参数  $q$  而且神经网络的实施又保证了模型的实时性, 从而表明该模型提供了一种更加准确且简单有效的预测方法.

## 5 结论

由于 VBR MPEG 视频无法用稳定的正态分布过程表示, 传统的线性 AR 预测模型难以获得良好的效果, 为了解决网络传输中的流量控制问题, 利用人工智能方法对实时性要求很高且无法准确描述问题的处理优势, 将模糊逻辑与神经网络的优点相结合, 建立了智能预测模型对其进行流量预测分析. 以标准视频序列作为实验对象, 测试结果表明, 与标准的 AR 预测模型相比, 该模型的预测误差不仅显著降低, 并且在预测的准确性与实时性方面具有明显的优势, 从而为 VBR MPEG 视频流量预测提供了一种新的更加有效的方法.

## 参考文献:

- [1] Farshchian M, et al Optimal error protection for real-time image and video transmission[ J]. IEEE Signal Processing Letters 2004, 11(10): 780- 783
- [2] Basir I et al A lightweight dynamic rate control scheme for video transmission over IP network[ J]. Pattern Recognition Letters 2004, 25(7): 817- 827.
- [3] Yoo S J Efficient traffic prediction scheme for real-time VBR MPEG video transmission over high-speed networks[ J]. IEEE Trans Broadcasting 2002, 48(1): 10- 18
- [4] Yoo S J et al Predictive and measurement-based dynamic

resource management and Qos control for videos[ J]. Computer Communications 2003, 26(14): 1651- 1661

- [5] Rikli N E. Modeling techniques for VBR video Feasibility and limitations[ J]. Performance Evaluation, 2004 57(1): 57- 68
- [6] Yao L. Real-time VBR video traffic prediction for dynamic bandwidth allocation[ J]. IEEE Trans Systems Man and Cybernetics-Part C: Applications and Reviews 2004, 34(1): 32- 47.
- [7] Chan W S et al Multiple forecasts with autoregressive time series models: Case studies[ J]. Mathematics and Computers in Simulation, 2004, 64(3-4): 421- 430

## 作者简介:



刘晓颖 女, 1973年9月生于河南郑州市, 清华大学电子系信息与通信工程学科博士后, 主要从事数字媒体网络的视频传输技术、流媒体技术、多媒体技术与应用、人工智能技术等方面的研究工作. E-mail: liuxy@sz.tsinghua.edu.cn



戴琼海 男, 1964年12月生, 博士, 清华大学教授, IEEE Senior Member 清华大学自动化系宽带网数字媒体实验室主任, 清华大学深圳研究生院宽带网多媒体研究中心主任. 已完成国家“九五”、“十五”课题及国家自然科学基金多项, 目前承担杰出青年基金和国家重点自然科学基金课题. 主要研究方向为视频处理及通信、光场及其立体视频、分布式编码、无线视频传输等和宽带信息网络业务资源模型及分析等. E-mail: qhda@tsinghua.edu.cn