

ATM 网络中全动态视频源的二次建模

赵海军¹, 崔梦天², 李明东¹

(1. 西华师范大学计算机学院, 南充 637002; 2. 中国科学院成都计算机应用研究所, 成都 610040)

摘要:通过对 ATM 网络视频业务传输性能及其建模分析, 针对高运动性(即全动态)视频源的建模, 给出基于块的 MPEG 运动估计算法, 提出一种非线性模型。基于自回归在 GOP 层对 MPEG 视频流进行二次建模。仿真表明, 该文提出的 NAR 模型能准确预测具有适度运动性的实际 MPEG 流, 对具有高运动性的流在不同的负荷和缓冲器的条件下提供可接受的预测。

关键词:运动估计; 自回归; 全动态; 非线性模型

Quadratic Modeling for Full Motion Video Source in ATM Networks

ZHAO Hai-jun¹, CUI Meng-tian², LI Ming-dong¹

(1. School of Computer Science, China West Normal University, Nanchong 637002;

2. Chendu Institute of Computer Applications, Chinese Academy of Sciences, Chengdu 610040)

【Abstract】 A motion estimation algorithm for block-based MPEG is given by analyzing video traffic transmission performance and modeling in ATM networks and modeling method of high degrees of motion video source, namely, full motion, and a non-linearity model is introduced. and used to AR-based quadratic modeling for MPEG video traffic at GOP layer. Simulation shows that NAR model can provide accurate prediction to real MPEG flow with moderate motion and acceptable prediction to traffic with high motion at various load and buffer sizes.

【Key words】 motion estimation; Auto Regressive(AR); full motion; non-linear model

1 概述

随着网络技术的飞速发展, 传输多种类型信息的宽带综合数字网是发展的必然趋势。ATM 把图、文、声、像数据转化为统一格式的信元序列, 从而为计算机网络的综合业务应用提供了基础。

近年来, 流建模已经成为一个活跃的研究领域, 文献[1]提出针对可变比特率(VBR)视频流的几种较好的模型, 如马尔可夫(Markov)模型, 自相似(Self-similr)模型和自回归(Auto Regressive, AR)模型。传统的马尔可夫模型只反应了实际序列的短时相关特性; 自相似模型与传统的马尔可夫模型的不同之处是它具有长时相关性; AR 模型有 2 个优点: (1)常见且计算上更容易; (2)存在多种随机过程来对自相关压缩 VBR 视频模型建模。针对这种视频类型已经提出了好几种 AR 模型, 文献[2]提出一种针对单一视频源的比特率编码的 AR(1)模型; 文献[3]提出一种无场景切换、变焦和追镜的视频会议 AR(2)模型和具有任意相关性的且基于生成 γ 分布的流模型。这种方法把 γ 过程分解成一个许多卡方序列的线性组合。该模型具有方便的解析特点而且兼容 AR 模型以及可以任意调整 γ 分布参数, 但该模型只能应用于视频会议而并不适合全动态视频的建模。对于全动态视频源序列的建模; 文献[4]提出更高阶的 AR(阶数 $p=5, 20, 30, 100$)来对画面组(GOP)层建模。这种模型能够模拟二阶统计值, 但过低估计了丢失率; 文献[5]提出一种基于嵌套的 AR 过程的更精确的模型, 即采用 2 个 AR 过程来解释场景的均值和波动, 能获得短时和长时相关性。但这种模型是基于场景级上的 MPEG 视频流建模而不是 GOP 建模。

本文提出基于块的 MPEG 运动估计算法, 引入一种非线

性模型, 并用它来基于自回归在 GOP 层对高运动性(全动态)VBR MPEG 编码的视频流进行二次建模并进行仿真。

2 MPEG 运动估计算法

在基于块的混合视频编码器中, 运动估计就是给当前帧中一个矩形块在时域上前面或后面的帧中找到一个最小化某个花费函数的过程, 这个花费函数可能是绝对差之和、均方差之和或联合考虑帧差码率和运动向量码率的一个花费函数^[6]。

对于一个宏块内任意一个部分, 它的运动向量表示为 MV , $MV.x$ 和 $MV.y$ 表示运动向量 MV 的 x 和 y 分量; 花费函数 $J(MV)=SAD(MV)+\lambda_{\text{motion}} \times QP \times R(MV)$; 最佳的运动向量(最小化花费函数 $J(MV)$ 的运动向量)表示为 MV_{best} ; $J(x, y)$ 表示运动向量是 (x, y) 时的花费; $SAD(x, y)$ 表示运动向量是 (x, y) 时的 SAD ; λ 是拉格朗日乘子; QP 表示宏块量化参数; R 是当前帧期望的比特数。一般认为, 在较小的范围内, 有如下关系:

如果 $\|MV_1 - MV_{\text{best}}\|$ 和 $\|MV_2 - MV_{\text{best}}\|$, 那么 $J(MV_1) < J(MV_2)$, 其中,

$$\|MV - MV_{\text{best}}\| = \sqrt{(MV.x - MV_{\text{best}}.x)^2 + (MV.y - MV_{\text{best}}.y)^2}$$

表示 2 个运动向量之间的距离。

将一个任意块分为 A, B 2 个部分, 这个块的最佳运动向量用 MV_{best} 表示, A 和 B 的最佳运动向量和花费分别用

基金项目: 四川省科技攻关计划基金资助项目(07GG006-014)

作者简介: 赵海军(1966 -), 男, 讲师、硕士, 主研方向: 无线数据通信; 崔梦天, 副教授、博士; 李明东, 教授

收稿日期: 2009-03-05 **E-mail:** zhaohai_jun@163.com

MV_{bestA} , MV_{bestB} , $J_A(MV)$ 和 $J_B(MV)$ 表示, 则有如下关系成立:

$$\min\{MV_{bestA \cdot x}, MV_{bestB \cdot x}\} \quad MV_{best \cdot x} \quad \max\{MV_{bestA \cdot x}, MV_{bestB \cdot x}\} \quad (1)$$

$$\min\{MV_{bestA \cdot y}, MV_{bestB \cdot y}\} \quad MV_{best \cdot y} \quad \max\{MV_{bestA \cdot y}, MV_{bestB \cdot y}\} \quad (2)$$

一个宏块的速度场可认为是分段连续函数 $f(x, y)$ ($x, y \in [1, 16]$), 因此, 宏块各点的位移 $g(x, y) = \Delta t \times f(x, y)$ 也是分段连续函数(其中 Δt 是帧的周期)。设 $g(x, y)$ 的取值范围是 Σ , $x_{\min} = \min\{x, |(x, y) \in \Sigma\}$, $x_{\max} = \max\{x, |(x, y) \in \Sigma\}$, $y_{\min} = \min\{y, |(x, y) \in \Sigma\}$, $y_{\max} = \max\{y, |(x, y) \in \Sigma\}$ 。

重复应用式(1)和式(2)可以推出, $MV_{best \cdot x} \in [x_{\min}, x_{\max}]$, $MV_{best \cdot y} \in [y_{\min}, y_{\max}]$ 。可以用宏块各部分的 SAD 之间的差别来表示运动的复杂程度。

首先在率失真框架下进行 16×16 运动估计, 得到相应的 SAD_{16} 和 MV_{16best} 。将一个宏块等分为 16 个 4×4 部分, 分别表示为 1, 2, ..., 16, 由 MV_{16best} 所对应的各部分的 SAD 表示为 SAD_i ; ($i=1, 2, \dots, 16$), 宏块所包含的 4 个块的 SAD、最佳运动向量和花费函数分别表示为 SAD_{ij} , MV_{ijbest} , J_{ij} , SAD_{ij} , MV_{ijbest} , J_{ij} , SAD_{jii} , $MV_{jii best}$, J_{jii} , SAD_{jv} , $MV_{jv best}$, J_{jv} 。

算法依次对每个块设置一个准 $MV_{best}(QMV_{best})$, 并假设这个块在 QMV_{best} 取得 SAD 最小(并不是花费函数最小)。设置的原则是 2 个水平相邻块的 QMV_{best} 的 x 分量相对于 MV_{16best} 的 x 分量的差值符号相反, 2 个垂直相邻块的 QMV_{best} 的 y 分量相对于 MV_{16best} 的 y 分量的差值符号相反。然后设置这个块在 QMV_{best} 下的 SAD, 它的设置必须反映由于 MV_{16best} 对当前块的 MV_{best} 的偏离而造成的 SAD 的增加, 在本算法中使用当前块各部分和临近的宏块部分之间的差别来计算, 并假设这个块的 SAD 在 QMV_{best} 附近水平方向和垂直方向都服从线性分布。比如说第 3 个块, 沿水平方向偏离 QMV_{best} 所导致的 SAD 的增量设为每像素 $abs(SAD_9 + SAD_{11} + SAD_{13} + SAD_{15} - 2 \times SAD_{10} - 2 \times SAD_{12}) / 14$, 沿垂直方向偏离 QMV_{best} 所导致的 SAD 的增量设为每像素 $abs(SAD_3 - SAD_4 + SAD_{11} + SAD_{12} - 2 \times SAD_{10} - 2 \times SAD_9) / 4$ 。然后先在水平方向搜索, 后在垂直方向搜索, 最后得到最小化花费函数 $J = SAD + \lambda_{motion} \times QP \times R(MV)$ 。

3 视频流的二次建模

非线性模型常见的典型描述是伏尔特拉级数和维纳展开。如果 $o(n)$ 是一个离散时不变伏尔特拉滤波器的响应, 输入为 $i(n)$, 则有:

$$o(n) = h_0 + \sum_p H_p[i(n)] \quad (3)$$

其中, $H_p[\bullet]$ 是 p 阶伏尔特拉算子。本文仅对基于 2 阶时不变滤波器的流数据建模问题进行研究, 滤波器算子用 H_2 表示, 称为 2 阶伏尔特拉(或二次)算子。二次算子表示式为

$$o(n) = H_2[i(n)] = \sum_{p_1} \sum_{p_2} h_2(p_1, p_2) i(n-p_1) i(n-p_2) \quad (4)$$

令 $q(n)$ 为用于建模的第 n 个 GOP, 序列 $i(n)$ 可表示为 C 个相互独立的加权序列 $\{o_k(n)\}$ 的和, 这些序列是把 C 个独立的高斯序列 $\{g_k(n)\}$ 通过 C 个相同的二次算子得到, 可表示为

$$q(n) = \sum_{k=1}^C \alpha_k H_2[g_k(n)] = \sum_{k=1}^C \alpha_k o_k(n) \quad (5)$$

其中, $\{\alpha_k\}$ 是均值为 0 的白高斯过程, 其方差为 σ^2 。目标就是确定最佳滤波器(或模型)来与 $q(n)$ 的统计值最佳匹配。

未知模型参数的确定, 是把模型输出的一阶和二阶统计值与真实数据的统计值进行匹配。首先计算 $o_k(n)$ 的均值:

$$r_1^o = E\{o_k(n)\} = \sigma^2 \sum_i h_2(i, i) \quad (6)$$

其中, $E\{\bullet\}$ 是求数学期望。 $o_k(n)$ 的自协方差序列 $r_2^o(\tau)$ 计算方法如下:

$$r_2^o(\tau) = 2\sigma^4 \sum_{i,j} h_2(i, j) h_2(i+\tau, j+\tau) \quad (7)$$

由于 $o_k(n)$ 是相互独立的随机过程, 则有

$$r_1^q = \left(\sum_{k=1}^C \alpha_k \right) r_1^o \quad (8)$$

r_1^q 是 $q(n)$ 的均值, $q(n)$ 的自协方差序列为

$$r_2^q(\tau) = \left(\sum_{k=1}^C \alpha_k^2 \right) r_2^o(\tau) \quad (9)$$

图 1 为这种建模方法的示意图。

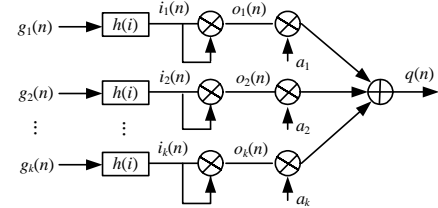


图 1 NAR 建模方法原理框图

其中, $o_k(n) = i_k^2(n)$, $i_k(n)$ 是具有脉冲响应 $H(i)$ 的线性模型的输出, 即:

$$i_k(n) = \sum_i h(i) g_k(n-i) \quad (10)$$

如果 $g_k(n)$ 为高斯过程, 则序列 $i_k(n)$ 也为高斯过程, 因此, $o_k(n)$ 是卡方变量。从统计学很容易得到, 独立的 χ^2 变量的线性组合可以准确地近似为一个 γ 变量。

未知模型参数 C 由下式决定:

$$C = \text{floor}(2e) + 1 \quad (11)$$

其中, e 是 γ 序列 $q(n)$ 的形状参数。权值 $\{\alpha_k\}$ 的计算可以通过把模型的输出均值和方差与真实数据的均值和方差进行匹配来得到, 前提是假设 $o_k(n)$ 是 $\chi^2(1)$ 或 $i_k(n)$ 的方差不变, 也就是 $E\{i_k^2(n)\} = \sigma^2 \sum_i h^2(i) = 1$ 。

序列 $i_k(n)$ 可以建模为一个级数为 b 的 AR 过程。给定 $i_k(n)$ 的自协方差, AR 系数可以用标准的 AR 参数估计算法确定。 $i_k(n)$ 的自协方差序列与 $q(n)$ 的自协方差序列的关系为

$$r_2^i(\tau) = \sqrt{c_2^q(\tau)} \quad (12)$$

其中, $c_2^q(\tau)$ 是归一化的自协方差序列, 它由 $r_2^q(\tau) = r_2^q(0) c_2^q(\tau)$ 决定。

4 模型仿真

4.1 模型跟踪的生成

对于生成合成的 GOP 序列来说, 其算法如下:

(1) 估计包括 AR 参数 $\{a_k\}$ 的估计和来自真实流的高斯过程的方差 σ^2 。AR 过程中适当的阶数 b 可以用样本自相关函数 ACF、最小描述长度(MDL)或其他方法来确定。这里采用样本 ACF, 定义为前向预测误差与后向预测误差的相关性。结果表明模型阶数 $b=5$ 对拟合试验视频序列是适当的。AR 参数值 $\{a_k\}$ 和 σ^2 的估计采用线性预测方法。对实验的视频序列而言, 分别得到 MOVIE2 和 STRWARS 的 $\{a_k\}$ 如下:

$$\{a_k\} = [0.85 \quad -0.02 \quad 0.04 \quad -0.06 \quad 0.10]$$

$$\{a_k\} = [0.78 \quad -0.01 \quad 0.05 \quad -0.03 \quad 0.08]$$

方差 σ^2 上面已得到, 分别为 $\sigma^2 = 0.24, 0.20$ 。

(2) 得到 $\{\alpha_k\}$ 。首先, 形状参数 e 用边缘概率密度 γ PDF 估计器来估计。其次, 未知模型参数 C 由式(11)确定。然后, $\{\alpha_k\}$ 就可以通过把模型输出的均值和方差与真实数据的均值

和方差进行匹配得到。结果得到 MOVIE2 和 STARWARS 的 e, C, α_k 分别为

$$1) e=4.00, C=7, \alpha_k=[47 \ 47 \ 47 \ 47 \ 47 \ 69.2 \ 60.4]$$

$$2) e=6.8, C=13, \alpha_k=[32 \ 32 \ 32 \ 32 \ 32 \ 32 \ 32 \ 32 \ 32 \ 32 \ 32 \ 55.70 \ 44.16]$$

(3)生成合成的 GOP 实现。可通过将一个高斯序列应用到 AR 模型,得到矩脉冲输出,再乘上适当的 α_k 并对所有 C 个实现相加。

4.2 信元丢失性能实验

4.2.1 实验模型

为了仿真 MPEG 视频流通过 ATM 网络的传输,把网络模拟为一个带有缓冲器(在 ATM 信元级上)有限容量排队系统和带有固定通道容量 a (或服务率)的(在信元/GOP 级上)一个服务器。由一个单一的 MPEG 视频源构成的输入序列以速率 λ 到达,输入可以是真实跟踪、NAR 跟踪或 CAR 跟踪。视频源以 GOP 速率 $t_g=(N \text{ 帧}/\text{GOP})/f \text{ 帧}/\text{秒}$ 在每个 GOP 中生成一个可变数量的信元,时间被划分成片,而且一片等于一个 GOP 的持续时间,即 $1/t_g \text{ s}$ 。对本文试验的视频序列来说, $t_g=440 \text{ ms}$ 或 460 ms 。在模拟器中,一个视频源的 GOP 被分成 46 Byte 有效载荷的 ATM 信元,而且这些信元被假设均匀分布在 GOP 持续时间 $1/t_g$ 上。服务率 a 按期望的负荷条件 δ 改变且定义为

$$a = \lambda / \delta \quad (13)$$

λ 的计算如下:

$$\lambda = \mu / t_g \quad (14)$$

假设为先进先出(FIFO)服务训练。当缓冲器存满且信元被抛弃时就出现信元丢失。用信源丢失率(Cell Loss Rate, CLR)来作为衡量指标且定义为:总的丢失信元数量与传输的总信元数量的比值。

4.2.2 实验结果

图 2 所示为 MOVIE2 序列的真实流、生成 NAR(非线性自动回归)流和生成 CAR(多级自动回归流)的 CLR。

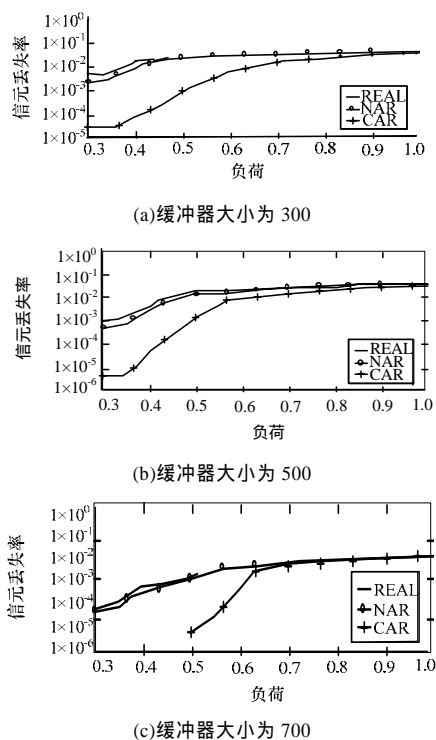


图 2 不同大小缓冲器的 MOVIE2 序列的信元丢失率

基于建模的 CLR 值估计如下:

(1)8 个独立样本的实现从各自的模型中生成;

(2)当 8 个实现输入到模拟器时,就计算每个实现的 CLR;

(3)计算出 8 个 CLR 的平均值,通过假设标准误差服从 t 分布来建立一个 96% 的置信区间。

对不同负荷情况的 CLR 都进行检验,范围从轻负荷到重负荷($\delta=15\% \sim 100\%$)。模拟针对 3 个最大的缓冲器:300, 500 和 700 信元来进行。

由图 2(a)可见,对缓冲为 300 信元而言,本文提出的 NAR 模型给出了接近真实流的 CLR 预测,尽管在轻负荷区有轻微的偏差。当缓冲器增大时,CLR 减小而且在不同系统负荷上的 2 条曲线匹配得非常好;在 CAR 模型下,在重负荷区域的性能与真实流非常一致,但在 $\delta > 80\%$ 的负荷条件下, CAR 模型的 CLR 开始减小且低于真实流的 CLR 2 个量级(当 $\delta=40\%$)。还可以观察到, CAR 模型的 CLR 对缓冲器的大小非常敏感,这可以从图 2(c)看到。至于 STARWARS 序列,本文所提出的 NAR 模型的性能在 $\delta < 58\%$ 对估计信元丢失稍微欠健壮。在最坏的情况下(当 $\delta=38\%$),CLR 相对于真实数据小一个数量级。原因是在某些视频序列部分存在高度运动性,但它提供了可接受的预测,而且大大优于 CAR 模型性能近 2 个数量级。于是得出结论,本文提出的 NAR 模型能准确预测具有适度运动的实际 MPEG 流的排队性能,而且对具有高运动性的流在不同的负荷和缓冲器的条件下能够提供可接受的预测。

5 结束语

视频流是 ATM 网络所支持的重要流类型。与其他多媒体流相比,视频流是动态的和相当复杂的,这导致源建模成为一件困难的工作。本文引入一种非线性模型,并用它对 GOP MPEG 视频流建模。此外,还将一个单一的视频流经过一个一般的缓冲器对模型的性能进行试验(缓冲器类似于一个网络节点)。实验结果表明,模型得到了与具有适度运动的真实视频非常好的匹配和与具有高度运动性的真实视频可接受的匹配。

参考文献

- [1] Adas A. Traffic Models in Broadband Networks[J]. IEEE Trans. on Communication, 1997, 35(7): 82-89.
- [2] Maglaris M, Anastassiou D, Sen P, et al. Performance Models of Statistical Multiplexing in Packet Video Communications[J]. IEEE Trans. on Communication, 1988, 36(7): 834-843.
- [3] Heyman D, Tabatabai A, Lakshman T V. Statistical Analysis and Simulation Study of Video Teleconference Traffic in ATM Networks[J]. IEEE Trans. on Circuits System Video Technology, 1992, 1(2): 49-59.
- [4] Alheraish A A, Alshebeili S A. AR-based Quadratic Modeling for GOP MPEG-encoded Video Traffic in ATM Networks[J]. Computer Communications, 2004, 27(1): 81-92.
- [5] Liu Derong, Sara E, Sun Wei. Nested Auto-regressive Processes for MPEG Encoded Video Traffic Modeling[J]. IEEE Trans. on Circuits System Video Technology, 2004, 11(2): 169-183.
- [6] 包红强. 基于内容的视频运动对像分割技术研究[D]. 上海: 上海大学, 2005.

编辑 金胡考