

网络环境下视频连续播放控制算法

许 延^{1,2}, 常义林¹, 刘增基¹

(11 西安电子科技大学 ISN 国家重点实验室, 陕西西安 710071; 21 中国网通集团研究院, 北京 100036)

摘 要: 网络的时延抖动会造成媒体播放的不连续, 本文从信宿端的角度来解决视频媒体的同步连续播放, 提出了一种自适应的动态媒体播放算法. 该方法在接收端设置缓冲区, 根据缓冲区的占用情况动态地调整媒体的播放速率, 播放速率是根据具体的媒体内同步要求选取的. 另外, 分析了马尔可夫调制的泊松到达情况下的排队模型, 给出了缓冲区门限的选取原则. 最后, 用不连续性和播放失真的方差衡量同步性能, 计算机模拟结果表明, 该算法的同步性能优于 Yuang 的算法. 通过选择合适的参数, 该算法可以有效地防止缓冲区下溢造成的播放停顿以及缓冲区上溢造成的播放跳跃, 从而实现同步平滑播放.

关键词: 媒体内同步; 缓冲区占用水平; 动态播放

中图分类号: TN915. 01 **文献标识码:** A **文章编号:** 0372-2112 (2004) 07-1078-04

A Control Scheme for Video Continuous Playback in Network Environment

XU Yan^{1,2}, CHANG Yi2lin¹, LIU Zeng2ji¹

(11 National Key Laboratory of ISN, Xidian University, Xi'an, Shaanxi 710071, China; 21 China Netcom Group Labs, Beijing, Beijing 100036, China)

Abstract: Network delay jitter often leads to the discontinuity of media playback. An improved dynamic playback algorithm is proposed in order to keep multimedia playing smoothly, which deals with the problem from the perspective of an end system. In the algorithm, playback rates of multimedia are variable according to the occupied buffer level with an intention of eliminating buffer underflow or overflow. The playback rates are chosen based on the requirement of the intramedia synchronization. To determine the thresholds of the occupied buffer level, a queuing model is established. Based on the analytic results, suitable thresholds can be determined according to the system parameters. Simulation results show the algorithm wins an advantage over Yuang's.

Key words: intramedia synchronization; buffer level; dynamic playback

1 引言

宽带通信技术和高性能计算机网络技术促进了多媒体应用的发展, 多媒体数据与传统数据的主要不同是引入了时间相关的媒体类型(如音频和视频). 这些媒体流的信息单元之间都存在着一定的时间关系, 在经过多媒体系统存储、传送而后播映多媒体数据时必须维持这种关系. 由于网络的时延抖动以及发送方和接收方时钟的差异等因素, 都会在接收方引起不同步^[1]. 如何维持单个媒体内部及相关媒体之间的时间关系, 即如何保持媒体的同步, 是远程多媒体应用中的一项关键技术.

同步控制可以在信源端进行^[2], 也可以在信宿端进行. 当在信源端实现同步控制功能时, 需要通过网络进行信息反馈, 对于在地域上很分散的系统就不太适宜. 本文将从信宿端的角度来解决媒体的同步播放问题.

在信宿端控制媒体的连续同步播放, 通常有两种方法, 一是以媒体单元为基础, 通过跳过或重复某个媒体单元来保持同步; 另一种是通过改变每个媒体单元的持续时间, 对于视频流来说, 就是改变播放帧率, 对于音频流, 则是改变静默期的

长短^[3,4]. 在低速率情况下(如 15 帧/秒), 用户对增加或减少一个媒体单元较为敏感, 而对播放速率的微小变化不敏感, 因此可以通过改变播放速率来避免缓冲区的上溢和下溢所造成的播放长期不连续, 从而改善同步播放系统的性能.

Yuang^[5]提出的播放速率调整方案的依据是缓冲区占用水平. 设定一个门限, 当缓冲区水平高于该门限时, 采用一个固定的最大播放速率; 当缓冲区水平低于该门限时, 采用线性减小的播放速率. 显然门限的选取, 对播放质量有很大的影响. 本文改进了文献[5], 设定两个门限值, 并着重于分析缓冲区动态播放参数的选取对播放质量的影响. 该方法在接收端设置缓冲区, 根据缓冲区的占用情况动态地调整媒体的播放速率. 缓冲区模型如图 1(b)所示, LL 是低门限, HL 是高门限. 当缓冲区的队列长度在[LL, HL]之间时, 采用正常帧速率; 当缓冲区队列长度小于 LL 时, 采用较小的帧速率以防止缓冲区播放空造成的播放停顿; 当缓冲区队列长度大于 HL 时, 采用较大的帧速率以防止缓冲区溢出造成的播放跳跃. 缓冲区大小的选取可参见文献[6], 本文不做讨论. 当选取了合适的缓冲区后, 门限 LL 和 HL 会直接影响播放质量. 本文建立了一个缓冲区的排队模型, 从分析该模型入手, 得出 LL 和 HL 的

选择依据. 文章最后就本文的算法与 Yuang 的算法进行了比较, 结果表明, 本文的算法在同步性能上优于 Yuang 的算法.

2 动态播放系统模型

接收端的动态播放系统模型如图 1 所示. 播放缓冲区的容量为 N , 该缓冲区有两个门限 LL 和 HL , 控制器根据缓冲区的占用情况控制播放速率. 当缓冲区队列长度在 $[LL, HL]$ 之间时, 播放速率为 L_1 ; 当缓冲区队列长度小于 LL 时, 播放速率正比例减小但不能小于最小允许播放速率; 当缓冲区队列长度大于 HL 时, 播放速率正比例增加, 但也不能大于最大允许播放速率.

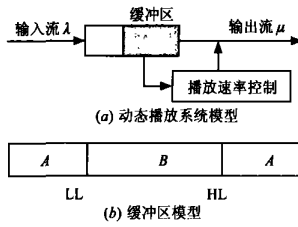


图 1 动态播放系统模型

2.1.1 确定最大和最小允许播放速率

最大允许播放速率 L_{max} 和最小允许播放速率 L_{min} 的选取主要是根据时延抖动要求确定. 例如, 对于 TV 品质的视频, 要求时延抖动不超过 10ms, 那么它的最大允许播放速率和最小允许播放速率就按如下方法确定. 如果正常播放速率为 20 帧/秒 (即每帧持续 $1/20 = 0.050s = 50ms$), 考虑时延抖动不超过 10ms 情况下, 每帧持续时间最短 40ms, 最长 60ms, 因此最大播放速率 $L_{max} = 1/0.040 = 25$ 帧/秒 (每帧持续 40ms), 最小播放速率 $L_{min} = 1/0.060 = 1.67$ 帧/秒 (每帧持续 60ms).

2.1.2 确定播放速率

根据缓冲区的当前占用情况决定合适的播放速率. 假设缓冲区中有 i 个数据帧 (不包括待播放的帧), 则播放速率由下式确定.

$$L(i) = \begin{cases} L_{min} + \frac{L_1 - L_{min}}{LL} \cdot i, & i < LL \\ L_1, & LL \leq i \leq HL \\ L_1 + \frac{L_{max} - L_1}{N - HL} \cdot (i - HL), & i > HL \end{cases} \quad (1)$$

Yuang 采用的播放速率如下^[5]:

$$L(i) = \begin{cases} \frac{\max(i, 1) \cdot L}{TH}, & i < TH \\ L, & i \geq TH \end{cases} \quad (2)$$

其中 L 为最大的播放速率, TH 为门限值.

3 排队模型分析

马尔可夫调制的泊松过程 (MMPP) 广泛用于研究和表示具有突发特性的业务模型. 视频流的到达过程可模型化为一个两状态 (突发状态和静默状态) 的 MMPP, 也称为开关泊松过程 IPP (Interrupted Poisson Process), 如图 2 所示. 图中状态 on 和 off 分别表示突发状态和静默状态, 由状态 on 转移到状态 off 和由状态 off 转移到状态 on 的概率分别为 A 和 B , 另外当系统处于状态 on 时, 数据的到达率为 K_{on} , 状态为 off 时, 没有数据到达, 即 $K_{off} = 0$. 则该模型的状态转移概率矩阵可表示为

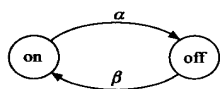


图 2 输入流的两状态模型

$$Q = \begin{bmatrix} Q_{on, on} & Q_{on, off} \\ Q_{off, on} & Q_{off, off} \end{bmatrix} = \begin{bmatrix} 1 - A & A \\ B & 1 - B \end{bmatrix}$$

其稳态概率分布 $x = [x_{on} \ x_{off}]$ 可由 $xQ = x$ 解得, 如下所示:

$$x = \begin{bmatrix} \frac{B}{A+B} & \frac{A}{A+B} \end{bmatrix} \quad (3)$$

考虑一个时间离散的服务系统, 时间轴被分成很多相等的间隔, 称为时隙. 整数时间 $\{0, 1, 2, \dots\}$ 为时隙的边界, 第 n 个时隙是指 $[n-1, n]$ 的时间间隔. 另外, 假设顾客的成批到达发生在一个时隙的开始, 顾客的离去发生在一个时隙的结束, 服务时间为时隙的整数倍.

用 X_n 表示在第 n 个时隙期间到达的顾客数, Y_n 表示第 n 个时隙数据源所处的状态, 定义

$$A_k = \Pr\{X_n = k \mid Y_n = i\}, \quad i \in \{on, off\}, k \in \mathbb{N}$$

$$\text{则有 } A_k = \begin{cases} 0, & k > 1 \\ 1 - K, & k = 0 \\ K, & k = 1 \end{cases}$$

为了求出稳态时的队长分布, 必须确定状态转移矩阵中的 p_{ij} .

服务时间为常数分布但依赖于队长. 假设当缓冲区中有 i 个帧时, 每帧的服务时间用 m_i 表示, 单位为时隙数. 则有

$$m_i = 71 / L(i) \delta$$

假设用 $M(n)$ 表示在 $(0, n]$ 期间到达的顾客数, 用 $a^{rs}(k, n)$ 表示从状态 r (时隙为 t) 转移到状态 s (时隙为 $t+n$) 期间, 即 $[t, t+n-1]$ 期间有 k 帧到达的概率. 即

$$a^{rs}(k, n) = \Pr\{M(n) = k, Y_n = j \mid Y_1 = i\}$$

因为在一个时隙内最多有一帧数据到达, 即 $k > 1$ 时, 有 $a^{rs}(k, 1) = 0$. 而且在一个时隙内不可能发生不同状态间的转移, 即当 $r \neq s$ 时, 有 $a^{rs}(k, 1) = 0$. 因此有:

$$a^{rs}(k, 2) = \begin{cases} 0, & r \neq s \text{ 或 } k > 1 \\ 1 - K, & r = s \text{ 且 } k = 0 \\ K, & r = s \text{ 且 } k = 1 \end{cases}$$

由 C2K 方程, 可得:

$$a^{rs}(k, n) = \sum_{i=0}^{k-1} a_{ri}(q_r a^{rs}(k-i, n-1) + q_r^* a^{rs}(k-i, n-1)), \quad r, s \in \{on, off\}$$

$$r^* = \begin{cases} on, & r = off \\ off, & r = on \end{cases}$$

为求得 p_{ij} , 考虑下面两种情况.

情况 1 当第 n 帧服务完毕离开缓冲区瞬间, 缓冲区中有 i 帧数据 (不包括第 n 帧本身).

第 n 帧 (位于状态 r) 播放完毕离开缓冲区瞬间, 开始播放第 $n+1$ 帧 (位于状态 s), 在第 $n+1$ 帧播放持续期间, 即 m_i 个时隙内, 继续有数据帧到达. 如果缓冲区未满, 则进入缓冲区; 如果缓冲区已满, 则丢弃该数据帧. 又设在第 $n+1$ 帧播放完毕离开缓冲区瞬间, 缓冲区中有 j 帧数据 (不包括第 $n+1$ 帧本身). 于是当 $i > 0$ 时, 有

$$p_{ij} = \begin{cases} \sum_{r \in \{on, off\}} \sum_{s \in \{on, off\}} P_r^{\#} E a^{rs}(j-i+1, m_i), & j < N \\ \sum_{k=1}^j \sum_{r \in \{on, off\}} \sum_{s \in \{on, off\}} P_r^{\#} E a^{rs}(j-i+k, m_i), & j = N \end{cases} \quad (4)$$

情况 2 当第 n 帧服务完毕离开缓冲区瞬间, 缓冲区中有 0 帧数据(不包括第 n 帧本身).

当第 n+1 帧数据播放完毕离开缓冲区瞬间, 缓冲区中的数据帧数目就等于在第 n+1 帧数据播放期间到达的数据帧数目. 因此, 当 i=0 时, 有

$$P_{ij} = \begin{cases} \prod_{r=1}^i P_r \prod_{s=1}^j a^{rs}(j, m_i), & j < N \\ \prod_{k=N-r+1}^i P_r \prod_{s=1}^j a^{rs}(k, m_i), & j = N \end{cases} \quad (5)$$

由式(4)和(5)可以得到 P, 通过解稳态方程 PP=P, 可以得到队长的平稳分布 P=[P₀P₁P₂, P_N].

与求 p_{ij} 一样, 在讨论帧丢失概率时, 也分两种情况: 缓冲区中有 i 帧数据情况下的丢失率(用 Lost_i 表示)和缓冲区为空时的丢失率(用 Lost₀ 表示). 则有

$$Lost_i = \prod_{k=N-r+1}^{m_i} P_r \prod_{s=1}^{m_i-1} a^{rs}(k, m_i)$$

$$Lost_0 = \prod_{k=N+1}^{m_1} q_{\alpha, \#} \prod_{s=1}^{m_1-1} a^{rs}(k, m_1-1)$$

综上, 可得帧丢失概率 P_{Lost}

$$P_{Lost} = \sum_{i=0}^N P_i List_i \quad (6)$$

缓冲区下溢的概率

$$P_{empty} = P_0 \quad (7)$$

4 门限选取

由上可以看出缓冲区大小的稳态分布是参数 LL 和 HL 的函数. 最合适 LL 及 HL 的选择是受缓冲区为空的概率、帧丢失概率(即缓冲区上溢的概率)和平均服务速率的约束. 为了得到上述三个变量与参数 LL 和 HL 的约束关系, 我们给出以下实验条件: 设帧的正常播放速率为 L=20 帧/秒, 为了保证媒体内同步的要求(相邻帧间的抖动小于 10ms), 选取最大播放速率为 L_{max}=25 帧/秒, 最小播放速率为 L_{min}=16.67 帧/秒. 此外, 为了方便, 选取 HL=N-LL, 这样就变为如何选择合适的 LL.

实验结果如图 3 至 5 所示. 需要说明的是图中业务强度指的是到达速率与播放速率的比值, 门限指的是实际门限与缓冲区容量 N 的比值, 平均播放速率比是平均播放速率与正常播放速率的比值. 当业务强度 pF 1 时, 系统有稳态解. Q=K/L, K 是平均到达率, K=(K_{off}B)/(A+B), 是 IPP 过程中处于 ON 状态下的平均到达率.

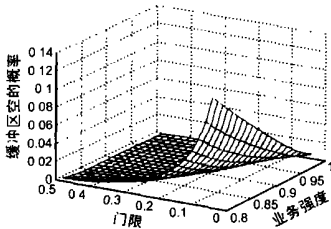


图 3 门限、业务强度与缓冲区为空的概率的关系

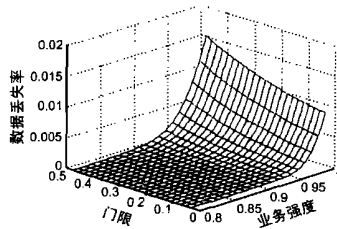


图 4 门限、业务强度与数据丢失率的关系

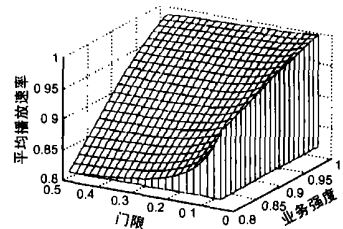


图 5 门限、业务强度与平均播放速率的关系

该模型采用了离散参数随机过程的分析方法, 用时隙的概念将连续时间参数离散化. 在本实验中, 网络接入速率为 71.2Mb/s, 帧大小为 15Kbytes, 也就是说每个时隙 TS=15*8*10³/(71.2*10⁶)=16.67ms, 这与 Yuang 的实验条件相同.

图 3 所示为业务强度、门限 LL 与缓冲区为空的概率之间的关系. 图 4 所示为业务强度、门限 LL 与数据丢失率之间的关系. 图 5 所示为业务强度、门限 LL 与平均播放速率之间的关系.

由图 3 可以看出, 门限 LL 越小, 缓冲区为空的概率越大, 为防止由于缓冲区播放空造成的播放停顿, 门限 LL 的选取应尽量大.

由图 4 可以看出, 门限 LL 越小, 数据丢失率越小. 为了防止缓冲区溢出造成的播放跳跃, 门限 LL 的选取应尽量小. 由图 5 可以看出, 随着门限的增加播放速率在减小, 考虑媒体内同步要求, 应使平均播放速率接近正常播放速率, 因此, 门限 LL 的选择应尽量小.

由上可以看出, 选择门限时, 应综合考虑缓冲区为空的概率、数据丢失率和平均播放速率这三方面的因素. 由图 3 可以看出, 当要求缓冲区为空的概率小于 10⁻³ 时, 门限应大于 0.30N. 由图 4, 数据丢失率小于 10⁻⁴ 时, 门限应小于 0.39N. 此外考虑平均播放速率, 由图 5 可以看出, 当门限在 0.30N 和

0.39N 之间时, 平均播放速率随门限增加而偏离正常播放速率, 但变化很小(从 19.34 帧/秒变为 19.36 帧/秒), 因此, 门限可以选择在 0.30N 和 0.39N 之间.

从以上的实验和分析可以看出, 门限 LL 直接影响着播放质量, 门限 LL 的选取应按照以下原则进行, 在一定的业务强度下, 首先考虑数据丢失率和缓冲区为空的概率这两个重要参数, 保证选择的门限满足这两个设计要求, 此外还需综合考虑门限对平均播放速率的影响, 从而得到满意的播放质量.

5 同步性能

为比较算法的同步性能, 这里采用了不连续性和播放失真的二阶特性, 即用它们的方差作为同步性能的客观指标.

5.1 不连续性

假设当第 n 个帧播放时, 缓冲区中有 i 帧, 第 n 帧的播放持续时间与正常帧的播放持续时间的差定义为第 n 个帧的播放不连续性, 用 d_{i, n} 表示. d(i) 定义为当缓冲区中有 i 帧时播放的不连续性. VoD(Variance of Discontinuity) 是不连续性的方差, 用来作为媒体流连续性的量度, VoD 反映了缓冲区下溢以及人为引入的播放间隙对播放质量的破坏作用.

$$d(i) = \begin{cases} 1/L(i) - T, & i > 0 \\ 1/L(0) + S - T, & i = 0 \end{cases} \quad (8)$$

式中: 本算法和 Yuang 算法的 $L(i)$ 分别由式(1)、(2)确定; T 是正常帧的播放持续时间, S 是一个随机变量, 表示从缓冲区下溢到第一个新数据到达的时间间隔, 假设 S 服从均值为 $1/K$ 的指数分布。

5.1.2 播放失真

Laoutaris^[7] 引入了一个紧凑公平的连续性测度 DoP (Distortion of Playout, 播放失真), 该标准综合了所有引起媒体异步的因素: 缓冲区下溢产生的播放间隙, 播放速度减慢引起的间隙以及缓冲区上溢产生的播放间隙。定义如下

$$\text{DoP}(i) = A \# d(i) + B \# l_i \# T, \quad A = B = 1 \quad (9)$$

假设最近一帧播放时, 缓冲区中有 i 帧, l_i 是下一帧播放期间丢失帧的平均值。A 和 B 分别是抖动引起的间隙和缓冲区溢出引起的丢失的加权系数, 抖动的存在和包的丢失对于可感知质量有同等程度的破坏作用^[8], 因此选择 $A = B = 1$ 。播放失真 (DoP) 的方差用 VDoP (Variance of distortion of playout) 表示。

定义 l_i^n 为第 n 个播放帧离开系统时缓冲区中有 i 帧情况下, 第 $n+1$ 帧播放期间丢失 k 帧的概率。 l_i^n 定义为当 n 趋近于无穷时 l_i^n 的稳态概率, l_i 是 l_i^n 的平均值。

需要说明的是本实验中时间是以时隙为单位的 (本实验中一个时隙相当于 16167ms), 因此 VoD 和 VDoP 的单位为时隙的平方。

图 6 是业务强度为 0.98 时, 采用 Yuang 算法和本文算法得到的门限与 VoD 的关系。可以看出本文算法得到的 VoD 总体上小于 Yuang 算法得到的结果, 在同样业务强度下, 采用本文算法得到的 VoD 的最小值较小, 因此在连续性上, 本文优于 Yuang 算法。

再考察两种算法得到的 VDoP。图 7 是业务强度下为 0.98 时, 采用 Yuang 算法和本文算法得到的门限与 VDoP 的关系。可以看出本文算法得到的 VDoP 小于 Yuang 算法得到的结果, 在同样业务强度下, 采用本文算法得到的 VDoP 的最小值明显小于 Yuang 算法。因此在播放失真方面, 本文也优于 Yuang 算法。

可以看出, 采用本文的算法得到的同步性能优于 Yuang 算法, 尤其是在播放失真方面明显优于 Yuang 算法。

6 小结

本文提出的是一种视频媒体流的动态播放算法。该方法是以缓冲区占用为基础, 根据缓冲区的占用情况动态地调整媒体的播放帧速率。根据媒体内同步的要求, 给出动态播放帧速率的选取范围。另外, 分析排队模型给出缓冲区门限的选取

原则。通过选择合适的参数, 该算法可以有效地防止缓冲区下溢造成的播放停顿以及缓冲区上溢造成的播放跳跃, 从而实现同步平滑播放。采用不连续性和播放失真的方差作为衡量同步的标准, 结果表明, 本文算法性能优于 Yuang 的算法。

为了简化问题, 本文在缓冲区门限的选取上假设 $HL = N - LL$, 也即只有一个门限变量需要确定, 而两个门限变量都需要确定的情况是一个需要进一步研究的问题。另外, 如何在线地选择最优门限也是一个值得研究的问题, 而研究此问题的一个关键是需要找到能够反映主观同步性能的评价指标。

参考文献:

- [1] Gerold Blakowski, Ralf Steinmetz. A media synchronization survey Reference model, specification and case studies[J]. IEEE Journal on Selected Areas in Communications, 1996, 14(1): 5- 35.
- [2] 许延, 常义林, 刘增基. 存储媒体的媒体内同步反馈控制算法[J]. 电子学报, 2002, 30(9): 1315- 1319.
- [3] Sue B Moon, Jim Kurose, Don Towsley. Packet audio playout delay adjustment performance bounds and algorithms[J]. ACM/ Springer Multimedia Systems, Jan, 1998, 6(1): 17- 28.
- [4] Ramachandran Ramjee, Jim Kurose, Don Towsley, Henning Schulzrinne. Adaptive playout mechanisms for packetized audio applications in wide area networks[A]. Proceedings of the Conference on Computer Communications (IEEE Infocom) [C]. Toronto, Canada: IEEE Computer Society Press, Los Alamitos, California, June, 1994. 680- 688.
- [5] Maria C Yuang, Shih T Liang, Yu G Chen. Dynamic video playout smoothing method for multimedia applications[J]. Multimedia Tools and Applications, 1998, 6(1): 47- 60.
- [6] Yan Xu, Yilin Chang, Zengji Liu. Calculation and analysis of competition buffer size in multimedia systems[J]. IEEE Communications Letters, 2001, 5(8): 355- 357.
- [7] Nikolaos Laoutaris, Ioannis Stavrakakis. Adaptive playout strategies for packet video receivers with finite buffer capacity[A]. Proceedings of the IEEE International Conference on Communications (IEEE ICC) [C]. Helsinki, Finland, June, 2001. 969- 973.
- [8] Mark Claypool, Jonathan Tanner. The effects of jitter on the perceptual quality of video[A]. ACM Multimedia Conference. 99[C]. Orlando, FL, USA, 1999. 115- 118.

作者简介:



许 延 女, 1971 年 3 月生于西安, 1992 年毕业于西安电子科技大学通信工程专业, 1999 年于该校获通信与信息系统学科工学硕士学位, 2002 年获通信与信息系统学科工学博士学位。现在在中国网通集团研究院工作, 主要研究方向为多媒体通信和网络管理。Email: hsyuan@163.com

常义林 男, 1944 年 12 月生于江苏省江都市, 西安电子科技大学教授, 博士生导师。主要研究方向为多媒体通信和网络管理。多次获得省部级科技进步奖励, 发表学术论文 100 多篇。