

一类 MPEG-4 FGS 流视频传输的质量定义及算法

黄天云

(西南民族大学计算机科学与技术学院 成都 610041)

摘 要 基于 PSNR(MSE)指出了 MPEG-4 FGS 流视频客观质量和质量变化指标的相关定义,讨论了视频分段场景基础上的质量最优化问题,通过状态迁移图和动态规划给出了 FGS 流视频的最优传输策略,并且用试验数据对 FGS 流视频的优化传输的控制粒度及实时传输算法进行了对比分析。

关键词 MPEG-4 FGS;质量定义;传输策略;视频分段;场景;状态迁移
中图法分类号 TP391

The Quality Definition and an Optimal Algorithm for MPEG-4 FGS Video Streaming

HUANG Tian-Yun

(College of Computer Science and Technology, Southwest University for Nationality, Chengdu 610041)

Abstract Since MPEG-4 FGS(Fine Granularity Scalability) coding has been dedicated to IP video streaming, a strict definition of its objective quality and quality variation is required, in order to provide efficient transmission of FGS-coded video and an attainable subjective quality perception for end users. Based on MSE(Mean Square Error) and PSNR(Peak Signal-to-Noise Ratio), an objective quality and quality variation definition for FGS-coded video is introduced in this paper, and a quality optimal problem of scene in video segments is formulated and solved using the state transfer graph and dynamic programming. The optimal transmission policy is also obtained and compared with a real-time transmission algorithm. Different aggregation levels(segmentation granularity)of video segments for optimal transmission are also examined by experimental data. Simulation results validate the proposition.

Keywords MPEG-4 FGS; quality definition; transmission policy; video segment; scene; state transfer graph

1 引 言

与传统的可扩展视频编码不同,MPEG-4 FGS(Fine Granularity Scalability)对 DCT 系数采取的位平面(bit-plane)编码^[1,2]可以根据网络带宽变化而动态地截断和解码,因此,近年来在 IP 上的流视频传输方法的研究中受到极大关注^[3~8]。人们对 FGS 编码视频在 IP 上传输的通用框架^[3,4]、基层和可扩展层算法的改进^[5]以及在组播及无线环境的传

输^[6,7]等均已进行了研究,对 MPEG-4 编码视频的流程特征也进行了分析^[8]。

然而,这些研究都未对 MPEG-4 FGS 编码视频的传输质量(quality)和质量变化(quality variation)及其关系进行定量分析。MPEG-4 FGS 流视频的终端表示质量既受位流速率的影响,也与连续帧间的速率变化有关。按文献^[9]的定义,视频传输质量的最大化即指保证在特定网络带宽下单帧图像有较好质量的同时,最小化连续帧间质量的急剧变化(一幅高质量的图像紧接着一幅低质量的图像,

对终端用户而言,其感受的主观质量(subjective quality)远远低于连续两幅中等质量的图像)。

尽管已有人对 MPEG-4 FGS 流量统计特性进行了分析^[10]并建立了相应的质量指标^[11,12],然而还存在以下问题:(1)算法假设 MPEG-4 FGS 视频基层 BL 和增强层 EL 都是 CBR 的.实际 FGS 基层编码一般是 VBR 的,而增强层通常是 CBR 的,以保证网络传输中最小视频的质量,同时根据网络带宽变化最大可能地提升增强层的传输率,从而提升 FGS 视频整体质量;(2)基于分段粒度的客户端缓冲回放策略将导致客户端极大的缓冲需求;(3)建立在分段的速率失真(rate distortion)理论基础上的优化控制策略,并不能保证分段的质量最优化.因为对多达几十甚至上百场景的一个分段,其中包含的视频内容和运动变化是多样的,视频质量可能会产生极大的波动.在对一个分段分配固定增强层速率的情况下,无法通过速率控制使得视频整体质量最好,同时质量变化(速率失真)最小;(4)即使在更细的控制粒度(如场景)上实施优化策略,假定控制策略采用固定间隔的时隙也是不合理的;(5)对最优速率和网络传输中实际的可达速率没有进行明确区分。

本文给出 MPEG-4 FGS 编码视频传输的客观质量(objective quality)和质量变化指标的严格定

义,并建立基于该质量指标的优化传输策略和算法。

本文第 2 节对 MPEG-4 FGS 的位平面编码技术进行简述;第 3 节给出相关客观质量指标的定义,并提出无丢失条件下的分段质量指标的最优化问题;该问题在第 4 节通过状态迁移图和动态规划得到解决,同时获得了在此优化问题下的分段传输策略和服务器最优调度策略;对可达速率算法和最优速率算法将进行对比;第 5 节讨论不同分段聚合(控制)粒度的选择,通过试验数据对相关质量指标和算法进行对比分析。

2 MPEG-4 FGS 位平面编码

位平面编码技术是 MPEG-4 FGS 与其它可扩展编码最大的不同之处^[1,2].对 8×8 的 DCT 系数(图 1),首先按 Z 形扫描构成 64 个连续整数,每个非 0 整数按二进制编码,如 9 编码成 1001,3 编码成 0011.64 个整数二进制编码后的值按列排布,而最大的整数 9 需要 4bits 编码,故构成 4×64 的矩阵(应用中允许多达 8bits 的深度).由此构成的每行称为一个位平面(bit-plane).其中第一个位平面是 MSB(Most Significant Bit-plane),其它依次进行,最后一个位平面是 LSB(Least Significant Bit-plane).

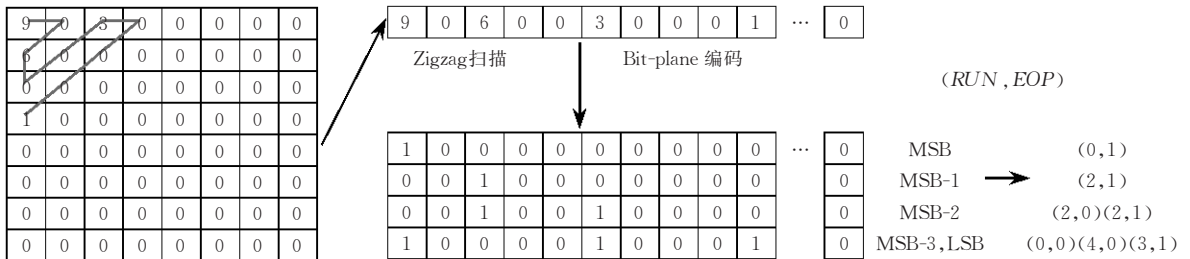


图 1 MPEG-4 FGS 位平面(bit-plane)编码过程示意图

对每个位平面按 (RUN, EOP) 方式编码,其中 RUN 表示 1 之前连续 0 的个数.如果位平面中还有 1 未编码, EOP 置 0;否则 EOP 置 1.最后 (RUN, EOP) 标记按 VLC 方式编码。

可见,按 MPEG-4 FGS 的位平面编码方式,增强层位流可以在传输过程中根据网络状况任意截断,而剩余部分仍然能够解码.因此,特别适合在网络带宽动态变化的环境(如 IP)中传输. MPEG-4 FGS 已被建议为 IP 网络上的流视频传输方式^[13]。

3 定义及定理

视频的主观感受质量因人而异,无法统一衡量。

参照 VQEG^[14],定义视频的客观质量指标为 PSNR。

定义 1. 对大小为 $X \times Y$ 像素的视频帧,视频序列 $l_1 \sim l_2$ 的 PSNR 定义为

$$PSNR(l_1, l_2) = 10 \log \frac{D^2}{MSE(l_1, l_2)},$$

其中, D 是像素的位深度(如 8 位灰度图像时, $D = 255$). MSE 定义为

$$MSE(l_1, l_2) = \frac{1}{X \cdot Y(l_2 - l_1 + 1)} \sum_{l=l_1}^{l_2} \sum_{y=1}^Y \sum_{x=1}^X [I(x, y, l) - \bar{I}(x, y, l)]^2,$$

其中, $I(x, y, l)$ 和 $\bar{I}(x, y, l)$ 分别是原始像素和解码后像素的灰度值.鉴于人类视觉系统 HVS 对亮度成份(luminance)更为敏感,在 PSNR 的计算中可忽

略色度成份(chrominance)。

假设 1. 假设连续视频序列共有 L 帧,且划分为 M 个分段(分段个数取决于具体的视频分段算法),其中每个分段包含的帧数目为 κ_m ,则显然有

$$\sum_{m=1}^M \kappa_m = L.$$

假设 2. 假设分段 m 包含 N 个场景,其中每个场景包含的帧数目为 $N_n (n=1,2,\dots,N)$,场景持续时间为 $C_k (k=1,2,\dots,N)$ (当前分段时,忽略下标 m),则显然有 $\sum_{n=1}^N N_n = \kappa_m$ 。

现定义图像帧、场景及分段的质量指标。

定义 2. 定义视频序列中图像帧 l 的质量指标为 $Q_l(r)$,基层质量指标为 Q_l^b ,增强层在编码速率为 r 时的质量指标为 $Q_l^e(r)$,则有

$$(1) Q_l(r) = PSNR(L, l);$$

$$(2) Q_l(r) = Q_l^b + Q_l^e(r);$$

$$(3) \bar{Q}(r) = \left(\sum_{l=1}^L Q_l(r) \right) / L;$$

$$(4) \sigma_Q^2(r) = \frac{1}{L-1} \sum_{l=1}^L [Q_l(r) - \bar{Q}(r)]^2 \\ = \frac{1}{L-1} \left\{ \sum_{l=1}^L [Q_l(r)]^2 - [\bar{Q}(r)]^2 \right\};$$

$$(5) Cov_Q(r) = \sigma_Q(r) / \bar{Q}(r);$$

$$(6) \rho_Q(r, k) =$$

$$\frac{1}{L-k} \sum_{l=1}^{L-k} \frac{[Q_l(r) - \bar{Q}(r)] \cdot [Q_{l+k}(r) - \bar{Q}(r)]}{\sigma_Q^2(r)}.$$

定义 3. 定义分段 $m (m=1,2,\dots,M)$ 中图像帧 $n (n=1,2,\dots,\kappa_m)$ 的质量指标为 $Q_{m,n}(r)$,并且整个分段的质量指标为 $\Theta_m(r)$,则定义 2 的相关结果同样适用. 并且有

$$(1) Q_{m,n}(r) = PSNR \left(\sum_{j=1}^{m-1} \kappa_j + n, \sum_{j=1}^{m-1} \kappa_j + n \right);$$

$$(2) \Theta_m(r) = PSNR \left(\sum_{j=1}^{m-1} \kappa_j + 1, \sum_{j=1}^m \kappa_j \right).$$

基于上述定义,可以通过以下定义有效地衡量图像帧和序列分段的质量变化等级。

定义 4. 定义分段 m 的图像帧的质量变化指标为 $Var_m(r)$,整个视频序列中分段的质量变化指标为 $V_m(r)$,则有

$$(1) Var_m(r) = \max_{i=2,\dots,\kappa_m} \{ |Q_{m,i}(r) - Q_{m,i-1}(r)| \};$$

$$(2) Var_{\min}(r) = \min_{1 \leq m \leq M} \{ Var_m(r) \};$$

$$(3) V_m(r) = |\Theta_{m+1}(r) - \Theta_m(r)| (1 \leq m \leq M-1);$$

$$(4) V_M(r) = \max_{1 \leq m \leq M-1} \{ V_m(r) \};$$

$$(5) V_{\min}(r) = \min_{1 \leq m \leq M-1} \{ V_m(r) \}.$$

并且其均值、方差、相关系数等也可以参照定义 2 进行定义。

定义 5. 定义分段 m 场景 $n (n=1,2,\dots,N)$ 中图像帧 i 的质量指标为 $\Phi_{n,i}(r)$,场景平均质量指标为 $\bar{\Phi}_n(r)$ (此处忽略分段下标 m),则有

$$(1) \Phi_{n,i}(r) =$$

$$PSNR \left(\sum_{j=1}^{m-1} \kappa_j + \sum_{j=1}^{n-1} N_j + i, \sum_{j=1}^{m-1} \kappa_j + \sum_{j=1}^{n-1} N_j + i \right);$$

$$(2) \bar{\Phi}_n(r) =$$

$$PSNR \left(\sum_{j=1}^{m-1} \kappa_j + \sum_{j=1}^{n-1} N_j + 1, \sum_{j=1}^{m-1} \kappa_j + \sum_{j=1}^n N_j \right) / N_n.$$

定义 1~定义 5 的相关指标均可用于视频序列的客观质量分析和评价。

定义 6. 定义分段粒度为 FGS 视频序列调度算法中的最小调度单元(一般情况下,分段粒度可以任意选择,如单帧、GoV、场景、固定长度或任意分段等. 试验中将做对比分析)。

按视频序列分解的层次结构,在逻辑上对视频序列做如下划分:每个视频序列可以按剪切的的不同划分为多个分段/镜头;而每个分段则包含不同的场景,这些场景则由图像帧按一定的 GoV 模式(GoV pattern)构成. 即存在如下关系:

$$\text{Sequence} \rightarrow \text{Segment} \rightarrow \text{Scene} \rightarrow \text{GoV} \rightarrow \text{Frame}.$$

这种方式更符合视频序列的统计特性^[15],也更有利于实施相应的实时调度算法。

以下分析特针对当前分段 m 进行。

假设 3. 假设视频序列按 FGS 方式进行编码. 其中基层编码是 VBR 的;而增强层是 CBR 的,其最大编码速率为 R_{\max} . 如基层最大速率 384Kbps;增强层速率 400~2000Kbps,间隔为 200Kbps^[10](我们仅考虑 SNR FGS 编码,此时运动估计和运动补偿仅在基层进行,增强层则包含原始图像与解码后的基层图像的量化误差)。

定义 7. 若 $r_e(n)$ 是场景 n 的增强层传输速率(按 MPEG-4 FGS 编码,假设基层的传输无丢失,则图像质量的提升就只与增强层传输的比特数有关),则可定义分段传输策略 \mathfrak{R} 如下:

$$\mathfrak{R} = \{r_e(1), r_e(2), \dots, r_e(N)\} \quad (1)$$

在此策略下,该分段增强层需要传输的总字节数为

$$B(\mathfrak{R}) = N_1 \cdot r_e(1) + N_2 \cdot r_e(2) + \dots + N_N \cdot r_e(N) \quad (2)$$

由此,对给定分段 m 的场景 n ,假设 r_{emax} 是该场

景所有图像帧增强层允许的最大传输率(由增强层最大编码率 R_{\max} 决定),该分段的增强层编码带宽限制为

$$B_{\max} = R_{\max} \cdot \sum_{k=1}^N C_k.$$

则分段最优传输策略 $\mathfrak{R}^* = \{r_e^*(1), r_e^*(2), \dots, r_e^*(N)\}$ 的求解转化为如下问题.

问题 1.

$$\begin{aligned} \Omega_m^* &= \text{minimize } \Omega(\mathfrak{R}) \\ &= \sum_{n=1}^N \left[\frac{1}{N_n} \sqrt{\sum_{i=1}^{N_n} |\Phi_{n,i}(r_e(n)) - \bar{\Phi}_n(r_e(n))|^2} \right] \\ \text{s. t. } &\begin{cases} B(\mathfrak{R}) = N_1 \cdot r_e(1) + N_2 \cdot r_e(2) + \dots + \\ \quad N_N \cdot r_e(N) \leq B_{\max}, \\ r_e(n) \leq r_{e\max}, n=1, 2, \dots, N \end{cases} \end{aligned} \quad (3)$$

分段最优传输策略的求解就是要在带宽限制条件下,使得分段中任意场景与原始场景有最为接近的平均质量指标,并且使该分段整体的质量变化指标(定义 4(1))最小化(任意场景中的图像帧质量指标 $\Phi_{n,i}(r_e(n))$,可以通过定义 5(1)计算,也可以根据增强层编码率 $r_e(n)$ 在统计意义上近似).

算法 1. MPEG-4 FGS 流视频的优化传输算法.

1. 服务器按层次结构对视频序列进行分解(定义 6),将视频序列的每个分段基于场景划分为不同的时隙(time slots) $[t_k, t_{k+1}) (k=0, 1, \dots, N-1)$. 最优传输策略将在分段基础上实施;

2. 在每个时隙 k , 服务器将根据当前的有效带宽 $X(t) (t_k \leq t < t_{k+1})$ 按平均速率 $r_s(k)$ 传输视频流,此时基层带宽为 $X_b(t) = (\bar{r}_b(k)/r_s(k)) \cdot X(t)$, 增强层带宽为 $X_e(t) = (r_e(k)/r_s(k)) \cdot X(t) = (\xi_{\max}/r_s(k)) \cdot X(t), \xi \in [0, 1]$;

3. 假设视频流在开始回放时均采用客户端缓冲^[3](client buffer, 称启动延迟,通常为 3~5s),记为 Δ_0 ; 同样其它时隙的回放延迟记为 $\Delta_k (k=1, 2, \dots, N-1)$. 由此,服务器将在当前时隙 k 客户端播放缓冲内容的同时,确定按多大的速率 $r_s(k+1)$ 传输时隙 $k+1$ 的视频数据.

考虑到基层数据的丢失将极大地影响视频的终端表示质量,需要分析基层数据的无丢失条件. 定义时隙 $[t_k, t_{k+1})$ 的最小传输率:

$$\beta_k(\Delta) = \min_{t \in (t_k + \Delta, t_{k+1})} \left(\int_{t_k}^t X(u) du \right) / (t - (t_k + \Delta)),$$

有如下结论.

定理 1. 给定时间 k 的长度 C_k 、服务器的平均传输速率 $r_s(k)$, 则在一个分段传输过程中,基层数据无丢失的充要条件是:对 $\forall k \in \{1, 2, \dots, N-1\}$, 当 $\Delta_k < C_k$ 时,满足 $r_s(k) \leq \beta_k(\Delta_k)$ (该定理给出了基层数据无丢失条件下,服务器传输率的一个上限. 它

取决于当前时隙的回放延迟 Δ_k 和可用的网络带宽 $X(k)$).

证明. 如果在每个时隙 $[t_k, t_{k+1})$ 中基层数据无丢失,则整个分段基层数据无丢失. 当 $\Delta_k \geq C_k$ 时,客户端缓冲有足够的的数据,可以保证在时隙 $[t_k, t_{k+1})$ 间完整回放,因此无丢失. 当 $\Delta_k < C_k$ 时,在时隙 $[t_k, t_k + \Delta_k)$ 可以保证完整回放;而在 $[t_k + \Delta_k, t_{k+1})$,客户端回放数据必须直接从服务器传输. 即在时隙 $[t_k + \Delta_k, t)$ 回放的数据量必须不大于服务器在时隙 $[t_k, t)$ 传输的数据量,因此对 $\forall t \in [t_k + \Delta_k, t_{k+1})$ 有

$$r_s(k)(t - (t_k + \Delta_k)) \leq \int_{t_k}^t X(u) du. \quad \text{证毕.}$$

4 问题求解及算法

问题 1 的求解可利用状态迁移图(state graph),通过动态规划实现. 同时,状态迁移过程中的每个时隙的回放延迟 $\Delta_k (k=1, 2, \dots, N-1)$ 也可由此确定. 值得注意:问题 1 的分段质量变化指标的最小化条件是分段中每个场景速率都与原始场景有尽可能小的偏差,而这必须以每个场景中的所有帧都有尽可能相同的质量指标为前提,这样才可以保证连续帧间的质量变化保持在一定范围,不至于过分突发. 同时,鉴于可获得服务器原始帧序列及采用最优传输策略后的视频质量,因此在无丢失条件下,最优传输策略的质量指标是可测的(measurable).

定义 8. 状态迁移图如图 2 定义,其中包括当前分段在每个时隙的开始时刻 t_k , 回放延迟 Δ_k 的所有可能取值. t_k 和 C_k 由聚合粒度(如场景或 GoV)决定. t_k 到 t_{k+1} 之间的迁移表示客户端缓冲的回放,以保证视频数据在传输过程中无丢失. 初始状态 Δ_0 表示客户端启动延迟. 终止状态 Δ_{N-1} 表示服务器完成当前分段最后一个时隙 t_{N-1} 的数据传输,进入下一个分段的调度.

图 2 中, $T = \sum_{k=1}^N C_k$ 是当前分段的持续时间. 底部时间数据 t 显示了分段中各场景时间的推进;而图中粗线箭头表示当前分段的一条可能的最短路径. 服务器将按此规则逐个分段调度,直至传输完毕.

定理 2. 使得系统调度中每个时隙传输的图像帧都取得最优的质量指标,同时保证连续帧间的质量变化保持在一定范围内的分段最优传输策略,可以通过在状态迁移图中寻找最短路径求解,从而也可得服务器优化调度策略.

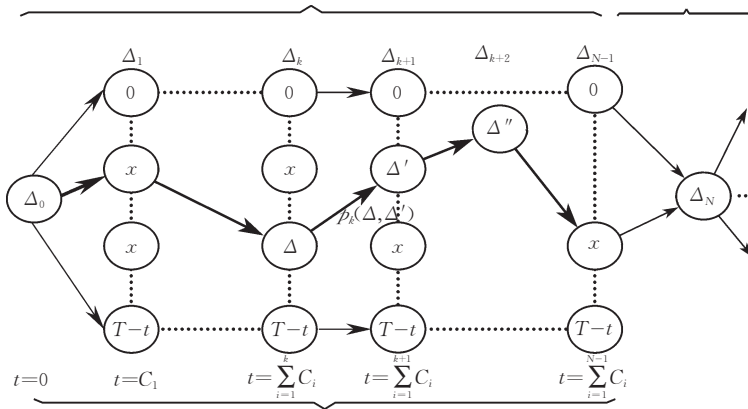


图 2 当前分段最优传输状态迁移图及服务器优化调度策略

证明. 考虑在时隙 $[t_k, t_{k+1})$ 间的视频流量传输, 此时对 $\forall k \in \{0, 1, \dots, N-1\}$, 有下面成立:

$$\Delta_{k+1} = \Delta_k + \left(\int_{t_k}^{t_{k+1}} X(t) dt \right) / r_s(k) - C_k \quad (4)$$

可见, 状态的迁移完全取决于当前时隙网络带宽参数和服务器传输率 $r_s(k)$. 因此, 从初始状态 Δ_0 到该分段最终传输完毕的所有可能路径代表了服务器调度的所有可能的策略.

记从状态 $\Delta_k = \Delta$ 到状态 $\Delta_{k+1} = \Delta'$ 迁移增加的损失(loss)为 $p_k(\Delta, \Delta')$, 则从初始状态到终止状态的最小损失可以通过动态规划和逐步修正求得. 定义 $r_c(k)$ 为使得系统从 $\Delta_k = \Delta$ 迁移到 $\Delta_{k+1} = \Delta'$ 的唯一可能取值; 而 $r_c(k+1)$ 为使得系统从 $\Delta_{k+1} = \Delta'$ 迁移到最短路径中下一状态 Δ'' 的唯一可能取值. 定义 $p_k(\Delta, \Delta')$ 为在传输率 $r_c(k)$ 下当前时隙的增强层平均质量变化指标, 从而问题 1 可得到求解. 而基层速率可测, 并且假设传输过程中无丢失, 因此服务器传输率 $r_s(t)$ ($t \in [t_k, t_{k+1})$) 可得, 并且可确定平均传输率 $r_s(k)$.

可以看出, 最短路径的求解与每个时隙的回放延迟 Δ_k 的取样精度有关. 证毕.

定义 9. 对于任意包含 M 个分段的 MPEG-4 FGS 存储视频序列, 若定义:

(1) 所有分段的最大场景数为

$$N = \max\{N_1, N_2, \dots, N_M\};$$

(2) 每个场景的速率行向量为

$$\mathbf{R}_i = [r_e^i(1) \ r_e^i(2) \ \dots \ r_e^i(N_i) \ 0 \ \dots \ 0]_{N \times 1};$$

(3) 每个场景的缓冲行向量为

$$\Delta_i = [\Delta_i(1) \ \Delta_i(2) \ \dots \ \Delta_i(N_i) \ 0 \ \dots \ 0]_{N \times 1},$$

则可以分别定义其速率迁移矩阵 \mathbf{R} 和缓冲控制矩阵 \mathbf{B} 如下:

$$\mathbf{R}_{M \times N} = [\mathbf{R}_1 \ \mathbf{R}_2 \ \dots \ \mathbf{R}_i \ \dots \ \mathbf{R}_M]';$$

$$\mathbf{B}_{M \times N} = [\Delta_1 \ \Delta_2 \ \dots \ \Delta_i \ \dots \ \Delta_M]';$$

推论 1(视频序列最优调度策略). MPEG-4 FGS 视频序列的速率迁移矩阵 \mathbf{R} 和缓冲控制矩阵 \mathbf{B} 唯一地确定了服务器的最优调度策略. 该策略仅受当前可用网络带宽参数的影响.

研究表明^[16], 当一个视频序列中的连续两帧发生切换(如明显的视频内容或运动的改变)时, 可以此作为视频场景(video scene)的分割点. 由于它是具体的视频分段和场景划分算法^[17,18], 一个场景中通常包含相近的视频内容和平稳运动, 因此有近似的视频质量指标和平滑的质量指标变化. 由定理 1 基层无丢失条件及定理 2 分段最优传输策略, 速率迁移矩阵 \mathbf{R} 在确定每个场景中图像帧的最低质量保证的同时, 按 GoV 模式和 MPEG-4 FGS 编码规则, 还将最大可能地提升 B 帧(其次 P 帧)的质量指标(注意到在 I 帧已有较好质量的情况下, 增强层速率的提高对 I 帧的质量提升并不明显). 而缓冲控制矩阵 \mathbf{B} 则根据当前的网络带宽参数, 确保了客户端回放过程的连续性. 因此, 矩阵 \mathbf{R} 和 \mathbf{B} 唯一地确定了服务器视频序列的最优调度策略.

然而, 注意到由定理 2 和推论 1 得到的最优调度策略是假设网络带宽已知条件下的优化结果(如果网络带宽可通过模型化参数或其它方法较好地估计). 由定理 2 得到的速率迁移矩阵 \mathbf{R} 并不保证实际传输过程中 FGS 基层和增强层的流量不超过可用网络带宽. 然而, 可以假设优化速率(optimal rate) $r_s(k)$ 能够达到(这可以通过状态迁移图求得), 并且回放延迟 Δ_k 可以通过客户端反馈(feedback)等机制得到, 以此来估计可用网络带宽, 反过来调整实时传输率.

定义 10. 可达速率(attainable rate)是指在优化速率假设下, 已知回放延迟 Δ_k 时, 根据当前网络状况实际允许达到的传输速率.

首先, 设在时隙 k 基层的均值速率为 $\bar{r}_b(k)$ (基

层是 VBR 编码的). 由式(4)有

$$\bar{X}(k-1) = r_s(k-1) \cdot (\Delta_k - \Delta_{k-1} + C_{k-1}) / C_k \quad (5)$$

在每个分段传输的初始阶段, 可以认为 $\bar{X}(k-1)$ 是时隙 k 的网络带宽估计(此时, Δ_{k+1} 尚未知). 然而考虑到网络流量的长时记忆和短时依赖, 在有限步后, 采用 AR 模型估计后续各时隙的带宽:

$$\bar{X}(k) = \phi_1 \bar{X}(k-1) + \phi_{l_1} \bar{X}(k-l_1) + \phi_{l_2} \bar{X}(k-l_2) + \phi_{l_3} \bar{X}(k-l_3) + Z(t), \quad Z(t) \sim N(0, \sigma^2) \quad (6)$$

可以通过 Yule-Walker 方法, 以最小化 AIC 值的方式得到式(6)的参数估计.

由此, 得到从优化速率到可达速率的实时调整算法.

算法 2. 可达速率算法.

对视频序列的每个分段, 算法首先由优化速率 $r_s(k-1)$ 和 Δ_{k-1} 估计可用网络带宽; 其次, 根据客户端缓冲延迟 Δ_{k-1} 和当前时隙长度 C_k 的关系, 动态地决定增强层实际可传输速率. 考虑到网络带宽利用率和视频流量的波动, 引入平滑参数 $\alpha \in [0, 1]$. 算法描述如下:

对视频序列的每个分段

初始化系统参数: $k=1; \alpha=0.5; \gamma=0.9$; 从客户端反馈接收缓冲延迟 Δ_{k-1} .

对当前分段的每个场景

利用式(5)和式(6)估计当前场景时隙可用网络带宽 $\bar{X}(k)$.

根据 Δ_{k-1} 和 C_k 的关系作如下检查:

1. 如果 $\Delta_{k-1} \leq C_k$, 则表明增强层无法传输, 置 $r_e(k)=0$;

2. 如果 $\Delta_{k-1} > C_k$, 则若 $k=1$, 置 $r_e(1) = (\bar{X}(k) - \bar{r}_b(k)) \cdot (\Delta_0 / C_1)$;

否则, $k > 1$, 置 $r_e(k) = \alpha \cdot (\bar{X}(k) - \bar{r}_b(k)) \cdot (\Delta_{k-1} / C_k) + (1-\alpha) \cdot r_e(k-1)$ (根据客户端缓冲状况, 最大可能地利用当前时隙带宽来传输增强层数据; 同时也考虑到连续场景间的速率尽可能平滑一致);

3. 确保步骤 2 计算的结果在增强层速率范围 $[0, r_{e\max}]$ 之内;

4. 传输当前时隙的场景数据, 并记录当前时隙的实际网络带宽 $X(k)$;

5. 按如下规则更新系统参数:

5.1. 若 $\bar{r}_b(k) + r_e(k) > X(k)$, 则表明当前时隙网络带宽估计过高, 置 $\alpha = \max\{\alpha \cdot \gamma, 0\}$;

否则, 置 $\alpha = \min\{\alpha / \gamma, 1\}$ (根据网络带宽估计值和实际测试值的关系, 动态调整下一时隙的平滑参数 α)

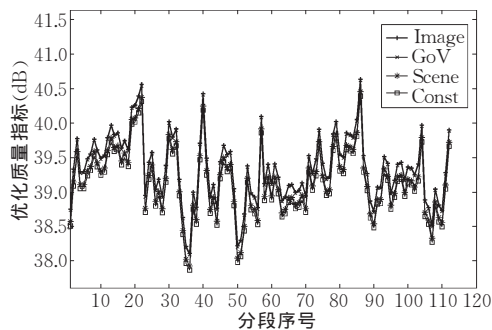
5.2. 置 $k=k+1$, 进入当前分段下一场景的调度.

5 试验分析

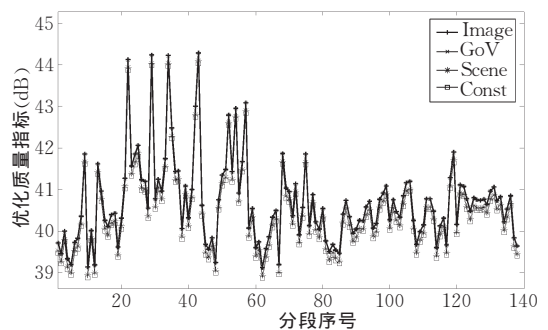
5.1 聚合粒度分析

从实践来看, 以场景作为分段调度中的最小粒度单位是直观的选择. 然而, 不同聚合粒度 (aggregation levels) (如 GoV、场景或任意分割) 下对优化调度策略影响的定量分析是必要的.

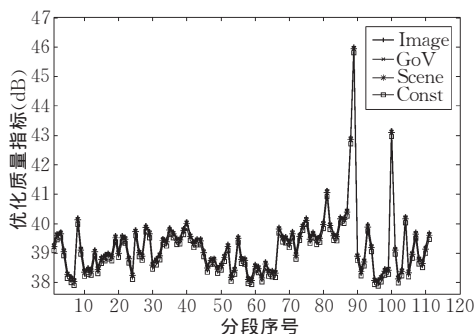
图 3 给出 4 个采用 Microsoft FGS 编码器^[19]的



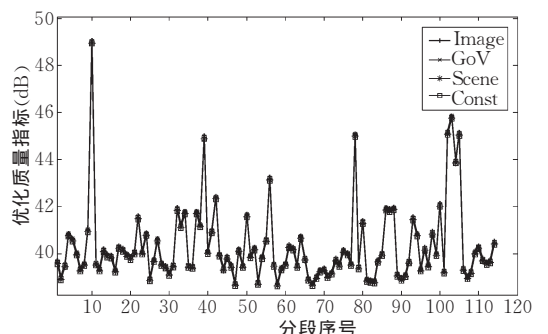
(a) The Firm



(b) Starwars



(c) News



(d) Football

图 3 不同分段聚合粒度下的视频序列优化质量指标 ($R_{\max}=800\text{Kbps}$)

视频序列,在不同分段聚合粒度下应用状态迁移图算法得到的视频序列优化质量指标 θ_m . 图 4 是在不同的增强层编码速率下(400~2000Kbps),每个视频序列的平均质量指标 \bar{Q} . 显然,一般情况下,基于帧的优化算法有最好的质量指标,而固定分段的质量指标相对较低. 然而,不同粒度下的总体质量指标

相差在 0.3~1dB 之间. 对要同时满足大量用户的视频服务器,采用较小的分段粒度如帧、GoV 进行优化调度,将极大地增加其负载. 采用固定分段则难以保证视频质量的突发稳定在一定范围内. 因此,以场景作为视频调度的粒度单位是一种直观而合理的选择.

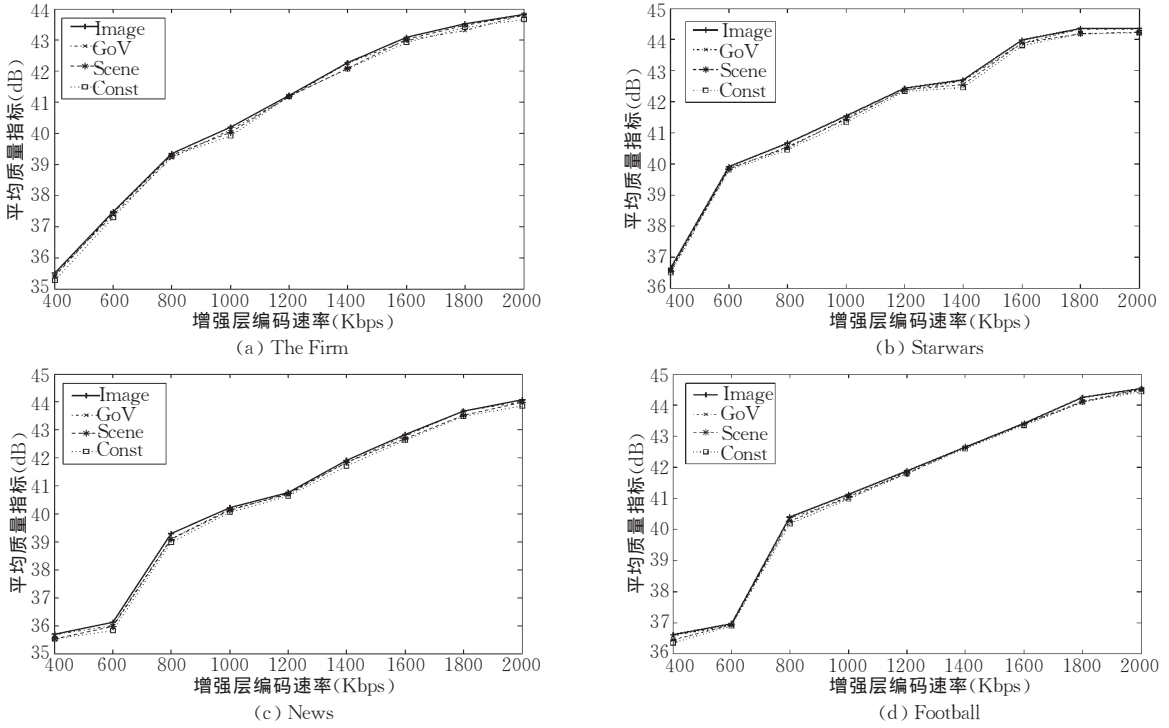


图 4 不同增强层编码速率(R_{max})下的视频序列平均质量指标

5.2 优化传输速率分析

基于对状态迁移图的分析 and 可达速率算法,利用 ns-2^[20] 对实时视频序列 Starwars 的分析结果如图 5~图 12 所示. 图 5 的节点 2~3 的链路采用 RED 队列,以避免 TCP 的周期性同步. 图 6 中序列 Starwars 的基层编码速率为 384Kbps,增强层最大编码速率为 1200Kbps. 图 7 给出了在 RED 队列下的 TCP 输出流量,同时在优化传输策略和可达速率算法下序列 Starwars 的第 1 分段(帧 1~6882)的传

输率也进行了对比分析. 可以看出,该分段中的每个场景按基本一致的速率调度,因此在保证基本视频质量的同时,可以保证质量变化指标保持在一定范围内. 而最优传输策略与实际序列有更加接近的速率趋势(图 7 的 180~210s). 图 8 是初始延迟 $\Delta_0 = 5$ 时,优化和可达速率的客户端回放延迟对比. 总体上可达速率算法的缓冲延迟在 6s 以内;而对最优调度算法,缓冲延迟可能达到 11s 甚至更大.

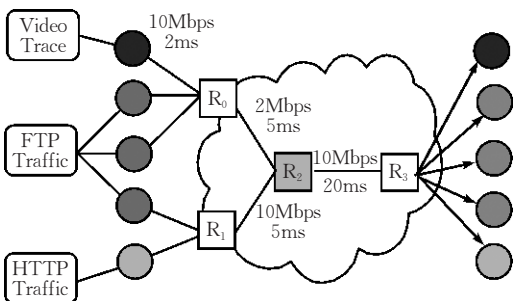


图 5 ns-2 下的 RED 队列仿真模型

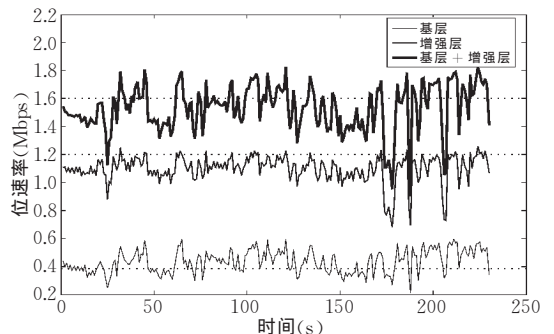


图 6 Starwars 分段 1 编码帧速率

对序列 Starwars 分段 1 场景 5~14(对应于帧 383~865)的原始编码率和可达速率算法的场景平均传输率在图 9 中给出;相应的场景中的帧质量指标在图 10 给出,其中垂直分隔线代表场景划分.图 11 是连续帧的质量变化指标(定义 4),总体上连续帧间质量变化在 $-2\text{dB} \sim 2\text{dB}$ 之间;对 Starwars 分段 1 的所有场景(123 个)的质量变化指标的分析(图 12)可以看出,连续场景间的质量变化平均值约为 2.2dB .以上分析表明:基于分段的可达速率调度算法可以保证视频序列场景的传输有较为接近的质量指标,同时其质量变化保持在一定的范围内,因此也就能更好地满足终端用户的主观质量要求(对其它序列的分析也得到类似的结果.限于篇幅,未列出).

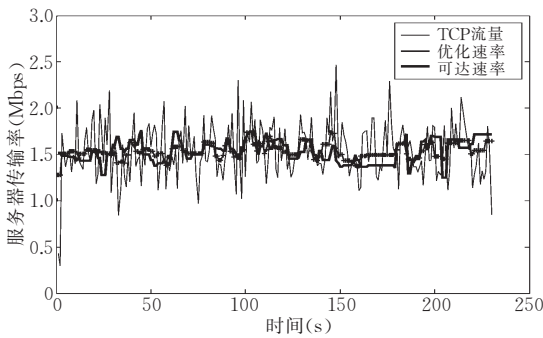


图 7 优化和可达速率对比分析图 (Starwars 分段 1 帧 1~6882)

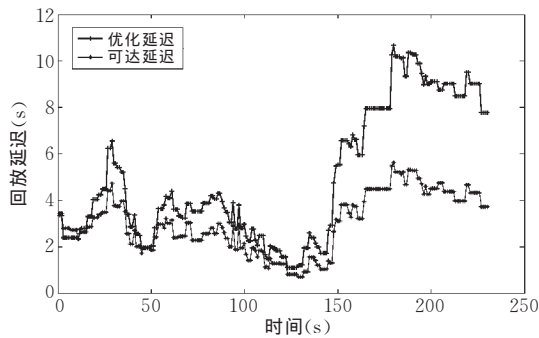


图 8 回放延迟对比分析图(Starwars 分段 1 帧 1~6882)

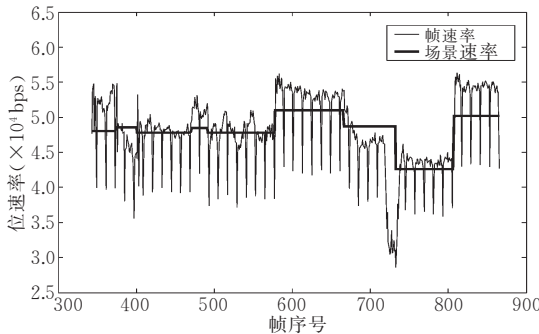


图 9 帧 383~865 的编码率与传输率(Starwars 场景 5~14)

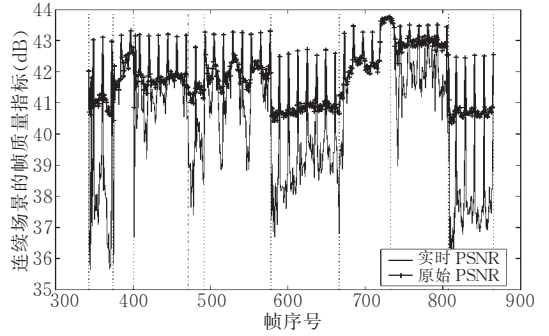


图 10 Starwars 连续场景(5~14)的帧质量对比图

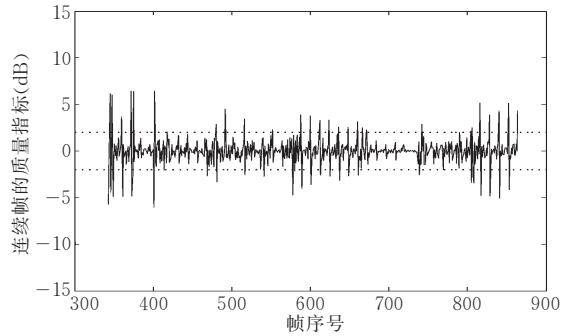


图 11 Starwars 连续帧 383~865 的质量指标变化

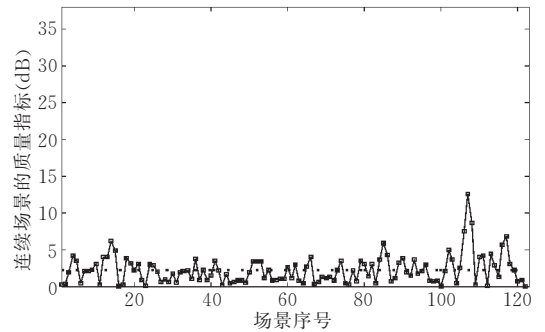


图 12 Starwars 分段 1 连续场景的质量指标变化

6 结 论

对流视频的主观质量进行严格的定义是非常困难的.基于客观质量指标如 PSNR(MSE)有其可测性,并且针对 FGS 视频的优化调度算法能取得较好的视频质量,因此在实践中可以得到应用.然而,对视频的质量变化指标(定义 4)尚无法很好地度量,并未应用到相应的优化算法中.这也是实践中需要深入考虑的问题.

参 考 文 献

1 ISO/IEC 14496-2. Generic coding of audio visual objects: Visual. ISO/IEC JTC1/SC29/WG11/N3095, Dec. 1999

- 2 Li W. P. . Overview of fine granularity scalability in MPEG-4 video standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 2001, 11(3): 301~317
- 3 Radha H. , Chen Y. , Parthasarathy K. *et al.* Scalable Internet video using MPEG-4. *Signal Processing: Image Communication*, 1999, 9(15): 95~126
- 4 Radha H. , van Der Schaar M. , Chen Y. . The MPEG-4 fine-grained scalable video coding method for multimedia streaming over IP. *IEEE Transactions on Multimedia*, 2001, 3(1): 53~68
- 5 Zhao L. , Kim J. W. , Kuo C. C. J. . Constant quality rate control for streaming MPEG-4 FGS video. In: *Proceedings of the International Symposium on Circuits and Systems (ISCAS)*, Scottsdale, Arizona, USA, 2002, 544~547
- 6 Vieron J. , Turletti T. , Henocq X. *et al.* TCP—Compatible rate control for FGS layered multicast video transmission based on a clustering algorithm. In: *Proceedings of the International Symposium on Circuits and Systems(ISCAS)*, Scottsdale, Arizona, USA, 2002, 453~456
- 7 van Der Schaar M. , Radha H. . Adaptive motion compensation fine granular scalability(AMC-FGS) for wireless video. *IEEE Transactions on Circuits and Systems for Video Technology*, 2002, 6(12): 360~371
- 8 Fitzek F. , Reisslein M. . MPEG-4 and H.263 video traces for network performance evaluation. *IEEE Network*, 2001, 15(16): 40~54
- 9 Zhang Q. , Zhu W. , Zhang Y.-Q. . Resource allocation for multimedia streaming over the Internet. *IEEE Transactions on Multimedia*, 2001, 3(3): 339~355
- 10 <http://trace.eas.asu.edu/fgs/index.html>
- 11 de Cuetos P. . Evaluating the streaming of FGS-encoded video with rate-distortion traces. Institut Eurecom, Sophia Antipolis: Technical Report RR-03-078, 2003
- 12 de Cuetos P. , Ross K. W. . Adaptive rate control for streaming stored fine-grained scalable video. In: *Proceedings of the NOSSDAV'02*, Miami, Florida, USA, 2002, 12(14): 3~12
- 13 Report on MPEG-4 visual fine granularity scalability tools verification tests. ISO/IEC JTC1/SC29/WG11/N4791. May 2002
- 14 VQEG. Video quality experts group; Current results and future directions. In: *Proceedings of the SPIE Visual Communications and Image Processing*, Perth, Australia, 2000, 772~783
- 15 Huang T. Y. , Sun S. X. . Statistical characteristics and model of MPEG VBR video stream. *Chinese Journal of Computers*, 2001, 24(9): 1002~1008(in Chinese)
(黄天云,孙世新. MPEG VBR 视频流——统计特性及其模型. *计算机学报*, 2001, 24(9): 1002~1008)
- 16 Saw Y. S. . *Rate Quality Optimized Video Coding*. Kluwer Academic Publishers, 1999
- 17 Huang C. L. , Liao B. Y. . A robust scene-change detection method for video segmentation. *IEEE Transactions on Circuits and Systems for Video Technology*, 2001, 11(12): 1281~1288
- 18 Dawood A. M. , Ghanbari M. . Scene content classification from MPEG coded bit streams. In: *Proceedings of the Workshop on Multimedia Signal Processing(MMSP)*. Copenhagen, Denmark, 1999, 253~258
- 19 Microsoft. ISO/IEC 14496 Video Reference Software. Microsoft-FDAM1-2. 3-001213
- 20 UCB/LBNL/VINT. Network Simulator — ns(version 2). <http://www.isi.edu/nsnam/>, Feb. 2005



HUANG Tian-Yun, born in 1973, Ph. D. . His current research interests include video flow modeling, analysis, video transmission, and media adaptation, etc.

Background

As a scalable coding algorithm for video streaming over the Internet, the transmission of MPEG-4 FGS video attracts more attention in recent years. Many researches of MPEG-4 FGS in wire(IP) and wireless environments are beginning to yield fruits. However, there is little work on exploring the quantitative analysis of the objective quality and quality variation of FGS-coded video.

Note that the attainable subjective quality of end user is affected not only by the quality of each frame in video sequence, but also by the variation between them. A strict def-

inition of its objective quality and quality variation is required, in order to provide efficient transmission of FGS-coded video and an attainable subjective quality perception for end user. This research purpose is trying to give a definition of objective quality and quality variation for FGS-coded video, based on MSE and PSNR. Since the definition is measurable, it can be generalized to the design of an optimal policy for real-time video transmission. This work can also be regarded as a part of research on video adaptation in the framework of MPEG-21 DIA(Digital Item Adaptation).