

一种鲁棒灵活的非平衡多描述视频编码和传输方案

李 彬 黄 峰 孙立峰 杨士强

(清华大学计算机科学与技术系 北京 100084)

摘 要 在网络上传输的视频经常因为丢包而影响终端接收到的视频的质量. 由于现有的视频编码器通常使用预测编码技术来减少时域冗余提高压缩率, 所以一个包的丢失都会引起错误的传播, 直到解码器收到帧内编码的帧为止. 针对这个问题, 文中提出一种灵活鲁棒的非平衡多描述编码和传输方案. 其编码器借助不同描述间的“同步帧”, 能够迅速从包丢失中恢复解码, 并保证终端视频的连续播放; 同时, 该方案能够灵活适应多路径传输和单路径传输两种情况. 文中还研究了单路径传输时不同描述间的码率分配问题, 并提出一种近似最优的快速码率分配方案. 实验结果充分验证了文中提出的非平衡多描述编码和传输方案的有效性、鲁棒性和灵活性.

关键词 多描述编码; 非平衡多描述编码; 码率分配; 视频编码; 传输方案

中图法分类号 TP393

A Robust and Flexible Unbalanced Multiple Description Video Coding and Transmission Scheme

LI Bin HUANG Feng SUN Li-Feng YANG Shi-Qiang

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084)

Abstract This paper propose a flexible and robust unbalanced multiple description coding and transmission scheme, which is capable of quickly recovering from packet losses and ensuring continuous playback, and further adaptive to both multiple path and single path transmission. Furthermore, this paper discusses the problem of rate allocation between different descriptions when only single path transmission is available and proposes an approximate while efficient rate allocation scheme with the help of two-state Markov link model and a representative rate-distortion model. The experimental results demonstrate the effectiveness of the proposed scheme.

Keywords multiple description coding; unbalanced multiple description coding; rate allocation; video coding; transmission scheme

1 引 言

随着网络技术和数字内容产业的迅猛发展, 作为丰富信息载体的视频流媒体逐渐成为互联网主流应用之一. 在现有的视频编码方法中, 人们通常采用

预测编码技术来减少视频信号间的时域相关性, 获得较好的压缩效率; 但是, 在获得高压压缩效率的同时, 预测编码技术也造成了丢失一个包也将导致解码过程中的错误沿时间轴方向传播的缺陷. 为了限制错误的传播, 前期研究提出了多种方案, 例如不时地插入采用帧内编码的帧来阻止错误传播, 或者使

用纠错码进行前向纠错(Forward Error Correction, FEC)来增加码流的鲁棒性并减少错误传播的长度^[1]. 然而,前者生成的码率较高,后者主要障碍在于, FEC 中当给定一个预先确定的冗余时,其纠错能力是固定的;当丢包率超出该阈值时, FEC 无法纠正错误,导致视频质量严重下降,即所谓的“悬崖效应(cliff effect)”. 尽管人们也提出了自动重传请求(Auto Retransmission reQuest, ARQ)方法,但是由于视频应用通常对包传送时间期限有严格的限制, ARQ 通常不适用^[2].

多描述编码(Multiple Description Coding, MDC)技术在提出后即被证明是一种处理错误传播的行之有效的办法^[3]. 低延迟的特性使得 MDC 在鲁棒的视频传输方面具有良好的应用前景. 一般来说,多数的 MDC 采用多个独立的视频描述进行传输,只要接收到一个或多个描述,那么解码器就可以恢复出相应质量的视频流. MDC 是根据各个描述之间的重要性关系被分为平衡多描述编码(Balanced Multiple Description Coding, BMDC)和非平衡多描述编码(Unbalanced Multiple Description Coding, UMD)的. 平衡多描述编码采用具有大致相同重要性和质量的描述,例如基于多相下采样(Poly-phase Downsampling)的方法^[4]、基于标量量化(Scalar Quantization)的方法^[5]、基于格子向量量化(Lattice Vector Quantization)的方法^[6]、基于成对相关变换(Pairwise Correlating Transform)的方法^[7]等. 这些方案通常都无法较好地解决如何控制多描述编码引入的冗余量问题,而且从带宽利用角度来看,当有多个路径可用时,平衡多描述编码不能充分利用各个路径的可用带宽. 与之对应,非平衡多描述编码^[8-11]使用多个不同重要性和质量的视频描述,为便于说明,这里采用两个描述,一个具有较高的重要性和质量,称为高分辨率(High-Resolution, HR)描述,另一个重要性和质量相对较低的描述称为低分辨率(Low-Resolution, LR)描述. 在这种情况下, LR 描述主要是作为冗余,用于隐藏 HR 描述由于包丢失而产生的错误. 与平衡多描述编码相比,非平衡多描述编码能够灵活控制引入的冗余量,同时充分利用不同路径的端到端可用带宽. 非平衡多描述编码器的核心问题是如何使用 LR 描述的信息恢复 HR 描述中由于丢包造成丢失的信息. 针对这个问题,文献[8-10]提出了局部错误隐藏技术,使用丢失区域附近的时域和空域的信息,但没有解决错

误传播到后续的帧间编码帧,减轻错误传播的能力有限. 文献[11]使用了基于序列的错误隐藏(sequence-based EC)算法,使用多帧恢复原则来最小化 HR 描述里的错误传播,然而为了恢复当前帧里的丢失信息,该算法必须等待并检查一定数量的后续帧,无法满足具有严格的延迟要求和突发丢包(burst packet losses)的视频应用.

因此,本文重点研究适应视频传输的非平衡多描述编码方案,利用 LR 描述和 HR 描述的相互协作提高视频传输的鲁棒性;同时探讨非平衡多描述方案在多路径和单路径条件下的传输问题,力求通过灵活的码率分配机制充分利用端到端的可用带宽. 针对以上问题,我们提出了一种新的非平衡多描述编解码器,使用“同步帧”来阻止 HR 描述里的错误,利用 HR 描述和 LR 描述的相互协作从包丢失中快速恢复,提高关键帧传输的鲁棒性. 在此基础上,我们还讨论了该方案在多路径和单路径条件下的传输性能,并借助两状态马尔可夫链路模型和简化的率失真模型提出了单路径条件下一种近似最优的码率分配方案. 仿真实验结果充分验证了所提出的非平衡多描述编码器以及传输方案的有效性、鲁棒性和灵活性.

本文第 2 节给出非平衡多描述编码器结构,重点分析借助“同步帧”来解决 HR 和 LR 间相互协作提高视频传输鲁棒性这一方法的特点;第 3 节讨论该编码器在多路径和单路径条件下的传输方案,继而在第 4 节提出近似最优的码率分配方案;第 5 节通过仿真实验验证所提编码器和传输方案的性能;第 6 节总结全文并讨论未来工作.

2 鲁棒的非平衡多描述编解码器

如上所述,非平衡多描述编码器的关键在于 LR 和 HR 描述间的相互协作,因此在所提出的非平衡多描述编解码器中,我们将 LR 编码器的输入与 HR 编码器的重构帧设计成关联的,取代其他非平衡多描述编码器方案中的原始视频帧,如图 1 和图 2 所示. 在编码器端,我们首先引入“同步帧”加强 HR 和 LR 描述之间的相关性. 所谓“同步帧”,即 LR 编码器重构的 I 帧,其作用主要是置于 HR 编码器的帧缓冲区并替代 HR 编码器的当前的重构的 P 帧. 这样, HR 编码器在编码下一个 P 帧时就以“同步帧”作为参考帧. 需要说明的是,“同步帧”的引入

增强了 HR 描述对包丢失进行恢复的鲁棒性. 在解码器端, HR 解码器必须取得“同步帧”来用于下一个 P 帧的运动补偿. 之所以考虑引入“同步帧”, 主要是因为 HR 描述的帧内刷新周期较长, 当 P 帧因为丢包缺失时, 必须等待下一个 I 帧到来才可以恢复, 而引入“同步帧”后, 丢失的 P 帧可以通过 HR 和 LR 的“同步帧”恢复, 提高了 HR 视频传输的鲁棒性. 一般来说, 在图 1 中所示的“同步帧”可以通过两种方式获得: (1) LR 解码器的输出帧作为“同步帧”; (2) 编码器的重构帧, 此时编码器的输入是 HR 解码器的输出.

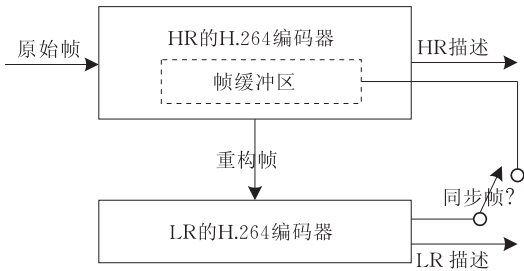


图 1 鲁棒的 UMDC 编码器结构

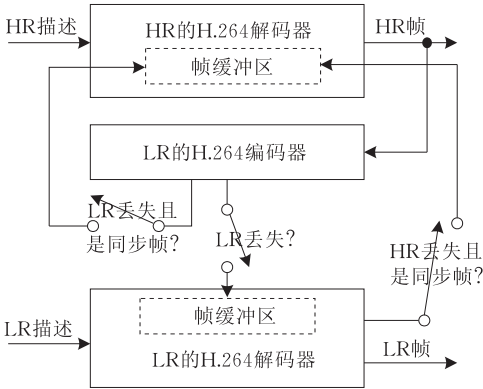


图 2 鲁棒的 UMDC 解码器结构

本文提出的非平衡多描述编解码器借助“同步帧”加强 HR 和 LR 描述的相关性, 能够处理多种包丢失情况, 下面分 4 种情况讨论编解码器的工作

方式:

(1) 当 HR 描述和 LR 描述都正确接收时, 两个描述都可以正确解码, 解码器的输出是 HR 描述的数据;

(2) 当 LR 描述在传输中发生包失, 而 HR 描述顺利接收时, 所提编解码器能够快速恢复所丢失的包. 如上所述, HR 编码器的重构帧作为 LR 编码器的输入, 此时只需把 HR 解码器的输出直接输入到一个标准编码器, 生成 LR 描述并正确重构 LR 描述中的因包丢失而损坏的帧, 继续 LR 描述的解码过程;

(3) 当 HR 描述在传输中发生包丢失, 而 LR 描述顺利接收时, 常规的非平衡多描述编解码器里通常只有两种选择: 要么 HR 描述的解码过程暂停直到接收下一个能够正确解码的 I 帧, 解码输出相应的 LR 描述的数据; 要么把相应的 LR 描述的数据作为 HR 描述的参考帧, HR 描述的解码过程继续. 然而前一种方法得到的 LR 描述质量较低, 后一种方法造成 HR 编码器和解码器的不匹配, 从而这两种方法都会降低解码视频的质量. 所提出的非平衡多描述编码器中, 由于 HR 描述的下一个 P 帧使用“同步帧”作为参考帧, 因此从正确接收的 LR 描述里取得“同步帧”并作为 HR 描述的参考帧时, 在 HR 编码器和解码器之间不存在不匹配的情况. 此时 HR 解码器只需等待从 LR 解码器输出得到下一个“同步帧”即可重新开始解码过程;

(4) 当 HR 描述和 LR 描述同时发生包丢失或者错误传播时, 解码器暂停解码等待下一个“同步帧”到达后重新开始解码. 在这个过程中解码器的输出为最近的已解码帧.

图 3 和图 4 分别说明了 LR 发生包丢失和 HR 发生包丢失时的示例, 其中 I_p 表示“同步帧”, P_x 表示在传输过程中丢失的 P 帧, P_d 表示丢弃的 P 帧.

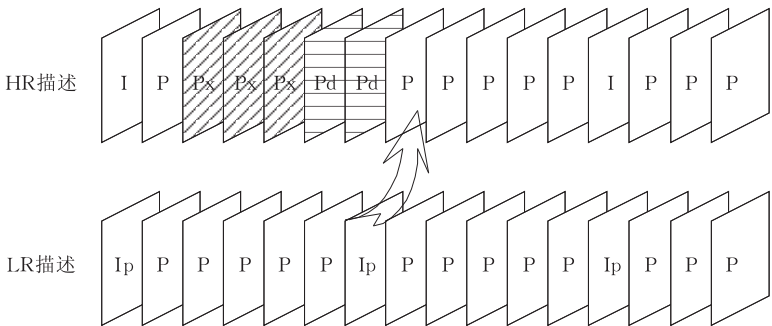


图 3 LR 描述的同步帧 I_p 重新启动 HR 描述的解码过程

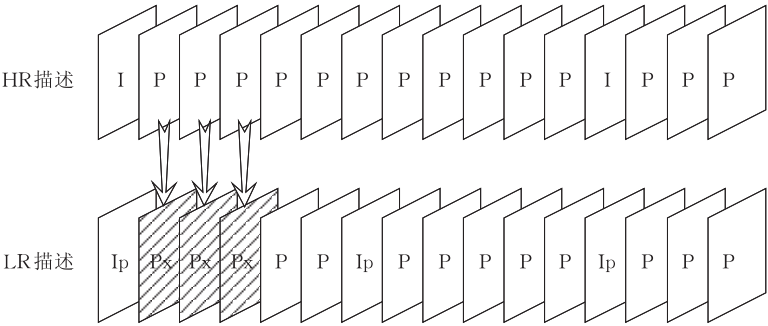


图 4 LR 描述丢失的 P 帧由 HR 的重构帧编码后重构

3 适应多路径和单路径的灵活传输方案

在提出了上述鲁棒的非平衡多描述编码器之后,我们重点研究多个描述的传输问题.多描述编码传输的性能主要取决于视频同一部分或者相邻部分的不同描述的包是否同时丢失.针对这个问题,通常人们采用分离的多个路径,即“路径多样性(Path Diversity)”来发送不同的描述,这种传输方案被称为多路径传输(Multiple Path Transport, MPT).大量的研究工作^[12-16]探讨了结合多描述编码和多路径传输进行可靠视频传输的方案,这些工作大多对比了多路径传输与单路径传输(Single Path Transport, SPT)的优劣.与使用单描述编码和单路径传输相比,结合多描述编码和多路径传输的性能得到有效提升.

由于多路径传输对于传输多个描述具有很内在的优越性,因此在传输过程中具有多条可用的路径时,所提出的非平衡多描述编码传输方案采用多条路径传输.然而,在一些情况下,由于种种原因无法征用多条路径,此时我们需要考虑多个描述在单一路径传输的问题.这种在多路径和单路径间灵活传输的机制也是本文区别于其他研究的主要特点之

一.当考虑单一路径传输时,如果只是简单地把每个描述的包放在一起,那么由于网络丢包的突发特性,对应视频同一部分或者相邻部分的不同描述的包同时丢失的可能性较大,会严重影响多描述编码传输的性能.文献[17]介绍了一种在单路径上传输非平衡多描述视频的方法.为了能够在单个路径上灵活地传输多个描述,我们借鉴该方法的一些思想,并对其进行了扩展和改进,主要方案如下:

在使用单路径传输多个描述时,为了确保 HR 和 LR 描述同时发生丢包的可能性较低,不能同时传输对应视频同一部分的 HR 和 LR 的包.因此所提方案将分别属于不同帧的 HR 和 LR 数据封装到同一个包,降低同一个数据包内的 HR 数据和 LR 数据之间的相关性.这样属于同一帧的 HR 和 LR 数据以一定间隔在同一条路径上传输.这种 HR 和 LR 数据间帧号偏移,提供了在单一路径上传输的灵活性.我们可以根据网络状况来决定相应的偏移量,当网络状况较好时,使用较小的偏移,而当网络状况较差时,使用较大的偏移.图 5 示意性地说明了在单一路径上传输的打包方式,其中 k 表示偏移.当然,偏移也会引入一定的延迟,在后面的实验中我们相应地讨论了这种偏移方案的性能.这里所提出的利用 HR 和 LR 间的偏移灵活适应不同网络状况的传输方案是本文区别其他研究的另一个特点.

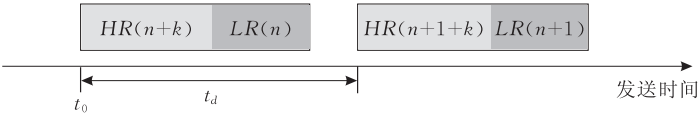


图 5 HR 和 LR 封装方法

4 近似最优的多描述码率分配方案

上面两节分别说明了所提出的鲁棒的非平衡多描述编码器和灵活的传输方案,这一节主要研究在

使用单路径传输多个描述时,如何根据网络状况分配 HR 和 LR 描述的码率,进而提高视频传输质量的问题.下面我们通过对网络丢包状况和率失真建模分析,求解在给定网络状况下,HR 描述和 LR 描述之间的近似最优的码率分配,从而使接收到的视

频质量近似最优。

4.1 两状态连续时间马尔可夫链路模型

我们首先采用文献[18]中提出的两状态连续时间马尔可夫链 $\{X_t\}$ 来模拟链路传输的状况, 其中 $X_t \in \{0, 1\}$ 表示链路上包丢失, $X_t=1$ 说明在 t 时刻发生了包丢失, $X_t=0$ 说明在 t 时刻包顺利传输。由此, 我们可以用式(1)表示马尔可夫链的微分生成矩阵, 从而对应该马尔可夫链的稳定分布为 $\pi = (\pi_0, \pi_1)$, 其中 $\pi_0 = \mu_1 / (\mu_0 + \mu_1)$, $\pi_1 = \mu_0 / (\mu_0 + \mu_1)$ 。在时间间隔 t 内, 状态 i 和状态 j 之间的转移概率可以通过式(2)计算。当给定平均丢包率 P_L 、平均突发丢包长度 \bar{b} (单位: 包) 和发送速率 λ 时, 参数 μ_0 和 μ_1 可以由式(3)得到。

$$Q = \begin{bmatrix} -\mu_0 & \mu_0 \\ \mu_1 & -\mu_1 \end{bmatrix} \quad (1)$$

$$\begin{cases} P(X_{t_0+t} = 0 | X_{t_0} = 1) = \pi_0 - \pi_0 \times \exp(-(\mu_0 + \mu_1)t) \\ P(X_{t_0+t} = 1 | X_{t_0} = 1) = \pi_1 + \pi_0 \times \exp(-(\mu_0 + \mu_1)t) \\ P(X_{t_0+t} = 0 | X_{t_0} = 0) = \pi_0 + \pi_1 \times \exp(-(\mu_0 + \mu_1)t) \\ P(X_{t_0+t} = 1 | X_{t_0} = 0) = \pi_1 - \pi_1 \times \exp(-(\mu_0 + \mu_1)t) \end{cases} \quad (2)$$

$$\begin{cases} \mu_0 = -P_L \lambda \left(1 - \frac{1}{\bar{b}(1 - P_L)}\right) \\ \mu_1 = \mu_0 (1 - P_L) / P_L \end{cases} \quad (3)$$

实际上, 可以通过将接收端测量出 P_L 和 \bar{b} 的值反馈回发送端, 能够持续计算并更新参数 μ_0 和 μ_1 的值。

4.2 简化的率失真模型

上一小节通过两状态连续时间马尔可夫链模拟了网络丢包状况, 为了进一步对码率分配问题进行研究, 我们考虑如何将码率和视频失真联系起来。对于压缩视频应用, 前期研究已经讨论了一些失真模型来估计给定码率情况下的失真状况, 然而这些率失真模型大多计算复杂, 缺乏实用性。结合视频传输严格的延迟要求, 我们需要选择一个相对容易实现又具有代表性的率失真模型。因此, 我们选择了文献[19]中提出的简化模型来估计失真, 如式(4)所示, 其中 K 是模型参数, R_i 和 $D(R_i)$ 分别表示码率和对应的失真。对于离线编码的视频序列, 参数 K 的值可以预先计算; 而对于实时编码, 先给 K 假设一个初始值, 然后随着编码的帧的增加不断改进 K 值。

$$D(R_i) \approx \frac{K}{R_i} \quad (4)$$

4.3 近似最优的码率分配模型

上面分别讨论了网络丢包模型和简化的率失真模型。当 HR 描述正确接收时, 接收端的失真为 D_{HR} , 当 LR 描述正确接收而 HR 描述发生包丢失时, 接收端的失真为 D_{LR} , 当 HR 和 LR 描述都发生包丢失时, 接收端的失真为 D_C 。若记链路的丢包率为 p , 当 HR 包丢失时正确接收 LR 的条件概率记作 p_{10} , 当 HR 丢失且 LR 也丢失的条件概率记作 p_{11} , 那么接收端的失真 D 可以表示成式(5)。

$$D = (1-p) \times D_{HR} + p \times p_{10} \times D_{LR} + p \times p_{11} \times D_C \quad (5)$$

从第 4.1 节中可以得到网络链路的丢包率 p 就是该马尔可夫链的稳定分布 π_1 , 条件概率 p_{10} 即是该马尔可夫链从状态 1 到状态 0 的转移概率 $P(X_{t_0+t} = 0 | X_{t_0} = 1)$, 条件概率 p_{11} 即为该马尔可夫链从状态 1 到状态 1 的转移概率 $P(X_{t_0+t} = 1 | X_{t_0} = 1)$ 。对于单路径传输方案而言, 属于同一帧的 HR 描述和 LR 描述之间的时间间隔为 $k \times t_d$, 如图 6 所示。这样, 式(5)可以重写为式(6)。

$$D = (1-p) \times D_{HR} + p \times P(X_{t_0+kt_d} = 0 | X_{t_0} = 1) \times D_{LR} + p \times P(X_{t_0+kt_d} = 1 | X_{t_0} = 1) \times D_C \quad (6)$$

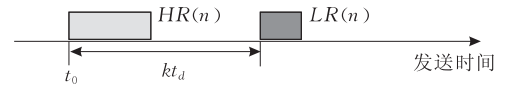


图 6 HR 和 LR 描述之间的延迟

记链路的带宽为 R_T , HR 描述的码率为 R_1 , LR 描述的码率为 R_2 , 则在 HR 描述和 LR 描述之间码率分配问题转化为在 HR 描述和 LR 描述的码率总和一定的条件下最小化接收端的失真的问题, 如式(7)所示。

$$\begin{cases} \text{Min } D \\ \text{s. t. } R_1 + R_2 = R_T \end{cases} \quad (7)$$

把式(2)和式(4)代入式(6), 可以得到式(8):

$$\begin{aligned} D \approx & (1-p) \times \frac{K}{R_1} + p \times \\ & (\pi_0 - \pi_0 \exp(-(\mu_0 + \mu_1)kt_d)) \times \frac{K}{R_2} + p \times \\ & (\pi_1 + \pi_0 \exp(-(\mu_0 + \mu_1)kt_d)) \times D_C \end{aligned} \quad (8)$$

显然, 当 $k \times t_d$ 足够大时, $\exp(-(\mu_0 + \mu_1)kt_d) \rightarrow 0$, 但是由于 $k \times t_d$ 关系到解码延迟, 所以 $k \times t_d$ 的值不能太大。通过选择一个合适的 k 值, 使得 $\exp(-(\mu_0 + \mu_1)kt_d) \approx 0$, 这样式(8)就可以近似成式(9)。

$$\begin{aligned} D \approx & (1-p) \times \frac{K}{R_1} + p \times (1-p) \times \frac{K}{R_2} + p \times p \times D_C = \\ & (1-p) \times K \times \left(\frac{1}{R_1} + \frac{p}{R_2}\right) + p \times p \times D_C \end{aligned} \quad (9)$$

而在给定网络丢包率和视频源条件下, p, K 和 D_c 与 R_1, R_2 的取值无关, 那么最小化失真 $\min(D)$ 的问题就转化成 $\min(1/R_1 + p/R_2)$ 的问题, 即 $\min(D) \Leftrightarrow \min(1/R_1 + p/R_2)$. 由函数极值理论, 我们可以知道当 $\partial f/\partial R_2 = 0$, 即 $R_2 = \sqrt{p}R_T/(1 + \sqrt{p})$ 时, $1/R_1 + p/R_2$ 取得最小值, 即 D 取得最小值.

5 实验研究及结果分析

我们对本文提出的非平衡多描述视频编码和传输方案的性能进行了模拟实验. 在第一部分实验中, 我们首先考察本方案在多路径传输情况下的性能, 主要比较了本文方案、传统的前向纠错码方案以及其他两种非平衡多描述编码方案. 这两种方案包括: ①独立的非平衡多描述编码方案(I-UMDC). LR 描述独立于 HR 描述, 即 UMDC 编码器以不同的目标码率独立地编码 LR 描述和 HR 描述; ②传统的非平衡多描述编码方案(C-UMDC)^[11]. LR 描述的生成只是简单地重新编码 HR 描述的重构帧. 在这两种方案里, 当 HR 描述遇到突发的包丢失时, HR 解码器等待 HR 描述的下一个 I 帧以重新开始解码过程, 这时 UMDC 解码器输出 LR 描述; 当 HR 描述和 LR 描述都遇到包丢失或错误传播时, UMDC 解码器暂停解码, UMDC 解码器的输出是最近解码的 HR 帧或者 LR 帧, 直到 UMDC 解码器取得下一个 I 帧以重新开始解码. 对于 FEC 方案, 我们采用 Reed-Solomon 码实现. UMDC 方案和 FEC 方案的性能比较是在同样码率, 同样冗余下进行. 由于数据交织直接影响 FEC 码在传输受到突发错误时的性能, 所以在实现 FEC 方案时, 采用了数据交织的方式.

我们通过修改 H. 264 参考软件 JM8. 6 来实现 UMDC 和 FEC 方案的编码器. 这里假定使用 RTP 负载格式打包 H. 264 视频流. 我们先后测试了不同的视频序列并在本文中给出了对应 Foreman 和 News 序列(大小: QCIF, 帧率: 15 帧/s, 帧数: 150 帧)的结果作为示例性说明. 下面的评价选择 3 个指标, 分别是平均 PSNR、暂停概率以及 PSNR 方差. 平均 PSNR 反映视频的客观质量, 另两个指标则说明视频的主观质量、暂停概率对应播放的流畅性, 而 PSNR 方差表示每帧视频质量的抖动. 实验参数设置如表 1 所示. 对于每一组模型参数, 实验模拟 30 次.

表 1 仿真实验参数设置			
UMDC		FEC	
高分辨率描述 码率/kbps	低分辨率描述 码率/kbps	原始码流 码率/kbps	FEC 冗余码 流码率/kbps
100	33.33	100	33.33

实验结果如图 7~图 12 所示. 图 7 和图 8 列出了 4 个方案的解码视频质量. 这里 HR 描述的帧内刷新(Intra Refreshment, IR)周期设为 50 帧, LR 描述的帧内刷新周期为 10 帧, SD 的帧内刷新周期为 10 帧. 实验结果表明, 本文提出的 P-UMDC 明显优于其他的 UMDC, 获得增益大约为 1.0dB~2.8dB. FEC 方案在链路状态好的时候性能优于本文提出的方案 P-UMDC, 但是当链路状态较差时, 方案 P-UMDC 优于 FEC. 图 9 和图 10 说明了 4 个方案的暂停概率, 对视频播放的连贯性要求暂停概率越小越好. UMDC 解码器暂停只有在 HR 和 LR 描述都遇到包丢失或者错误传播才会发生, 本文所提出的 P-UMDC 暂停概率最低. 当链路的突发错误变长时, FEC 方案视频流的播放经常中断. 除了解码视频质量和暂停概率, 我们还使用 PSNR 方差来评价 UMDC 和 FEC 的性能, 图 11 和图 12 说明了 4 个方案的 PSNR 方差情况. 与前两个指标一样, 本文所提出的 P-UMDC 性能好于其他方案, 说明应用 P-UMDC, 视频流抖动较小.

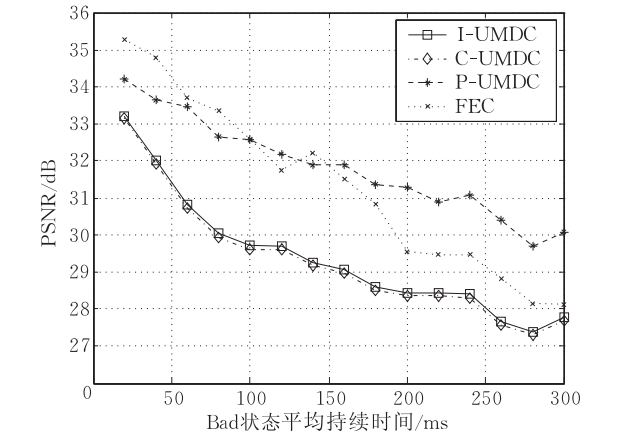


图 7 Foreman QCIF 序列的 4 种方案在不同 Bad 状态平均持续时间下的 PSNR 对比

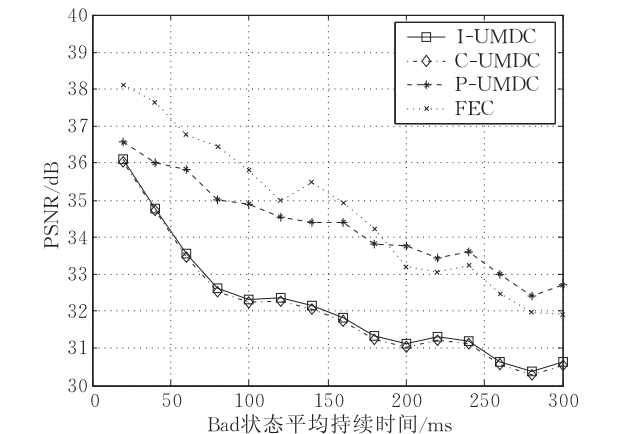


图 8 News QCIF 序列的 4 种方案在不同 Bad 状态平均持续时间下的 PSNR 对比

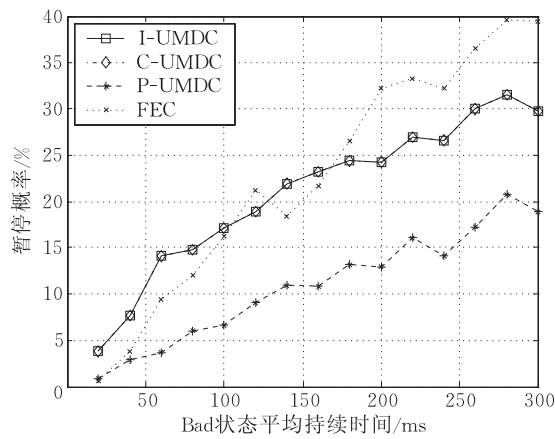


图 9 Foreman QCIF 序列的 4 种方案在不同 Bad 状态平均持续时间下的暂停概率对比

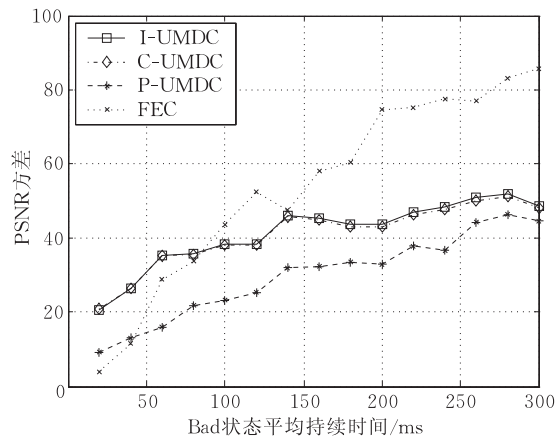


图 11 Foreman QCIF 序列的 4 种方案在不同 Bad 状态平均持续时间下的 PSNR 方差对比

正如第 3 节中提及,在一些情况下,多路径传输无法实现,所以我们继而通过模拟实验来考察单路径传输方案的性能.在后面的实验中,我们比较了多路径传输的 P-UMDC,单路径传输的 P-UMDC 和 FEC 3 个方案的性能,主要选择平均 PSNR、暂停概率以及 PSNR 方差 3 个指标,实验结果如图 13~图 18 所示.在图 13 和图 14 中,我们给出了多路径传输的

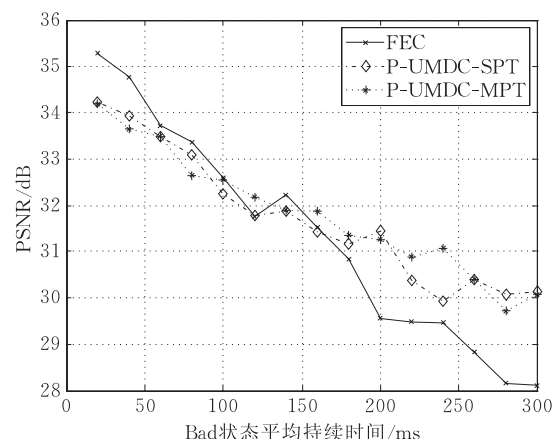


图 13 Foreman QCIF 序列的 3 种方案在不同 Bad 状态平均持续时间下的 PSNR 对比

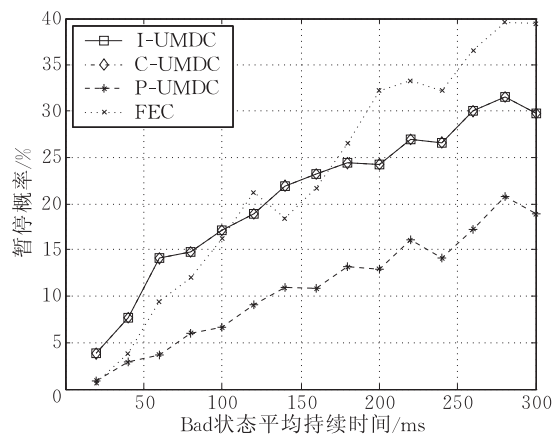


图 10 News QCIF 序列的 4 种方案在不同 Bad 状态平均持续时间下的暂停概率对比

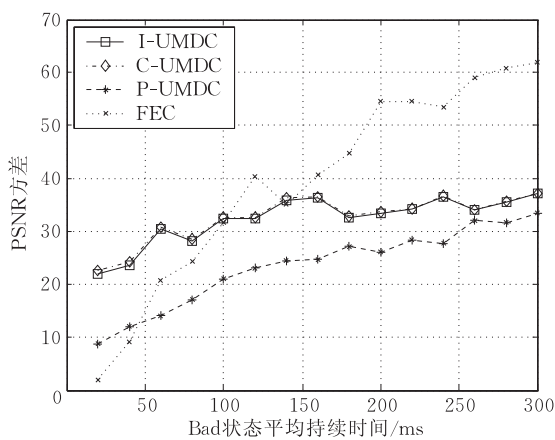


图 12 News QCIF 序列的 4 种方案在不同 Bad 状态平均持续时间下的 PSNR 方差对比

P-UMDC-MPT、单路径传输的 P-UMDC-SPT 和 FEC 系统的解码视频质量,图 15 和图 16 给出了暂停概率的对比情况,图 17 和图 18 给出了 PSNR 方差的曲线.从图 13~图 18 中,我们可以明显发现单路径传输的 P-UMDC-SPT 和多路径传输的 P-UMDC-MPT 的性能很接近,这说明了我们的单路径传输的 UMDC 方案在降低 HR 和 LR 描述丢包相关性方面具有很好的效果.

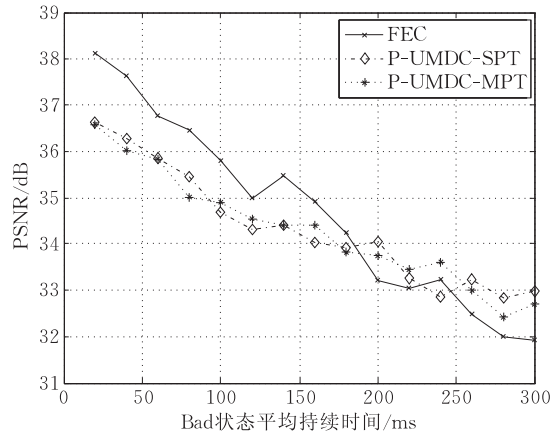


图 14 News QCIF 序列的 3 种方案在不同 Bad 状态平均持续时间下的 PSNR 对比

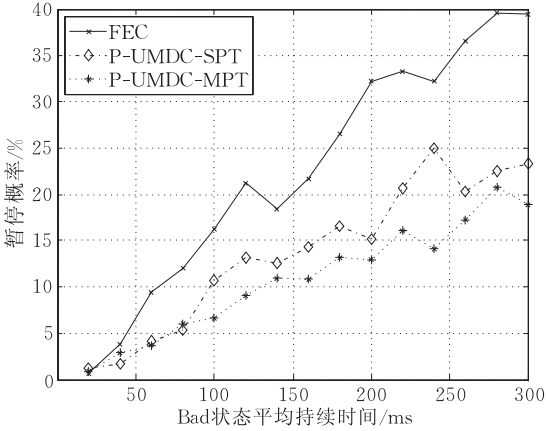


图 15 Foreman QCIF 序列的 3 种方案在不同 Bad 状态平均持续时间下的暂停概率对比

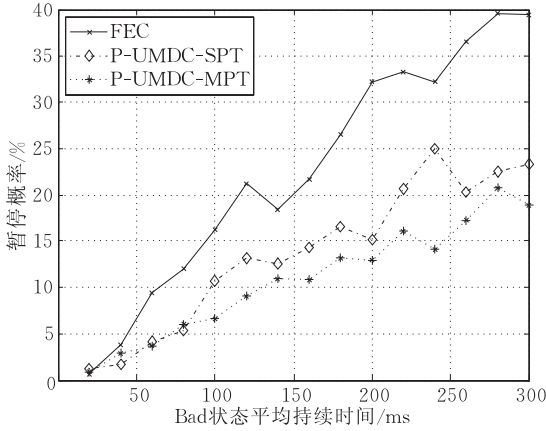


图 16 News QCIF 序列的 3 种方案在不同 Bad 状态平均持续时间下的暂停概率对比

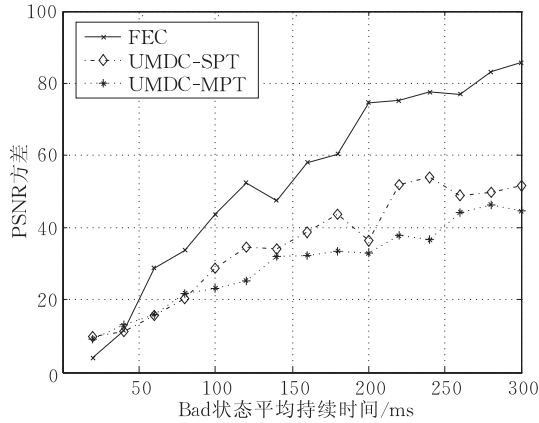


图 17 Foreman QCIF 序列的 3 种方案在不同 Bad 状态平均持续时间下的 PSNR 方差对比

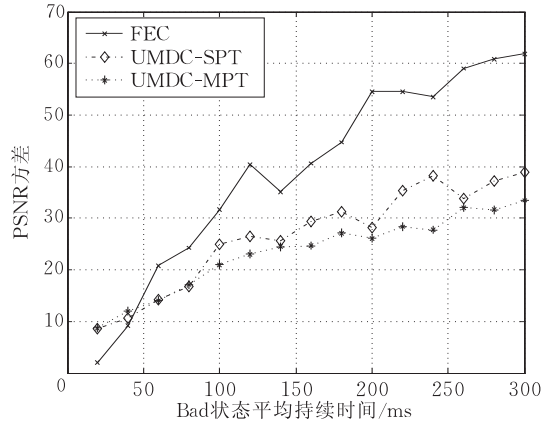


图 18 News QCIF 序列的 3 种方案在不同 Bad 状态平均持续时间下的 PSNR 方差对比

下面,我们通过模拟仿真实验结果来验证第 4 节所提出的近似最优的码率分配方案的效果.在仿真实验里,我们设置平均突发丢包长度为 1.5 和 2.0,包丢失概率为 15%,20%,25%和 30%,分别模拟不同的网络状态.我们设置视频流的总码率为 300kbps.在图 19 中,我们列出了平均突发丢包长度为 1.5,包丢失概率为 20%的网络状态下的模型结果和仿真结果,我们可以看到模型结果的曲线和仿真结果的曲线大体走势类似,符合得相当好.从图 19(b)中我们可以看出,对于 News QCIF 序列来说,当低分辨率描述的码率为 90kbps 时,接收到的视频的 PSNR 达到最大值,而对于 Foreman QCIF 序列来说,当低分辨率描述的码率为 75kbps 时,接收到的视频的 PSNR 达到最大值.而在第 4.3 节中,我们的近似最优的码率分配模型得出的结论为:当低分辨率描述的码率 $R_2 = \sqrt{p}R_T / (1 + \sqrt{p})$ 时,接收到的视频的失真最小(也就是 PSNR 最大),把包

丢失概率 $p=20\%$ 代入该式,可以得出 $R_2 = \sqrt{0.2} \times 300 / (1 + \sqrt{0.2}) = 92.7\text{kbps}$ 时,接收到的失真最小.从以上论述可以看出,我们的近似最优的码率分配模型得出的结果与仿真实验的结果匹配得非常好,这说明了我们的近似最优码率分配模型的有效性.

最后,我们通过仿真实验结果来比较本文方案(P-UMDC)和其他两种非平衡多描述编码方案(I-UMDC 和 C-UMDC)在引入不同冗余度情况下的 PSNR 性能.在仿真实验里,HR 描述的帧内刷新周期设为 50 帧,码率为 100kbps,LR 描述的帧内刷新周期为 10 帧,LR 的码率为 10kbps~80kbps,Bad 状态的持续时间为 100ms.图 20 是 3 种 UMDC 方案在引入不同冗余度下的 PSNR 比较,横坐标代表 LR 描述的码率.从图 20 可以看到引入的冗余度越小,P-UMDC 的性能优势越大,这进一步说明了我们的 P-UMDC 方案优于其他两种 UMDC 方案(I-UMDC 和 C-UMDC).

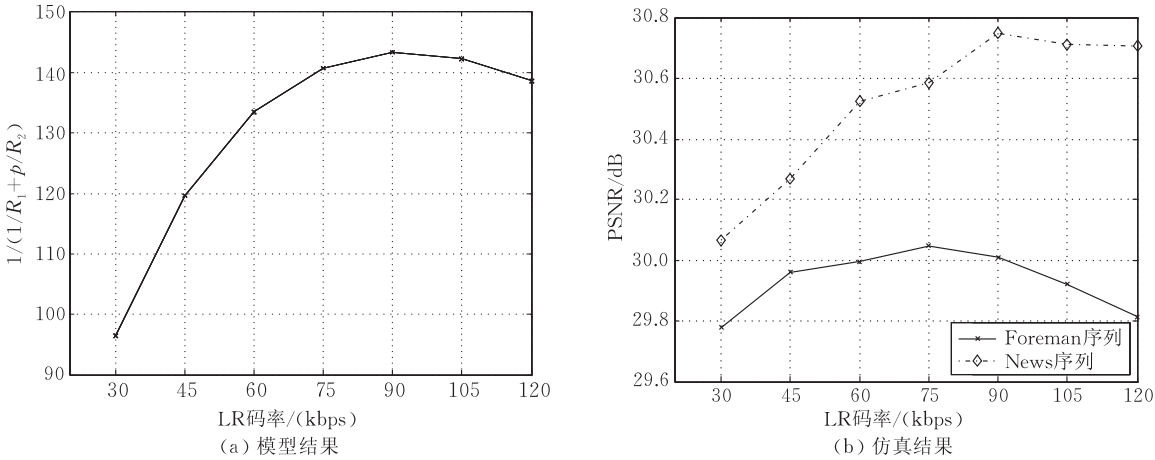


图 19 平均突发丢包长度为 1.5,包丢失概率为 20%且总码率为 300kbps 的网络状态下的模型结果和仿真结果

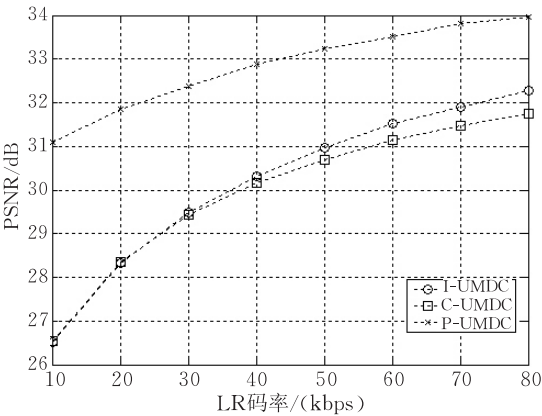


图 20 3 种 UMDC 方案在引入不同冗余度下的 PSNR 对比(Foreman QCIF 序列,平均 Bad 状态持续时间为 100ms)

6 结论及展望

本文提出了一种灵活鲁棒的非平衡多描述编码和传输方案,能够有效地适用于存在包丢失状况和严格传输时限要求的视频流媒体应用. 文章详细讨论了编解码器架构、多路径和单路径传输技术的结合以及单路径传输时借助两状态马尔可夫丢包模型和视频的率失真模型的不同描述之间码率分配的模型. 从文中叙述和实验结果可以看出,本文提出的非平衡多描述编码与其他几种常见的非平衡多描述编码方法相比,能更快地从丢包错误中恢复过来且保证更连续地播放,具有良好的鲁棒性. 同时本文提出的方案结合了多路径传输和单路径传输,而使用单路径传输的性能与使用多路径传输的性能差异不大,传输更加灵活.

尽管本文提出的非平衡多描述编码技术容错性能良好,但它的编码效率仍有待提高. 同时虽然本文

讨论了非平衡多描述编码技术与多路径传输和单路径传输技术的结合,但是如何在单路径传输和多路径传输之间自适应切换尚未在本文中考虑. 这些都将在今后的研究加以解决.

参 考 文 献

[1] Rizzo L. Effective erasure codes for reliable computer communication protocols. ACM Computer Communication Review, 1997, 27(2): 24-36

[2] Begen A C, Altunbasak Y, Ergun O, Ammar M H. Multi-path selection for multiple description video streaming over overlay networks. Signal Processing: Image Communication, 2005, 20(1): 39-60

[3] Goyal V K. Multiple description coding: Compression meets the network. IEEE Signal Processing Magazine, 2001, 18(5): 74-93

[4] Fumagalli M, Lancini R, Stanzione A. Video transmission over IP by using polyphase downsampling multiple description coding//Proceedings of the IEEE International Conference on Multimedia and Expo. Tokyo, Japan, 2001: 1095-1098

[5] Vaishampayan V A. Design of multiple description scalar quantizers. IEEE Transaction on Information Theory, 1993, 39(3): 821-834

[6] Servetto S D, Vaishampayan V A, Sloane N J A. Multiple description lattice vector quantization//Proceedings of the IEEE Data Compression Conference. Snowbird, UT, 1999: 13-22

[7] Wang Y, Orchard M T, Vaishampayan V A, Reibman A R. Multiple description coding using pairwise correlating transform. IEEE Transactions on Image Processing, 2001, 10(3): 351-366

[8] Apostolopoulos J G, Wee S J. Unbalanced multiple description video communication using path diversity//Proceedings of the IEEE International Conference on Image Processing (ICIP'01). Thessaloniki, Greece, 2001: 966-969

[9] Comas D, Singh R, Ortega A. Rate-distortion optimization in a robust video transmission based on unbalanced multiple description coding//Proceedings of the Workshop on Multimedia Signal Processing. Cannes, France, 2001: 581-586

- [10] Ekmekci S, Sikora T. Unbalanced quantized multiple description video transmission using path diversity//Proceedings of the SPIE Image and Video Communications and Processing 2003. Santa Clara, CA, 2003; 301-311
- [11] Fumagalli M, Lancini R, Tubaro S. A novel error-concealment algorithm for an unbalanced multiple description coding architecture//Proceedings of the Packet Video 2004. Irvine, CA, 2004; 21-30
- [12] Apostolopoulos J. Reliable video communication over lossy packet networks using multiple state encoding and path diversity//Proceedings of the Visual Communications Image Processing. San Jose, CA, USA, 2001; 392-409
- [13] Gogate N, Chung D, Panwar S S, Wang Y. Supporting video/image applications in a mobile multihop radio environment using route diversity and multiple description coding. IEEE Transactions on Circuits System for Video Technology, 2002, 12(9): 777-792
- [14] Mao S, Lin S, Panwar S S, Wang Y, Celebi E. Video transport over ad hoc networks: Multistream coding with multipath transport. IEEE Journal on Selected Areas in Communications, 2003, 21(10): 1721-1737
- [15] Begen A C, Altunbasak Y, Ergun O. Multi-path selection for multiple description encoded video streaming//Proceedings of the IEEE International Conference on Communication. Anchorage, Alaska, USA, 2003, 3; 1583-1589
- [16] Chakareski J, Setton E, Liang Y, Girod B. Video streaming with diversity//Proceedings of the IEEE International Conference on Multimedia and Expo. Baltimore, Maryland, USA, 2003, 1; 9-12
- [17] Vilei A, Convertino G, Oliva S, Cuppone R. A novel unbalanced multiple description scheme for video transmission over WLAN//Proceedings of the 3rd ACM International Workshop on Wireless Mobile Applications and Services on WLAN Hotspots. Cologne, Germany, 2005; 39-47
- [18] Lee Yen-Chi, Altunbasak Y, Mersereau R M. Optimal packet scheduling for multiple description coded video transmissions over lossy networks//Proceedings of the IEEE GLOBECOM'03. San Francisco, USA, 2003, 6; 3569-3573
- [19] Begen A C, Altunbasak Y, Ergun O, Ammar M H. Multi-path selection for multiple description video streaming over overlay networks. Signal Processing: Image Communication, 2005, 20(1): 39-60



LI Bin, born in 1981, Ph. D. candidate. His research interests are mainly in the areas of video compression and communication.

HUANG Feng, born in 1979, Ph. D. candidate. His re-

search interests are mainly in the areas of video coding and transmission for wireless networks.

SUN Li-Feng, born in 1972, Ph. D., associate professor. His research interests are mainly in the areas of interactive multi-view video and media streaming.

YANG Shi-Qiang, born in 1952, professor, Ph. D. supervisor. His research interests mainly include multimedia signal processing and video content analysis.

Background

Video transmission over networks is hampered by packet losses. Even a single packet loss may cause error propagation until an intra-coded frame is received. Indeed, packet losses greatly degrade the video quality. To address the issue, many methods are proposed. A general method to limit error propagation is to insert intra coding frames. However the high bitrate makes it particularly difficult to apply in many applications. Forward Error Correction (FEC) codes can also be deployed to increase the robustness of the coded bit-stream and thus reduce the length of error propagation. Nevertheless, the obvious obstacle thrown in the way of FEC codes is that its correction capability is fixed when given a predetermined redundancy. A pattern of excessive packet losses essentially results in incapability to correct errors, thus causing severe quality degradation. Another well-known error-protection method, i. e. Automatic Repeat reQuest (ARQ) technique is not suitable for live video applications because of the stringent constraints on the packet delivery deadlines. Fortunately, an alternative approach to tackle the problem of error propagation is applying Multiple Description Coding

(MDC) techniques. Attractive feature of low delay property makes MDC most promising especially for live video applications. As long as the descriptions do not encounter packet loss concurrently, an acceptable quality could be obtained.

This paper proposes a novel flexible and robust unbalanced multiple description coding and transmission scheme, which is capable of quickly recovering from packet losses and ensuring continuous playback, and further adaptive to both multiple path and single path transmission. Furthermore, it discusses the problem of rate allocation between different descriptions when only single path transmission is available and propose an approximate while efficient rate allocation scheme with the help of two-state Markov link model and a representative rate-distortion model. The experimental results demonstrate the effectiveness of our proposed scheme.

This research is supported by the National Natural Science Foundation of China under grant No. 60503063. It is also partly supported by the National Basic Research Program of China (973) under grant No. 2006CB303103.