

面向 Internet 传输的多描述可分级视频编码

朱 红 吴成柯 方 勇 王养利

(西安电子科技大学 ISN 国家重点实验室 西安 710071)

摘 要: 带宽变化和丢包错误是当前 Internet 视频传输面临的主要问题, 解决的有效途径是对信源采用可分级编码和多描述编码。研究表明, 这两种编码方法具有互补性, 联合起来使用能提供更好的“质量自适应”视频传输。传统的联合方案是用多描述编码保护可分级编码的基层信息, 该文对此进行了拓展, 提出了一种改进的多描述可分级联合视频编码方案。在该方案中, 首先用多描述编码方法处理第一级可分级编码的基层信息, 然后对得到的多描述进行第二级可分级编码。与传统方案相比, 该方案的优点是可以同时兼顾压缩视频流的效率和鲁棒性。

关键词: Internet, 视频传输, 可分级编码, 多描述编码

中图分类号: TN919.8, TP393 **文献标识码:** A **文章编号:** 1009-5896(2005)01-0112-04

Joint Multiple Description and Scalable Coding for Robust Transmission of Video over Internet

Zhu Hong Wu Cheng-ke Fang Yong Wang Yang-li

(National Key Lab. on ISN, Xidian University, Xi'an 710071, China)

Abstract Current Internet video transmission is characterized mainly by variations in throughput and packet loss. Scalable Coding(SC) and Multiple Description Coding(MDC) have been proposed as source coding techniques that are robust to such channel problems. Studies showed that combination of the two approaches(MDSC) can guarantee better “quality adaptation” video transmission. Traditional MDSC improves SC by introducing redundancy in base layer so that the chance of receiving at least one description of base layer is greatly enhanced. This idea is further expanded, and a new improved MDSC scheme is proposed in this paper. In the proposed approach, MDC is applied to the base layer of the first order SC, and then the derived descriptions are encoded by second order SCs. The advantage of this approach is that both the coding efficiency and robustness of the coding stream can be simultaneously considered.

Key words Internet, Video transmission, Scalable coding, Multiple description coding

1 引言

在 Internet 上传输视频已成为多媒体业务发展的一个方向。视频与一般数据不同, 它对带宽、时延和丢包等因素非常敏感, 但由于异构性和“尽力而为”的传输特点, Internet 无法对这些基本要求保证服务质量(QoS)。所以, Internet 视频传输面临着诸多挑战, 如网络带宽存在较大的波动, 因阻塞造成的丢包现象不可避免等。为此, 人们提出了多种基于应用层的视频传输方案^[1,2]。

目前, 解决网络带宽波动的合适方法是可分级视频编码(如 MPEG-4 FGS)^[3], 其基本原理如图 1 所示。可分级视频编码方法通常产生两个视频流: 基层码流和增强层码流。基层码流是必须传输的, 并且码率一般比较低, 增强层码流可以根据实际带宽进行任意截断。由于该方法产生的视频流可以在一个很大的码率范围内自适应调整, 因而能够适应复杂的网络带宽波动。

而解决丢包错误的有效办法是多描述视频编码^[4]。多描述方法也生成多个码流(也称作描述(MD), 一般为两个, 基本原理见图 2), 各描述的相互关系表现为: 同等重要、独立解码、相互增强。各个描述通过相互独立的信道进行传输。如果两个信道的状态都正常, 则中央解码器工作, 恢复的视

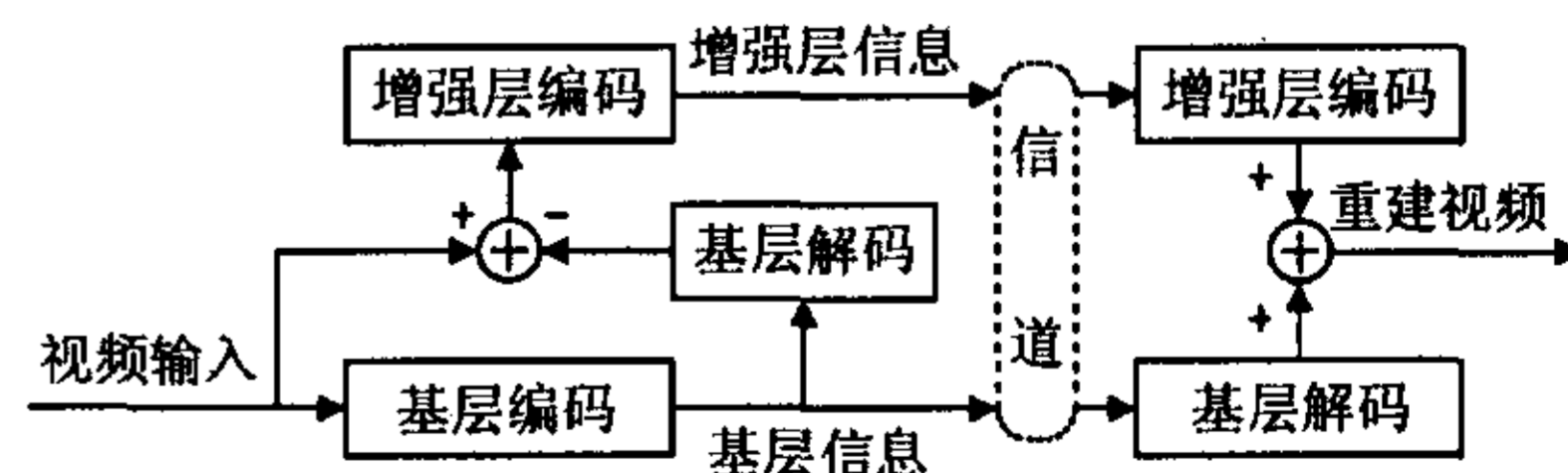


图 1 可分级视频编解码系统框图

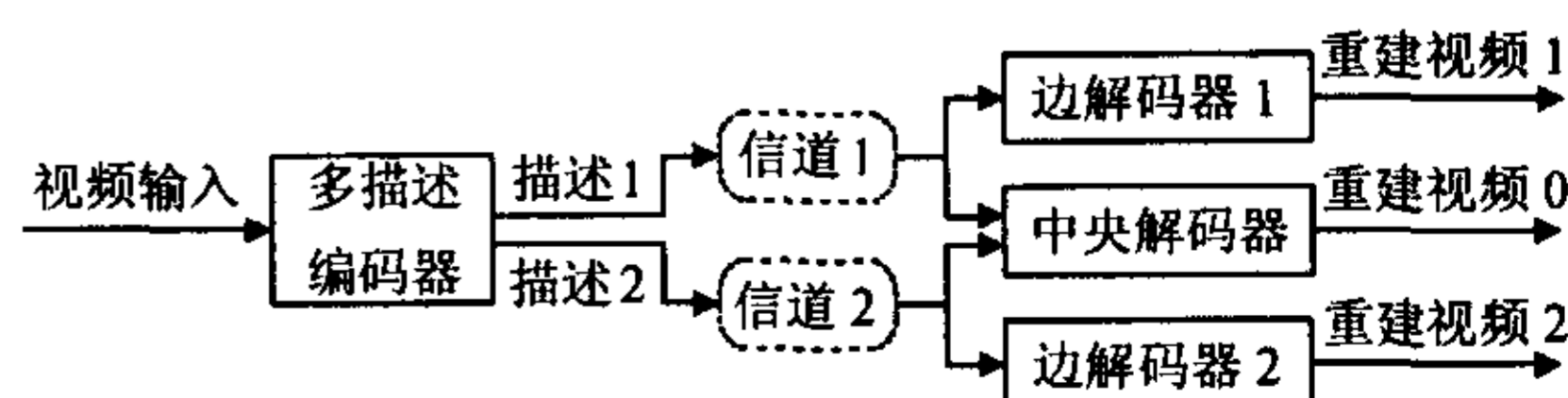


图 2 多描述视频编解码系统框图

频质量最好; 如果只有一个信道的状态正常, 则相应的边解编码器工作, 仍能恢复出可以接受的视频质量。

文献[5]详细给出了两种编码方法的性能比较结果。研究表明, 可分级视频编码适合运用在支持优先级传输机制的网络或提供不等错误保护的场合, 因为生成的码流是按重要性编排的, 基层数据比增强层数据重要, 增强层中的高比特平面数据比较低比特平面数据重要。所以说, 可分级码流的解码过程是顺序级联的, 如果基层信息在传输中由于丢包错误而造成不可恢复, 即使后续增强层数据仍被正确接收, 也不能被解码而只能丢弃。可见, 保证基层信息的可靠传输, 是可分级视频编码成功应用于当前 Internet 视频传输的关键。多描述视频编码则非常适合运用在突发丢包错误明显和“尽力而为”的 Internet。综合可分级编码和多描述编码的优点, 人们提出了多描述可分级编码(MDSC)^[6]的思想。MDSC 的一般思路是运用多描述方法保护可分级编码的基层信息^[7], 这构成了一个单向保护关系。实际上, 还可以采用可分级编码方法进一步编码各个描述, 这就构成了一个两级的交叉互补关系, 使得两种方法可以充分协作, 在提高压缩视频编码效率的同时, 仍能兼顾考虑它的鲁棒性, 这是本文突出的主要思想。

2 多描述可分级视频编码

在文献[7]中, Wang Huisheng 和 Ortega 提出了一种 MDSC 方案, 方案框图见图 3, 工作原理如下。

编码端: 用一个多描述基层编码器(MD base layer encoder)对基层信息编码, 产生两个基层描述 BL1 和 BL2, 进而多描述基层编码器模拟接收端可能的 3 种解码情况(任意一个基层描述被正确接收或两个基层描述同时被正确接收): 如果 BL1 和 BL2 同时接收, 则重建基层信息为 \hat{S} , 然后用标准的可分级编码框架 MPEG-4 FGS(增强层编码 0 模块)对原始视频 S 和 \hat{S} 的差值图像 $S - \hat{S}$ 编码, 产生增强层信息 EL0; 如果只有一个描述被正确接收, 则重建的低质量基层信息为 \hat{S}_1 或 \hat{S}_2 , 同样使用相应的 FGS 编码器(增强层编码 1, 2 模块)对 $S - \hat{S}_1$ 和 $S - \hat{S}_2$ 编码, 产生增强层信息 EL1 或 EL2。

层; (2) 如果信道的丢包率较高, 则同时发送两个基层描述, 并根据信道的变化情况决定是否发送增强层 EL0, EL1 和 EL2。

解码端: 解码系统由两部分构成, 多描述基层解码器和增强层解码器。取决于 BL1 和 BL2 的接收情况, 多描述基层解码器可以恢复出 \hat{S} , \hat{S}_1 或 \hat{S}_2 的重建信号 \tilde{S} , 同时增强层切换开关选择相应的增强层码流(EL0、EL1 或 EL2)进行解码。最后, 通过合成解码的基层信息 \tilde{S} 和增强层信息 \tilde{e} , 即为最终的视频输出。

该方案通过在视频编码的 DCT 变换域使用一个粗量化器获得基层信息, 在对下一个视频帧进行预测和编码时, 以该基层信息为参考。多描述的获得是通过将基层信息分解到 BL1 和 BL2, 分解方法: (1) 重要信息如帧内编码模式时的 DCT 直流系数和帧间编码模式时的运动矢量重复放入 BL1 和 BL2; (2) 其它 DCT 系数交替的放入 BL1 和 BL2。

该 MDSC 方案的基本思想是运用多描述编码方法优先保护可分级编码的基层信息, 这样就构成一个多描述编码支撑可分级编码的单向关系。实际上, 该方案存在需要改进之处, 分析如下。

(1) FGS 编码时, 一般假设网络能保证一个最小的传输带宽 C_{\min} , 所以 BL1, BL2 的合成码率(一般较低, 且小于 C_{\min})基本上在编码阶段就被固定了。于是, 运动补偿总是参考一个最低质量的重构基层, 因而编码效率较低, 造成 EL0, EL1 和 EL2 的码率偏高;

(2) 当网络性能较差甚至恶化(主要是丢包率上升)时, 由于缺乏保护, 增强层信息(EL0, EL1 和 EL2)被发送或成功接收的几率自然降低, 而数据量较少的基层信息由于受到多描述机制的保护, 被成功接收到的几率自然升高, 但接收端的恢复视频质量却无法再有效提高, 这就间接地弱化了多描述编码机制在抗高突发丢包错误方面的优势;

为此, 我们提出了如下改进方法, 采用多描述编码保护一个码率在一定范围内可以调节的基层视频流。也就是说, 在图 3 中多描述基层编码器的后端, 再加入一级可分级编码模块(称为第二级可分级编码, 相应地, 把原来的可分级编码称为第一级可分级编码)。详细的编码过程在第 3 节中给出。这种改进的优点分析如下:

(1) 从第一级可分级编码的角度看, 可以适当地减小基层编码的量化步长, 基层信息的码率和重建质量相应提高, 于是运动补偿更有效, 从而改善增强层的编码效率;

(2) 从多描述编码和第二级可分级编码的角度看, 在网络丢包率较高同时网络带宽波动的情况下, 既可以充分发挥多描述优良的鲁棒性保护尽量多的基层信息, 又可以适应一定的网络带宽变化;

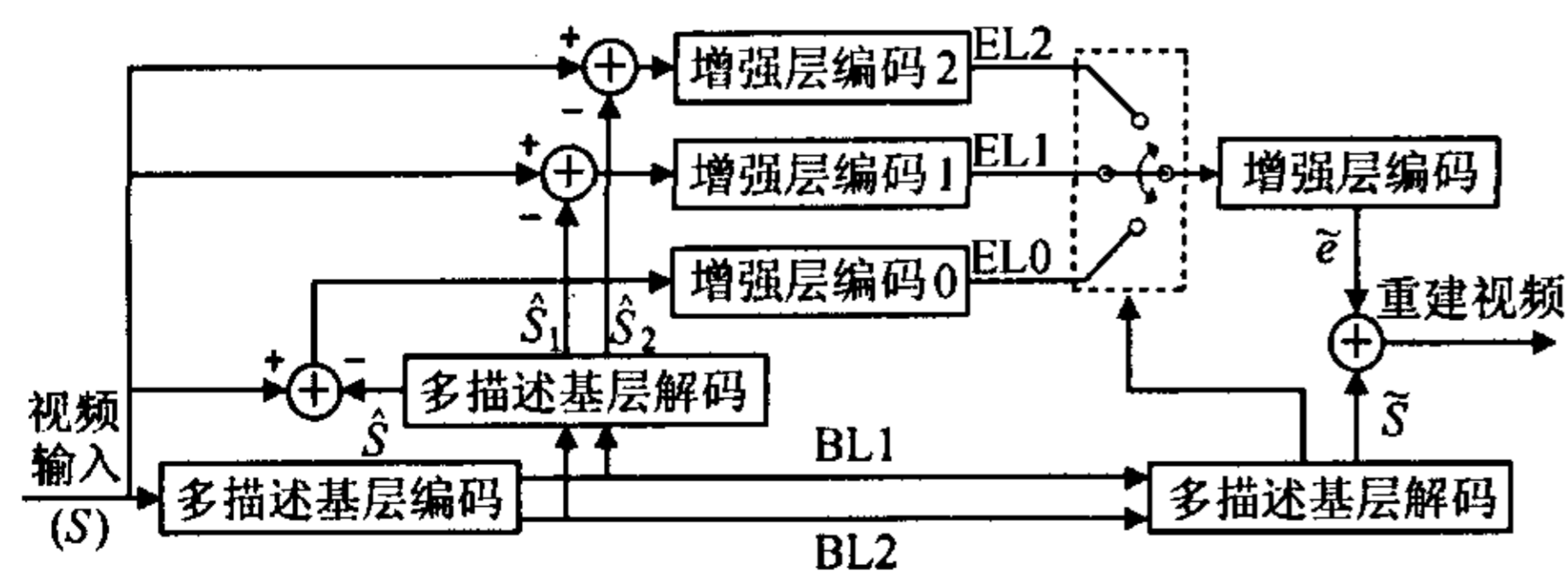


图 3 文献[7]中提出的多描述可分级视频编码方案

编码的结果是产生了 5 部分信息(描述 BL1, 描述 BL2; 增强层 EL0, EL1 和 EL2), 相应的发送策略: (1)如果信道的丢包率很低, 则只发送其中一个基层描述和相应的增强

3 改进的多描述可分级视频编码

通过以上分析, 我们知道: 首先用多描述编码方法处理第一级可分级编码的基层信息, 然后对得到的多描述进行第二级可分级编码的优点是可以同时兼顾压缩视频流的效率和鲁棒性, 真正实现可分级编码和多描述编码的优势互补。本文提出的改进多描述可分级视频编码方案的流程图如图 4 所示, 图中的“第一级 FGS 增强层编码”模块的结构以及信号名称 (BL1, BL2, \hat{S} , \hat{S}_1 , \hat{S}_2 , EL2, EL1, EL0) 的含义和相互关系与图 3 基本相同, 而 BL1_BL, BL2_BL 分别表示 BL1 和 BL2 的基层信息, BL1_EL, BL2_EL 分别表示 BL1 和 BL2 的增强层信息。编解码过程和功能叙述如下。

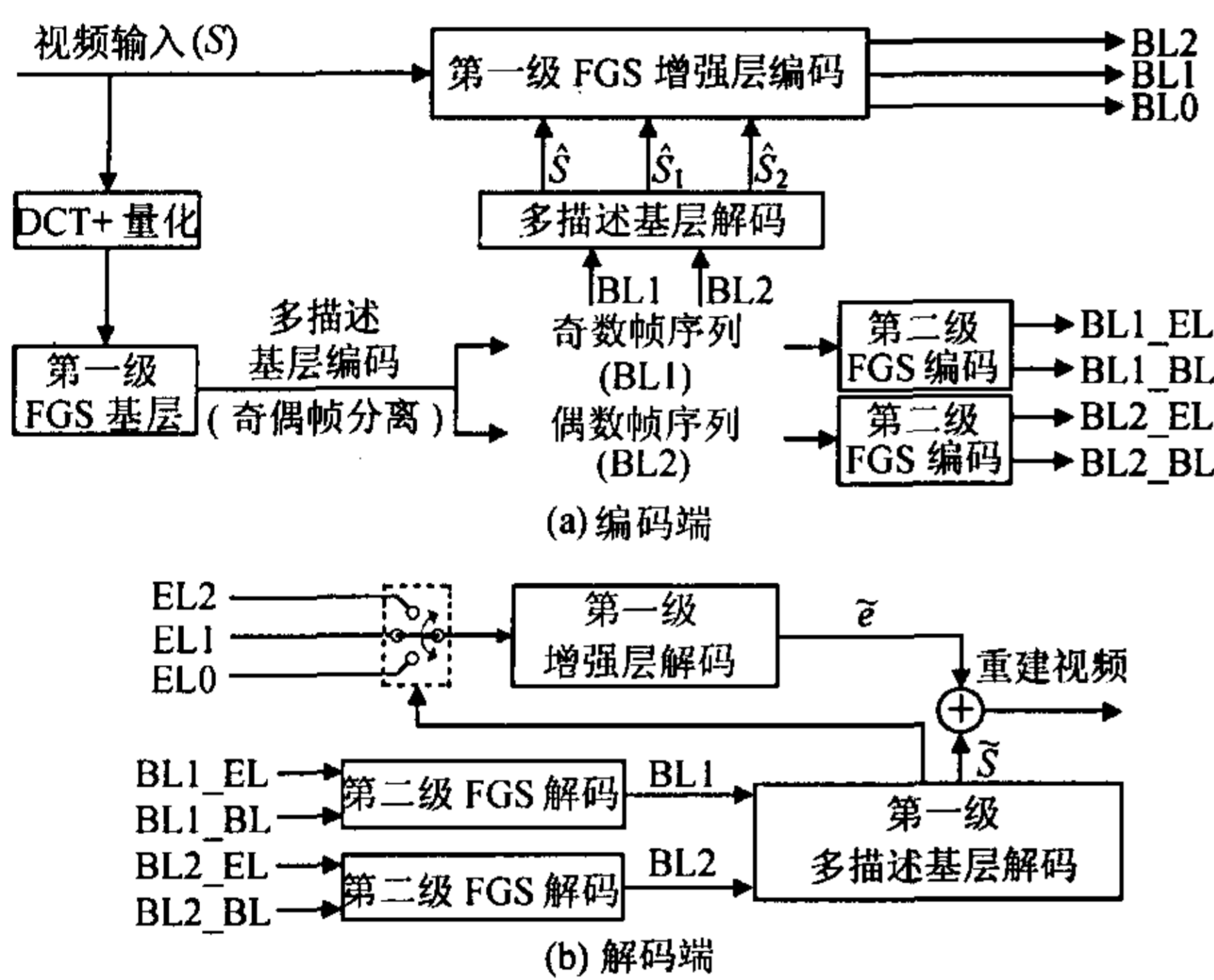


图 4 改进的多描述可分级视频编码方案

(1) 模拟 \hat{S} , \hat{S}_1 , \hat{S}_2 的 3 种接收情况, 采用 FGS 完成第一级可分级编码。仍然通过对 DCT 系数的量化方式得到第一级 FGS 的基层, 但量化参数 (QP1) 可以根据需要适当减小。减小 QP1, 可提高第一级 FGS 的编码效率, 多描述环节保护的基层信息更多, 同时由于该基层信息又被第二次分层压缩, 所以能很好地适应高突发丢包错误和各子信道的网络带宽变化; 反之, 增大 QP1, 则第一级 FGS 的编码效率降低, 多描述环节保护的基层信息减少, 如果网络丢包率很小并且可以保证信息量较少的基层被可靠的传输, 则通过相应的截取第一级 FGS 增强层码流, 就能很好地适应网络带宽的波动。可见, 通过调整 QP1, 可以兼顾考虑视频的压缩效率和对传输错误的鲁棒性。

(2) 对于多描述编码, 我们采用 H.263+视频编码标准中的视频冗余编码技术 (Video Redundancy Coding, VRC)^[8] 将编号为奇数的帧序列编码为一个描述, 将编号为偶数的帧序列编码为另一个描述。为防止帧间错误扩散, 在对两个描述进行压缩编码时, 采用独立分段预测 (Independent segment prediction)^[9] 编码方式, 即对一个描述的时间/空间预测编码被限制在本描述内进行。由于视频图像序列相邻帧之间存在

高度的时间冗余性, 所以在两个描述之间必然存在一定程度的相关性, 利用这种相关性, 在接收端可以从一个描述 (正确接收) 估计另一个描述 (未被正确接收), 从而即使在丢包率较高的网络传输环境下, 仍能保证一定的视频恢复质量。

(3) 通过对描述 BL1 和 BL2 的重建和第二次量化 (量化参数设为 QP2), 就可以实现第二次可分级编码 (仍采用 FGS 框架)。减小 QP2, 可以提高第二级 FGS 的编码效率, 随之基层信息量也增加了, 这就要求网络要保证一个最小传输带宽 C_{\min} 。实际上, 由于访问流量分布的不均匀和网络本身的异构性, Internet 的带宽波动可能导致局部时间内无法满足 C_{\min} 的条件, 这足以对恢复视频的质量造成严重影响, 如果这类事件频繁 (C_{\min} 越大, 则几率越高) 发生, 就会造成视频传输总体可靠性的降低。反之, 增大 QP2, 虽然第二级 FGS 的编码效率降低了, 却对 C_{\min} (与 QP2 成递减关系) 的要求放宽了。可见, 通过调整 QP2, 也可以在编码效率和视频传输的可靠性之间取得折衷。

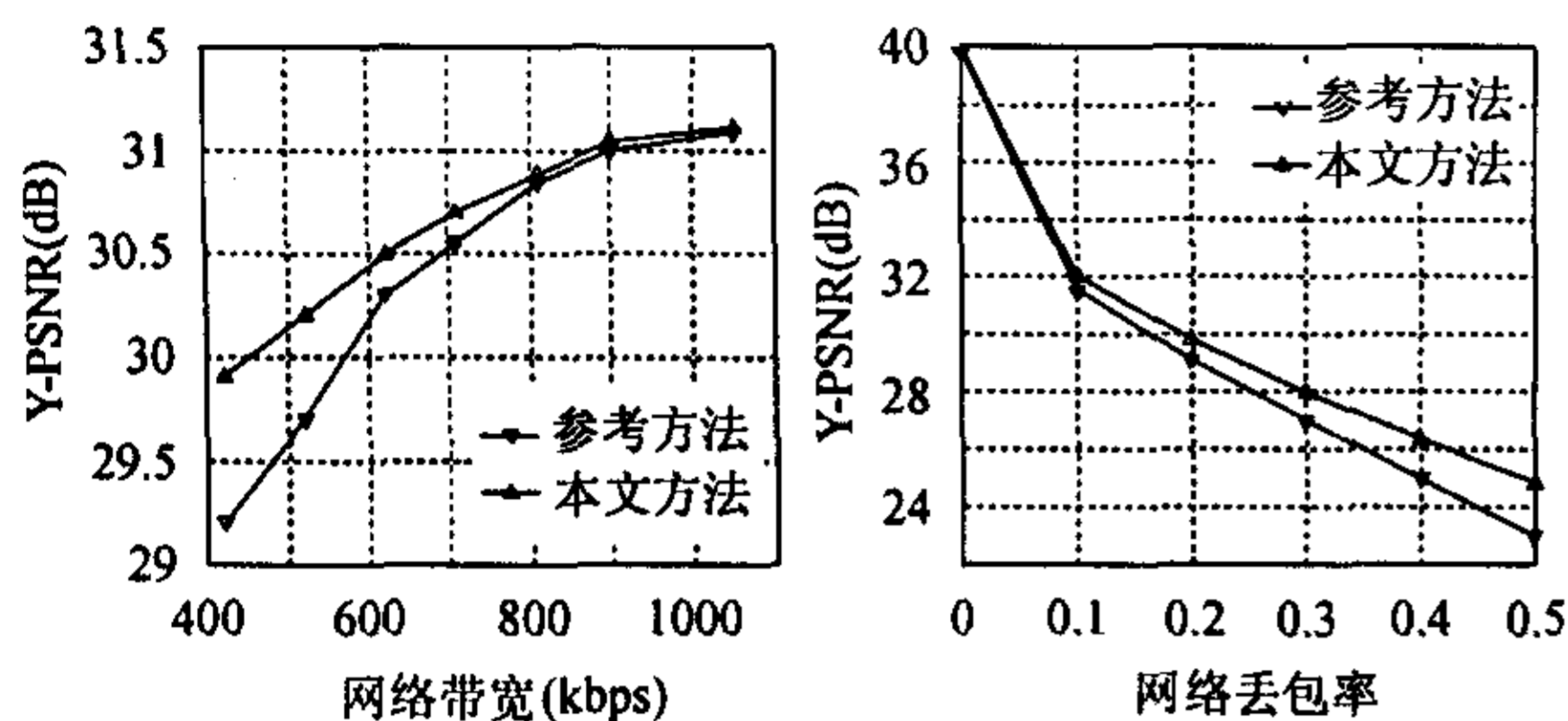
(4) 根据网络丢包率和带宽的变化情况, 可以选取适当的发送和解码策略。码流的发端发送优先顺序以及收端解码顺序依次为: ① BL1_BL (描述 1 的基层), BL2_BL (描述 2 的基层); ② BL1_EL (描述 1 增强层), BL2_EL (描述 2 增强层); ③ EL2, EL1, EL0。其中, EL2, EL1, EL0 的发端发送方法和收端切换方法与图 3 基本相同。

由于对第一级可分级编码基层的两个描述分别做了进一步的分级编码, 所以与文献[7] MDSC 方案相比, 本文方案的编码器将产生更多的基层流和增强流, 这无疑增加了编解码器的复杂度。它的优点是可以提高视频传输的灵活性和可靠性, 但一定程度上也增加了编解码器的复杂度。可见, 性能的改善是以牺牲系统的复杂度为代价的。

4 实验结果

本文使用 ns2 网络仿真器模拟 Internet 视频传输过程, 对本文方法和参考文献[7]中提出的多描述可分级视频编码方法 (以下称“参考方法”) 的性能做了比较实验。包的大小定为 512 byte。使用标准图像序列“Stefan”进行实验研究。图像序列为 CIF 格式, 4: 2: 0, 250 帧, 25 帧/秒。奇数帧和偶数帧序列中的第一帧编码为 I 帧, 其余全为 P 帧。参考方法基层编码时的 QP 和本文方法的 QP1, QP2 的选择方法是: (1) 编码过程中保持不变且使最终的码率尽量接近预期值; (2) $QP1 \leq QP$, 并相应的确定 QP2 的值, 使第一个条件得到满足。使用亮度分量的峰值信噪比 (PSNR) 做为衡量重建图像质量的客观标准。因为 Internet 网络丢包错误图样及其对压缩视频流的影响具有随机性, 所以每次实验均重复 20 次, 然后取实验结果的平均值。考虑两种典型情况: (1) 网

络丢包率一定(15%),重建视频质量(Y-PSNR)随网络带宽的变化情况,仿真和比较结果参见图 5(a),而编码参数随网络带宽的变化情况见表 1 所示(表中给出了 3 组典型数据);(2)网络带宽一定(716 kbps),重建视频质量随丢包率的变化情况,仿真和比较结果参见图 5(b),而编码参数随丢包率的变化情况见表 2 所示(表中给出了 3 组典型数据)。



(a) 实验条件:网络丢包率为 15% (b) 实验条件:网络带宽为 716 kbps

图 5 本文方法与参考方法的性能比较

表 1 丢包率 15%, 编码参数随网络带宽的变化

网络带宽(kbps)		420	518	623
编码	QP	21	17	14
参数	(QP1,QP2)	(19, 13)	(14, 12)	(11, 10)

表 2 网络带宽 716 kbps, 编码参数随丢包率的变化

丢包率(%)		0	10	20
编码	QP	12	12	12
参数	(QP1,QP2)	(12, 0)	(9, 18)	(9, 23)

从图 5(a)和表 1 可以看出,在网络丢包率一定的条件下,如果网络带宽较窄,本文方法可以通过减小 QP1,使得多描述机制可以保护更多的第一级 FGS 的基层信息,并且由于引入了第二级 FGS 编码对多描述进行更细的分层,可以根据需要调整被保护信息的码率,从而比参考方法性能要好,这从低带宽部分的性能相差较大即可看出。另一方面,随着带宽的增加,两种方法的编码重心均趋向于可分级编码,两种方法的性能渐趋接近,但本文方法的性能仍稍优于参考方法。而从图 5(b)和表 2 可以看出,在网络带宽一定的前提下,随着丢包率的增加,两种方法的编码重心均趋向于多描述编码,由于本文方法可以通过调整 QP1 和 QP2 以保护更多的基层信息,所以在丢包率较大时,本文方法性能明显优于参考方法。

5 结论和进一步的工作

多描述和可分级两种视频编码方法具有互补性,联合起来使用能提供更好的“质量自适应”视频传输。本文对文献[7]的多描述可分级视频编码方法进行了拓展,提出了一种改进的多描述可分级视频编码方案。该方案除了用多描述编码

方法处理第一级可分级编码的基层信息外,还对得到的多描述进行了第二级可分级编码,该方案在网络丢包率较高或网络带宽波动较大的情况下,既可以充分发挥多描述优良的鲁棒性保护尽量多的基层信息,还可以适应一定的网络带宽变化,因而更适合面向 Internet 的视频传输。

进一步的工作需要研究根据网络带宽和丢包率的变化自适应调整和优化编码参数(QP1 和 QP2)的方法,使得在多描述与可分级两种编码模块之间合理分配码率资源,达到压缩视频流的整体性能最优。

参考文献

- [1] Wang Yao, Zhu Q F. Error control and concealment for video communications: a review. *Proc. IEEE*, 1998, 86(5): 974 - 997.
- [2] Wu D P, Hou T, Zhu W, Zhang Y Q, Peha J. Streaming video over the Internet: Approaches and directions. *IEEE Trans. on Circuits and Systems for Video Technology*, Special Issue on Streaming Video, 2001, 11(3): 282 - 300.
- [3] Li W. Overview of fine granularity scalability in MPEG-4 video standard. *IEEE Trans. on Circuits and Systems for Video Technology*, 2001, 11(3): 301 - 317.
- [4] Goyal V. Multiple description coding: compression meets the network. *IEEE Signal Processing Magazine*, 2001, 18(5): 74 - 93.
- [5] Lee Y C, Kim J H, Mersereau R M, Altunbasak Y. Layered coded vs. multiple description coded video over error-prone networks. *EURASIP Signal Processing: Image Communication*, May 2003, vol.18: 337 - 356.
- [6] MPEG2003/N5540, Applications and requirements for scalable video coding. Pattaya, March 2003.
- [7] Wang Huisheng, Ortega. Robust video communication by combining scalability and multiple description coding techniques. in *Proceeding of SPIE*, San Jose, CA, Jan. 2003.
- [8] Wenger S, Knorr G, Ott J, Kossentini F. Error resilience support in H.263+. *IEEE Trans. on Circuits and Systems for Video Technology*, 1998, 8(7): 867 - 877.
- [9] Wang Yao, Ostermann J, Zhang Y Q. *Video Processing and Communications*. New Jersey: Prentice Hall, 2002: 472 - 511.

朱红: 男, 1971 年生, 博士生, 主要研究方向: 视频编码与无线/Internet 传输。

吴成柯: 男, 1938 年生, 教授, 博士生导师, 目前主要从事计算机视觉、计算机图形学、图像处理和图像通信的教学和科研工作。

方勇: 男, 1979 年生, 博士生, 主要研究方向: 视频编码与无线/Internet 传输。

王养利: 男, 1965 年生, 副教授, 博士, 主要研究方向: 数字视频传输与处理。