

文章编号:1004-4213(2010)05-0935-5

基于帧的多描述视频编码及其错误隐藏*

励金祥¹, 郑淑红², 杨任尔¹, 邵洁¹

(1 宁波大学 信息科学与技术学院, 浙江 宁波 315211)

(2 浙江育英职业技术学院, 杭州 310018)

摘要:为了解决视频流在不可靠网络上的错误传播问题,使用基于帧的多描述视频编码,提出采用预测的预处理和后处理过程方案,实现了描述间的冗余插入;实现了几种不同复杂度,不同性能的错误掩盖算法以适应多样化的网络传输环境.仿真实验结果表明,这种编码系统能有效控制视频流的错误传播,并且编码后的数据流更能适应各种网络传输状况.

关键词:视频流;多描述视频编码;多路径传输;错误隐藏

中图分类号:TN919.81

文献标识码:A

doi:10.3788/gzxb20103905.0935

0 引言

视频流在不可靠网络上传输,由于网络带宽资源有限,网络传输状况不稳定,存在数据包丢失或误码等问题,这会引起视频的整帧丢失,其影响会传递到后续帧,进而导致整个视频序列的质量严重恶化.目前往往采用错误掩盖技术解决视频传输错误问题^[1].错误掩盖技术利用了视频信号在时间和空间上的相关性,可在一定程度上恢复丢失的视频图像信息,但通常这些算法都假设一帧的大部分宏块被成功接收到^[2],只有一个或者少量宏块的数据丢失,因此它并不能真正解决整帧丢失的问题.陈建乐等在文献[3-4]中提出基于光流场不变性的整帧图像错误掩盖算法,该算法使用有理插值对前帧图像的运动矢量场做预处理来得到稠密的运动矢量场,然后采用前向投影技术将前一帧图像内容沿着运动矢量方向投影到当前图像的对应位置上;曹宁等在文献[5-6]中也提出几种恢复整帧丢失视频图像的算法,这些算法都假设前后帧被成功接收到,但由于视频压缩编码采用帧间编码,由于参考帧丢失,即使当前帧的信息被完整接收到,也不能解码恢复当前视频图像,因此这类方法也不能根本上解决视频传输的错误传播问题.

为此,本文提出采用基于帧的多描述视频编码算法,采用多路径传输方式^[7],能较好地解决视频流传输过程中因整帧丢失而引起的视频序列质量严重恶化的问题.此外,为了适应多样化的网络传输环

境,系统实现了几种不同复杂度,不同性能的解码算法以达到不同的错误掩盖效果.

1 多描述编码介绍

多描述编码方法采用多个相互独立但同时又具有一定相关性的子信号(描述)来刻画同一视频信号,其中各个描述可以分别独立解码获得重建值.系统将这些描述独立编码后再通过多路径信道传输到客户端,如果某些信道的数据包丢失或出错,接收端的解码器仍可从各个正确接收到的描述中恢复出视觉上可接受的重建信号,收到的描述个数越多,重建信号的效果越好.

多描述编码技术通过引入描述间的冗余来提高传输鲁棒性.具有两个描述的多描述编码器可由 5 个参量来刻画:两个描述的码率 R_1 和 R_2 ,两个描述都收到情况下的失真率 D_0 (即中央通道的失真率)以及只收到单个描述情况下的失真率 D_1 和 D_2 . 设 $R=R_1+R_2$ 为总码率, $D=(D_1+D_2)/2$ 为只收到一个描述时的平均失真率(即边通道的失真率), R^* 为获得同样失真率 D_0 的常规编码(实验中称为“单通道”)所需的参考码率,则多描述编码的冗余即为 $\rho=R-R^*$. 对于固定的 D_0 (或 R^*),多描述编码的目标是在给定冗余度 ρ 的情况下,使边沿失真 D 达到最小.而为适应网络环境和服务质量要求的变化,多描述编码器还应能动态调整描述间的冗余度,当网络传输差错率低时,减小 ρ 的值可以提高编码效率,反之,网络传输差错率大时,提高 ρ 的值以保证传输的鲁棒性^[8].

多描述编码方法的关键技术是如何产生平衡的若干个数据流,最经典的方法是 Vaishampayan 等提出的基于标量量化多描述编码方法(Multiple

*国家自然科学基金(60672063)、浙江省教育厅科研项目(Y200909912、Y200906750)资助

Tel:0574-66098738

Email:lijinxiang@nbu.edu.cn

收稿日期:2008-12-31

修回日期:2009-04-02

Description Scalar Quantizers, MDSQ)^[8],其需要设计一个复杂的标号分配函数,对信源进行不同准确度的量化,单个描述进行大步长的量化,而多个描述结合时则可得精细的量化.其他的还包括基于对变换的多描述编码方法^[9](Pairwise Correlating Transform, PCT),基于空间扩展等的多描述编码^[10].这类基于像素或者基于块的多描述编码方法具有不同的编码复杂度、编码效率和描述间冗余调节控制灵活性,但是它们与目前常用的图像或视频压缩标准都不相兼容,即集成后的复合编码系统已改变了原来压缩编码内部算法的结构.

本文采用的基于帧的多描述编码,与上述经典的基于像素或者基于块的多描述编码方法相比,其优势表现为,首先能较有效地解决因整帧丢失而引发的视频流错误传播的问题.其次能与现有复杂的压缩编码标准相兼容,且编码算法复杂度低.

2 基于帧多描述视频编码及其冗余度插入

典型的视频传输系统中,视频数据包通过 RTP/UDP/IP 协议传送.在实际应用中,由于以太网的 MTU 大小限制,数据包大小一般在 1 000 字节左右,至多不超过 1 500 字节;无线网络中情况也是类似的.在 IEEE802.11 标准中,一个数据帧有 34 字节的帧头信息,净荷大小在 0~2 312 字节之间可变,为了保证传输效率,净荷大小一般需大于 500 字节.在 3G 环境一个典型应用:QCIF, 10 frame/s, 64 kbit/s信道传输,每个编码后的视频帧平均大小为 800 字节.那么在这样的应用下,一个编码后的视频帧用一个 RTP 包来传送,一个 RTP 包丢失将会导致整个视频帧的丢失^[5].由于视频编码采用帧间编码,那么在后续帧重建过程中,由于参考帧丢失,即使当前帧宏块的残差以及运动矢量信息被完整接收到,也无法正确地恢复后续帧的信息,在传统方法中,一帧的丢失将使后面的帧冻结,直至收到新 I 帧,这会导致视频序列质量严重恶化.如图 1.

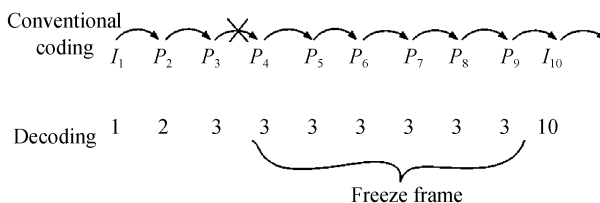


图 1 常规编码引起的错误传播
Fig. 1 Error propagation of video stream caused by conventional coding

为了解决这个问题,John Apostolopoulos 在文献^[11]提出基于帧的多描述视频编码的初步思路,

它的基本思想是采用多路径传输技术,将原始信号编码序列分成奇数帧与偶数帧两个描述.在通过网络多路径传输过程中,其中一部分的丢失将不影响另一部分的成功传输,而丢失部分将利用收到的码流进行重建,如图 2.

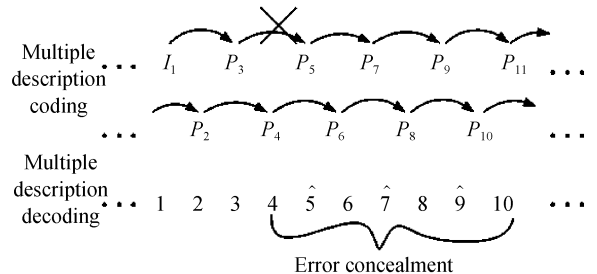


图 2 基于帧的多描述视频编码有效控制视频流错误传播
Fig. 2 Frame-based multiple description video coding control error propagation effectively

这种多描述编码初步方案实际是通过时域上的下抽样方法实现的,并不能实现描述间冗余的调节控制,而多描述编码器应能动态地调整描述间的冗余度,以适应网络环境和服务质量要求的变化.为了解决这个问题,作者在文献^[7]中提出了基于预处理和后处理多描述图像编码冗余插入方案,在此把这种算法原理推广到多描述视频编码中.其基本思路:为了在两个描述间插入一定冗余,在实现基于帧的多描述编码过程中,在每个描述中插入额外的增强帧,为了有效增加描述间的冗余,可以选择那些估计值与原始值绝对差和比较大的帧为增强帧,其编码系统的原理如图 3.

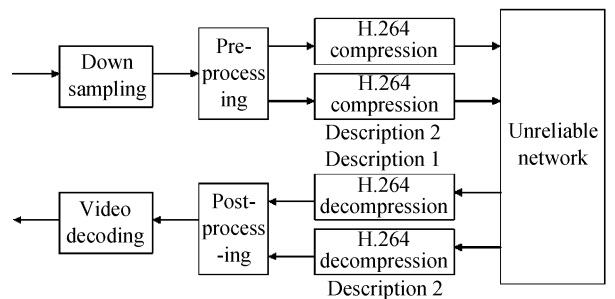


图 3 基于帧的多描述视频编码系统
Fig. 3 Frame-based multiple description video coding system

其中基于预测的预处理方法如图 4.对于原始视频 $f_0(x, y, t)$,其中 $x=1, 2, \dots, r, y=1, 2, \dots, c, t=0, 1, \dots, n-1$,首先采用在时间域上的下抽样方法把它拆分成两个描述,其中描述 1 为偶帧 $f_1(x, y, t)(x=1, 2, \dots, r, y=1, 2, \dots, c, t=0, 1, \dots, n/2-1)$,描述 2 为所有的奇帧 $f_2(x, y, t)(x=1, 2, \dots, r, y=1, 2, \dots, c, t=0, 1, \dots, n/2-1)$.对于描述 1 中任何帧 $f_1(x, y, t)$,使用插值方法由 $f_2(x, y, t-1), f_2(x, y, t)$ 估计出它的预测值 $\hat{f}_1(x, y, t)$,同样地也

由 $f_1(x, y, t)$ 预测估计出描述 2 的值 $\hat{f}_2(x, y, t)$, 然后对 $f_1(x, y, t)$ 与 $\hat{f}_1(x, y, t)$ 以及 $f_2(x, y, t)$ 与

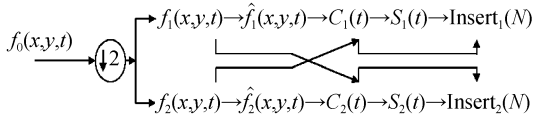


图 4 多描述视频编码的预处理过程

Fig. 4 Pre-processing of multiple description video coding

$\hat{f}_2(x, y, t)$ 分别进行比较, 计算出每帧的绝对差 $C_1(t)$ 和 $C_2(t)$, 其中

$$C_1(t) = \sum_{x=0}^r \sum_{y=0}^c |f_1(x, y, t) - \hat{f}_1(x, y, t)| \quad (1)$$

$$C_2(t) = \sum_{x=0}^r \sum_{y=0}^c |f_2(x, y, t) - \hat{f}_2(x, y, t)| \quad (2)$$

接着对所有的 $C_1(t)$ 和 $C_2(t)$ 分别进行排序产生 $S_1(t)$ 和 $S_2(t)$ (其中 $t \in [0, n/2]$)。假如要插入 N 帧增强图像, 可由 $C_1(t)$ 和 $S_1(t)$ 结合得出绝对差最大 N 帧的帧号信息 $\text{Insert}_1(N)$, 用同样的方法得到 $\text{Insert}_2(N)$ 。至此, 描述 1 可由原图像的偶帧以及 $\text{Insert}_1(N)$ 中指定奇帧形成, 描述 2 也用类似的方式产生。在每个描述中, 插入的这些额外的帧号信息在传输时作为附带信息, 携带在每个描述的数据包中。

3 解码错误掩盖的多种算法

视频流在传输过程中, 由于网络信道的不可靠性, 部分描述可能会丢失; 此外因各个信道是相互独立的, 同时发生传输失败的可能性极小, 那么就不用考虑二个描述同时没有接收到的这种可能性, 即总存在成功传送的描述。而接收端只要有描述收到, 就能重建不同视觉效果图像。

在接收端, 首先对接收到的视频流进行压缩解码, 然后分情况进行多描述解码。如果两个描述都收到, 那么对于那些在两个描述中共同接收的帧, 其像素值可取这两帧的平均数, 其它的帧进行按序合并, 并能得到整个视频序列。如果只收到一个描述时, 使用这个描述的附带信息 $\text{Insert}_1(N)$ 或者 $\text{Insert}_2(N)$, 把那些不存在的帧的像素值由另一个描述中与其相邻帧的像素值进行插值估计。插值方法选用可以根据接收端设备的性能, 对于低性能接收设备, 可采用简单快捷的前向复制或后向复制收到的帧, 对于一般性能的, 使用线性插值方法恢复视觉效果可接受的图像; 对于高性能接收设备可采用运动估计插值方法^[12], 恢复出高质量的视频图像。

对于前面三种插值方法实现过程: 设共有 $2n$ 帧序列, 其中描述 1 为偶数帧序列, 表示为 $f_1(x, y, t)$, 描述 2 为奇数帧序列, 表示为 $f_2(x, y, t)$, $t=0,$

$1, \dots, n-1$ 。又设当描述 1 或描述 2 丢失时, 由另一描述估计出来的值分别为 $\hat{f}_1(x, y, t)$ 和 $\hat{f}_2(x, y, t)$, 则

$$\hat{f}_1(x, y, t) = \begin{cases} f_2(x, y, t) & t=0 \\ \omega f_2(x, y, t-1) + (1-\omega) f_2(x, y, t) & \text{其它} \end{cases} \quad (3)$$

$$\hat{f}_2(x, y, t) = \begin{cases} f_1(x, y, t) & t=n/2-1 \\ \omega f_1(x, y, t+1) + (1-\omega) f_1(x, y, t) & \text{其它} \end{cases} \quad (4)$$

式中的 ω 是插值算法的加权系数, 其取值范围为 0 与 1 之间, 当 $\omega=1$ 时, 解码实现了前向复制错误掩盖, 当 $\omega=0$ 时, 解码实现了后向复制的错误掩盖, 而当 $\omega=0.5$ 时, 解码实现了普通的前后帧双向线性插值的错误掩盖。

本文三种方法对于基本静止的视频序列能取得较好的错误隐藏效果, 但对运动强度较大的视频序列却会产生巨大的失真。为减少运动产生的影响, 系统实现基于帧间运动估计的插值方法, 其实现的步骤:

Step1 计算运动矢量(图 5): 把 $n-1$ 帧作为参考帧, 对 $n+1$ 帧进行分块后, 对每块进行运动估计, 其中块的坐标为 (x, y) , 计算得到运动矢量为 $(\Delta x, \Delta y)$;

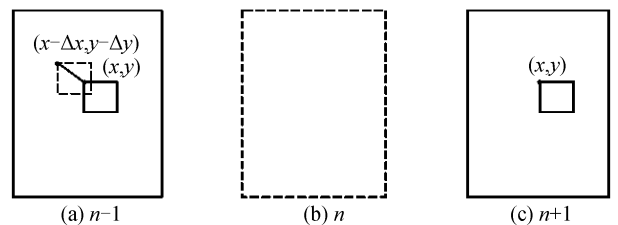


图 5 $n+1$ 帧和 $n-1$ 帧之间的运动矢量

Fig. 5 Motion vector between $(n+1)$ th frame and $(n-1)$

Step2 进行双向运动矢量估计(图 6): 对 n 帧进行分块(其块坐标为: (x, y) 。相对于 $n-1$ 帧, 它的运动矢量估计为 $(\Delta x/2, \Delta y/2)$, 而相对于 $n+1$ 帧, 其运动矢量为 $(-\Delta x/2, -\Delta y/2)$ 。

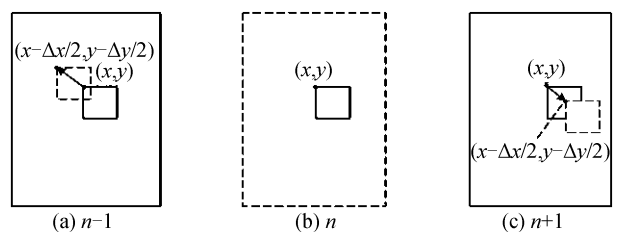


图 6 n 与前后两帧之间的运动矢量

Fig. 6 Motion vectors between n th frame and $(n-1)$ th, $(n+1)$ th frame

Step3 进行双向运动补偿: 对 n 帧中块 (x, y)

的某一位置 (i, j) 的像素估计值为

$$\hat{f}(i, j, n) = (f(i + \Delta x/2, j + \Delta y/2, n-1) + f(i - \Delta x/2, j - \Delta y/2, n+1))/2 \quad (5)$$

式中 $f(i, j, n-1)$ 、 $f(i, j, n+1)$ 为 $n-1$ 、 $n+1$ 帧相应位置的像素值。

4 实验结果及分析

4.1 基于帧多描述视频编码鲁棒性及其冗余插入

实验采用 100 帧 CIF 格式的 Foreman 和 Vectra 序列, 帧率都为 30 Hz, 使用 H. 264 压缩工具, 其中帧间编码的模式没有 B 帧, 参考帧为 I 帧, 运动估计的搜索范围为 16×16 像素点, 熵编码是基于上下文的算术编码。

实验分别对未插入增强帧、插入 8 帧、12 帧、18 帧的几种情况进行比较。实验中需要分析的“单通道”是指对原图像直接进行常规的压缩编解码的情况; “双通道”是指将原视频按“常规编码”编码后, 把数据流在两条不同的信道中重复传输时的情况; “边沿通道”为接收端只收到一个描述时的情况; “中央通道”为同时接收到两个描述时的情况, N 为预处理过程中插入的增强帧的帧数。

图 7 是 Foreman 和 Vectra 视频序列在不同的 N 值情况下, 边沿通道的率失真函数的比较。从图

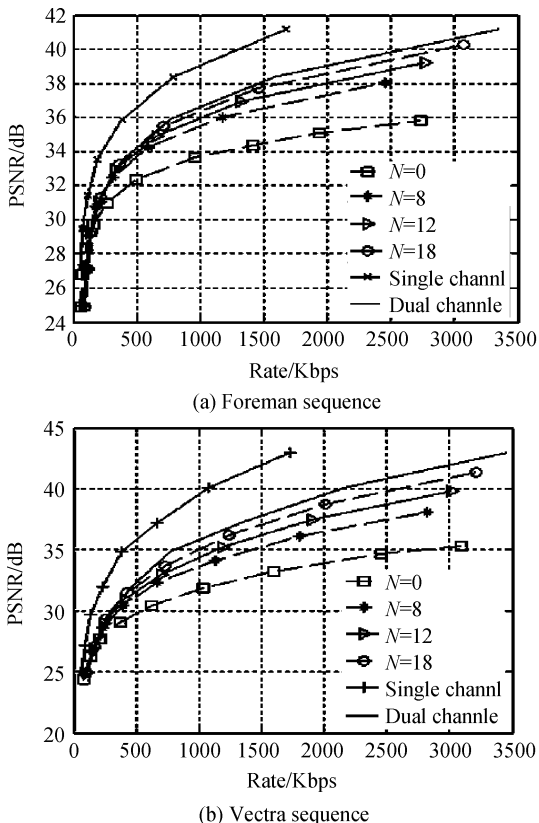


图 7 Foreman 和 Vectra 边沿解码器率失真函数
Fig. 7 Rate-Distortion functions of side decoder for Foreman and Vectra sequence

中可以观察到不管有无插入增强帧, 边沿通道的信噪比能保持一定的水平, 这说明这种多描述编码鲁棒性, 因为边沿通道重建的视频图像相当于丢包率达到 50% 下重建图像, 这也说明它能有效控制视频流的错误传播。另外, 随着插入帧数的增加, 边沿通道的编码效率也提高了, 也就是意味着多描述编码器抗干扰性能提高了。

图 8 是 Foreman 和 Vectra 视频序列在不同的 N 值情况下, 中央通道的率失真函数。从图中可以看出, 无论什么情况, 插入冗余后的多描述的中央信道要比双通道中传输的效果要好, 这说明这些多描述编码方案的有效性, 否则图像经过压缩编码, 只要在两个不同的信道传送重复的数据。另外随着 N 的增加, 相应的中央通道的编码效率也降低。

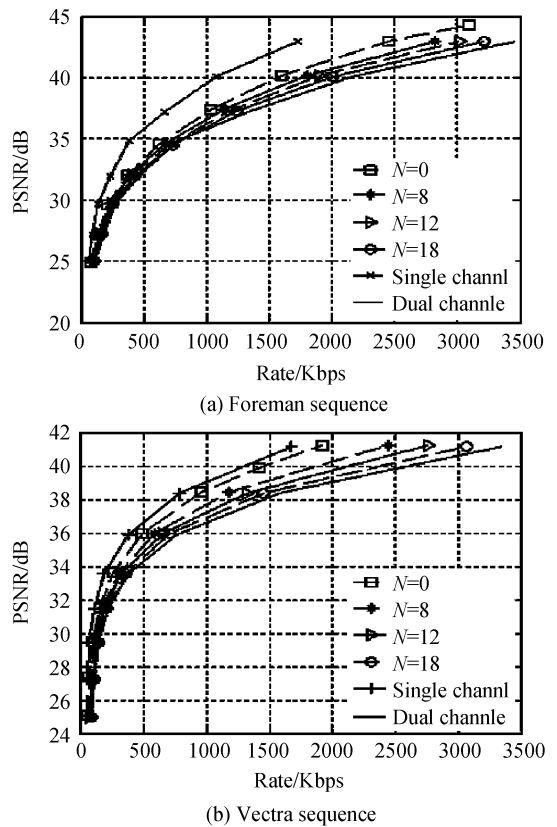


图 8 Vectra 中央解码器率失真函数
Fig. 8 Rate-distortion functions of central decoder for Foreman and Vectra sequence

结合图 7、图 8 可以看出, 这种编码方法能在编码的有效性与传输的鲁棒性之间进行折衷控制, 插入的帧数越多, 中央通道的编码效率降低, 但是边沿通道的编码效率提高, 编码器抗干扰性能提高, 即传输的鲁棒性提高了。因此这种多描述编码器可以根据不同的网络状况实现灵活的冗余度控制: 如果在网络带宽比较大、数据也较易丢失的情况, 可以在描述间插入较多的冗余以提高传输的鲁棒性; 反之, 在网络带宽较小、数据不易丢失的情况下, 可插入较少

的冗余以提高视频传输效率。

4.2 解码错误掩盖的多种算法

实验采用 200 帧 CIF 格式的 Foreman 和 Vectra 序列,使用 H. 264 压缩工具,量化步长为 28,而其中帧间编码的模式、参量选定,熵编码方式与上面的实验相同。另外实验中假设只收一个描述,另外一个描述(100 帧)由收到的描述,采用运动估计、线性插值、前向复制、后向复制四种错误掩盖的解码算法估计得到,其解码性能如图 9。

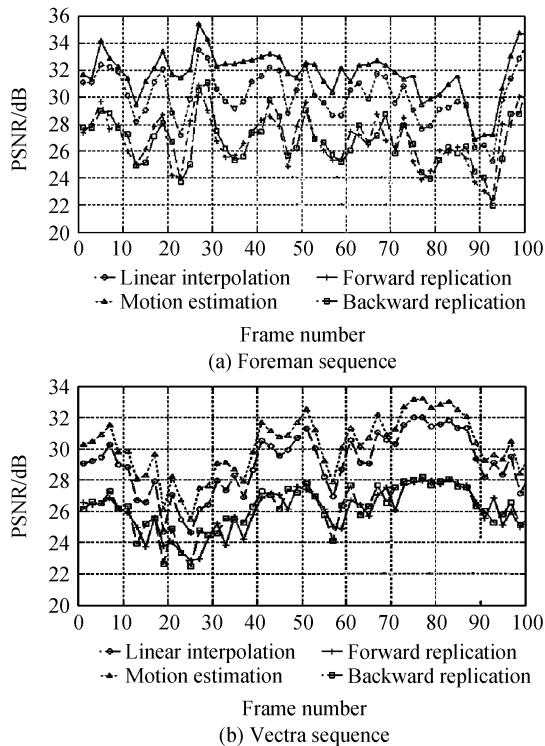


图 9 Vectra 序列各种解码算法的重建视频质量对比
Fig.9 Comparison of reconstruction quality of various decode algorithms for Foreman and Vectra sequence

从图 9 可以看出,采用基于运动估计错误掩盖的解码算法,解码后视频图像的信噪比远远要比其它三种方法高,线性插值方法的也比前向复制、后向复制好,而后面的两种方法之间不能比较,这与前后帧之间的相关性有关系,而这种相关性也是随机的。

对于上述几种解码方法,接收设备可以根据本身的性能选择解码算法。如果接收设备是高性能计算机,那么可以选择基于运动估计的错误掩盖解码算法,解码器可重建高质量的视频图像;而对于接收设备是手机来说,因其电池性能低,计算处理能力低,那就应该采用线性插值算法,甚至复杂度更低的前向或者后向复制方法进行错误掩盖的解码算法,解码器能重建视觉效果能接受的视频图像。

5 结论

视频在网络中传输,因网络传输带宽不稳定,出

现丢包、延时等问题,导致整个视频序列质量严重恶化,本文采用基于帧的多描述视频编码系统,能有效地解决因整帧丢失的错误传播问题。同时为了适应网络环境和服务质量要求的变化,使得这种多描述编码应能动态调整描述间的冗余度,系统采用预测的预处理和后处理过程方案实现了描述间冗余的插入。另外针对只收到部分视频信息的情况,解码器实现了几种不同复杂度、不同性能的错误掩盖解码方法,这种解码算法的可选择性使得编码系统更能适应多样化的网络传输环境。

参考文献

- [1] KUNG Wei-ying, KIM Chang-su, KUO C C J. A spatial-domain error concealment method with edge recovery and selective directional interpolation [A]. Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing [C]. Hong Kong, 2003, 5: 6-10.
- [2] ZHANG Rong-fu, ZHOU Yuan-hua. An error concealment method for video communications considering both frequency and spatial domains [J]. *Journal of Shanghai Jiaotong University*, 2004, 38(4): 606-609.
张荣福,周源华. 视频通信中频域和空域相结合的错误掩盖技术[J]. 上海交通大学学报, 2004, 38(4): 606-609.
- [3] CHEN Jian-le, LIU Ji-lin, CHEN Guo-bin, et al. An error concealment algorithm for whole frame loss [J]. *Journal of Image and Graphics*, 2007, 12(10): 37-42.
陈建乐,刘济林,陈国斌,等. 一种针对整帧图像丢失的错误隐藏算法[J]. 中国图象图形学报, 2007, 12(10): 37-42.
- [4] BELFIORE S, GRANGETTO M, MAGLI E, et al. Concealment of whole-frame losses for wireless low bit-rate video based on multiframe optical flow estimation [J]. *IEEE Transactions on Multimedia*, 2005, 7(2): 316-329.
- [5] CAO Ning, HU Jian-rong, MA Yin-song. Error concealment algorithm for entire frame loss based on optical flow estimation [J]. *Journal on Communications*, 2007, 28(5): 127-140.
曹宁,胡建荣,马银松. 基于光流估计的整帧恢复算法[J]. 通信学报, 2007, 28(5): 127-140.
- [6] CHEN Yu, YU Kim, LI Jiang, et al. An error concealment algorithm for entire frame loss in video transmission [C]. *Proceedings of Picture Coding Symposium*, 2004: 15-17.
- [7] YANG Ren-er, CHEN Ken, HE Jia-ming. Redundancy insertion of multiple description image coding based on prediction [J]. *Opto-Electronic Engineering*, 2007, 34(10): 108-113.
杨任尔,陈恳,何加铭. 基于预测的多描述图像编码冗余插入的研究[J]. 光电工程, 2007, 34(10): 108-113.
- [8] VAISHAMPAYAN V. Design of multiple description scalar quantizers [J]. *IEEE Transactions on Information Theory*, 1993, 39(3): 821-834.
- [9] WANG Y, ORCHARD M, VAISHAMPAYAN V, et al. Multiple description coding using pairwise correlating transforms [J]. *IEEE Transactions on Image Processing*, 2001, 10(3): 351-366.
- [10] ZHANG Wei, JIANG Gang-yi, WANG Zeng-fu, et al. Research on image multiple description coding [J]. *Journal of Image and Graphics*, 2004, 9(3): 257-264.
张炜,蒋刚毅,汪增福,等. 图像位置的多描述编码方法[J]. 中国图象图形学报, 2004, 9(3): 257-264.
- [11] APOSTOLOPOULOS J. Reliable video communication over lossy packet networks using multiple state encoding and path

- diversity[C]. *Proceedings of Visual Communications and Image Processing*, 2001:392-409.
- [12] WANG Zheng-ning, PENG Qiang, ZHU Chang-qian. Implementation and analysis of fast search algorithms for block motion estimation in H. 264 test model[J]. *Journal of Computer Applications*, 2004, **24**(9):91-93.
王正宁,彭强,诸昌铃.几种快速运动搜索算法在 H. 264 中的实现及分析[J]. *计算机工程与应用*, 2004, **24**(9):91-93.

Frame-based Multiple Description Video Coding and Error Concealment

LI Jin-Xiang¹, ZHENG Shu-hong², YANG Ren-Er¹, SHAO Jie¹

(1 *Faculty of Information Science and Engineering, Ningbo University, Ningbo, Zhejiang 315211, China*)

(2 *Zhejiang Yuying Colledge of Vocational Technology, Hangzhou 310018, China*)

Abstract: To solve the error propagation of video transmission on unreliable networks, a new frame-based multiple description coding scheme with flexible redundancy insertion between descriptions via prediction is presented. To meet the varied environments of networks, several decode algorithms with various complexities and efficiency are also proposed. Simulation result shows that the presented method is demonstrated its adequate efficiency in preventing error propagation of video stream. And, the system is more robust in the case of packet-loss over different networks.

Key words: Video stream; Multiple description video coding; Multi-path transmission; Error concealment



LI Jin-xiang was born in 1966. He is currently a lecturer, and his research interests focus on mobile communication and image processing.