

一种基于 H.263 的视频多描述编码方法^{*}

李世平^{1,2}, 郁梅², 蒋刚毅^{1,2}, Tae-Young CHOI³

(1. 中国科学院 计算技术研究所 中国科学院研究生院, 北京 100080; 2. 宁波大学 信息科学与工程学院, 浙江 宁波 315211; 3. Division of Electronics Engineering, Ajou University, Suwon 442-749, Korea)

摘要: 网络上传输图像与视频数据流所遇到的一个重要问题是信息包在传输过程中可能发生丢失或过度延时, 从而造成解码信号的视觉质量严重下降。多描述编码旨在通过对同一个信息源编码成多个具有一定相关性的比特流, 并在多个不同的信道上传输, 以改善不可靠信道上传输的图像与视频信号质量。提出了一种基于 H.263 的视频信号多描述编码方法, 实验结果显示了所提出方法的有效性。

关键词: 多描述编码; H.263; 视频多描述编码

中图法分类号: TN762 文献标识码: A 文章编号: 1001-3695(2005)03-0178-03

H.263-based Multiple Description Video Method

LI Shi-ping^{1,2}, YU Mei², JIANG Gang-yi^{1,2}, Tae-Young CHOI³

(1. Insitute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China; 2. School of Information Science & Engineering, Ningbo University, Ningbo Zhejiang 315211, China; 3. Division of Electronics Engineering, Ajou University, Suwon 442-749, Korea)

Abstract: Packet loss and delay in image or video communication on Internet is a big problem, as it finally results in degradation of the reconstruction image quality. MDC (Multiple Description Coding) is an alternative solution in solving this problem. In the MDC method, the same information source is encoded into multiple correlative streams, which are transferred in several channels separately. So it can improve the image and video quality transferred in the irresponsible networks. A MDC method based on H.263 is proposed, and experimental results show that the proposed algorithm is efficient.

Key words: Multiple Description Coding; H.263; Multiple Description Video Method

1 引言

在不可靠 IP 网络中传输视频流数据存在数据包可能丢失或者在实时系统中延时过长等问题^[1]。通常的解决方法是采用纠错编码(如 FEC), ARQ 等, 但对于网络视频实时传输情况, 有时难以达到令人满意的效果。对于有较高延时要求的实时传输需求而言, ARQ 中信息包由于重发的原因延迟时间太长同信息包丢失的情形没有两样; 另外, ARQ 通常要求有稳定的反馈传输通道。

多描述编码假设在发送端和接收端之间有多个独立的并行信道, 各信道中误码的发生是相互独立的, 即所有信道同时发生误码的概率很小。多描述编码是将信源分解成多个描述, 系统将这些描述独立编码后再通过多径信道传输到客户端^[2~4]。在接收端, 根据正确传输码流的不同, 选择不同的解码恢复方案。只要有任意一个描述被正确传送到接收端, 多描述解码器就可以恢复出一定质量的视频信号。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的描述中恢复出视觉上可接收的重建信号, 而如果多个描述都同时收到, 则可获得更好的重建效果。采用多描述编码方法是阻

止噪声信道上的图像视频通信因数据包误码或丢失而导致视觉质量严重下降的一种有效途径^[5~7]。与分层编码不同, 其各个描述没有等级, 因此, 多描述编码适合没有优先权保护的信道或网络。本文就视频图像的多描述编码进行研究, 给出了一种基于 H.263 的多描述编码方法, 实验结果说明了文中方法的有效性。

2 多描述编码

多描述编码的实现方法有多种, 最初 Bell 实验室的 Jayant 用来对语音信号进行编码, 利用 DPCM 编码器将语音信号的奇数采样值和偶数值分离并各自编码成两个独立的比特流。如果偶数(奇数)流丢失, 则可从奇数(偶数)流中通过插值获取。图 1 为两个描述情况下的多描述编码原理图, 图 1(a) 中解码器 1 或 2 分别用于描述 1 或 2 的解码, 称为边沿通道, 解码器 0 则用于当两个描述都收到情况下的信号解码, 称为中央通道。 D_0, D_1, D_2 分别为各通道信号失真。若只有一个信道的描述被正确传输到接收端, 重建信号应该满足不小于最小限度的失真值; 而如果两个信道的信号都收到, 则其中一个描述可用来增强另一信号, 以获得更高质量的重建值。当考虑高传输率(R_1, R_2) 时, 尝试在 $E(d_1) \leq D_1$ 和 $E(d_2) \leq D_2$ 的条件下找到最好的方案使 $E(d_0)$ 最小, 如图 1(b) 所示。目前图像多描述编码方法主要有基于量化、基于变换和基于扩展的方法。基于量化的方法主要是通过对单个描述进行大步长的量化, 而多

收稿日期: 2004-03-08; 修返日期: 2004-06-12
基金项目: 浙江省青年人才基金(RC01057), 浙江省自然科学基金(601017), 人事部留学归国人员科研基金, 宁波市重点博士基金(2003A61001, 2004A), 宁波市留学归国人员科研基金。

个描述结合时可得到更为精细的量化。Vaishampanyan 提出了基于标量量化的多描述编码 (MDSQ): 信源进行标量量化 ($Q_1 \& Q_2$) 后输入标号分配函数, 如图 2 所示, 按对应的区间进行标号分配后直接形成两个描述。在解码端若只收到一个描述则根据量化器 Q_1 或 Q_2 重建质量仍可接收的信号。该方法关键之处在于标号分配函数的设计, 图 3 为两种不同的标号分配方法: 嵌套式和线性标号分配。

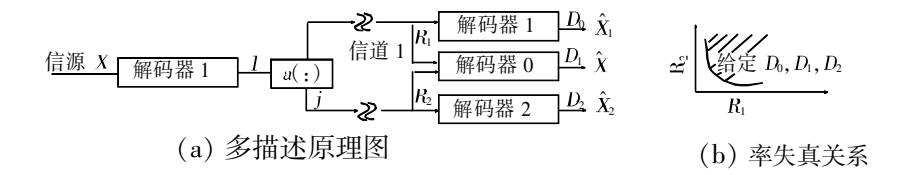


图 1 多描述编码结构框图

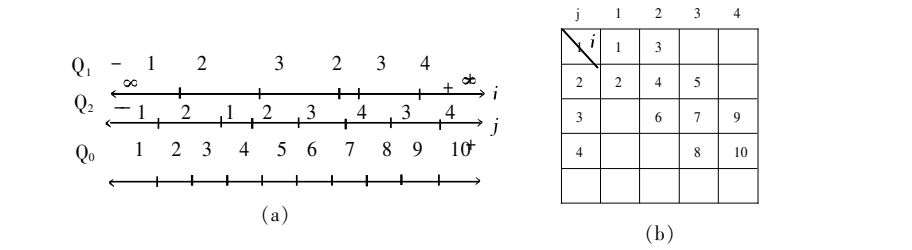


图 3 MDSQ 的量化形式

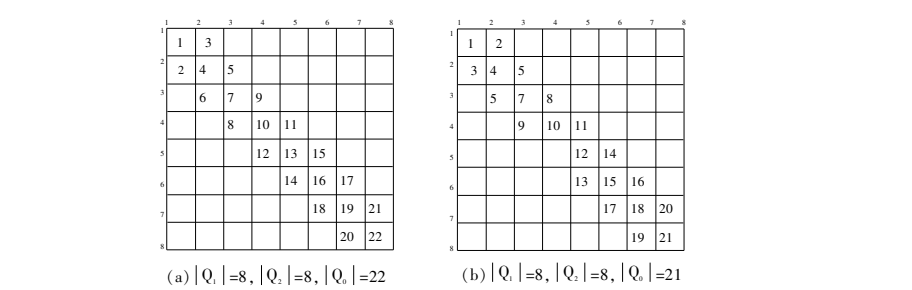


图 3 MDSQ 两种标号分配形式

3 基于 H. 263 的视频多描述编码方法

多描述编码中的视频编码与图像编码的关系并非像单描述中的 MPEG 和 JPEG 那么直接。这主要是由于运动估计时各描述所取的参考帧可能不一致, 如果其中某个描述的数据包丢失, 则可能导致解码端在运动补偿时发生误匹配, 即编解码器所用参考帧不同而导致解码信号的质量下降。

视频多描述编码的方案很多, 但由于编码方案的不同, 在运动补偿时所采用的方法也因而各异。本文采用的是基于 H. 263 的多描述编码方案。这里, 我们采用基于嵌套式标量量化的方法, 研究两个描述的情况: 仅对基于 H. 263 视频信号压缩中的 DCT 系数编码生成两个描述, 其余的数据和单描述编码的数据保持相同(但编码块样式 CBP 除外, 因为它和 DCT 系数是相关联的)。

表 1 是对 Foreman 和 Silent 序列进行 H. 263 编码产生的流文件统计分析的结果。由表中数据可知, 对于 I 帧编码方式, Foreman 序列和 Silent 序列的流文件结构基本相同, 其中大部分数据(即 DCT 系数等)都可以进行基于嵌套式标量量化的多描述编码, 然后分别存放于两个独立的流文件中, 从而使得输出比特率尽量小。而对于 P 帧编码方式, Foreman 序列和 Silent 序列的流文件结构差别极大, 其原因在于两个序列属于两类不同的场景, Foreman 序列的运动较为剧烈, 图中的人物和背景的不规则运动较多, 而 Silent 序列中人物运动较少, 背景静止。所以, 对于画面运动较丰富的 Foreman 序列, 运动矢量数据 (Vectors) 占的比例较大, 有 12.5%, 如果不对它进行多描述编码, 这部分数据只能冗余地存放在两个流文件中, 从而导致输出比特率增加。然而, 若对运动矢量进行多描述编码, 单

通道的解码图像质量将因为运动矢量误差太大而变得非常之差, 因而尽管可降低比特率但也是得不偿失的。而对于画面运动不丰富的 Silent 序列, 已编码宏块指示位 (COD) 和头信息 (Header) 的比例比较大, 分别占了 37.5% 和 18.9%, 但由于该两部分数据对于解码非常重要, 稍许的误差都可能导致几个块组甚至整帧图像的解码同步错误, 故不适宜采用多描述编码。

表 1 Foreman 和 Silent 序列进行 H. 263 编码产生流文件中各信息所占比例

序列	类别	Coeff_Y	Coeff_C	Vectors	CBPY	MCBPC	COD	Header	可多描述编码
Foreman	帧内 (I 帧)	87.8%	10%	0	1.3%	0.7%	0%	0.2%	99.8%
	帧间 (P 帧)	72%	2.8%	12.5%	6.8%	2.5%	2.2%	1.2%	84.1%
Silent	帧内 (I 帧)	85.9%	11.7%	0	1.2%	1.0%	0%	0.2%	99.8%
	帧间 (P 帧)	27.7%	0%	5.3%	8.3%	2.3%	37.5%	18.9%	38.3%

4 实验结果

多描述编码是以增加输出比特率来换取传输可靠性的提高。所以, 衡量一个多描述编码方案的标准是: 增加的输出比特率要尽可能少、提高的可靠性要尽可能高(当通道误码率为一定的情况下, 描述的数目越多, 则可靠性越高), 且每个通道的解码质量应是可接受的。选择主观质量、亮度峰值信噪比、输出比特率和可靠性等作为评价指标。定义误帧率为传输流中错误帧的比例, 而错误帧指对应流数据中有误码的帧, 因而误帧率 = 误码率 $\times$ 该帧对应流数据长度。传输可靠度指帧能成功传送的概率, 传输可靠度 = 1 - 误帧率。

本文所提出的基于 H. 263 的多描述编码方案是在 ITU-T 校验模型 TMN1.7 的基础上实现的。实验所采用的 GOP(图像组) 长度为 15, 采用 IPPP 编码方式。实验没有采用码率控制方案, 而是通过设定特定的量化步长来控制码率。其中高码率采用的 I 帧量化步长为 6, P 帧量化步长为 5, 低码率采用的 I 帧量化步长为 12, P 帧量化步长为 10。

表 2 为对测试序列 Foreman 和 Silent 进行压缩产生的平均亮度信噪比和输出比特率。由表中数据可得, 对于低码率和高码率的情况, 两个描述的总输出比特率相对于单描述的输出比特率的冗余度分别为 20% 和 10% 左右, 是可以接受的。在表 2 中, 若多描述编码方式中的两个描述都收到, 则其图像重建的信噪比等同于表 2 中右侧单描述编码方式时的信噪比。

表 2 对 Foreman 和 Silent 序列压缩产生的输出比特率 (kbps) 和峰值信噪比 PSNR (dB)

序 列	码 率	基 于 H. 263 的 多 描 述 编 码 方 式				H. 263 单描述 编 码 方 式		冗 余 率
		描 述 1		描 述 2				
		输 出 比 特 率	PSNR	输 出 比 特 率	PSNR	输 出 比 特 率	PSNR	
Foreman	低 码 率	103.2	25.36	102.6	25.43	166.8	31.70	23.38%
	高 码 率	227.4	27.66	227.4	27.70	409.2	35.83	11.14%
Silent	低 码 率	57.6	25.41	57.6	25.44	94.4	31.73	22.03%
	高 码 率	124.8	28.19	124.8	28.27	220.0	36.18	13.45%

图 4 给出了测试序列 Foreman 在高码率下的 I 帧和 P 帧解码图像。从主观图像质量来看, 描述 1 和描述 2 的解码序列的失真和块效应比较严重, 但该画面已能提供大致信息, 是可接受的; 若进一步对由描述 1 和描述 2 重建的图像进行后处理, 能获得良好的视觉效果。而如果两个描述都收到, 则重建后的图像质量很好。而图 5 给出了测试序列 Foreman 在高码率下各帧多描述编码方式的信噪比。图中显示, 描述 1 和描述 2 的信噪比曲线基本重合, 这说明两个通道的输出

比特率是均匀的。若多描述编码方式中的两个描述都收到, 则其图像重建的信噪比等同于图 5 (a) 中单描述编码方式时的信噪比。



图 4 视频多描述编码后由不同通道解码得到的序列图像

图 6 为当存在误帧情况下, 测试序列 Foreman 在高码率下各帧的信噪比。在误帧率为 5% 的情况下, 得到解码序列的信噪比和零误帧率下得到的解码序列信噪比很接近, 而且其信噪比曲线的波动不大。而在误帧率为 10% 的情况下, 得到解码序列的信噪比相对于零误帧率下得到的解码序列信噪比, 下降比较明显, 但其信噪比曲线的波动不大。在误帧率为 50% 的情况下, 得到解码序列的信噪比更低了, 而且其信噪比曲线的波动也很大, 该波动主要是因为在一个 GOP 内, 信噪比因为误帧率高而发生了急剧的下降导致的。

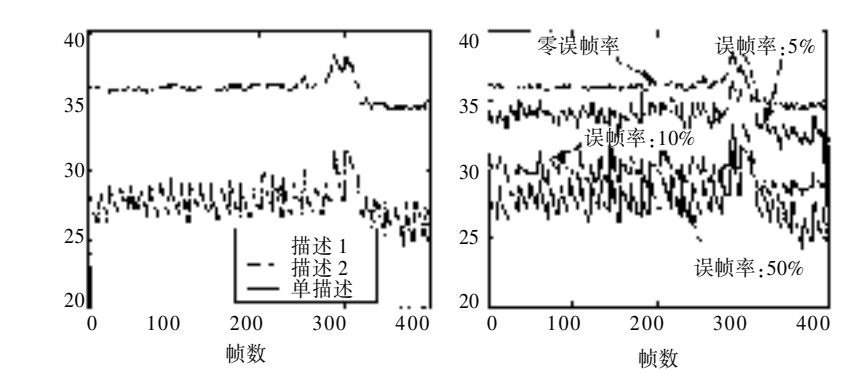


图 5 Foreman 序列多描述编码信噪比

图 6 存在误帧情况下的多描述编码信噪比

事实上, 根据国标对于线路误码率在 $10^{-6} \sim 10^{-7}$ 之间的规定, 我们可以推导出, 通常情况下误帧率不会超过 5%, 由此

(上接第 177 页) 而且层位交点的检测也是准确无误的, 这为程序的自动追踪提供了条件, 进而可以实现高效率的地质模型输入。在我们研制的地质模型输入系统中, 由于该技术的使用, 使一个普通的地质模型输入只需要几十次简单的鼠标操作, 在几分钟内即可完成, 模型输入环节不再是制约提高工作效率和勘探质量的瓶颈。

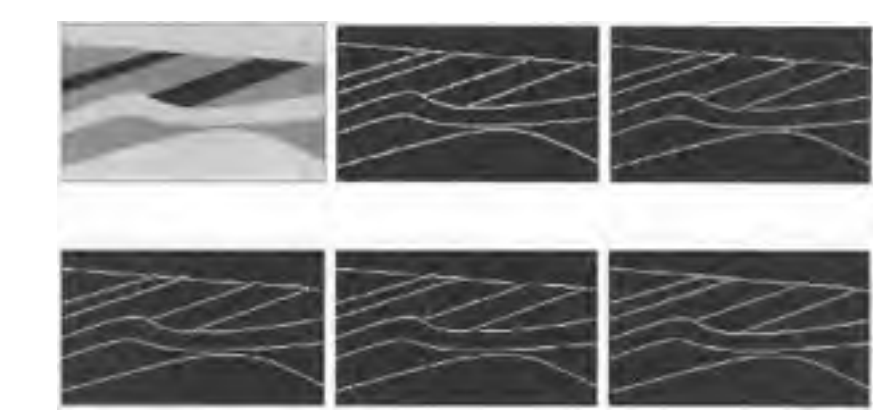


图 4 地质层位提取效果对比图

实验结果表明, 将基于最简算子的边缘检测方法用于地质建模, 不仅在精度和速度上都取得了其他方法无可比拟的效果, 而且给地质层位的自动追踪提供了条件, 进而使地质模型

可得, 本文所实现的多描述编码方案是有效的。

5 结论

本文提出并实现了基于 H. 263 的多描述编码方案。实验结果显示, 该方案在比单描述方案增加输出比特率 10% ~ 25% 的情况下, 可以使两个信道的单独解码质量达到可接受的程度, 从而大大提高了信道的传输可靠性。另外, 实验还表明两个信道的输出比特率是均匀分布的, 从而进一步验证了本文所提出的多描述编码方案的有效性。

参考文献:

[1] K Goyal, J Kovacevic. Generalized Multiple Description Coding with Correlating Transforms[J]. IEEE Transactions on Information Theory, 2001, 48(7): 2199-2224.

[2] Yao Wang, M Orchard, V Vaishampayan, et al. Multiple Description Coding Using Pairwise Correlating Transforms[J]. IEEE Trans on IP, 2001, 10(3): 351-366.

[3] 张炜, 蒋刚毅, 汪增福, 等. 图像信号的多描述编码方法[J]. 中国图像图形学报, 2004, 9(3): 257-264.

[4] V K Goyal, J Kovacevic, R Arian, et al. Multiple Description Transform Coding of Images[C]. Proc. Data Compression Conference, 1998.

[5] R Singh, A Ortega, L Perret, et al. Comparison of Multiple Description Coding and Layered Coding Based on Network Simulations[C]. Proceedings of SPIE, 2000. 929-939.

[6] D Chung. Multiple Description Image Coding Using Signal Decomposition and Reconstruction Based on Lapped Orthogonal Transforms[J]. IEEE Trans. on CSVT, 1999, 9(6): 895-908.

[7] ITU-T Rec. H. 263. tmn1. 7[S]. 1996.

作者简介:

李世平(1980-), 男, 硕士研究生, 研究方向为数字通信处理、视频编码与传输; 郁梅(1968-), 副教授, 博士, 研究方向为数字通信处理、视频编码与传输。

输入的自动化过程得以实现, 这在整个地质建模研究中是一个巨大的突破, 对能源的勘探和开发有重要意义。

参考文献:

[1] 张献民. 论地质 - 地球物理模型的某些问题[J]. 河北地质学院学报, 1995, 18(1): 33-37.

[2] Wang Shuqin, Zhang Jinhai, Wang Yong. Simplest Operator based Approach for Efficient and Accurate Edge Detection of Binary Image [C]. The ICAMT2003 Conference Proceedings, World Scientific Publishing Company, 2003.

[3] William K Pratt. Digital Image Processing-third Edition[M]. Los Altos, California: PixelSoft, Inc. , 2001.

[4] 刘海龙. 实时 Sobel 算子图像边缘检测器[J]. 仪器仪表学报, 1991, 12(1): 15-22.

[5] 邓湘金, 云日升, 吴一戎. 一种新边缘检测算子——正弦算子[J]. 电子与信息学报, 2002, 24(11): 1462-1469.

作者简介:

王淑琴(1977-), 女, 山东德州人, 工学学士, 在读硕士, 研究方向为数字水印、图像处理。