

容错视频编码技术研究^{*}

胡 飞 龙庆麟

(华中科技大学电信系 武汉 430074) (广州金鹏集团公司 广州 510665)

摘 要 本文对容错视频编码技术进行了分类概述。文章首先指出 Internet 视频通信中容错视频编码的必要性,并对容错视频通信系统的典型结构和基本思想进行了描述。然后文章对近年来出现的容错熵编码、容错预测编码和分层视频编码及区分保护等容错视频编码技术进行了分类介绍。

关键词 容错编码,差错控制,视频通信,Internet

Review of Error Resilient Video Coding Techniques

HU Fei LONG Qing-Lin

(Huazhong University of Science and Technology, Wuhan 430074) (Guangzhou Jinpeng Group Co., Ltd, Guangzhou 510665)

Abstract This paper reviews error resilient video coding techniques. Firstly the paper demonstrates the necessity of error resilient coding for video communications over Internet and introduces the typical structure and essential principle of error resilient video communication system. Then the error resilient coding techniques including robust entropy coding, error-resilient prediction and layered coding with unequal error protection are classifiedly reviewed.

Keywords Error resilient coding, Error control, Video communication, Internet

1 引言

TCP/IP 协议使得 Internet 能够建立于结构简单、成本较低的通信设施基础之上,具有广泛包容多种通信链路、多种网络技术和多种终端设备的能力,能够提供更加灵活的服务模式,可以在同一通信网络上提供多种应用,因此具有带宽利用率高、支持多种业务、操作灵活、易于管理、成本低等特点。然而目前 Internet 只提供尽力服务(Best-effort)的服务模式,不提供服务质量(QoS)保证。因此在 Internet 中,时延及时延抖动以及数据包丢失等传输差错不可避免,是 Internet 通信协议及其应用必须解决的重要问题之一^[1]。

在 Internet 迅速发展和普及的同时,近年来视频通信技术的研究和应用也取得了巨大发展^[2,3]。随着网络带宽增长,通讯成本下降,视频通信逐渐成为 Internet 的一类重要应用。然而 Internet 的 Best-effort 服务模式与视频通信对传输网络的 QoS 要求存在本质上的矛盾,成为阻碍视频通信应用进一步发展的重要因素。这一矛盾也是视频通信领域的一个重要研究内容,目前国内外研究者已经提出大量解决或缓解该矛盾的方法,如容错编码、信道可靠性控制^[4~6]、差错隐藏^[7~9]等。

其中信道可靠性控制只单纯地从提高信道传输质量出发,没有考虑到传输内容的特性和要求;而差错隐藏本质上是对多媒体流中残留或预留的冗余信息进行进一步挖掘。这两类方法虽然具有一定的独立性,但在网络资源利用效率和差错控制效果上很难实现全局优化。容错编码是目前 Internet 视频通信中广泛使用的差错控制方法。容错编码是指针对 Internet 传输的差错特性,在对视频信息编码时增加一定的容错信息,使得传输差错对视频信息的解码和回放质量的影响最小,从而实现 Internet 视频通信差错控制。

本文分析了 Internet 视频通信的特点,并对近年来出现

的 Internet 视频通信容错视频编码技术进行了分类概述。文章第 2 节指出了 Internet 视频通信中容错视频编码的必要性;第 3 节对容错视频通信系统进行了描述;第 4 节在大量查阅国内外相关文献的基础上分别对容错熵编码、容错预测编码和分层视频编码及区分保护等容错视频编码技术进行了分类介绍;最后对全文进行了总结。

2 Internet 传输差错与视频通信

目前 Internet 只提供 Best-effort 服务模式,不提供 QoS 保证。传输差错是 Internet 通信协议及其应用必须解决的重要问题之一。对于 Internet 视频通信而言,传输差错主要表现为数据包丢失。作者在文[9]中分析指出,数据包丢失是 Internet 通信中不可避免的现象,Internet 技术的设计思想以及存储转发机制、应用流量的突发性、通信网络的异构性、新应用的不断涌现、网络资源需求不断增长等等,是导致 Internet 传输中发生数据包丢失的主要原因。

另一方面,视频信息与电子邮件、文件传输等其它网络应用相比,具有传输带宽要求高、持续时间长、实时性强、流量突发性强、对数据包丢失敏感等独特的特征。文[9]中描述了数据包丢失对采用 H. 263 协议^[10]压缩的视频流的影响。当数据包丢失率达到 5%时,可以明显感觉到视频回放质量的失真;当数据包丢失率达到 10%时,视频回放质量已经令人无法接受;另一方面,由于人的生理特征,视频通信能够容忍一定的包丢失率。当包丢失率为 3%时,几乎看不出数据包丢失对视频回放质量的影响。

3 容错视频通信系统

通过以上分析可知,在数据包丢失不可避免的 Internet 上传输视频时要求视频编码具有一定的容错性。图 1 描述了典型的容错视频通信系统。视频首先在发送端被视频编码器

^{*} 本课题得到国家教育部重点科学技术项目基金(项目编号 2000175)、国家科技部技术创新基金(项目编号 00C26224210641)、广州市博士后基金资助。胡 飞 博士后,主要研究方向为信息技术。龙庆麟 硕士,工程师。

进行容错编码,同时进行压缩以降低数据率;压缩后的比特流被分割成变长的数据包(Packetization)并与其它类型的数据如音频、控制信息等复用(Multiplexing);编码后的视频数据包通过网络传输,可能发生时延抖动或数据包丢失等传输差

错,并经过一定的时延后到达接收端;接收端首先对复用数据包进行解复用(Demultiplexing),对视频流进行容错解码,然后得到重建的视频信息。

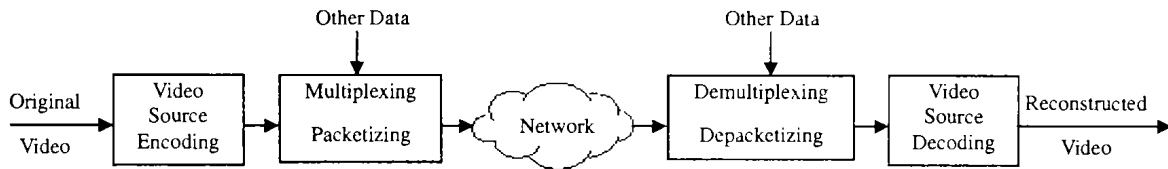


图1 视频通信系统

容错视频编码本质上是在纯粹的信源编码(尽可能地去掉冗余)的基础上预留或加入一定的冗余信息。但由于通信网络提供给应用的带宽往往是有限的,增加冗余信息的同时意味着原始信息的可用带宽减少,从而要求视频编码器增加压缩率以降低视频流的码率,这样反过来又会降低视频质量。图2描述了不同传输信道的接收端视频回放质量和冗余度的关系。显然,随着通信链路数据包丢失率的增加,要求视频流提供更多的冗余信息,以保证最佳的视频解压质量。同时还可从图中看出,过度增加冗余信息反而会导致视频回放质量下降。

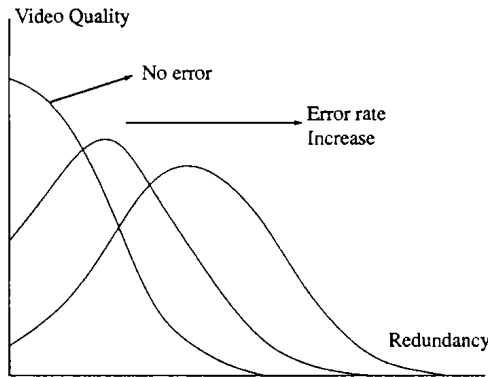


图2 不同传输信道接收端视频回放质量和冗余度的关系

4 容错视频编码

容错视频编码由视频编码器进行,目的是使视频接收端对收到视频流的解码尽量不受传输差错的影响。与只追求编码效率的编码器相比,容错编码器的编码效率要低,因为必须使用一定的冗余比特对视频流进行保护,用以提高视频流对传输差错的健壮性(robust),使得在有传输差错的传输环境下获得尽可能好的视频质量。容错视频编码的研究目标就是如何用最少的冗余比特获得最多的容错能力。

国内外相关文献提出了多种在视频流中增加或预留冗余信息的容错视频编码技术。有的技术通过增加冗余阻止传输差错扩展;有的冗余用来为视频解码器提供更多信息用以进行差错恢复;还有的技术利用冗余对视频流的部分着重保护,当传输差错发生时保证基本的视频质量。根据视频信源编码过程,本节分别对容错熵编码、容错预测编码、分层编码和区分保护以及可看作为一类特殊的分层编码的多描述编码等近年来出现的容错视频编码技术进行分类概述。本文以研究容错编码为主要内容,视频编码的基本过程不再论述,相关知识可参阅文[2,3]等。

4.1 容错熵编码

压缩视频流对传输差错敏感的一个主要原因是视频编码器在熵编码过程中采用 VLC(Variable Length Coding; 变长

编码)技术。VLC 码字中任何一个比特错误或丢失不仅使该码字不可解码,而且使随后即使被正确收到的码字也由于失同步而全部不可解码^[3]。在熵编码阶段增强视频流容错能力的方法有:

插入重同步标记(Inserting Resynchronization Markers):周期性地插入重同步标记是增强视频流容错能力的一个简单而有效的方法。如 H. 261^[11]、H. 263^[10]以及 MPEG-4^[12]都采用这类方法。这些标记被设计为可以很容易地从其它码字区分开来的符号。通常头信息(视频帧中宏块的时间和空间位置信息和其它预测信息)紧跟在重同步标记后面以得到更好的保护。这样,解码器检测到重同步标记后就可以重新开始正确的解码。显然,插入重同步标记将降低视频编码效率:首先,标记越长,出现频率越高,需要使用越多的比特。其次,重同步标记的使用中断了预测的连续性,如 MV 或 DC 的预测,因而需要使用更多的比特表示 MV 或 DC。然而,重同步标记越长,出现频率越高,解码器就越容易,而且能够更快地恢复同步,使得传输差错在重建的视频帧中影响的区域更小。因此,相对较长的同步码字在实际应用的容错视频编码系统中得到广泛使用。

可逆变长编码(RVLC: Reversible Variable Length Coding)^[14, 15]:如前文所述,在通常 VLC 技术中,一个比特的错误会导致解码器丢弃其后的所有比特,直到获得下一个同步标记。RVLC 技术使解码器不仅能够解压重同步码字后面的比特(图3中1所示),而且当差错发生时(图3中2所示)能够从逆向方向解压错误比特后下一个同步标记前尚未损坏的比特(图3中3所示),这样,RVLC 技术能够保留更多的有效比特,使差错影响的范围进一步减小(图3中4所示)。RVLC 技术已被 MPEG-4 和 H. 263 标准采纳,同插入重同步标记方法一起使用。

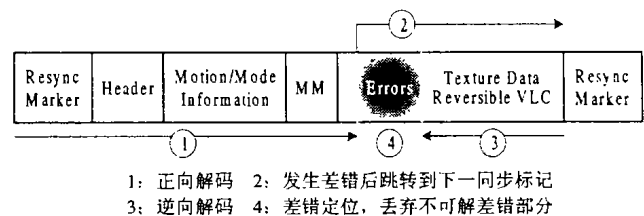


图3 RVLC 容错解码过程

4.2 容错预测编码

压缩视频流对传输差错敏感的另一重要原因是预测技术的使用,包括时间预测和空间预测。一旦传输差错使得重建的视频帧不同于正确的解码结果,随后的视频帧由于参考帧不同于编码时使用的参考帧,解码时就会产生错误,而且错误将沿着时间预测路径延续下去。同样,尽管 DC 和 MV 的预测

限制于一个视频帧内,传输差错也会沿着空间预测路径传播。在大多数视频编码标准中,空间预测范围一般进一步限制在一个 GOB 或一个片 (slice) 内,以限制传输差错空间传播范围。

I 块或 I 帧插入 (Insertion of Intra-Blocks or Frames)^[15~17]: 周期性地插入用 I 模式编码的宏块 (MB) 或视频帧能够阻止传输差错时间维上的传播,这种方法非常有效而且能够灵活使用。但由于 I 模式编码复杂度高,压缩效率低,其使用的数量 (即频率) 和 (MB) 的空间位置需要仔细设计。

I 模式编码的使用频率显然依赖于通信信道的质量。许多视频系统通过其他方式获得通信网络的信息,如无线通信系统中的天线信号强度和有线系统中的 RTCP 协议。I 模式编码的宏块 (MB) 的空间位置确定也有多种方法。如随机选择 MB; 根据 MV 反映出的运动强度选择 MB; 或者结合考虑 MB 上一次 I 模式更新的时间。这些方法都能够有效控制传输差错的空间扩展。而基于率失真函数的方法可以用来同时优化选择 I 模式编码的数量和空间位置^[17]。

基于段的独立预测编码 (ISD: Independent Segment Prediction): 这种方法将视频数据划分为多个独立的段 (segment), 空间预测和时间预测仅限于段内进行。这样一个段内出现的差错不会影响到其它段内的视频数据,这种方法也可以有效限制传输差错的影响范围。如 H. 263 协议将视频帧划分为多个区域,每个区域可以是一个 GOB 或一个 Slice, 帧间预测只能限制在对应的区域内进行^[18]。将视频序列划分为奇数帧序列和偶数帧序列,帧间预测只能在各自的序列内进行,也属于独立段预测技术。

4.3 分层编码和区别差错保护

分层编码技术将视频信号分解为一个基本层和多个增强层 (图 4 所示)。基本层提供分辨率较低但基本可以接受的视频质量,每个增强层可以对视频质量进行进一步改善。因此,分层编码方法可以使拥有不同链路带宽 (带宽异构性) 和解码能力 (终端异构性) 的多个用户以不同的质量级别获取同一视频信息。这种编码方法也称为分级编码 (Scalable coding)、渐进式编码 (Progressive coding)、嵌入式编码 (Embedded coding), 是群组通信中广泛应用的一种重要视频编码方法。

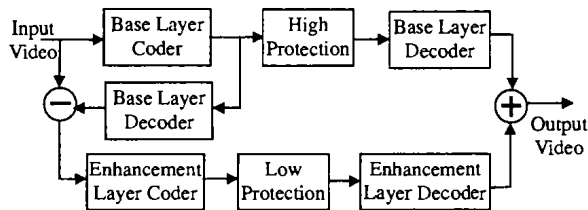


图 4 分层视频编码与区别差错保护

分层编码技术通过与区别差错保护 (UEP, Unequal Error Protection) 技术结合使用以进一步提高容错视频通信能力。UEP 本身也有多种实现方式,如通过传输网络为重要的层分配可靠性高的信道、使用不同保护强度的 FEC 编码^[19] 或对重要的层允许更多的重传提供更强保护等。但 Internet 在日前和将来一段时间内不大可能提供 UEP 支持,因此 UEP 更多的是在应用层实现,如采用数据包层的 FEC 技术为不同层的数据包提供区别差错保护。

MPEG 和 H. 263 等基于块的复合视频编码标准提供多种分层编码的途径,包括: 1) 时间 (temporal) 分层: 在时间维上对视频信号进行亚采样,使基本层只包含较低帧率的视频

流,而补充的视频帧包含在增强层; 2) 空间 (spatial) 分层: 与时间分层原理相似,只是亚采样在空间维上进行,使基本层只包含较小尺寸的视频; 3) 信噪比 (SNR) 分层: 基本层包含量化步长较大的 DCT 系数,增强层则对原始系数和基本层编码系数的误差进行编码; 4) 数据划分 (data partitioning): 基本层包含头信息和运动信息,其它信息留给增强层编码。

4.4 多描述编码

多描述编码 (MDC, Multiple Description Coding) 可看作为一类特殊的分层编码。同上述分层编码类似, MDC 也将原始视频流分解为多个流,但分解方式不同。MDC 分解的每个子流称为描述 (description), 各个描述之间是相关的而且具有相同的重要性。任何一个描述可以提供一个基本级别的视频质量,多个描述则可以对视频质量进一步改善。因此 MDC 也能够拥有不同链路带宽和解码能力的多个用户以不同的质量级别接收同一视频信息 (如图 5 所示)。

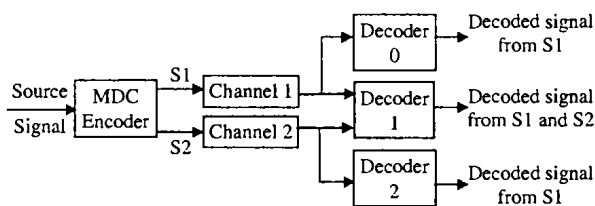


图 5 MDC 编解码过程

由于每个描述都可以提供一定级别的视频质量,因此所有描述必须共享视频源的一些基本信息,即彼此具有相关性。描述间的相关性使解码器能够通过正确收到的描述去估计出现差错的描述的相关信息。但相关性使视频质量得到改善的同时也增加了视频编码的冗余,降低了编码效率。同分层编码相比, MDC 的一个优点是不需要传输网络为特定的层提供可靠性更高的信道。实现 MDC 分解的方法有重叠量化 (Overlapping Quantization)^[20]、重叠正交变换 (Lapped Orthogonal Transform)^[21]、相关预测 (Correlated Predictor)^[22]、相关线性变换 (Correlating Linear Transform)^[23,24] 以及 H. 263 标准采用的交叉时空采样 (Interleaved Spatial-temporal Sampling) 等。

总结 尽管基于 Internet 的视频通信有着普遍的应用需求和广泛的应用前景,然而 Internet 服务模式与视频通信 QoS 要求之间存在的矛盾阻碍着视频通信应用的进一步发展。由于视频通信应用的实时性要求,认为信源编码和信道编码可独立进行的经典香农理论 (Shannon separation theorem) 并不适合于 Internet 视频通信差错控制。对于基于 Internet 的视频通信而言,针对 Internet 传输的差错特性,通过在编码视频流中增加一定的容错信息,使得传输差错对视频信息回放质量的影响最小,从而实现差错控制的容错编码是目前广泛使用的有效的差错控制方法。

本文在作者所阅及文献的基础上对近年来国内外出现的主要容错视频编码技术进行了概述。然而随着通信技术和信息处理技术的不断发展,新的应用需求将不断涌现;同时容错视频编码作为视频通信领域的一个研究热点,研究成果也不断取得新的进展。要了解最新研究成果,还必须及时跟踪国际最新动态。但总的来讲,Internet 视频通信技术将不断完善,最终得以实现高质量、高可靠性视频通信的普遍应用。

(下转第 152 页)

4 已有工作与比较

我们在设计这样一个通用的审计框架以及实现的过程中,已经注意到国内已经有了相关方面的工作。《安全操作系统审计的设计与实现》^[6]、《Linux Shell 安全审计机制的扩展》^[9]中所描述的工作(特别是前者)与我们的工作有很多类似的地方,但是我们的工作不是简单地重复前人的工作,从前面的叙述不难看出我们工作相对国内前人工作的特色之处:

- 开放的通用框架的设计。通用框架的设计基于流行的国际标准--POSIX,与具体的实现平台和具体针对的安全级别无关。框架中从数据结构到函数接口的国际化使得该框架能够面向更广泛的应用。审计实现要求遵从 TCSEC,进而从实现上也达到与国际接轨。

- 审计不仅仅面向系统,审计函数库和系统审计的用户接口使得应用程序能够自主地利用审计机制。这对于审计是个必要的且有力的扩充。

- 实现方面的审计子模块与系统其他模块之间的相对高独立性,以及审计截获对系统切入口的统一化,使审计模块和系统内核的独立升级成为可能。

结论与展望 本文所叙述的是一个符合 POSIX1003.1e 标准的通用的审计框架的设计工作以及该框架在 GB 17859-1999 第三级别的 Linux 安全增强系统上的实现工作。审计框架设计部分的工作,希望能对类似的 POSIX 的兼容系统的审计实现都能有所启发和帮助,实现部分的工作也成为了 SOFTOS 的一个子模块。然而这部分的工作主要集中在审计

数据的获得和对所获得的审计数据的简单约简之上,面向安全应用(如 IDS)的审计记录分析的工作尚且不足。这方面的工作有望成为以后工作的一个重点,从而进一步完善这个审计框架。

参考文献

- 1 Draft Standard for Information Technology-Portable Operating System Interface (POSIX)-Part 1: System Application Program-Interface (API)- Amendment #: Protection, Audit and Control Interfaces [C Language] IEEE Standards Department 1997
- 2 Orange Book Parts I and II: THE CRITERIA and RATIONALE AND GUIDELINES NCSC DOD: NIST Dec. 1985
- 3 Anderson J P. Computer Security Threat Monitoring and Surveillance. James P. Anderson Co., Fort Washington, PA, 1980
- 4 Bonyun D. The Role of a Well-Defined Auditing Process in the Enforcement of Privacy Policy and Data Security. In: Proc. of the 1981 IEEE Symposium on Security and Privacy, 1981. 19~26
- 5 A Guide to Understanding Audit in Trusted Systems, The Rainbow Books. National Computer Security Center 1987
- 6 茅兵. 基于 Linux 的安全操作系统的开发. 2000 年中国自由软件发展战略研讨会暨第一届中国自由软件应用论坛会刊. 中国国家高技术智能计算机系统专家组与中国共创软件联盟主办, 北京, 2000
- 7 Linux kernel 2.4.2. <http://www.kernel.org>
- 8 刘海峰, 卿斯汉, 刘文清. 安全操作系统审计的设计与实现. 计算机研究与发展, 2001, 38(10)
- 9 汪立东, 方滨兴. Linux Shell 安全审计机制的扩展. 软件学报, 2002, 13(1)
- 10 Bishop M. A standard audit trail format. In: Proc. of the 18th National Information Systems Security Conf. Baltimore, Maryland, USA, 1995

(上接第106页)

参考文献

- 1 Comer D E. Computer Networks and Internets. Prentice Hall International, Inc., 1997
- 2 Steinmetz R, Nahrstedt K. Multimedia Computing, Communications and Applications. Prentice Hall International, Inc., 1995
- 3 Tekalp A M. Digital Video Processing. Prentice Hall International, Inc., 1996
- 4 胡飞, 朱耀庭, 朱光喜. Internet 视频点播差错控制. 计算机研究与发展, 2002, 39(1): 28~34
- 5 Hu Fei, Zhu Guangxi, Zhu Yaoting. Enhanced ARQ-based Packet Loss Recovery for Real-time Communication. In: Proc. of Intl. Conf. on Info-tech & Info-net (ICIT'2001), Beijing, Oct. 2001
- 6 Girod B, Faerber N. Feedback-based error control for mobile video transmission. Proc. IEEE, Special Issue on Video for Mobile Multimedia, 1999, 87(10), 1707~1723
- 7 Zhu Q, Wang Y. Error concealment in visual communications. Compressed Video over Networks. Narcel Dekker, Inc., 2000
- 8 Hong M C, Kondi L, Scwab H, Katsaggelos A. Video error concealment techniques. Signal Processing: Image Communications, special issue on Error resilient Video, 1999, 14(6-8): 437~492
- 9 胡飞, 朱光喜, 朱耀庭. Internet 视频通信差错隐藏技术研究. 计算机科学, 2002, 19(4): 124~127
- 10 Video coding for low bitrate communication. ITU-T Recommendation H. 263, 1998
- 11 Video codec for audiovisual services at p×64kbits. ITU-T Recommendation H. 261, 1993
- 12 Information technology - Generic coding of audio-visual objects - Part 2: Visual, Final Proposed Draft Amendment 1. ISO/IEC JTC 1/SC 29/WG 11 N2802 (MPEG-4), July 1999
- 13 Wen J, Villasenor J. A class of reversible variable length codes for robust image and video coding. In: Proc. IEEE Intl. Conf. on Image Proc., Santa Barbara, CA, Oct. 1997
- 14 Wen J, Villasenor J. Reversible variable length codes for robust image and video transmission. In: 1997 Asilomar Conf. Pacific Grove, CA, Nov. 1997
- 15 Naka N, Adachi S, Saigusa M, Ohya T. Improved error resilience in mobile audio-visual communications. In: IEEE Intl. Conf. on Universal Personal Communications, Tokyo, Japan, Nov. 1995, 1, 72~706
- 16 Coete G, Shirani S, Kossentini F. Optimal mode selection and synchronization for robust video communications over error prone networks. IEEE Journal on Selected Areas in Communications, May 1999
- 17 Zhang R, Regunathan S L, Rose K. Video coding with optimal Inter/Intra-mode switching for packet loss resilience. IEEE Journal on Selected Areas in Communications, June 2000
- 18 Wenger S, Knorr G, Ott J, Kossentini F. Error resilience support in H. 263+. IEEE Trans. on Circuits and System for Video Technology, 1998, 8(6): 867~877
- 19 Kondi L, Ishtiaq F, Katsaggelos A K. Joint source-channel coding for scalable video. In: Proc. SPIE conf. On Visual Communications and Image Processing, San Jose, CA, Jan. 2000
- 20 Vaishampayan V A. Design of multiple description scalar quantizers. IEEE Trans. Info. Theo. 1993, 39: 821~834
- 21 Chung D-M, Wang Y. Multiple description image coding using signal decomposition and reconstruction based on lapped orthogonal transforms. IEEE Trans. on Circuits and Systems for Video Technology, 1999, 9(6): 895~908
- 22 Ingle A, Vaishampayan V A. DPCM system design for diversity systems with applications to packetized speech. IEEE Trans. Speech and Audio Processing. 1995, 3: 48~57
- 23 Goyal V K, Kovacevic J, Areal R, Vetterli M. Multiple description transform coding of images. IEEE Int. Conf. Image Proc. (ICIP98), Chicago, 1998, 1: 674~678
- 24 Wang Y, Orchard M, Reibman A. Optimal pairwise correlating transforms for multiple description coding. IEEE Int. Conf. Image Proc. (ICIP98), Chicago, 1998, 1: 679~683