

面向异构网络的 视频编码技术

白慧慧 张萌萌 赵耀 著

SHIPIN

BIANMA JISHU



中国水利水电出版社
www.waterpub.com.cn

面向异构网络的 视频编码技术

白慧慧 张萌萌 赵耀 著



中国水利水电出版社
www.waterpub.com.cn

内 容 提 要

随着网络技术的发展,图像/视频业务在多媒体通信中占据越来越重要的地位。然而,目前的互联网和无线网络存在网络拥塞、网络异构性等问题,严重影响了多媒体数据流实时可靠的传输,因此,设计兼具压缩性和鲁棒性的编码方案成为目前图像/视频编码研究的重点内容。作为一种效率高容错性能强的编码,多描述编码已经引起了国内外学者的广泛关注。

本书全面介绍了多描述编码的基本原理和最新的研究成果。旨在帮助读者多描述编码获得全面深刻的理解和认识,从而使读者更好的把握异构网络下图像视频编码算法的特点和设计重点,为读者以后进一步的研究和开发打下坚实的基础。

本书适合作为图像视频编码方面研究人员和工程技术人员的参考资料,也适合作为本科高年级及研究生等的学习教材和辅导教材。

图书在版编目(CIP)数据

面向异构网络的视频编码技术 / 白慧慧, 张萌萌,
赵耀著. — 北京: 中国水利水电出版社, 2011.3
ISBN 978-7-5084-8332-0

I. ①面… II. ①白… ②张… ③赵… III. ①视频编
码—研究 IV. ①TN762

中国版本图书馆CIP数据核字(2011)第042742号

书 名	面向异构网络的视频编码技术
作 者	白慧慧 张萌萌 赵耀 著
出版发行	中国水利水电出版社 (北京市海淀区玉渊潭南路1号D座 100038) 网址: www.waterpub.com.cn E-mail: sales@waterpub.com.cn 电话: (010) 68367658 (营销中心)
经 售	北京科水图书销售中心(零售) 电话: (010) 88383994、63202643 全国各地新华书店和相关出版物销售网点
排 版	中国水利水电出版社微机排版中心
刷 印	北京市兴怀印刷厂
规 格	184mm×230mm 16开本 7.5印张 159千字
版 次	2011年3月第1版 2011年3月第1次印刷
印 数	0001—2000册
定 价	20.00元

凡购买我社图书,如有缺页、倒页、脱页的,本社营销中心负责调换

版权所有·侵权必究

序

20 世纪末，随着多媒体技术和网络通信技术的飞速发展，基于有线网络的多媒体信息传输已经拥有了广泛的应用范围，涵盖了视频会议、可视电话、电子商务、远程教学以及远程医疗等领域。

进入 21 世纪以来，由于第三代移动通信技术的突飞猛进，在无线网络中实时可靠的传输图像和视频也已成为必然需求，众多新兴的无线视频业务应运而生，如视频短信、手机电视，移动影院等。

人们正试图通过各种终端设备，以有线或无线接入的方式，随时随地实现多媒体通信的需求。由此可见，现实生活中的网络已经发展成为有线链路和无线信道并存的异构网络。然而，日益严重的网络拥塞和网络异构性给传统的视频编码带来了巨大的挑战。

在传统的视频编码中，视频数据是经过高度压缩的数据流，其比特之间的数据相关性很小，对错误十分敏感。此外，为了获得较高的压缩比，通常采用时域运动补偿技术和变长的统计熵编码技术。这些技术的采用使得一旦发生误码，一方面，当前帧的错误将会影响到后续帧的解码，使得解码出来的视频质量严重下降；另一方面，传统的视频编码更着重于提高压缩性能，在产生数据传输错误后主要被动依赖于后期信道编码的纠错能力。新近制定的视频编码国际标准如 MPEG-4 中的精细颗粒分级编码（Fine Granular Scalable，简称 FGS）等，也开始试图采用新的编码框架，来更好地适应网络传输。FGS 的基本思想是基本层采用较强的错误保护措施，如较强的前向纠错技术（Forward Error Correction，简称 FEC）和自动重传请求（Auto Repeat Request，简称 ARQ），以保证传输的可靠性，但是这种方法也存在如下问题：首先，基本层错误会导致解码重建质量的严重下降；其次，增强层依赖于基本层，若基本层丢失，增强层也将无用；再者，反复的 ARQ 会引起延

时过度,强 FEC 也会因其复杂度带来额外延时,严重影响视频的实时播放。

在异构网络下,用户需要随时随地通过有线或无线途径来接入网络。不同用户不仅在计算机处理能力和外设性能方面存在很大差异,而且网络接入类型和连接各个用户的不同子网也会在物理通道、带宽资源、传输延时等方面有所不同,例如宽带到窄带的传输或有线到无线的传输,这种带宽的剧烈波动极易造成包丢失、传输延迟等问题。另一方面,由于多径、衰落、噪声影响和信道干扰等原因,无线信道的传输错误率较高,有可能导致所传输的视频数据的大量字段失效或彻底丢失。这些问题对于压缩视频流数据是致命的,因为压缩后的数据一般是由不等长码字构成的码流,如果出现错误或数据包丢失,都会引起错误扩散等一系列的问题,这不但严重影响视频业务的服务质量,而且甚至会导致整个视频通信系统完全失效,成为限制网络实时视频技术发展的瓶颈。

针对不可靠网络的实时视频传输问题,近年来多描述编码(Multiple Description Coding,简称MDC)技术,得到了越来越多国内外学者的关注。多描述编码假设在信源和信宿之间有多个信道,各个信道同时出错的概率非常低,通过生成多个同等重要、可独立解码的关于编码的描述(即比特流),从而保证在其中一些描述丢失的时候,仍可以得到可接受的图像质量,而随着描述的增加,图像质量也随之提高。与传统视频压缩技术不同,多描述编码的研究目的不是单纯追求高的压缩性能,而是通过其描述间的相关性在保证有效压缩的前提下,提高传输的可靠性。与分层编码相比,多描述编码的优势就在于不需要网络提供一个专门的可靠信道保障基本层的传输。在丢失情况严重的网络中,分层编码可能需要更多的重传或者加入更多的前向纠错码来实现基本层的无错传输,在这种情况下使用多描述编码更有效。

多描述编码理论在20世纪70年代就开始建立,但是直到近年来才逐渐引起人们的重视。目前该领域的研究主要集中在多描述编码本身的实现算法上,因此,为了更有效地解决其在异构网络上高效、可靠、实时传输的问题,更好地满足在异构网络上不同用户的需求,急需探索新的适应网络传输的多描述视频编码算法。

本书中的研究成果可以广泛地应用于在复杂异构网络下的视频通信、视

频监控、视频点播等领域。对于探索新的视频编码技术，解决日趋复杂的在异构网络上高质量的视频传输具有重要的理论意义和实用价值，有助于促进我国在视频通信领域的发展。

本书由北京交通大学白慧慧博士、赵耀教授以及北方工业大学张萌萌副教授共同完成，此外该书受到了国家自然科学基金（No. 60903066，No. 60972085）、北京市自然科学基金（No. 4102049）、北京市属高等学校人才强教深化计划（PHR201008187）项目的资助。

本课题的研究攻关和本书的编写出版得到了国家863项目、国家自然科学基金、高等学校博士点基金、新世纪优秀人才支持计划和北京交通大学十五重大专项基金的资助。

编者

2011年1月1日

前言

随着网络技术的发展,图像/视频业务在多媒体通信中占据越来越重要的地位。然而,目前的互联网和无线网络存在网络拥塞和网络异构性等问题,严重影响了多媒体数据流实时可靠的传输,因此,设计兼具压缩性和鲁棒性的编码方案成为目前图像/视频编码研究的重点内容。作为一种效率高、容错性能强的编码,多描述编码已经引起了国内外学者的广泛关注。

本书以设计高效的多描述图像视频编码器为主要研究目的,融合现有的差错控制技术,在静止图像编码方面,提出了基于格型矢量量化的多描述图像编码方案和基于平移格型矢量量化的多描述图像渐进编码方案;在视频编码方面,提出了率失真优化的时域亚采样的多描述视频编码方案和基于优先级传输的多描述视频编码方案。

(1) 提出了基于格型矢量量化的多描述图像编码方案。该方案将经典的多描述格型矢量量化的理论应用于图像编码中,对方案中的各个模块逐一进行了改善:充分利用不同子带的小波系数的方向相关性,将小波系数组织成矢量;对编码器的两个参数即格型矢量量化的步长和子格的索引值进行了优化设计;改进了索引分配策略和相应的交替传输机制,有效地利用了矢量之间的相关性,使得解码端能够更好地进行错误隐藏。与基于标量量化的多描述图像编码相比,在同样比特率下,中心重建质量有 0.2~0.7dB 的提高,单路重建质量有 2~5dB 的提高。它与基于相关变换的多描述图像编码相比,在同样比特率下,中心重建质量提高了约 2dB,单路重建质量提高了约 5dB。

(2) 提出了基于平移格型矢量量化的多描述图像渐进编码方案。该方案的主要特点如下:一方面,平移格型矢量量化算法充分利用了格型矢量量化器的几何结构,与传统的多描述格型矢量量化算法相比,不但具有较好的率失真性能,而且有效地减小了中心失真和单路失真之间的差距,SVS 方案在

子格点索引值为 $N=7$ 时的落差高达 $9\sim 11\text{dB}$ ，而所提出的多描述平移格型矢量量化算法中，两者落差只有 4dB 左右。另一方面，在传统的小波零树编码的基础上，设计了基于格型矢量量化的零树编码，进而实现了中心路和单路的渐进传输。它与基于标量量化的多描述渐进编码方案相比，在同样比特率下，中心重建质量提高了 $1\sim 1.5\text{dB}$ ，单路重建质量提高了 $0.2\sim 0.5\text{dB}$ 。

(3) 提出了率失真优化的时域亚采样的多描述视频编码方案。该方案利用代价函数优化算法对联合的帧内插帧拷贝模式进行合理的选择，从而实现了率失真优化的设计目标。同时，该方案也可以很好地兼容现有的视频编码标准，如 H.26X 系列和 MPEG-X 系列。它与传统的时域亚采样方案相比，在同样比特率下，整个视频序列的平均单路重建质量提高了 $0.4\sim 0.8\text{dB}$ ，对于存在运动变化较大的帧，其单路重建质量提高了约 8dB ，并且在主观视觉效果方面都取得了明显的改善；在包丢失网络环境下，所提出的方案比传统方案的重建质量提高了 $0.3\sim 1\text{dB}$ ；与一维的空域亚采样的多描述视频编码方案相比，无论是中心重建质量还是单路重建质量，所提出的方案都有明显的改善，分别提高了 $0.2\sim 1.4\text{dB}$ 和 $0.3\sim 1.2\text{dB}$ ；与二维的空域亚采样方案相比，所提出的方案具有更为明显的优势，在同样比特率下，中心重建质量提高了 $2.5\sim 3\text{dB}$ ，单路重建质量提高了 $2\sim 3\text{dB}$ 。

(4) 提出了基于优先级传输的多描述视频编码方案。该方案突破了传统的基于 FEC 的多描述编码方法中必须使用可伸缩编码的限制，具备了与现有标准视频编解码器的良好兼容性。在构造消息的过程中，充分利用原始视频序列的运动变化特征，使得每个消息内部都能够保持较好的时域相关性，有利于恢复丢失的信息；同时，采用灵活的 GOP (Group of Picture) 组织形式匹配视频序列的运动情况，保证发生显著运动变化的帧被编码为优先级较高的 I 帧，从而使优先级的分布更为合理。在优先级编码传输方案中，按照数据段的重要性以及实际的包丢失率，提出了优先级的设定方案。它与基于平等保护的方案相比，该方案有效缓解了当包丢失率达到一定程度时重建视频质量直线下降的问题，重建质量也有了明显的改善，其性能的最大差距达到了 8dB 左右。

编者

2011 年 1 月 1 日

目 录

序

前言

第 1 章 绪论	1
1.1 引言	1
1.2 差错控制技术的研究现状	3
1.2.1 差错控制技术的研究动机	3
1.2.2 差错控制技术的分类	4
1.2.3 MPEG-4 和 H.264/AVC 标准中的差错控制	9
1.3 多描述编码的研究现状	10
1.3.1 多描述编码的研究意义和应用	10
1.3.2 多描述编码的国内外研究现状	13
1.4 本书的主要工作和组织结构	15
第 2 章 多描述编码的基本理论和方法	18
2.1 引言	18
2.2 多描述编码的信息论基础	18
2.2.1 传统的率失真函数	18
2.2.2 多描述编码的率失真函数	19
2.3 多描述编码的实现方法	20
2.3.1 基于亚采样的多描述编码	20
2.3.2 基于量化的多描述编码	22
2.3.3 基于相关变换的多描述编码	23
2.3.4 基于 FEC 的多描述编码	24
2.4 本章小结	25
第 3 章 基于格型矢量量化的多描述图像编码	26
3.1 引言	26
3.2 格型矢量量化器概述	27
3.2.1 格的定义和性质	27
3.2.2 格型矢量量化的算法	28

3.3	经典的 MDLVQ 方案	30
3.3.1	SVS 方案的框架	31
3.3.2	索引分配	32
3.4	改进的 MDLVQ 方案	35
3.4.1	小波变换编码基础	35
3.4.2	整体框架设计	37
3.4.3	编码端的优化方案	40
3.4.4	解码端的优化方案	42
3.4.5	实验结果和分析	43
3.5	本章小结	46
第 4 章	基于平移格型矢量量化的多描述图像编码	47
4.1	引言	47
4.2	平移格型矢量量化算法	47
4.2.1	SLVQ 算法的原理	48
4.2.2	实验结果和分析	50
4.3	改进的小波零树编码	51
4.3.1	传统的小波树结构	52
4.3.2	基于格型矢量量化的零树编码	55
4.4	基于 SLVQ 的多描述图像渐进编码	57
4.4.1	设计框架	57
4.4.2	实验结果和分析	57
4.5	本章小结	62
第 5 章	率失真优化的时域亚采样的多描述视频编码	63
5.1	引言	63
5.2	设计框架及实现	64
5.2.1	预处理的理论分析	66
5.2.2	预处理方法	67
5.2.3	运动补偿帧插值算法	70
5.2.4	预处理的率失真优化方法	72
5.2.5	开关信道下的后处理方法	74
5.2.6	包丢失信道下的后处理方法	76
5.3	实验结果和分析	77
5.4	本章小结	86

第 6 章 基于优先级传输的多描述视频编码	87
6.1 引言	87
6.2 优先级编码传输算法	88
6.2.1 纠删码简介	88
6.2.2 数据分割	89
6.2.3 打包算法	91
6.3 所提出的方案	94
6.3.1 设计框架	94
6.3.2 优先级的设定	96
6.4 实验结果和分析	97
6.5 本章小结	101
第 7 章 总结	103
7.1 工作总结	103
7.2 工作展望	104
参考文献	105

第 1 章 绪 论

1.1 引言

长期以来，人类在自然界感受到的最主要信息是视觉信息，与其他形式的信息相比较，视觉信息更加直观、生动和形象。随着信息时代的到来和多媒体技术的发展，人们已不再满足于传统的语音及文本数据的通信方式，包含图像或视频的视觉信息通信（Visual Communication）受到越来越广泛的关注^[1]。它与文字信息不同，图像视频信息需要较大的存储容量和较宽的传输信道，如表 1.1 所示是几种常见的视频图像源未经压缩的原始数据率。鉴于目前的通信网络性能，如表 1.2 所示，直接传输未压缩的图像视频是不现实的。因此，对图像视频信息进行高效压缩是图像视频通信需要首先解决的问题。

表 1.1 几种常见的视频图像源的未压缩数据率^[1]

视频源	每秒帧率	分辨率（帧内像素数）	未压缩数据率（Mbit/s）
NTSC	30	720×480	124.416
PAL	25	720×576	124.416
VCR	25	352×288	30.4128
HDTV	30	1920×1080	746.496

注 这里计算是基于 YUV4:2:0 格式，每个像素 12bits，未压缩数据率=每秒帧率×帧内像素数×12。

表 1.2 通用的通信网络的性能^[2]

通 信 网 络	传 输 速 率
公共交换电话网（PSTN）	56kbit/s
非对称数字用户环路（ADSL）	6Mbit/s
综合业务数字网（ISDN）	64 - 1536kbit/s
Ethernet	10Mbit/s
GPRS 无线接入	170kbit/s
3G	户外：144 - 384kbit/s；室内：2Mbit/s

图像压缩的基本理论起源于 20 世纪 40 年代末 Shannon 的信息理论。在这个理论框架下，出现了一些无失真信源编码算法，原始图像可以由压缩数据完全恢复出来，这种压缩方法的基本思想是对于小概率的输入使用长码字表示，而对于大概率的输入采用短码字表



示。常见的无失真压缩技术有：Huffman 编码、算术编码、词典编码等，但是这些算法的压缩性能是有限的，难以满足大多数图像视频存储和传输的需要。根据应用的需求，人们开始了对有失真压缩的广泛研究，有失真压缩的目的是去除图像数据中的冗余信息和对视觉不重要的细节分量，以尽可能少的码字来表示所处理的图像，其原始图像不能由压缩数据完全恢复出来，恢复图像只是某种失真度下的近似。目前大多数压缩方法都属于这一类，如预测编码、变换编码等。为了使图像/视频压缩算法更加实用化，ISO/IEC, ITU-T 等国际组织先后制订了许多成功的图像和视频编码标准，如以 JPEG 和 JPEG-2000 为代表的静止图像压缩标准；以 MPEG-1 和 MPEG-2 为代表的中高码率多媒体数据编码标准，其中视频图像压缩标准是其主要内容；以 H.261, H.263, H.263+, H.263++，H.264/AVC 等为代表的低码率、甚低码率运动图像压缩标准，以及面向对象应用的 MPEG-4 标准。

近年来，随着通信网络带宽和存储媒介容量的不断增加，人们对于进一步提高压缩比的图像视频编码算法的探索兴趣逐渐下降，更多地开始关注怎样更好地将编码算法应用到各种实际的系统中。例如，为了在互联网上灵活地提供可分级的图像分辨率，分层编码应运而生；为了能够更好地观赏三维立体效果的图像/视频节目人们开始研究三维图像/视频编码，等等。同时，人们对包括实时图像/视频在内的网络多媒体业务的需求不断增长，网上图像和视频传输已经拥有极广泛的应用范围，涵盖了交互式视频服务、视频会议、视频电话、电子商务、远程教学以及远程医疗等领域。

然而，互联网和无线通信网络的传输信道并不十分可靠。一方面，互联网是一个庞大的网络，具有很强的异构性（Heterogeneity），互联网用户可以随时通过有线或无线途径接入网络，不同用户不仅在 CPU 处理能力和外设性能方面存在很大差异，而且网络接入类型和连接各个用户的不同子网也会在物理通道、带宽资源、传输延时等方面有所不同，甚至对于同一个用户在不同时间段内的数据传输，实际所能够得到的传输带宽也会存在较大的波动。互联网的异构性和多用户共享信道容量的特点，决定了其“尽力而为”（Best-effort）的服务方式，即互联网本身并不确保数据传输的可靠性，数据包通过互联网传输时由于网络拥塞等原因而导致的包丢失、传输延迟等不可避免。另一方面，无线信道由于其固有的多径干扰、衰落、高误码率等特性使得无线传输环境更加恶劣，造成所传输图像/视频数据的大量字段丢失或失效，甚至会导致整个图像/视频通信完全失效，成为限制网络实时图像/视频技术发展的瓶颈。

国外从 20 世纪 90 年代开始了针对图像/视频网络传输的差错控制技术（Error Control）的研究，国内对该课题的研究才开始不久，这一领域还有许多的理论和实际问题有待进一步解决。因此，对图像/视频网络传输进行广泛深入的研究，设计兼具高效压缩和可靠传输性能的图像/视频编码方案，是一项非常具有挑战性和实用价值的研究课题，无论对学术界还是工业界都有着十分重要的意义。



1.2 差错控制技术的研究现状

1.2.1 差错控制技术的研究动机

一个典型的图像/视频通信系统^[2]包括 5 个部分,如图 1.1 所示。原始图像/视频首先经过图像/视频编码器压缩以减少数据率,压缩后的比特流分割成定长或者变长的数据包,并且与其他的数据类型如音频数据等进行复用。如果网络能够保证无错传输,则这些数据包可以在网络上直接传送。否则,就要经过信道编码的阶段,典型的是使用前向纠错(FEC, Forward Error Correction)来纠正传输中的错误。在接收端,收到的数据包经过解数据包和信道解码后,得到的比特流输入图像/视频解码器中进行解码来重建原始图像/视频。

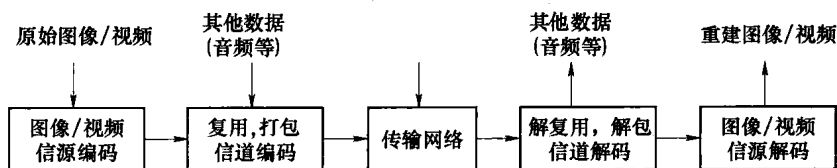


图 1.1 典型的视频通信系统

网络传输中的差错可分为两类:随机比特错误和删除错误。随机比特错误是由物理信道的不完善而引起的比特反转、插入和删除。随机比特错误对编码码流的影响取决于编码方法和受影响的信息内容。当使用定长编码时,随机比特错误仅仅影响一个码字,所引起的错误通常是可以接受的。但是如果使用变长编码(如 Huffman 编码),随机比特错误可能导致接收到的码字与实际的编码信息失去同步,从而其后续的码字也无法解码,直到下一个同步码字出现才可能重新解码。删除错误是由于包交换网络(如 Internet)中的数据包丢失、突发性差错或者短时间的系统故障所引起的。在变长编码中的随机比特错误也可能引起删除错误,因为一个比特的错误可能导致后续的比特都无法解码,从而变成无用的码字。通常,由于删除错误造成的是连续比特的丢失或损坏,因此,比随机比特错误更具有破坏性。

通过上述对典型图像/视频通信系统和网络常见传输差错的分析,可以发现图像/视频通信中的差错控制是非常具有挑战性的,其原因如下:第一,压缩后的视频流由于使用了时间域的预测编码和变长编码,更易遭遇严重的传输错误。如图 1.2 所示,一个单一的错误样点将导致时间域和空间域中错误的不断扩散。第二,由于使用了变长编码,随机比特错误的后果可能相当于删除错误,从而损坏了视频帧的大部分信息。例如,基于块的混合编码中,如果某个宏块出现随机比特错误,则可能导致该宏块完全不可解码,解码器将根据重同步标记解码下一个宏块,并且考虑使用其周围的宏块来进行错误隐藏。这时,如果

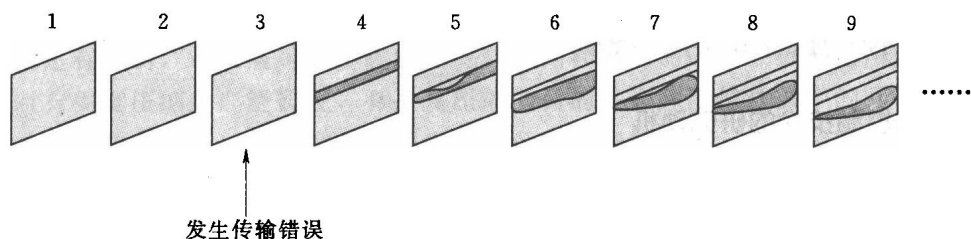


图 1.2 传输错误在空域和时域扩散

破坏的宏块数量较多, 则可能产生较为严重的失真。图 1.3 给出了重建宏块的视觉效果, 与原始帧相比, 重建的视频帧中许多标注的宏块出现了严重的失真。

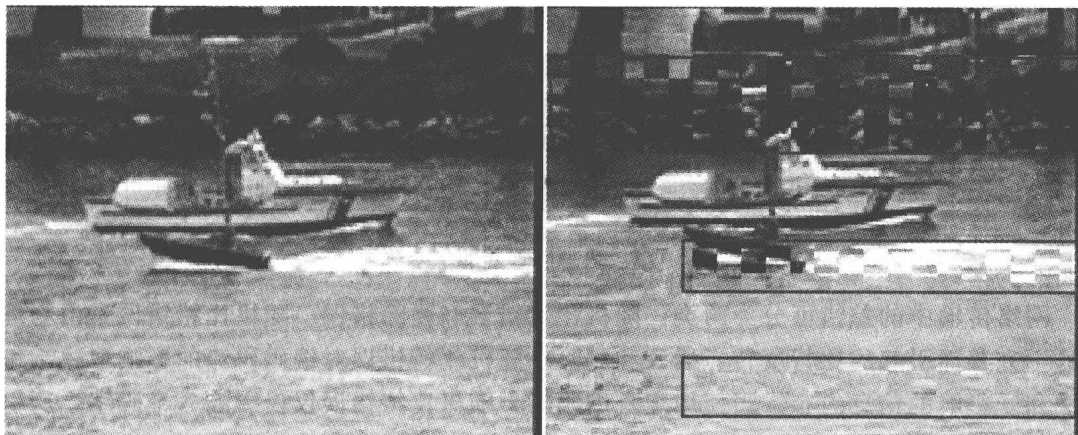


图 1.3 压缩后的视频流在数据包丢失情况下的重建效果
(左: 原始帧; 右: 丢包率 40%)

1.2.2 差错控制技术的分类

现有的视频传输的差错控制技术^[2,4]可以分为以下 4 类:

- (1) 传输层的差错控制, 包括信道编码, 打包复用以及传输协议等。
- (2) 信源编码器端的差错控制, 即抗误码编码, 目的是使得编码后的比特流对于潜在的传输错误具有一定的鲁棒性。
- (3) 解码器端的差错控制, 即在解码器端检测错误, 隐藏错误造成的后果。
- (4) 编解码器交互的差错控制, 即编码器能够根据解码器检测的错误反馈自适应地进行调整。

1. 传输层的差错控制

传输层的差错控制技术是在编码后的比特流中进行检错和纠错, 以及在必要的时候重



传被破坏的数据。在实际应用中, 这些方法很难完全纠错, 因此, 在解码器收到的比特流中仍然存在比特错误或包丢失的错误。其具体控制方案主要有: 前向纠错、交织打包机制、限制延迟的重传机制、不平等的错误保护等。

(1) 前向纠错。FEC 是数据通信中用于检错和纠错的最常见方法。将 FEC 直接应用于压缩比特流时, 对随机比特错误比较有效。在 H. 261 和 H. 263 标准中都有相应的使用方案。

对于基于包交换的传输来说, 应用 FEC 有一定的困难, 主要原因是当出现一个数据包丢失时, 必须恢复几百个比特。因此, FEC 通常跨越数据包使用, 使得包丢失将只导致 FEC 块中的一个或几个字节的丢失。文献 [3] 中, 用 RS 编码与块交织相结合来恢复丢失的 ATM (Asynchronous Transfer Mode) 单元, 如图 1.4 所示。对于存储器中每行 28 字节的视频数据分别使用 RS (32, 28, 5) 编码, 从而得到每行 32 字节的数据块。然后从存储器逐列读出数据形成 32 个 ATM 单元的有效载荷, 并附加一个字节指明序列号。在接收端, 先进行去交织操作, 如果解码器中检测到丢失一个 ATM 单元, 则表明每个 32 字节的行中丢失了一个字节。根据 RS (32, 28, 5) 编码的纠错能力, 32 个 ATM 单元中最多丢失 2 个单元都能够完全恢复原始的视频数据。

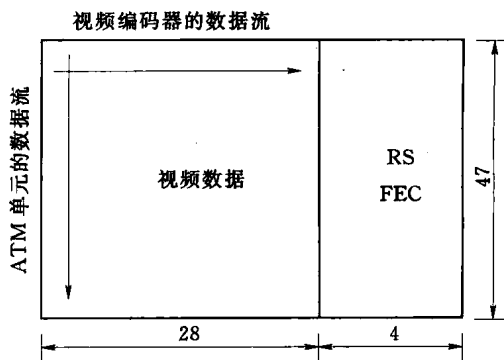


图 1.4 使用 FEC 实现 ATM 单元的丢失恢复 (图中数据单位为字节)^[3]

(2) 交织打包机制。打包机制的重要性就是要使得传输错误能够被孤立在一个小的范围内。这意味着压缩后的比特流不能被依次绑定成为相同大小的数据包, 而是应该根据所使用的编码算法进行打包, 使得一个数据包包含一个或者几个相互独立的编码数据块。为了防止一个数据包丢失而引起连续的块丢失, 可以使用交织打包的机制, 即把相邻块或相邻行的数据放入不同的数据包中。如图 1.5 所示, 将每个视频帧的奇数行和偶数行数据分别打包, 在接收端, 如果奇数行数据损坏, 则可以利用相邻的偶数行数据进行估计, 实现错误隐藏。

数据包 1	数据包 2	数据包 3	数据包 4	...
帧 1 偶数行数据	帧 1 奇数行数据	帧 2 偶数行数据	帧 2 奇数行数据	...

图 1.5 交织打包的一个实例

(3) 限制延迟的重传机制。使用传输协议来实现差错控制, 典型的的就是 ARQ (Auto-



matic Error Request) 技术, 即在检测到数据包丢失或者延迟超时后进行重传。重传机制主要适用于非实时的数据传输, 通常情况下很难满足实时视频的需求。再有, 重传机制也可能进一步恶化网络情况, 造成更多的网络拥塞, 进而增加数据包的丢失率。因此, 设计的重传机制应该将延迟限制在可接受的范围内。

(4) 不平等的错误保护。不平等的错误保护是根据二进制比特流中各比特的不同重要性, 对重要比特设定较高的优先级, 进行重点保护以降低其误码率。不平等的保护可以由重传机制和 FEC 来实现。例如, 在基于块的混合编码中, 帧的头信息和一些边信息比块的数据更重要, 因此, 这些重要的比特就需要增加纠错性能较强的 FEC 来重点保护。对于分层编码而言, 重传机制和纠错性能较强的 FEC 可应用于基本层, 而纠错性能较弱的 FEC 可应用于增强层。另外, 不平等的保护也可以采用不同的传输协议来实现, 例如, 用 TCP 协议传输极少量的高优先级数据, 而低优先级则采用 UDP 协议^[5]。

2. 编码端的差错控制

编码端的差错控制, 即抗误码的信源编码方法, 这些方法能够产生有一定鲁棒性的比特流, 从而在解码端, 即使有错误出现, 依然能够重建可接受的图像视频质量。相对于为了压缩效率而优化设计的信源编码器, 这些考虑传输可靠性的信源编码器在没有发生传输错误时, 要重建同样的图像视频质量, 其编码效率会有一定的损失。这些额外付出的比特, 也称作冗余, 引入它们可以在有传输错误发生的时候, 增强图像视频的重建质量。编码端差错控制的设计目标是在假设的信道环境下, 用给定数量的冗余来获得最佳的解码重建质量, 或者是以保证所需的重建质量为前提, 实现冗余的最小化。向比特流中引入冗余有许多的方法, 主要有错误隔离、鲁棒的二进制编码、限制预测循环、分层编码、多描述编码以及联合信源信道编码等。

(1) 错误隔离。压缩比特流对于传输错误具有敏感性的主要原因之一是视频编码器中变长编码的使用。在变长编码的码字中, 任意比特发生错误或者丢失不仅使得该码字无法解码, 而且也会使得其后续码字即使在正确接收的情况下也可能无法解码或者产生误码。错误隔离技术就是试图将一个传输错误的影响隔离在所限制的区域内。实际应用中经常在压缩后的比特流中设置周期性的“重同步标记 (Resynchronization Markers)”及采用“数据分割 (Data Partitioning)”来实现。具体的讲, 就是把两个同步标记中间的数据进一步分割成更精细的独立可解码的逻辑单元来获得更好的错误隔离。

(2) 鲁棒的二进制编码。该方案也可以直接改进二进制编码, 使得编码后的比特流对于传输错误更加鲁棒。可逆的变长编码 (RVLC, Reversible Variable - Length Coding) 和错误恢复的熵编码 (EREC, Error - Resilient Entropy Coding) 就是这样的技术。如图 1.6 所示, RVLC 算法^[6], 与重同步标记共同使用, 不仅可以前向解码, 也可以后向解码, 使得传输错误影响的区域进一步缩小。EREC 不使用同步码字, 而是利用编码比特的