

面向物联网的 视频编码算法和系统

MIANXIANG WULIANWANG DE
SHIPIN BIANMA SUANFA
HE XITONG

王安红 李志宏 著



國防工業出版社

National Defense Industry Press

面向物联网的 视频编码算法和系统

王安红 李志宏 著

国防工业出版社

· 北京 ·

前言

目前,美国、欧盟各国正紧锣密鼓地加速自己的物联网建设,我国也将物联网建设纳入战略规划。2009年8月7日,温家宝总理在江苏调研无锡传感网中心时指出:“尽快建立中国的传感信息中心,或者叫‘感知中国’中心。”

除了传统的传感器感知的温度、湿度等简单信息外,将来的物联网将同时传送内容丰富的多媒体信息。因此,为了实现“感知中国”,首要任务就要解决多媒体信息的获取、高效编码和无线传输,其中,视频、图像等视觉信息是感知的最直接最重要的信息,同时也是数据量最大的一种信息,为了能在将来的物联网上进行视频传输,必须对其进行有效压缩,这就是视频编码问题。另外,将来的物联网中,物物互联,天罗地网,因此,方便的无线通信将成为视频通信的主要模式。然而,物联网的种种特性和要求,对传统的视频编码和传输算法提出很多挑战,传统的视频编码由于在编码端进行复杂的运动估计等算法,造成很高的编码复杂度,同时,传统算法更加注重压缩效率的提高,对抵抗信道丢包和干扰等情况缺少足够的重视,造成算法的鲁棒性差,不能满足物联网需求。因此,目前,研究面向物联网的视频编码传输算法是物联网专业和视频压缩领域的关键问题,已成为研究热点。

为了解决传统视频编码中编码复杂度高的问题,分布式视频编码(Distributed Video Coding, DVC)得到越来越多学者的关注。DVC基于信息论中的分布式信源编码(Distributed Source Coding, DSC)理论,抛弃了传统的只能由编码端利用信源统计特性的原则,提出在解码端利用信号统计特性同样可以实现有效的压缩,这首先为简化编码端打下了理论基础。同时,由于DVC采用信道码原理来进行信源压缩恢复,因此,DVC具有一定的信道传输的鲁棒性。正因为低复杂度编码和一定程度的鲁棒性,DVC特别适合于现在新兴的低能耗网络终端的视频传输要求,适合将来物联网的视频传输需求。

本书面向将来的物联网传输环境,讲述DVC编码算法和系统,内容包括DVC的原理、算法和主要方法,着重反映近几年来国际上在这一领域的研究水平和动态,书中还包含了作者所在课题组近几年在DVC方面的研究成果。全书共分为11章。第1章为绪论,讲述视频编码的传统方法和物联网传输环境的特点。第2章扼要介绍DVC的相关理论基础——DSC原理,并对目前常用的两种DSC方法从性能上进行了仿真对比。第3章~第8章分别介绍重要的DVC方法,包括目前应用最为广泛的Stanford的DVC编码框架、格型量化边信息编码、DVC中的边信息生成方法、高效

的混合 DVC 方法、可伸缩的 DVC 和鲁棒 DVC 方法。第 9 章介绍我们构建的一个基于 DVC 的手机视频通信系统。第 10 章介绍一种结合分布式压缩感知的 DVC 算法。第 11 章总结全书。

本书可为物联网的视频压缩技术提供很好的理论基础,可为物联网、信号与信息处理、通信工程、计算机等领域中关注视频压缩和传输的工程技术人员、科研教学人员、研究生和大学本科高年级学生使用。

本书凝结了课题组全体成员历年来研究工作的心血。第 1 章、第 4 章、第 5 章、第 6 章、第 7 章、第 11 章的部分内容由王安红执笔;第 2 章、第 3 章、第 8 章、第 9 章、第 10 章由李志宏执笔。王安红负责全书的组织、整理和统稿。北京交通大学的赵耀教授在 DVC 理论研究方面给予大量无私的指导和帮助,林春雨在 DVC 仿真算法中进行了大量的研究,为本书提供了重要的素材。张雄、白慧慧等参加了研究工作并提出了许多有价值的见解。

本书内容涉及的研究项目得到国际科技合作计划(2010DFA11010)、国家 863 计划(2007AA01Z175)、国家自然科学基金(61073142)、太原科技大学博士启动基金项目(20092011)资助。

由于作者水平有限,书中不足和疏漏之处在所难免,敬请同行专家和广大读者不吝指导。

作者

2011 年 3 月于太原

目 录

第1章 绪论	1
1.1 背景	1
1.1.1 传统视频图像编码技术	1
1.1.2 物联网传输环境	2
1.2 分布式视频编码及研究现状	4
1.2.1 高效率 DVC 的研究现状	6
1.2.2 可伸缩 DVC 的研究现状	9
1.2.3 鲁棒 DVC 的研究现状	10
1.3 本章小结	11
参考文献	11
第2章 相关理论	18
2.1 引言	18
2.2 DSC 原理介绍	18
2.2.1 基于 Turbo 码的 Slepian-Wolf 编码器的实现	21
2.2.2 RCPT 编解码原理	26
2.2.3 实验结果及分析	27
2.2.4 基于 LDPC 码的 Slepian-Wolf 编码器	28
2.2.5 LDPCA Slepian-Wolf 编解码算法	30
2.2.6 实验结果及分析	31
2.3 本章小结	32
参考文献	32
第3章 基于 LVQ 的像素域 Wyner-Ziv 编码算法	34
3.1 引言	34
3.2 格的定义和性质	35

3.3	LVQ 算法	36
3.4	经典的 LVQ 边信息编码方案	38
3.5	边信息生成方法	39
3.5.1	运动补偿内插	39
3.5.2	运动补偿外插	45
3.6	像素域 LVQ-WZ 编码方案	46
3.6.1	编码框架设计	46
3.6.2	编解码步骤	46
3.6.3	基于边信息的 LVQ 精细重构	47
3.6.4	实验结果及分析	48
3.7	本章小结	52
	参考文献	52
第 4 章	基于运动补偿精细的小波域 Wyner-Ziv 编码	54
4.1	引言	54
4.2	编码框架设计	54
4.3	主要模块介绍	55
4.4	实验结果及分析	58
4.5	本章小结	61
	参考文献	61
第 5 章	基于分块 SW-SPIHT 的可伸缩 DVC 算法	62
5.1	引言	62
5.2	小波变换和 SPIHT 算法介绍	62
5.2.1	SPIHT 中的小波树结构	63
5.2.2	SPIHT 算法原理	66
5.3	基于 SW-SPIHT 的可伸缩 WZ 方案	68
5.3.1	SW-SPIHT 算法	68
5.3.2	基于边信息的小波系数重构	69
5.4	基于小波块 SW-SPIHT 的可伸缩 WZ 视频编码方案	70
5.4.1	编码框架	70
5.4.2	实验结果及分析	72
5.5	本章小结	75
	参考文献	75

第 6 章 基于 LQR-Hash 的残差 DVC 算法	76
6.1 引言	76
6.2 经典方案分析	76
6.3 残差 LQR Hash DVC 算法	77
6.4 实验结果及分析	79
6.5 本章小结	85
参考文献	86
第 7 章 基于帧内模式决策的混合 DVC 算法	87
7.1 引言	87
7.2 编码框架及算法	87
7.3 实验结果及分析	91
7.4 本章小结	95
参考文献	95
第 8 章 鲁棒 DVC 算法	96
8.1 引言	96
8.2 多描述编码思想	97
8.3 基于优化零填补的鲁棒多描述 DVC 编码框架	98
8.3.1 编码框架	99
8.3.2 优化的前处理和后处理过程	99
8.3.3 实验结果及分析	104
8.4 高效鲁棒的 2D-DVC 方案	108
8.4.1 基本思想	108
8.4.2 编码框架	109
8.4.3 实验结果及分析	111
8.5 本章小结	116
参考文献	116
第 9 章 基于 DVC 的手机视频通信系统	117
9.1 引言	117
9.2 系统框图	117
9.3 软硬件环境	118

9.3.1 硬件配置	118
9.3.2 软件环境	118
9.4 智能手机端 WiFi 通信连接的建立	120
9.4.1 WiFi 简介	120
9.4.2 网络通信的具体实现方法	120
9.5 系统性能	121
9.6 本章小结	123
第 10 章 残差分布式压缩感知视频编码	124
10.1 引言	124
10.2 相关工作	125
10.2.1 DVC 的相关模型	125
10.2.2 分布式压缩感知编码	125
10.2.3 梯度投影稀疏重建	126
10.3 残差分布式视频编码	126
10.3.1 RDCVS 的基本思想	127
10.3.2 残差联合稀疏模型	127
10.3.3 RDCVS 编码	128
10.3.4 RDCVS 解码	128
10.4 实验结果	130
10.5 本章小结	132
参考文献	132
第 11 章 总结和将来工作展望	134
11.1 总结	134
11.2 工作展望	135

第1章 绪 论

1.1 背 景

1.1.1 传统视频图像编码技术

在家庭影院、VCD、DVD 等多媒体应用和可视电话、会议电视等可视化通信中,如何有效减少所需的数据量和占用的频带是必须解决的重要问题。在这些应用场合中,图像和视频占用的数据量最多,因此,如何用尽可能少的数据来尽量不失真地表示图像和视频已成为这些应用的关键,这就是图像和视频的编码压缩问题。

图像压缩已研究了几十年,提出了诸如 DPCM、DCT、VQ 等压缩方法,ISO/IEC、ITU-T 等国际组织也先后制定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^[1],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^[2]是其主要内容),以 H.261、H.263、H.263+、H.263++、H.264/AVC^[3]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 标准等。

上述方法统称为传统编码算法,传统的视频编码算法如图 1-1 所示。对关键帧,编码器直接采用正交变换、量化及熵编码等手段进行帧内编码,去除帧内相关性;

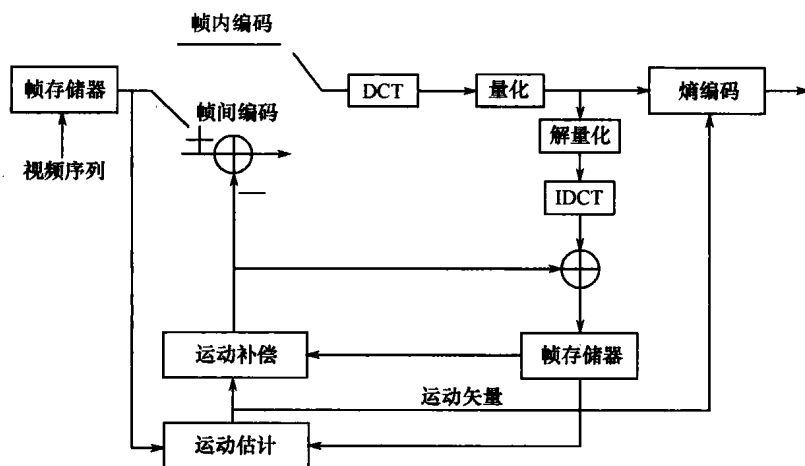


图 1-1 传统的视频编码器框架

对非关键帧,除了采用正交变换、量化及熵编码等手段去除帧内相关性以外,更主要的是采用基于运动估计和运动补偿的帧间预测编码技术来去除帧间相关性,降低传输码率,提高压缩性能。而传统的视频解码算法首先采用帧内解码来恢复关键帧,然后根据接收到的运动矢量、预测误差、量化参数等进行帧间解码,重构非关键帧。

1.1.2 物联网传输环境

2005 年,国际电信联盟(ITU)在一份报告中首次为我们描绘了一个神奇的“物联网”(The Internet of Things)时代:公文包会提醒主人忘带了什么东西;衣服会“告诉”洗衣机对颜色和水温的要求……物联网将传感器(Sensor)嵌入到各种物体中,将实现感知信息与现有互联网(Internet)的整合。一个简单的物联网例子如图 1-2 所示,主要包括中心计算机、互联网和各类传感器。

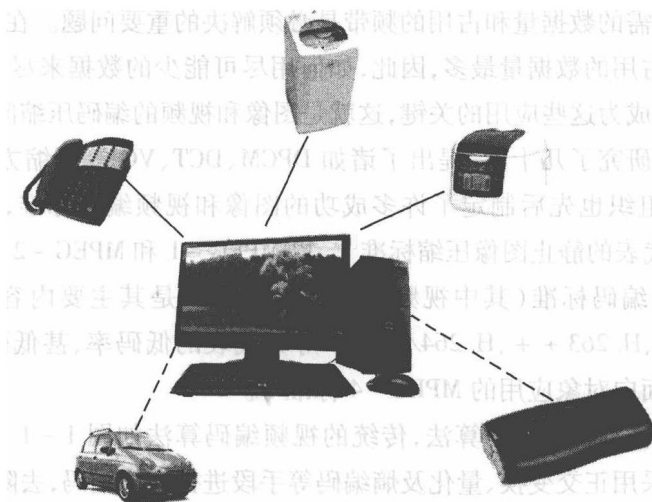


图 1-2 简单的物联网例子

目前,美国、欧盟各国正紧锣密鼓地加速自己的物联网建设,我国也将物联网建设纳入战略规划。2009 年 8 月 7 日,温家宝总理在江苏调研无锡传感网中心时指出:“尽快建立中国的传感信息中心,或者叫‘感知中国’中心。”

除了传统的传感器感知的温度、湿度等简单信息外,将来的物联网将同时传送内容丰富的多媒体信息。因此,为了实现“感知中国”,首要任务就要解决多媒体信息的获取、高效编码和无线传输,其中,视频、图像等视觉信息是感知的最直接、最重要的信息,同时也是数据量最大的一种信息。为了能在将来的物联网上进行视频传输,必须对其进行有效编码压缩。另外,将来的物联网中,物物互联,天罗地网,因此,方便的无线通信将成为视频通信的主要模式。然而,物联网的种种特性和要求,对传统的视频编码和传输算法提出很多挑战。

(1) 低速采样和高质量的视频图像获取。由于设备内存小和无线网络带宽较窄的特殊性,将来的物联网将要求低速的信号采样频率以减少数据量。而传统方法一般按照 Nyquist 定理对视频信息进行高速采样,造成采样数据量大的问题,成为设备内存和后期压缩编码的负担。在带宽较窄的无线通信中,这种传统采样方法很难取得高质量的视频图像,不能满足物联网通信的需求。

(2) 低复杂度视频图像编码。物联网通信将使用大量的移动视频摄取设备,由于这些设备运算能力低和电能的有限性,已经远远不能适应传统编码算法的高复杂性。

传统的视频编码方法,如 MPEG、H.26X 系列标准,在编码端由于采用了运动估计、运动补偿、正交变换、标量量化和熵编码,造成了较高的计算复杂度。其中运动估计是去除视频帧间相关性的主要手段,但同时也是运算量最复杂的操作,因为对于每一编码块,要和参考帧的每一块作相似度计算;相对而言,在解码端,由于没有了运动估计的搜索操作,其复杂度比编码端简单了 5 倍~10 倍。因此,传统的视频编码方法适用于编码端运算能力较强或非实时的一次压缩多次解码的应用场合,如广播、流媒体的点播服务等。而物联网的通信中将出现大量的移动视频摄取设备,如移动照相机、大规模传感器网络、网络视频监控等。这些设备都具有图像和视频的摄取功能,需要进行现场的视频编码,并把码流传送到中心节点,进行解码播放。这些设备比较简单,本身的运算能力和电能十分有限,在功率、显示处理能力以及内存、硬件的支持等方面与传统计算设备存在很大差异,已经远远不能适应传统视频编码中运动估计等算法的高复杂性,而其解码端(如基站、中心计算机等)拥有较多的计算资源,可以进行复杂的计算,与传统视频编码适用的场合恰恰相反。

(3) 高密鲁棒信息传输。物联网将实现人和物体、物体和物体间的全方位互联,这样造成微小区域内基站连接节点的高度密集性。高度密集的通信节点之间干扰严重,将带来传输误码率高、丢包严重等不可靠传输问题,对视频编码算法的传输鲁棒性提出很苛刻的要求。

物联网络中存在信道干扰、网络拥塞和路由选择延迟等问题,将导致数据错误及分组丢失等现象,而无线通信网络信道的随机比特错误和突发性差错等问题进一步恶化了信道状况,导致所传输的视频数据的大量字段失效或彻底丢失。这些问题对于压缩数据是致命的,因为压缩后的数据一般是由不等长码构成的码流,如果出现错误或数据包丢失,都会引起错误扩散等一系列问题,不但严重影响视频业务的服务质量,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。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 Granularity Scalability, FGS)^[4]、吴枫等人提出的更高质量的渐进

精细粒度可分级编码(Progressive Fine Granularity Scalability, PFGS)^[5]等,也开始试图采用新的编码框架来更好地适应网络传输。在 FGS 编码中,基本层采用较强的错误保护措施,如较强的 FEC 和 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,当网络包丢失严重时,系统质量将严重下降,特别是基本层错误会导致解码重建质量的严重下降;其次,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。

综上所述,为了在将来的物联网上向用户提供高质量的视频服务,就必须克服终端运算能力低以及现有网络中不可靠传输带来的问题,因此,需要设计同时具有编码复杂度低和容错能力强的视频编码。对此,我们课题组从 2005 年,开始关注面向物联网的视频编码技术,在深入分析 DVC、分布式压缩感知等理论的基础上,从信号采集、压缩和传输等环节的简化进行研究,以实现低复杂度、鲁棒的视频编码传输算法,并提出了一些可行的编码方案,构建了一套适合低复杂智能手机无线通信的视频编码系统,这些工作将在本书第 3 章~第 10 章中详细讲解。下面首先介绍分布式视频编码的基本原理和研究现状,有关的一些信息论理论将在第 2 章介绍。

1.2 分布式视频编码及研究现状

为了解决传统视频编码中编码复杂度高的问题,DSC(Distributed Source Coding)得到越来越多学者的关注。DSC 基于 20 世纪 70 年代的信源编码理论:无失真情况下的 Slepian-Wolf^[6,7]理论和有失真情况下的 Wyner-Ziv^[8-10]理论,包括后面的解码端边信息 Wyner-Ziv 编码理论^[11,12];这些理论抛弃了传统的只能由编码端利用信源统计特性的原则,提出在解码端利用信号统计特性同样可以实现有效的压缩,这首先为简化编码端打下了理论基础。

分布式视频编码(DVC)^[13]是 DSC 理论在视频压缩中的成功应用,其基本思想是将视频的相邻帧看作相关信源,对相邻帧采用“独立编码和联合解码”的编码框架,这与以往视频编码标准 MPEG 对相邻帧的“联合编码和联合解码”结构有着本质的区别。典型的 DVC 如图 1-3 所示,在待编码的图像序列中等间隔地抽取一组帧,称为关键帧,其编码和解码都采用传统的帧内方式,如 H.264 帧内编解码技术。在关键帧之间的其他帧,称为 WZ 帧,这些帧则采用帧内编码帧间解码的 WZ 编码方式,由于 WZ 编码将传统视频编码算法中计算量巨大的运动估计部分或全部转移到解码端,DVC 实现了低复杂度编码;其次,WZ 编码器中,Slepian-Wolf 编码器由信道码形成,其解码端采用信道码纠错算法,当信道码的纠错能力较强时,即使 WZ 码流在传输过程中出现错误,此错误也可以被纠正,因此,DVC 具有一定的信道传输的鲁

棒性。正因为低复杂度编码和一定程度的鲁棒性,DVC 特别适合于现在新兴的低能耗网络终端的视频传输要求,适合将来物联网的视频传输需求。

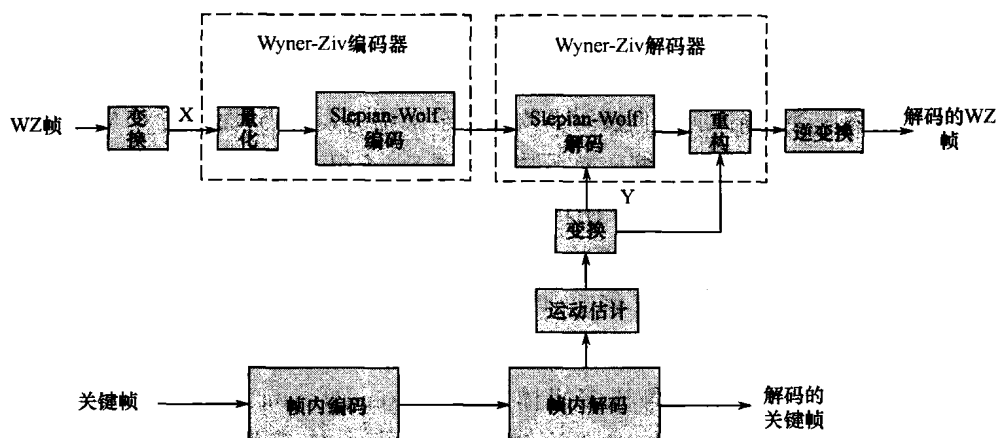


图 1-3 典型的 DVC 框架

一个 DVC 在低能耗手机通信中应用的例子如图 1-4 所示,它采用转码(Transcoding)方式实现了低运算能力手机间的视频通信。以手机 A 到手机 B 的通信为例,手机 A 用低复杂度编码的 DVC 方式进行压缩,然后将压缩码流传输到网络中心处理器,中心处理器将分布式视频码流转码为 MPEG 码流,然后传给手机 B,手机 B 用较低复杂度的 MPEG 解码算法得到恢复的视频。这种通信模式融合了 DVC 和传统 MPEG 编码方式的优势,移动终端只需简单计算量,而将大批计算集中到网络中某特定设备进行,满足了低能耗视频设备的“低复杂度编码”需求。

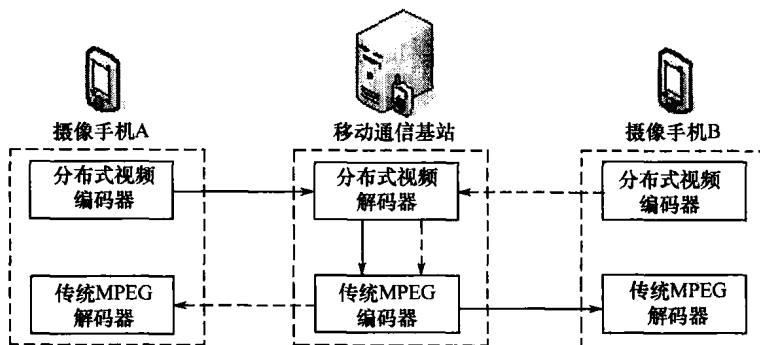


图 1-4 无线视频的转码结构

综上所述,DVC 将 DSC 原理应用在视频编码中以实现低复杂度的视频编码框架,然而,DVC 作为一种有别于传统编码的全新的编码框架,它在压缩性能、适应网络传输的鲁棒性和可伸缩性能等方面还有许多需要改进的空间,以下各节将从各个方面分析其研究现状和不足。

1.2.1 高效率 DVC 的研究现状

再来分析图 1-3 的编码框架,一般而言,WZ 编码器由量化器和一个基于信道码的 Slepian-Wolf 编码器组成,就 WZ 的输入 X (也称为主信息)而言,DVC 可分为像素域和变换域两类方案,前者直接对 WZ 帧像素采用 WZ 编码,而后者对 WZ 帧先进行变换,然后用 WZ 编码器压缩变换系数。美国 Stanford 大学的 Girod 等较早实现了像素域 DVC^[14-16],对每个像素采用均匀标量量化,然后对量化序列用基于 Turbo 码的 Slepian-Wolf 编码器压缩。像素域 WZ 编码取得了介于传统帧内和帧间编码的率失真性能;进而,Girod 课题组又将 DCT 变换用于 DVC,提出了基于 Turbo 码的 DCT 域 DVC^[17,18]。美国 California 大学的 Ramchandran 课题组也提出了 DCT 域 DVC 方案,即基于伴随式的高效率鲁棒多媒体编码(Power-Efficient, Robust, hIgh-compression, Syndrome-based Multimedia Coding, PRISM)^[19-21]方案,他们对 8×8 块的 DCT 变换系数进行标量量化,用 Trellis 码压缩量化的 DCT 系数。由于变换编码进一步去除了图像的空间冗余,因此,DCT 域 DVC 在编码效果上要优于像素域的方法。在上述所提方案的基础上,又有人提出了一些改进算法来提高 DVC 的性能,如小波域的 PRISM^[22]、欧洲 DISCOVER 组织基于 Girod 框架所提出的一系列 IST-DVC 算法^[23,24]。国内许多高校和科研院所也从各个方面对 DVC 理论和方法进行了大量的有价值的研究,如微软亚洲研究院、中国科学院对 DVC 的优化的量化方案^[25]、哈尔滨工业大学提出的小波域 DVC^[26,27]方案、仅在解码端利用空间相关性的像素域 DVC^[28]、WZ 多视点编码研究^[29]、清华大学的基于感兴趣区域的 DVC 算法^[30]、宁波大学的 DVC^[31]、山东科技大学的 DVC^[32]、上海交通大学的 DVC^[33]、南京邮电大学的空间域 DVC 改进方案^[34]、西安交通大学的边信息改进方案^[35]、本课题组在 Girod 框架之上提出的像素域^[36]和小波域^[37]的 LVQ-DVC 框架等。然而,目前的研究结果表明,DVC 的性能介于传统帧内编码和帧间编码之间,与传统的帧间视频编码标准相比还是具有较大的差距,因此,如何提高 DVC 的压缩性能是当前的研究热点之一,下面按照 DVC 框架的各个模块作些分析。

首先,在量化器模块设计方面,WZ 编码器中的量化器对信源进行一定压缩,同时将信源表示为一个索引序列,以方便解码端利用边信息进行索引恢复。为实现简单,几乎所有的 DVC 都采用均匀标量量化,如像素域 DVC 中将 SQ 直接应用到各个像素、DCT 域 DVC 将标量量化应用于 DCT 系数等,但简单标量量化的特性不能令人满意。一些文献对 WZ 编码中量化器做了理论和应用研究。Zamir 和 Shamai^[12]证明,高信噪比时,当主信息和边信息是联合高斯信源时,则嵌套的线性格型量化可以接近 WZ 率失真函数,于是,参考文献[38,39]做出了启发式的设计方案,而 Xiong 等^[40,41]研究提出了一个嵌套格型量化和 Slepian-Wolf 编码器,后来又运用了 Trellis

和格型结合的方案^[42]。对于 DSC 量化器的优化问题,参考文献[43]考虑用 Lloyd 算法^[44]取得局部最优的 WZ 矢量量化器的设计问题,实现了固定码率的优化。参考文献[45]采用率失真优化的矢量量化,将码率看成是量化索引的函数,但是方案效率低下并且复杂。参考文献[46]实现了在相邻区域寻找局部最优的量化问题,但是参考文献[47]说明参考文献[46]的方案由于码书的相邻,限制了其全局的最优化。参考文献[48]考虑了将 Lloyd 应用于没有边信息的 Slepian-Wolf 编码。Girod 课题组将 Lloyd 方法应用到一般情况的理想 Slepian-Wolf 编码器,其码率依赖于量化器的索引和边信息^[49],而其随后的参考文献[50]表明,在高码率时,在某些约束下,最优化的量化器是格型量化器,这验证了参考文献[49]的实验结果。另外,参考文献[51]中提出将小波变换和嵌入式的 SPIHT 量化方法应用于多谱图像的压缩中。哈尔滨工业大学^[27]和我们的方案^[52-54]中,将小波变换和 SPIHT 应用于 DVC 中,小波系数和 SPIHT 量化的完美结合使得方案取得较好的率失真性能。总之,追求简单实用而优化的变换和量化方法将是提高 DVC 性能的一个关键。

其次,在 Slepian-Wolf 编码器模块方面,许多研究者提出了许多改进方法。Slepian-Wolf 编码器是 DVC 中的另一个关键技术,尽管 20 世纪 70 年代的理论就表明 Slepian-Wolf 编码和信道编码息息相关,但仅在近几年随着高性能的信道码,如 Turbo 码、LDPC 码的出现,才逐渐出现了一些实用的 Slepian-Wolf 编码器。1999 年,Pradhan 和 Ramchandran 较早提出使用 Trellis^[55-59]作为 Slepian-Wolf 编码器,后来参考文献[60]中提出了嵌入式的 Trellis Slepian-Wolf 编码器,从此,更高性能的信道编码技术被利用到 DSC 中,如基于 Turbo 码^[61-68]的压缩方案。后来的研究中发现,基于低密度奇偶校验码(Low Density Parity Check, LDPC)的 Slepian-Wolf 更能接近理想限,参考文献[69-71]利用 LDPC 编码器来压缩二进制信源,自此,引发了人们对这一领域的研究兴趣和广泛关注。Slepian-Wolf 编码器的码率和理想 Slepian-Wolf 限的差值反映了其性能的好坏,目前较为常用的 Turbo 码 Slepian-Wolf 编码器距离理想界限 3% ~ 11%^[62],而基于 LDPC 的 Slepian-Wolf 编码器距离理想界限仍然有 5% ~ 10%^[71],在主边信息相关性低和码长较短时差距较大,因此,追求更高的压缩率、缩减与 Slepian-Wolf 限的差距将是 Slepian-Wolf 编码器在很长一个时间内的研究目标。

再有,实现码率的自适应控制也是 Slepian-Wolf 编码器实用化的关键问题,Slepian-Wolf 编码器在 DVC 中的作用与熵编码在传统编码中的作用类似,但是在传统编码中,由于编码端知道信源的统计相关性,因而,编码端可以根据相关性准确地发送比特从而实现无损恢复。但是在 DVC 中,由于编码端不知道边信息的相关性,从而不能准确知道解码端无损恢复所要求的比特数,导致码率控制较为盲目。目前,常用解码端和反馈信道来实现 DVC 的码率自适应,如 Girod 方案等,但反馈给实际应用

带来了局限性。PRISM 方案提出通过在编码端进行简单的时间相关性估计来发送伴随式,方案虽然没有使用反馈,但因为编码端和解码端相关性的差距,导致伴随式比特的不准确。后来的参考文献也研究了如何去掉反馈信道问题,如参考文献[72]提出用编码端的码率控制来去掉反馈,但是比解码端控制的率失真性能降低了约 1.2dB。参考文献[73]提出在编码端利用位平面的交叉概率来估计 DVC 要发送的校验位,去掉反馈信道。参考文献[74]提出用 Fold 函数对小波系数进行处理,利用 Fold 函数的周期性、边信息的相关性来去掉反馈,进而,参考文献[75]又利用一个 Laplacian 模型来估计量化后主边信息的交叉概率,然后根据交叉概率来发送合适的 WZ 比特以去掉反馈信道。参考文献[76]提出了像素域的无反馈 DVC 方案。参考文献[77]提出了多传感器 DVC 中自适应码率分配和变参数量化问题,可根据视频的运动情况和实际的信道统计特性来分配码率。

编码端的运动估计是传统视频编码成功的一个重要因素,相比之下,现有 DVC 方案将运动估计移到解码端,利用解码端的恢复帧来进行运动估计以产生边信息。然而,不准确的恢复帧造成了运动估计的不准确,这将引起边信息的性能下降,最终造成 DVC 性能的下降,因而,提高运动估计性能是提高 DVC 性能的一个关键环节。DVC 中的运动估计最早是由 PRISM^[19-21] 小组提出的,他们对 DCT 块的量化系数计算校验码 CRC (Cyclic Redundancy Check) 并发送到接收端,通过比较边信息中的参考块和当前块的校验码来辅助运动估计,但方案较为复杂。Stanford 的 Girod 等人^[14-17] 首先用解码端的运动估计内插来产生边信息,但由于此方法没有利用当前帧的任何信息,因而性能较低。为了保持编码器的简单特性并能获取当前帧的信息, Girod 在参考文献[78]提出了一种基于 Hash 的运动估计方法,即编码端采用 DCT 量化系数的一个子集作为 Hash 并发送到解码端,解码器利用接收到的 Hash 信息,在解码帧的参考块中进行运动估计以获得好的边信息。实际上,PRISM 方案的 CRC 也可以看作是一种 Hash 信息。参考文献[79]在 Girod 基础上提出自适应的 Hash,参考文献[80,81]提出一种低质量参考 Hash,即 WZ 帧用零运动矢量 H.264 压缩的版本。但是,Hash 信息的码率、复杂性及其对运动估计性能提高的有效性等还有待进一步研究。参考文献[82]提出解码端的多边信息运动估计生成方法,但解码端复杂度有所升高。另外,有些参考文献提出一种编、解码端共同分担运动估计的方法来提高性能,如参考文献[83]用光流场估计法在编码端快速得到块的运动情况,解码端根据此情况选择合适的边信息生成方法,但是这些方法一定程度地增加了编码端复杂度。另外,参考文献[23,37]提出了解码端边信息逐渐精细方法,该方法不需要增加额外的 Hash 比特,而是利用部分解码的当前帧信息来逐步更新边信息,因而,也是一种较好的提高编码性能的方法。可喜的是,从 2006 年开始,Stanford 大学用期望最大算法 (Expectation Maximization, EM)^[84] 来学习当前帧和其他帧的视差,形成所谓视差非

监督学习法^[85],为提高边信息的性能提供了很好的思路。参考文献[86-88]将视差非监督法应用于分布式多视点视频编码,取得很好的效果。2008年,他们将此方法应用于单视点DVC的边信息生成^[89],实验表明,基于EM的解码端运动估计取得了很好的效果,提高了边信息性能,使得DVC的性能随着GOP的升高而得到提高。在之前的研究中,因为边信息的性能差,GOP变大时DVC的性能反而变差。后来的参考文献[90]将视差非监督法和格雷码等新技术用在多视点DVC的研究中并取得了明显效果。

另外,DSC和DVC中主、边信息的相关模型在很大程度上影响Slepian-Wolf编码器的性能,目前有参考文献对此进行了理论和应用研究。如参考文献[91]针对高斯信源定义了两个实用的相关性模型。参考文献[92]针对变换域不同子带的主信息和边信息提出不同的相关性模型,并提出动态在线噪声模型以提高相关性估计。对量化后的主、边信息的相关性的表示方法也会在很大程度上影响DVC的性能,格雷码可以对欧几里得距离较小的量用较小的Hamming距离来表示,从而提高了量化二进制序列的相关性表示,并最终提高Slepian-Wolf编码器的压缩率。参考文献[93]用理论和实验证明了格雷码在DVC中的有效性,参考文献[94]提出用格雷码、Zero-Skip、编码DCT系数符号等技术来有效地表示相关性,最终提高了性能。

最后,在DVC的量化重构方面,目前许多参考文献采用量化序列在给定边信息时的条件期望来进行重构。参考文献[95]对重构函数进行了精细,特别是当边信息和解码的量化值不在一个区间时,使用训练和回归法得到误码率和重构值之间的回归线,从而提高了重构的性能。

1.2.2 可伸缩DVC的研究现状

由于网络带宽和客户端网速的异构性,视频编码技术对编码算法的可伸缩性提出了一定的要求。MPEG等标准算法都有自己的可伸缩算法框架,如MPEG-4的FGS、PFGS算法等。为提高DVC对网络带宽的适应能力,有必要研究可伸缩的DVC方案。现有的一些研究对这方面做了一定工作,参考文献[96]首先将连续精细(Successive Refinement)的编码思想应用在WZ编码中,提出了WZ连续精细思想。接着,参考文献[97-99]将此思想应用在DVC中,提出了分层的DVC方案,其中在基层用H.264压缩,增强层用WZ的位平面编码。参考文献[100]将连续精细编码思想应用在MPEG的FGS框架中,在增强层位平面灵活地选用LDPC或者熵编码,实现了比FGS高3dB~4.5dB的可伸缩编码。但是,这些方案在基本层都采用了较为复杂的帧间预测编码,影响了DVC的低复杂度编码;再者,传统的帧间预测编码影响了基本层的鲁棒性。为了克服这些缺点,参考文献[101]在基本层采用H.264帧内编码,增强层采用WZ编码,实现了空间、时间、质量可分级的可伸缩DVC,实现了低