



# 分布式视频编码 算法与系统

● 朱金秀 曹宁 张瑶 著



中国工信出版集团



电子工业出版社  
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY  
<http://www.phei.com.cn>

# 分布式视频编码算法与系统

朱金秀 曹宁 张瑶 著

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

## 内 容 简 介

本书从无线多媒体通信的特点和传统视频编码的原理入手,简要分析了传统视频编码在移动环境下的多媒体通信业务中面临的瓶颈问题。然后介绍了在无线网络技术的发展和推动下目前主流研究的视频编码技术、分布式视频编码的原理及研究现状和分布式视频压缩感知理论及研究现状,并着重介绍了作者在分布式视频编码以及分布式压缩感知视频编码研究中所提出的一些算法和系统。

本书主要分为以下几个部分:绪论、分布式视频编码相关理论、基于相关能量的相关噪声模型研究、基于时空相关性的 DCVS 多假设预测重构算法、基于块测量值的自适应边信息生成算法、分布式视频压缩感知中码率控制算法、基于云平台的视频编码与传输系统等内容。

本书可作为物联网工程、通信工程、计算机科学与技术、电子信息工程等相关专业高年级本科生、研究生和研究人员科研用书,也可作为物联网、视频处理等领域研究开发人员的参考书。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

### 图书在版编目(CIP)数据

分布式视频编码算法与系统 / 朱金秀, 曹宁, 张瑶著. —北京: 电子工业出版社, 2016.5  
ISBN 978-7-121-28670-4

I. ①分… II. ①朱… ②曹… ③张… III. ①视频编码—研究 IV. ①TN762

中国版本图书馆 CIP 数据核字(2016)第 089326 号

策划编辑: 任欢欢

责任编辑: 任欢欢

印 刷: 三河市双峰印刷装订有限公司

装 订: 三河市双峰印刷装订有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编: 100036

开 本: 720×1000 1/16 印张: 9.25 字数: 186 千字

版 次: 2016 年 5 月第 1 版

印 次: 2016 年 5 月第 1 次印刷

定 价: 42.00 元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010)88254888, 88258888。

质量投诉请发邮件至 [zltz@phei.com.cn](mailto:zltz@phei.com.cn), 盗版侵权举报请发邮件至 [dbqq@phei.com.cn](mailto:dbqq@phei.com.cn)。

本书咨询联系方式: 192910558(QQ 群)。

# 前 言

视觉能以直观、生动的方式反映世界，是人类观察和认识世界最为重要的手段，人们获取的信息大约有 70% 来源于图像。随着信息技术和互联网的飞速发展，多媒体特别是视频已成为人们获取信息最主要的载体，同时是信息领域开发和研究的热点之一。然而，由于数据量极其庞大，视频信息如果没有经过高效的数据压缩，会给信息的存储和传输带来很大的困难，因此研究视频的压缩编码方法具有重要意义。

近年来，分布式视频编码（DVC）实现系统开始受到关注，20 世纪 70 年代从理论上证明了多个相关信源独立编码、联合解码同样可以达到联合编码、联合解码的编码效率，虽然上述理论已经提出近 50 年，但由于缺乏具体的实现方法，导致 DVC 的发展较为缓慢。直到近年来，随着无线视频传感器网络和无线视频移动设备的应用需求增多，DVC 重新成为研究热点。目前，分布式编码技术的主要问题在于，分布式编码的性能和传统视频编码标准 H.264/AVC 之间仍存在很大的差距，要改进分布式编码技术性能上的不足，需要多方面的理论支持。对于视频压缩，DVC 是一种全新的框架，该框架虽包括一些传统信源信道编码中常用的环节，如变换、量化、纠错编码、运动估计、码率控制等，但这些环节需要在新的框架下重新设计和优化。另外，DVC 中还有一些特有的环节，如生成边信息的质量、相关噪声模型参数的估计和码率的估计和控制等，需要进一步深入研究。

本书从无线多媒体通信的特点和传统视频编码的原理入手，简要分析了传统视频编码在移动环境下的多媒体通信业务中面临的瓶颈问题，对当前分布式视频编码技术领域的研究热点和难点进行了研究，并着重介绍了作者在分布式视频编码以及分布式压缩感知视频编码研究中所提出的一些算法和系统。全书共 7 章，第 1 章为绪论部分，主要讲述传统混合视频编码技术、分布式视频编码技术和分布式压缩感知视频编码及研究现状。第 2 章阐述了分布式视频编码相关理论，包括分布式视频编码和分布式压缩感知视频编码的原理及系统。第 3~6 章详细介绍了课题组在分布式视频编码方面的研究：DVC 中基于相关能量的相关噪声模型研究、基于时空相关性的 DCVS 多假设预测重构算法、基于块测量值的自适应边信息生成算法、分布式视频压缩感知中码率控制算法。第 7 章介绍了两种分布式视频编码的应用——基于云平台的分布式压缩感知视频编码系统与传输系统及灌区巡视中的飞行器视频传输系统。

本书可作为物联网工程、通信工程、计算机科学与技术、电子信息工程等相关专业高年级本科生、研究生和研究人员科研用书，也可作为物联网、视频处理等领域研究开发人员的参考书。

本书的主要内容来源于作者主持承担的常州科技支撑计划(工业)项目(编号:CE20140079),常州市传感网与环境感知重点实验室建设,江苏省江苏高校品牌建设工程一期项目立项专业(序号:PPZY2015B141),及“世界水谷”与水生态文明协同创新中心的研究成果。本书由朱金秀、曹宁、张瑶著。在本书即将出版之际,由衷感谢本课题组其他成员历年来对研究工作的付出。他们是:汤一彬、程浩、裴颖、刘骏。

刘峰教授对全书进行了审阅,电子工业出版社任欢欢编辑为本书的出版和文字编辑做了大量工作,在此一并表示感谢!

本书是一部针对分布式视频编码关键技术进行研究的学术著作,书中很多内容还处于学术探索阶段,期望其对分布式视频编码理论深入研究具有推动作用,并最终应用于实际的视频编码领域。由于作者的水平有限,书中的不足及疏漏之处在所难免,恳请同行专家及广大读者不吝赐教。

作 者

2016年3月

# 目 录

第 1 章 绪论	1
1.1 传统混合视频编码技术	1
1.2 分布式视频编码技术及研究现状	3
1.2.1 分布式视频编码技术	3
1.2.2 DVC 系统中相关噪声模型研究现状	4
1.3 分布式压缩感知视频编码及研究现状	5
1.3.1 分布式压缩感知视频编码	5
1.3.2 DCVS 系统中关键技术研究现状	7
1.4 本章小结	12
参考文献	12
第 2 章 分布式视频编码相关理论	21
2.1 引言	21
2.2 分布式视频编码原理介绍	21
2.3 分布式视频编码系统	23
2.3.1 基于像素域的分布式视频编码系统	23
2.3.2 基于频域的分布式视频编码系统	24
2.4 分布式压缩感知视频编码原理介绍	25
2.4.1 压缩感知基本理论及主要应用	25
2.4.2 分布式压缩感知理论	30
2.4.3 分布式压缩感知视频编码系统	32
2.5 本章小结	36
参考文献	36
第 3 章 基于相关能量的相关噪声模型研究	42
3.1 引言	42
3.2 DVC 系统中相关噪声模型	42
3.2.1 基于变换域的 DVC 系统结构	42
3.2.2 DVC 中的相关噪声模型	43
3.3 基于相关能量的 DCT 域相关噪声模型	44
3.3.1 相关噪声分布参数估计方法分析	44
3.3.2 噪声 DCT 系数分布的数学分析	45

3.3.3	基于相关能量噪声模型的建立	47
3.3.4	考虑量化噪声的参数估计	49
3.4	实验与分析	50
3.4.1	帧级别 DCT 系数相关噪声模型的实验与分析	50
3.4.2	帧级别 DCT 系数带相关噪声分布的实验分析	55
3.4.3	基于提出相关噪声模型的 DVC 系统实验与分析	57
3.5	本章小结	59
	参考文献	59
<b>第 4 章</b>	<b>基于时空相关性的 DCVS 多假设预测重构算法</b>	<b>61</b>
4.1	引言	61
4.2	视频压缩感知中多假设预测重构算法分析	61
4.2.1	基于块的图像压缩感知算法原理	61
4.2.2	多假设预测压缩感知重构算法分析	63
4.3	基于时空相关性的压缩感知多假设预测重构算法	67
4.3.1	帧内压缩感知多假设预测关键帧重构算法	67
4.3.2	帧内/帧间压缩感知多假设预测 CS 帧重构算法	69
4.3.3	基于时空相关性的自适应多假设预测重构算法	70
4.4	实验与分析	73
4.4.1	关键帧帧内多假设预测重构实验与分析	73
4.4.2	CS 帧三种多假设预测重构实验与分析	74
4.4.3	自适应多假设预测重构系统实验与分析	76
4.5	本章小结	80
	参考文献	80
<b>第 5 章</b>	<b>基于块测量值的自适应边信息生成算法</b>	<b>82</b>
5.1	引言	82
5.2	基于块测量值的自适应边信息生成算法	82
5.2.1	基于块测量值的自适应边信息生成系统结构	82
5.2.2	编码端 CS 块分类及测量值提取	83
5.2.3	运动补偿时域内插生成初步边信息	85
5.2.4	测量域运动估计与残差重构算法	87
5.2.5	测量值矢量重组的边信息修正算法	89
5.2.6	自适应修正边信息算法	89
5.3	实验与分析	90
5.3.1	三种方式产生的残差实验结果比较	91
5.3.2	测量域运动估计与残差重构的实验与分析	92

5.3.3	基于测量值边信息生成算法的实验与分析	95
5.3.4	自适应修正边信息生成实验与分析	97
5.4	本章小结	100
	参考文献	101
<b>第 6 章</b>	<b>分布式视频压缩感知中码率控制算法</b>	<b>102</b>
6.1	引言	102
6.2	分布式视频压缩感知中码率控制算法研究	102
6.2.1	基于压缩感知的量化失真分析与研究	102
6.2.2	压缩感知的压缩重构失真分析与研究	108
6.2.3	DVC 中基于压缩感知的率失真模型研究	113
6.2.4	分布式视频压缩感知中码率控制算法	116
6.3	实验与分析	117
6.3.1	基于压缩感知率失真模型实验与分析	117
6.3.2	基于反馈的码率控制算法实验结果分析	119
6.4	本章小结	122
	参考文献	123
<b>第 7 章</b>	<b>基于云平台的视频编码与传输系统</b>	<b>124</b>
7.1	引言	124
7.2	一种基于云平台的分布式压缩感知视频编码系统	124
7.2.1	基于云平台的分布式压缩感知视频编码系统设计	124
7.2.2	基于云平台的分布式压缩感知视频编码系统详细设计	125
7.2.3	视频转码技术介绍	129
7.3	灌区巡视中的飞行器视频传输系统	131
7.3.1	无人机灌区巡视视频监视系统设计	132
7.3.2	飞行器视频传输系统的设计	133
7.3.3	基于 RTMP 协议的视频系统设计	134
7.3.4	基于 FFmpeg 搭建流媒体服务器	135
7.4	本章小结	139
	参考文献	139
<b>附录 A</b>	<b>FFmpeg 参数说明</b>	<b>140</b>



# 第 1 章 绪 论

## 1.1 传统混合视频编码技术

自 1948 年提出视频数字化概念至今，视频编码技术经历了近 70 年的发展，不仅在理论研究上取得了重大进步，在实际应用中也获得了巨大成果。伴随着现代网络技术和多媒体技术的发展，远程监控、视频会议、视频流媒体、远程教学、远程医疗等服务无处不在，这些以点对点或点对多点视音频通信为主要形式的应用都迫切需要高质量的往来视频传输的支持。高效率、高质量的视频压缩和实时传输技术越来越受到社会各界人士的广泛关注。

视频数据具有庞大的数据量，对于当前有限的传输带宽和存储空间都是难以承受的，这使得视频压缩（编码）技术成为国内外研究和应用的热点<sup>[1]</sup>。为了使不同厂商生产的各种终端之间能够互相兼容和通信，20 世纪 80 年代末，视频压缩编码走向了标准化的路程。国际标准化组织和国际电信联盟先后制定了一系列的视频压缩编码标准<sup>[2]</sup>，从图 1-1 中可以了解到各种视频编码标准的发展及其应用简史。视频图像标准大致分为三个系列：MPEG 系列、H.26X 系列和 AVS 系列。在制定和发布视频图像压缩标准的过程中，这三个系列标准既有各自独立研究和开发的方面，也有相互借鉴和合作的方面。

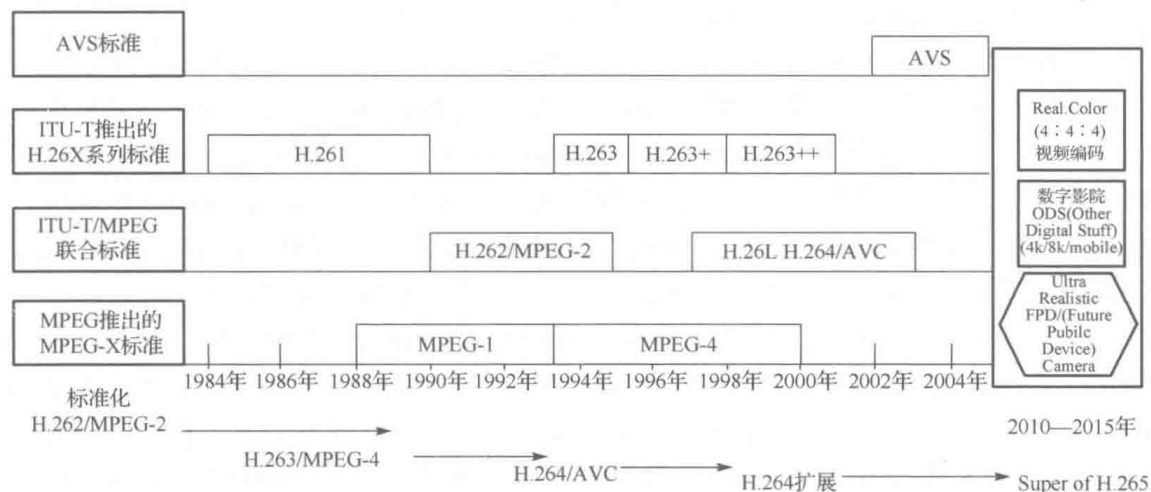


图 1-1 视频编码标准的发展及其应用简史

下面对这些视频编码标准进行简要的阐述。

1991年, ISO/IEC 制定了第一个视频压缩编码标准 MPEG-1<sup>[3]</sup>, 对视频进行压缩编码后的码率为 1.5 Mbit/s, 可以被用于视频的传输与存储。1994年, ITU-T VCEG 和 ISO/IEC MPEG 这两大组织联合颁布了 MPEG-2 标准<sup>[4]</sup>。该标准主要包括系统、视频、音频以及复合性(检验和测试视音频及系统码流) 4 个文件, 并且按照不同的压缩比被分成 5 个档次, 最常用的有 11 种组合, 分别应用于标准数字电视、高清晰度电视等, 码率为 4~100 Mbit/s。1998年, ISO 通过了“视听对象的编码标准”——MPEG-4<sup>[5]</sup>, 该标准首次提出了一种基于对象的编码思路, 并且它在定义视频压缩编码标准的同时, 还强调了多媒体通信的交互性和灵活性, 主要应用于可视电话和视频会议等。

1988年, 由于会议电视和可视电话的需要, CCITT 发布了码率为  $p \times 64$  kbit/s ( $p = 1, 2, 3, 4, \dots, 30$ ) 的 H.261 建议<sup>[6]</sup>。这个视频编码方案的提出对以后出现的各种视频编码标准都产生了极大的影响力。1995年, ITU-T SG15 VCEG 提出了 H.263 的建议<sup>[7]</sup>, 规定传输的最初码率低于 64 kbit/s, 主要适用于低码率图像范围, 但在实际应用中也可适于高速率图像编码。随后, 为了适应低码率传输要求, 进一步提高图像的质量, ITU-T 于 1998 年和 1999 年相继推出了 H.263+ 和 H.263++。2003年, 经 ITU-T SG16 VCEG 和 ISO/IEC MPEG 合作, 正式颁布了 H.264/AVC 视频编码标准<sup>[8]</sup>。H.264/AVC 标准集中以往的视频编码标准的优点, 在编码性能上得到了提高: 在同等图像质量的前提下, H.264/AVC 的压缩比可以达到 MPEG-4 的 2 倍。通常情况下, H.26X 系列标准侧重于视频信息的数据压缩效率, 调整在特定的码率下的传输, MPEG-X 标准则倾向于控制视频编码质量。AVS<sup>[9]</sup>是我国具备自主知识产权的第二代信源编码标准。AVS 目前主要面向于 HDTV、HD-DVD 和宽带流媒体等视频应用, 在 H.264 编码框架的基础上发展自己的编码技术。

目前, “高质量视频”已成为了数字视频应用发展的趋势, 在这种情况下, 急需一种更高效的视频编码标准来解决当前所面临的问题。为此, 制定视频编码标准的两大国际组织纷纷展开了基于下一代视频压缩编码标准的研究工作。ISO/IEC MPEG 于 2009 年的 2 月启动了高性能视频编码 (High-performance Video Coding, HVC)<sup>[10]</sup>的需求草案, 同年 4 月, ITU-T VCEG 启动了下一代视频编码的需求草案<sup>[11]</sup>。2010 年 1 月, 在召开的第 39 次 VCEG 会议暨第 91 次 MPEG 会议上, ITU 提出与 MPEG 再次合作, 并成立了联合工作组 (Joint Collaborative Team, JCT), 共同参与制定下一代视频编码标准 H.265 或 HEVC (High Efficiency Video Coding)。2012 年 8 月, 爱立信公司推出了首款 H.265 编解码器<sup>[12]</sup>, 而在仅仅六个月之后, 国际电联 (ITU) 就正式批准通过了 HEVC/H.265 标准, 标准全称为高效视频编码 (High Efficiency Video Coding, HEVC), 相较于之前的 H.264 标准有了相当大的改善, 中国华为公司拥有最多的核心专利, 是该标准的主导者。

H.265 旨在在有限带宽下传输更高质量的网络视频, 仅需之前带宽的一半即可播放相同质量的视频。这也意味着, 我们的智能手机、平板等移动设备将能够直接在线播放 1080p 的全高清视频。H.265 标准也同时支持 4K (4096×2160) 和 8K (8192×4320) 超高清视频。可以说, H.265 标准让网络视频跟上了显示屏“高分辨率化”的脚步。

## 1.2 分布式视频编码技术及研究现状

### 1.2.1 分布式视频编码技术

虽然分布式视频编码的基础理论早已提出, 但直到近十几年, 随着无线网络技术的发展和推动, 分布式信源编码的实现技术才逐步得到学者们的重视和研究, 做出了很多有意义的成果, 也出现了一些实用的分布式视频编码方案。例如, 斯坦福大学的 Bernd Girod 等人在 2002 年提出了基于像素域的 Wyner-Ziv 视频编码系统<sup>[13]</sup>, 在该编码方案中系统将视频序列分为关键 (K) 帧和 Wyner-Ziv (WZ) 帧, 对关键帧采用传统的 H.264/AVC 帧内编解码技术进行处理, WZ 帧则采用基于信道编码的 Slepian-Wolf 编解码方法; 为了提高系统的率失真性能, 2004 年, Bernd Girod 等人又提出了基于变换域的分布式视频编码方案<sup>[14]</sup>, 在编码端 WZ 帧首先通过离散余弦变换 (Discrete Cosine Transform, DCT) 消除帧内的空间相关性, 获得更高的压缩性能。而美国加州大学伯克利分校 Ramchandran 等人提出了一种 Berkley 视频编码方案, 即著名的 PRISM 系统<sup>[15-16]</sup>, 其中 P 表示高效率 (Power-efficient), R 表示鲁棒性 (Robust), I 表示高的压缩率 (High-compression), SM 表示综合多媒体编码 (Syndrome-based Multimedia Coding), 该方案在 DCT 变换域中以块为单位进行编/解码, 根据视频帧与预测帧中块间的相关性确定该块是否需要编码, 并根据编码所需比特平面数估计相应码率。

Z.Xiong 等人提出了分级 Wyner-Ziv 视频编码<sup>[17-18]</sup>, 该编码方案由采用 H.26X 编码的基本层和基于 LDPC 编码的增强层两个部分组成, Wyner-Ziv 编码器则主要由量化、DCT 和基于 LDPC 码的增强层组成, 基本层和增强层的编码独立进行, 解码端则采用联合解码, 这就避免了因为传输错误而引起的编解码不匹配, 而且具有很好的抗误码特性; 而欧盟 ISTFET 的 DISCOVER 小组提出了 DISCOVER (Distributed Coding for Video Services) 分布式编码框架<sup>[19]</sup>, 该系统基于斯坦福大学提出的频域编码系统而建立, 但其支持变长的图像组 (Group of Picture, GOP), 并且在编码端采用 LDPCA 对 WZ 帧进行编码, 而解码则采用软译码, 编码效率更高。其他的分布式视频编码方案还有 Tagliasacchi 提出的基于小波变换的编码方案<sup>[20]</sup>, 其框架中编码端首先对视频帧进行离散小波变换, 将变换系数分成 8×8 块, 根据当前帧

与前一帧系数之间的相关性将系数块分为跳过块、帧内块和帧间块三类,只对帧间块的 16 个低频系数进行 Wyner-Ziv 编码; Sehgal 等学者提出的“state-free”编码方案<sup>[21]</sup>;以及 Pradhan 等提出的借助信道编码技术实现信源压缩编码的 DISCUS 编码方案<sup>[22]</sup>等。

在这些方案的实现中,编码器都是将每个视频帧看成“独立信源”,即每个视频帧进行独立的编码,而解码器则利用视频帧之间的相关性和信道解码技术进行解码,重构视频帧。其中,解码 WZ 帧涉及两个过程,解码器首先利用已解码的关键帧通过运动估计和运动补偿生成边信息,将边信息与原始 WZ 帧之间的关系视为一个虚拟的噪声信道,利用边信息、信道编码的校验码恢复出解码符号流,然后再结合边信息重构出原始 WZ 帧,所以边信息的生成精度和相关噪声模型的准确度很大程度上决定了分布式视频编码的率失真性能。生成的边信息越精细,所需传输的 WZ 帧的校验比特越少,解码后的 WZ 帧质量也越好;相关噪声模型越精确,边信息与原始 WZ 帧差距的拟合越精确,系统的编码效率越高。所以如何提高边信息的生成质量、改善相关噪声模型以优化解码成为了解码端的两个关键技术。

### 1.2.2 DVC 系统中相关噪声模型研究现状

DVC 中原始 WZ 帧数据和边信息之间相关噪声模型的好坏是影响编码效率的关键因素之一。如果模型精确地描述了 WZ 帧和 SI 帧的关系,则编码效率高,反之,若模型失败或者与真正的相关噪声分布存在较大的不匹配,则会造成编码效率的损失。现有的文献通常认为边信息和当前解码 WZ 帧的概率分布近似符合零均值单参数拉普拉斯分布<sup>[23-25]</sup>。为了使模型更精确地描述实际的噪声分布,研究者们近年来对空间域和频率域两类相关噪声模型提出了改进方法,这些方法大致可以分为四类:第一类是考虑序列时间和空间的变化调整模型参数<sup>[26-28]</sup>;第二类是通过了解码端部分已解码信息渐进学习来逐步修正后续解码的噪声分布模型<sup>[29-33]</sup>。第三类是考虑相关噪声在不同频率带系数分布的变化<sup>[34-38]</sup>;第四类是考虑量化噪声的影响以适合不同的编码器设置的噪声分布模型<sup>[39-40]</sup>。

在第一类的方法中,研究者们考虑序列时间和空间的变化对相关噪声模型的影响。文献[26]考虑序列时间的变化调整模型参数,提出了一种帧级别的解码端估计拉普拉斯分布参数的算法。文献[27-28]考虑序列时间和空间的变化调整模型参数:文献[27]提出了分帧、块、像素 3 种不同级别在解码端估计拉普拉斯分布参数的算法;文献[28]分别在像素域提出了基于序列、帧、块和像素不同级别相关噪声模型及参数估计算法,在变换域提出了基于序列的子带、基于帧的子带以及 DCT 系数不同级别相关噪声模型及参数估计算法。

在第二类的方法中,通过在解码过程中由解码端部分已解码信息得到更接近实际的残差,用残差信息来迭代更新分布参数,以提高后续解码噪声分布模型的准确

性<sup>[29-33]</sup>。文献[29]对当前子带解码后,利用已解码当前 DCT 子带,重新估计当前子带所有的系数方差,根据空间域和变换域残差方差具有线性关系<sup>[41]</sup>,由重新估计的系数方差来更新块方差,新的块方差被用来估计下一子带的系数方差,依次更新至所有子带。文献[30]和[31]提出在解码端利用已解码频率子带信息,对每个 DCT 频率子带噪声方差进行渐进细化来提高噪声分布参数的估计。文献[32]提出基于已解码信息的边信息依赖相关信道参数在线估计算法,该算法首先利用已解码前一位比特面信息,对每个比特面信道进行渐进细化估计来提高估计的准确性,其次,该算法还提出一个新颖的 DCV 结构,该结构采用基于哈希运动估计技术在解码端产生高质量的边信息。文献[33]提出了变换域 Wyner-Ziv 视频编码交叉子带噪声模型精细方法,在文献[28]提出的在线估计残差 DCT 变换系数分布参数的基础上,设计了一个在解码过程中从频率子带和比特面两个角度不断更新残差,以提高噪声估计准确性的自适应渐进精细化方案。

在第三类的方法中,针对相关噪声在不同频率子带的系数分布不同以及重要性不同,研究对不同频率子带系数进行不同方式的参数估计方法[34-38]。文献[34]对变换域不同残差系数带设置不同的参数进行相关噪声模型参数的估计;文献[35]提出了一种带校正系数的虚拟相关信道参数估计方法,其中修正系数通过数据统计试验得到。文献[36]考虑如果视频信号的 Wyner-Ziv 编码要达到相同精度,DCT 变换系数中高频分量对解码端的运动估计精确度要求更高,提出了把低频子带系数作为基本层,采用文献[28]的相关噪声模型参数估计方法;把高频子带系数作为增强层信息,采用期望最大化算法进行相关噪声模型参数估计的方法。文献[37]针对变换域中不同的频率子带对时域相关噪声的敏感度不同,提出了基于小波变换域的虚拟依赖信道模型。文献[38]则认为噪声分布参数不仅仅与频率子带有关,还与双向运动估计残差的大小有关,首先根据块的运动估计残差能量对块进行分类,假定分为  $m$  类,则对块进行  $4 \times 4$  DCT 变换产生 16 个不同频带系数,则对属于第  $k$  类 ( $1 \leq k \leq m$ ) 的第  $j$  子带系数的分布参数通过查表可得,表格通过对不同序列在不同量化参数 QP 下训练得到。

另外,针对量化噪声对相关噪声分布的影响,在建立噪声分布模型时加入了量化噪声分量<sup>[39]</sup>。文献[40]中用在编码端关键帧和快速重构关键帧之间的误差作为 WZ 帧的量化误差发送到解码端,在解码端噪声分布参数估计时考虑量化噪声的影响,以适合不同的编码器设置的噪声分布模型,取得了较好的编码效率。

## 1.3 分布式压缩感知视频编码及研究现状

### 1.3.1 分布式压缩感知视频编码

采样定理是通信与信号处理学科中的一个重要基本结论,它的重要性尤其体现

在现代信号获取、采样、感知和模数转换设备上。1949年,香农在通信和信号处理领域提出了采样定理<sup>[42]</sup>,它可以追溯到报务员和数学家的早期工作<sup>[43]</sup>上。采样定理指出,当采样速率达到信号带宽的2倍以上时,才能由采样信号精确重建原始信号,可见带宽是采样定理对采样的本质要求。然而随着社会信息化的不断深入,人们对信息量的需求也在不断上升,需要信号的带宽也越来越宽,在信息获取中对采样速率和处理速度等要求也越来越高。目前解决这些问题的常见方案是信号压缩,于是引出了一个问题:能否利用其他变换空间描述信号,建立新的信号描述和处理的理论框架,使得在保证信息不损失的情况下,用远低于Nyquist采样定理所需的采样数就能精确或近似精确地恢复信号。传统的采样理论可以应用到任何满足带宽有限约束条件的信号中,通常我们对具有更多约束条件的信号感兴趣:如具有结构特性和更少的自由度,而不仅仅是带宽有限约束特性的描述<sup>[44]</sup>。

近年来,由美国科学院院士D. Donoho、E. Candes及T. Tao等人提出的压缩感知理论(Compressive Sensing, CS)<sup>[27-32]</sup>为解决这一问题开创了新的思路,同时构造了相关的信号恢复算法,成功实现了信号采样与压缩的同步,突破了传统的Nyquist采样定理的限制。压缩感知理论证明:如果信号在某个变换域是稀疏或可压缩的,就可以用一个与变换基不相关的测量矩阵将信号变换得到的高维信号投影到一个低维空间上,这种投影包含了原信号绝大部分的有用信息,采样率要远低于Nyquist采样频率,然后通过求解一个最优化问题就可以从这些少量的投影中以很高的精度重构出原信号。在该理论框架下,采样速率不再取决于信号的带宽,而在很大程度上取决于两个基本准则:稀疏性(可压缩性)和非相干性,或者说是稀疏性和等距约束性。

由于视频序列连续的图像帧之间有很大的相关性,而且视频图像在某些变换域上近似稀疏,因此可以利用视频图像信号的帧内、帧间相关性,将压缩感知理论应用到分布式视频编码中,就形成了分布式视频压缩感知(DCVS)理论。分布式视频压缩感知系统的研究从2009年才开始陆续进行,也取得了不少成果。Josep Prades-Nebot等人于2009年首次提出了具体的DCVS系统方案<sup>[45]</sup>,在该系统中将视频帧组分为关键帧和CS帧,对关键帧进行传统的帧内编解码,CS帧采用基于块的CS测量方法,但是该方案并没有充分利用视频帧之间的相关性。2009年,Thong T. Do等人提出了另外一种DCVS方案<sup>[46]</sup>,与前一种方案不同的是,该方案在编码端先将CS帧分别进行基于帧和基于块的测量、量化后再进行传输;解码端利用关键帧和CS帧基于块的测量值通过多假设预测重构法得到边信息,然后结合CS帧基于帧的测量值重构出CS帧,该方案通过对边信息的利用更加充分地挖掘了视频帧之间的相关性,提升了系统的性能。干宗良等人在文献[47]中提出了一种基于CS的Wyner-Ziv空域可分级视频编码算法,但因为其基本层仍采用传统帧内编码方式,所以系统的率失真性能虽然提高了,但编码端的复杂度并没有降低。上述几个编码

框架只对 CS 帧进行了 CS 测量, 而关键帧仍然采用传统的帧内编码方法。文献[48]提出了一个基于自适应训练学习字典的 DCVS 方案, 该方案在编码端对关键帧采用基于块的测量, 对 CS 帧采用通过解码端反馈的块测量率自适应分配的方式, CS 帧重构采用基于训练学习的冗余字典作为稀疏基, 该字典由边信息和前后相邻的已解码关键帧共同训练构成。文献[49]也提出了对 CS 帧采用解码端反馈的块测量率自适应分配方式。文献[50]在文献[49]的基础上, 对分布式编码压缩感知系统利用自适应测量率进行码率控制。目前关于 DCVS 解码方面的研究比较少。文献[51]利用残差帧修改了联合优化目标函数, 有效提升了图像的重建质量。文献[52]利用图像内部的稀疏部分与非稀疏部分之间的相关性修正了 POCS 算法<sup>[53]</sup>的初始化和迭代过程, 得到了更好的重建质量。文献[54]则根据边信息的稀疏性修正正交匹配追踪 (OMP)<sup>[55]</sup>算法, 从而实现了联合重建。文献[56]提出了利用视频序列的时空相关性, 在压缩感知测量域进行运动估计生成字典的多假设预测重构算法。文献[57]利用主成分分析算法提取帧间运动信息构建稀疏字典, 在重构过程中引入非局部相似性正则化方法以保持视频帧的边缘和纹理特征。文献[58]采用最大似然算法进行字典学习, 将相关性和稀疏约束融合到一个模型中, 在字典学习的同时重构信号。文献[59]利用当前图像的结构信息, 在编码端采用基于 KL 变换的稀疏基进行压缩感知测量和重构。

### 1.3.2 DCVS 系统中关键技术研究现状

#### 1. 边信息生成算法研究现状

最典型的边信息生成方法是帧插值法, 帧插值法分为内插插值法 (MC-I)<sup>[60-63,67-74,77-106]</sup>和外推插值法 (MC-E)<sup>[64-66,75-76]</sup>。MC-I 的主要原理是: 在解码端利用已解码的前后关键帧进行运动估计, 认为边信息帧的待插值块到两个已解码帧预测块的运动矢量大小相同而方向相反, 从而获得当前 WZ 帧像素块的运动矢量, 根据运动矢量和前后关键帧进行线性插值生成边信息。MC-E 的主要原理是: 利用前面已解码视频帧, 包括已解码关键帧和已解码 WZ 帧来估计边信息帧中待插值块的运动矢量。由于 MC-E 算法在生成边信息时利用前面已解码的视频帧, 所以它具有实时性强和对视频帧组的大小不敏感的特点, 适用于实时性要求高的场合, 但是由于 MC-E 算法没有充分考虑前后相邻帧之间的相关性, 当视频帧组数值较小时, 性能远远小于 MC-I 算法。MC-I 和 MC-E 算法都是通过对已解码的相邻帧进行运动估计来寻找相邻帧之间的运动轨迹, 再根据“线性运动的一致性”假设来生成边信息帧。事实上物体并不都是沿线性轨迹运动的, 这在一定程度上导致了产生的边信息不够准确。为了获得更精确的边信息, 国内外研究学者在上述理论的基础上, 主要从以下几个方面进行了研究: 提高对物体运动场估计精度的运动补偿插值算法方面的研究; 从编码端向解码端传送 WZ 帧的某些特征信息, 辅助解码端生成边信

息方面的研究；利用得到的部分解码图像中所携带的关于原始 WZ 帧的信息迭代地改进边信息的研究。

(1) 运动补偿插值算法改进。传统视频编码标准中的运动估计目的是得到失真代价最小的运动矢量，而 DVC 边信息生成中的运动估计目的是尽量估计出真实的运动场。因此，对边信息生成中运动补偿插值法的改进研究的主要目的是能够估计出接近“真实”的运动场。为了改善运动矢量集获得局部运动的一致性，文献[67]提出对前后两帧已解码的图像进行双向运动估计，并对运动场采用加权矢量中值滤波器进行平滑来改善运动域信息，由经过改善后的运动矢量产生的边信息的准确度比 MC-I 等算法要高，文献[68—70]都采用这一方法产生边信息。为了得到连贯和更平滑的运动矢量场，学者们又提出了复杂度比较高的光流分析法[71, 72]。为了处理非零加速度的运动物体，文献[73]提出了高阶运动内插的方法。为了得到较真实的运动轨迹，文献[74]采用线性回归分析法对几个连续的关键帧中每个块运动轨迹进行估计，得到的有效非线性运动矢量对初始线性运动估计得到的线性运动矢量进行修正，产生更准确的边信息。文献[75]提出了外推产生边信息的线性回归模型，文献[76]在文献[75]的基础上提出了基于线性回归模型的外推产生边信息方法，首先对前面已解码的两帧视频分别进行外推预测和基于线性回归模型的前向、后向预测，并对这三种预测结果进行基于概率融合产生边信息。针对单纯的 SAD 块匹配准则中没有考虑相邻像素块对搜索结果的影响，导致运动估计后获得的运动矢量场的空间连续性不好，而真实的运动矢量场在非边缘区域的块间是平滑变化的问题，文献[77]提出了在运动估计搜索过程中考虑相邻块失真对边信息进行优化的算法，该算法除了考虑当前块的失真外，还考虑相邻块在运用当前块的运动矢量后能否获得较小的失真，综合得出最优的运动矢量。文献[78]则提出了一种加权的 SAD 准则，但使用加权 SAD 准则最小化每个像素块的匹配误差，最终得到的运动矢量是独立而不平滑的；文献[79]则综合考虑了视频帧之间的空间相关性和时间相关性，提出一种联合 SAD 和边界绝对误差 (BAD) 的边信息生成算法，其中 BAD 是指搜索块与当前块的空间相邻块边界像素的绝对误差和。为了提高运动估计的精度，文献[80—82]将 H.264/AVC 中的分数像素运动估计、文献[77, 83—84]将 H.264/AVC 中的分层运动估计、文献[80, 85—88]将 H.264/AVC 中的多参考帧运动估计，应用于分布式视频编码中的运动估计。

(2) 基于编码端 WZ 帧特征信息的算法改进。研究表明，由于解码端不能获得当前 WZ 帧的信息，单纯对运动估计算法进行改进的边信息生成方法所能带来的增益有限，特别是当视频运动较为复杂、剧烈的时候，就需要更加有效的算法以估算更准确的边信息。所以为了得到更精确的边信息，可从编码端向解码端传送 WZ 帧的某些特征信息，以辅助解码端完成帧的运动估计和运动补偿。文献[89, 90]在提出的 PRISM 系统中，编码端向解码端传送 WZ 帧进行变换量化后的 CRC 数据，解



码端则利用这些数据改善运动估计的结果从而生成更精确的边信息。文献[91]提出的基于哈希码的运动补偿(MCH)算法中,编码端将原始WZ帧的某些特征信息转换成哈希码传送到解码端,辅助解码端进行运动估计以生成更精确的边信息。文献[92]在此基础上,在编码端对WZ帧中的块进行分类,仅计算并传输那些运动比较剧烈的块的哈希码,辅助解码端进行双向运动估计以及最优块模式选择等来进一步提高边信息的质量。文献[93]提出的基于哈希码边信息生成算法中,解码端检测出初始边信息中的低质量区域,对低质量区域向编码端请求发送其哈希信息以增强质量,并使用遗传算法从多个候选块中选出最好的解决方案。文献[94]提出了基于图像相关性的重叠块双向运动补偿算法,在编码端计算原始WZ帧和前后关键帧的相关性,并将相关性数据发送到解码端以辅助运动估计,得到较精确的运动矢量,结合重叠块双向运动补偿生成边信息。

(3) 解码端迭代改进算法。除了利用编码端传送的WZ帧特征信息外,还可以进一步在解码端充分利用部分已解码边信息帧中所携带的关于原始WZ帧的信息,迭代地改进边信息,这就是迭代边信息生成算法的基本思想。该算法能在不增加码率及编码复杂度的条件下,进一步提高边信息的质量。文献[95]利用部分已解码WZ帧的条件概率密度函数(PDF)来检测初始运动矢量中的可疑矢量,并对可疑运动矢量进行重新估计来修正初始边信息。文献[96]则提出了一种基于比特平面更新的边信息迭代改进算法,即利用已译码的比特平面进行重构得到的部分解码WZ帧,作为更新的边信息来对后续的比特平面译码。文献[97]也提出了一种基于比特平面更新的边信息迭代改进算法,并在迭代过程中采用多假设运动补偿去噪,然后利用每个比特平面的解码信息对边信息进行迭代更新。文献[98]则提出了一种基于DC域运动估计的迭代算法,首先对编码端传输的图像块DC系数进行译码,得到WZ帧的解码DC帧,通过在DC帧和已解码的关键帧之间进行运动估计,用得到的运动矢量修正初始运动估计生成的运动矢量来进行迭代解码。文献[99]在此基础上提出了一种新的改进算法,用得到的DC域运动矢量来修正部分解码WZ帧,再用修正后的部分解码WZ帧和已解码关键帧之间运动估计生成的运动矢量来进行迭代解码。文献[100]提出了一种多分辨率迭代改进方法,解码器从已解码的低分辨率图像中学习,来改进高分辨率图像运动估计,从而有效地提高边信息的质量和系统的编码性能。文献[101]提出在DCT域的迭代算法,在解码出所有的DCT子带后更新边信息,而文献[102]在文献[101]基础上,对更新后边信息采用去方块滤波器,文献[103, 104]同样提出DCT域的迭代算法,不同的是利用前一个已解码频率子带进行重构得到的部分解码WZ帧,作为更新的边信息来对后续的频率子带解码。文献[105]在文献[103, 104]的基础上,在迭代更新过程中采用了根据WZ帧与相邻参考帧的帧距进行自适应设置搜索窗的运动估计算法。

文献[106]提出了在解码端基于视觉可感知的块模式判断方法,该文献从修正边