

# 从单视点 DVC 到多视点 DVC 的研究进展

胡晓飞 朱秀昌

(南京邮电大学 图像处理与图像通信江苏省重点实验室, 南京 210003)

**摘要** 多视点视频编码对编码器的功耗、复杂度、存储能力和传输能力有较高的要求,而分布式信源编码理论恰适应了这一需求。为了使人们对这一技术现状有一概略了解,首先总结了分布式视频编码的基本理论和关键技术,然后介绍了分布式信源编码理论从单视点到多视点的拓展,接着论述了分布式多视点视频编码的研究现状,最后分析了分布式信源编码理论在多视点视频编码领域的发展前景。

**关键词** 分布式视频编码 多视点视频编码 Wyner-Ziv PRISM

中图法分类号: TN919. 81 文献标识码: A 文章编号: 1006-8961(2009)10-1925-09

## Research Progress from Single View DVC to Multi-view DVC

HU Xiao-fei, ZHU Xiu-chang

(Image Processing and Image Communication Key Laboratory, Nanjing University of Posts and Telecommunications, Nanjing 210003)

**Abstract** Multi-view video coding has higher demands on power consumption, complexity, storage capacity, and transmission capacity of the encoder, whereas the distributed source coding theory adapts to the needs. The basic theory and the key technology for distributed video coding are summarized firstly as the background information. Second, the expansion from single view to multi-view for distributed source coding theory is introduced. Then, the research status of distributed multi-view video coding is exposted. Finally, the development prospect for distributed source coding in the field of multi-view video coding is analyzed.

**Keywords** distributed video coding, multi-view video coding, Wyner-Ziv, power-efficient, robust, high compression syndrome based multimedia coding(PRISM)

## 1 引言

传统的视频编码标准,如 MPEG 或 H. 26X,都是采用基于离散余弦变换(DCT)和运动估计帧间预测技术的混合编码结构,即在编码端通过充分挖掘视频信号的空间和时间相关性来降低信号的冗余度,以达到压缩视频信号数据量的目的。在这种混合编码框架中,编码器承担了运动估计这一高运算量的计算任务,且率失真性能完全由编码端决定,这使得编码器的设计复杂度远高于解码器。尽管它们

在当今的视频编码领域得到了普遍应用,然而在一些特定的场合,如无线视频监控、无线 PC 相机和传感网络等,这种编码框架将面临严峻的挑战,因为这些应用与传统的视频编、解码系统的要求不同,即它们对编码器的功耗、复杂度、存储能力和传输能力有较高的要求,于是一种新的视频编码框架——分布式视频编码(DVC)应运而生。

在 DVC 中,具有相关性的视频序列是使用相互独立的编码器进行编码,而解码时则进行联合解码。由于其只在解码端使用了所有信源的信息,因此编码端的复杂度可以很低,这正适应了上述低处理能

基金项目:江苏省高等学校研究生创新计划项目(CX07B\_107z)

收稿日期:2008-06-02;改回日期:2008-09-01

第一作者简介:胡晓飞(1975 ~ ),女。南京邮电大学讲师,信号与信息处理专业在读博士研究生。研究方向为视频信号处理、分布式信源编码。E-mail:fhuxf@126.com, huxf@njupt.edu.cn

力的发送终端和高处理能力的接收服务器的需求。分布式信源编码具有编码复杂度低、抗误码性能强、适用于多终端多视点视频编码等特点,已引起了国内外众多学者的关注,美国斯坦福大学、伯克利大学都在此领域的研究中取得了显著的成就。国内清华大学、微软亚洲研究院、中国科学技术大学和南京邮电大学等高校及科研机构也相继展开了 DVC 视频编码的研究工作。

随着人们对视频逼真度和清晰度要求的提高和视频编解码技术的进步,普通的 2 维视频不断显现了其在表征场景深度信息方面的局限性,而能够提供 3 维视觉效果的立体或多视点视频则越来越受到学术界和工业界的重视。由于分布式信源编码的特点更适合于相机阵列、无线传感网络和自由视点电视等多视点视频编码的应用场合,近几年国内外的研究者们已逐渐把目光转向多视点分布式视频编码。

## 2 单视点 DVC

分布式信源编码的理论基础是 Slepian-Wolf 理

论<sup>[1]</sup>和 Wyner-Ziv 理论<sup>[2]</sup>。2002 年,斯坦福大学的 Aaron 等人,基于 Slepian-Wolf 理论和 Wyner-Ziv 理论,首先构建了 Wyner-Ziv DVC 的框架,并给出了这一方案实现的视频编解码的结果<sup>[3]</sup>。

Aaron 等人提出的分布式编解码器的性能明显优于传统的帧内编码,但与传统的帧间编码相比,还有差距。这就需要利用视频帧的空间与时间的统计相关性来进一步压缩数据量,以提高编码器的性能。于是产生了基于变换域的 Wyner-Ziv(WZ)视频编码器<sup>[4]</sup>,在文献[3]的基础上采用 DCT 变换来作用于视频序列帧,从而降低了空间冗余度,在相同条件下,编解码结果的峰值信噪比(PSNR)比文献[3]算法编解码结果的 PSNR 高出 0.5 ~ 2.5 dB。图 1 所示即为 WZ 视频编码器的基本框架,由于其编码端的“优化设计”模块可以完成编码模式判决、GOP 长度自适应选择、分析帧间相关性等功能,解码端可以设计算法优化“内插或外推”模块和“Turbo 解码”模块,从而改进了 WZ DVC 编码器的性能(详见第 4 节和第 5 节)。

在斯坦福大学着手研究 Wyner-Ziv 视频编码方

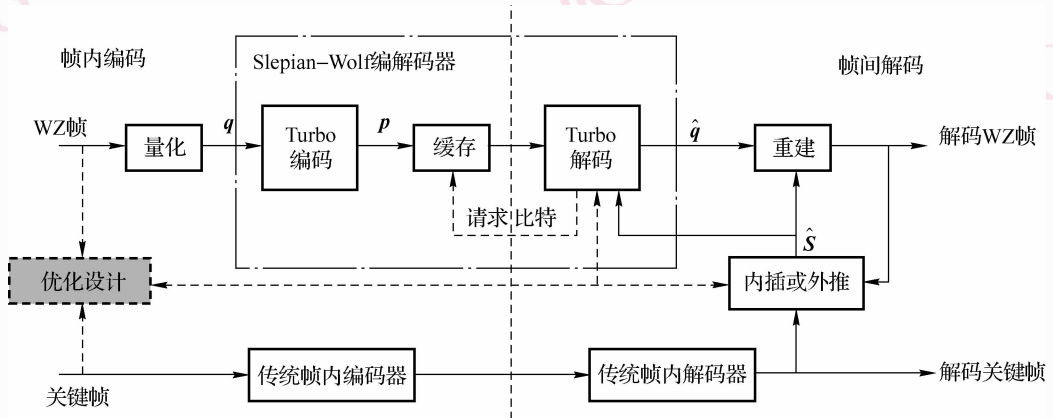


图 1 帧内编码/帧间解码的 Wyner-Ziv 视频编解码器

Fig. 1 Intra coding/inter decoding Wyner-Ziv video codec

案的同时,伯克利大学的 Ramchandran 等人也提出了一种变换域的 DVC 方法<sup>[5]</sup>,即 PRISM(power-efficient, robust, high compression syndrome based multimedia coding)。由于两者以相同的信息理论为基础,因此文献[4]中的关键功能模块,如变换、量化、纠错编码等,在 PRISM 编码的过程中也是不可或缺的,即 PRISM 编码与文献[4]算法的基本原理是一致的,同时也有其独特之处:PRISM 以子块

为单位编码,边信息也是以子块为单位的,既没有利用关键帧,也没有反馈信道,而且利用信道编码的特性来使抗误码性能可以随着量化精度的改变而提高。

实验结果表明,PRISM 编码器的率失真性能介于 H. 263 + 帧内编码与帧间编码之间,同时展示了其卓越的抗误码性能——随着误码率的提高,H. 263 + 的率失真性能急剧下降,而 PRISM 的率失

真性能则呈缓慢的线性下降的变化趋势。

基于 Wyner-Ziv 的视频编解码器与 PRISM 视频编解码器是 DVC 研究领域的典型代表,它们为 DVC 的研究奠定了基础,随后大部分的 DVC 研究基本上都是在此基础上展开的。

### 3 多视点 DVC

近几年,由于多视点视频的应用前景逐渐明朗,因此针对多视点视频的编码研究已成为该领域学者们的研究重点,国际视频标准化组织——MPEG 成立了 3 维音视频特别研究小组,着手讨论多视点视频编码技术的实现方案。由于当今大多数多视点视频编码方案都是从单视点视频编码标准扩展而来的,如编码端执行视点间和时域预测,致使编码器承担了沉重的运算负担。然而,在很多多视点视频编码应用领域,由于视频编码的特点是编码节点处理能力较低、能源供应受限、通常节点间不进行通信,而是共同和一个信息处理能力较强的中心处理器通信,因此普通的多视点视频编解码方案就不符合此

类应用需求,而分布式信源编码方案恰恰符合此要求,因此已成为实现多视点视频编码的解决方案之一而受到国内外学者的关注,多视点 DVC 的特点将在此领域得到广泛的应用。

文献[6]和文献[7]从信息论的基本理论出发分析了分布式信源编码的本质,认为其实际上是信源-信道编码问题,并提出了分布式信源编码,用于具有多个传感节点的传感网络的实施方案,同时指出了传感网络中的分布式信源编码所面临的问题。

多视点 DVC 可以由单视点 DVC 推广而来,同时其网络结构特点决定了其不同于单视点 DVC 的编解码特点:它既可以挖掘同一个相机捕获的视频序列间的相关性,又可以在相邻相机捕获的图像之间挖掘相关性。作为多视点 DVC 视频编码的特例,文献[8]给出了立体视频编码的框架(如图 2 所示),图中“第 2 视点”的视频信号  $S_1$  的编码即为单视点 DVC 方案的推广,其中的功能模块——“立体边信息的建立”既利用了第 2 视点捕获的图像的时域相关性,又利用了和第 1 视点捕获的图像间的空域相关性,这充分体现了多视点 DVC 的特点。

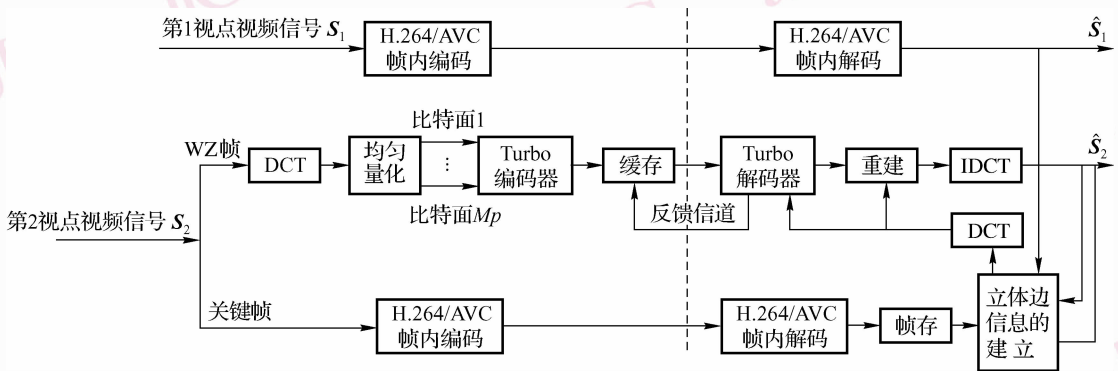


图 2 立体视频编解码结构

Fig. 2 Stereo video codec architecture

既然多视点 DVC 可以由单视点 DVC 推广而来,那么单视点 DVC 的关键技术在多视点 DVC 中同样适用,同时还需兼顾视点间的视频压缩处理环节。

## 4 DVC 的关键技术

### 4.1 挖掘视频信号相关性

传统的视频编码都是在编码端通过充分挖掘视频信号的空间和时间相关性,以压缩数据,由于一些特殊的要求,致使 DVC 的这个任务不可能完全由编

码器承担,编码器只能起到辅助的作用。文献[9]提出的基于 Wyner-Ziv 空间域的视频编码框架是以块为基本单位,并以每个块的 SAD (sum of absolute differences) 和块的空间平滑性为判决依据来决定块的编码模式,即同一帧的块可以是帧内编码,也可以是 WZ 编码。这样,通过 SAD 的度量去除了时域冗余,通过空间平滑性的度量去除了空域冗余,使率失真性能比文献[3]的算法有了明显的提高。小波变换的独特优势,使其在 DVC 领域中占有一席之地,文献[10]就是采用小波变换来去除视频信号的空间相关性,在不过多地增加编码器复杂度的前提下,

提高了编码器的性能,同时利用小波变换的特点取得了视频流的可分级特性,这种特性在信道带宽有限的多视点 DVC 中尤为重要,许多多视点 DVC 方案都是基于小波变换的<sup>[11-13]</sup>。对于多视点 DVC,文献[11]在编码节点间传送 Hash 信息,而通过适当挖掘视点间的相关性,也能取得意想不到的效果,文献[11]利用的功耗、码率和失真之间的关系模型为

$$D_E(R_E, P_E, P_H) = \sigma^2 e^{-\gamma[R_E \cdot g(P_E) + P_H(R_H)]} \quad (1)$$

式中,  $D_E$  (下角 E 代表 encoding) 是非关键帧有损编码的失真度,  $R_E$  是编码的码率,  $P_E$  是编码器的功率,  $P_H$  是接收到的 Hash 数据的功率,  $R_H$  是 Hash 信息的码率,  $P_H(R_H) = \alpha R_H$  ( $\alpha$  表示收到的 Hash 码率和功率之间的关系参数),  $\sigma$  和  $\gamma$  是模型参数,  $g(P_E)$  是反映编码效率的函数。实验证明,如果编码器的功率  $P_E$  和在编码节点间传输的 Hash 信息的功率  $P_H$  太低,就会影响编码器的性能。但这并不是说,  $P_E$  和  $P_H$  越高越好,而是存在一个最优的功耗,使失真  $D_E$  最小。设总的功耗  $P_T$  (下角 T 代表 total) 由编码器的功率  $P_E$ 、接收到的 Hash 信息的功率  $P_H$  和传输功率  $P_{tr}$  (下角 tr 代表 transmit) 构成,即

$$P_T = P_E + P_H + P_{tr} \quad (2)$$

其中,

$$P_{tr} = \eta(d) \cdot R_E \quad (3)$$

式中,  $d$  是传输距离,  $\eta(d)$  是传输功耗和距离  $d$  的关系函数,那么就有

$$R_E = \frac{P_{tr}}{\eta(d)} = \frac{P_T - P_E - \alpha R_H}{\eta(d)} \quad (4)$$

于是,式(1)就变成

$$D = D_E\left(\frac{P_T - P_E - \alpha R_H}{\eta(d)}, P_E, \alpha R_H\right) \quad (5)$$

式中,如果把  $P_T$ ,  $\alpha$  和  $\eta(d)$  固定,那么式(5)就反映了失真度、Hash 码率和编码器功耗之间的关系。因此编码节点间传送的 Hash 信息从一定程度上影响了编码器的率失真性能。

综上所述,在挖掘视频信号相关性方面的研究上,它们挖掘了视频序列的空域或时域相关性,或者是二者兼顾,从而在一定限度内,以增加编码器的复杂度为代价,换来了编码性能的提高。

#### 4.2 GOP 长度和关键帧编码

与传统的视频编码类似,DVC 中的 GOP (group

of pictures) 的长度定义为视频序列中两个相邻 I 帧之间的帧数。对于两类 DVC 方案,只有 WZ 编码的过程中存在确定关键帧的问题。在文献[3]提出的 DVC 方案中,视频帧编码结构的 GOP 长度是确定的,文献[14]研究了 GOP 长度对 Wyner-Ziv 视频编码器性能的影响(包括编码复杂度、解码复杂度和率失真性能)后指出, GOP 长度越小,编码用时越长,解码用时越短,越有利于解码端的运动估计和产生高质量的边信息,从而可得到高质量的 WZ 帧的重建图像。然而 GOP 长度越小,序列中就会有越多的 I 帧,这将影响编码效率。文献[15]提出了 GOP 大小可自适应改变的 Wyner-Ziv 视频编码方案,即编码端利用时域相关性,通过分析运动特征来自适应选择 GOP 大小,解码端则根据 GOP 大小灵活进行运动估计,实验结果表明,空域和频域的自适应 GOP 长度和固定 GOP 长度相比,率失真性能分别提高了 0.8 dB 和 0.4 dB。

在关键帧的编码方法方面,由于解码端边信息帧由重建的关键帧产生,因此关键帧的解码图像质量对编码器的性能起着决定性的作用。文献[16]把解码端运动估计产生的运动矢量信息通过反向信道传送到编码端,关键帧用反向预测的方法进行编码,从而提高了编码效率。

对于多视点 DVC,为了解决由关键帧帧内编码造成的高码率问题,可利用同一时刻不同视点的的关键帧之间的几何关系进行仿射变换来进一步压缩码率<sup>[11]</sup>,在 GOP 长度较小的情况下,这个仿射变换参数还可以用于该 GOP 后续的非关键帧的编码。

#### 4.3 边信息帧的利用

DVC 中边信息帧对编码性能起着决定性的作用,边信息帧与 WZ 帧越接近,需要传输的“纠错”比特数就越少,DVC 编码器的性能就越高,它决定了解码重建图像的质量和编码器的压缩效率。近年来,为提高边信息的图像质量,关于 DVC 编码中边信息帧的产生方法的研究层出不穷。文献[17]用迭代改进算法对初始边信息进行改进,同时结合增强解码算法、运动估计算法和多个边信息帧的选择方法进行编码。文献[18]用期望最大化(expectation maximization, EM)算法来更新运动矢量场分布和信源模型参数,结果显示,其不但具有良好的率失真性能,且随着 GOP 长度的增加,编码器的率失真性能不降反升。

诸多的边信息产生方法无非都是为了提高边信

息帧的质量,以便提高重建 WZ 帧的质量和降低传输码率。建立高质量的边信息帧不外乎以下两种途径:一是直接在解码端充分挖掘信号的相关性;二是从编码端传输当前 WZ 帧的辅助信息,以帮助解码端建立准确的边信息帧。

目前,单视点 DVC 的率失真性能还不能与传统的混合编码策略相媲美,接收端获得的边信息帧的不准确性是造成这一结果的主要因素之一。多视点 DVC 能从一定程度上解决这个问题,因为对于多视点 DVC 来说,除了时域相关性以外,还有视点间的相关性可以利用,这不仅有利于建立更准确的边信息帧,尤其对于大运动量的序列,当时域预测失效时,就可以由视点间的预测来弥补这个不足。因此,多视点 DVC 与传统的混合编码器的编码效率的差距要小于单视点 DVC 与传统的混合编码器编码效率的差距,即 DVC 理论用于多视点视频编码比用于单视点视频编码有优势。

鉴于多视点视频编码的结构特点,文献[19]通过研究不同视点之间的几何关系和进行视差补偿,以及挖掘视点间的相关性,提高了不同视点重叠部分的压缩率。文献[20]在编码端通过挖掘每个视点捕获的视频序列的时域相关性和空域相关性来进行 WZ 编码,在解码端则利用中心处理器进行视差补偿来获得边信息,并研究了视差补偿对编码性能的影响,结果表明,利用经过视差补偿的边信息能使码率降低 10%。

文献[11]把多视点 DVC 应用于无线视频传感网络,并将各个视点捕获的图像序列用单视点 DVC 进行编码,同时,编码端通过在不同视点间传输有限的 Hash 比特来挖掘视点间的相关性,解码端则利用多个边信息帧作参考,并通过仿射变换进行视差补偿,同时分析了运算复杂度,结果表明,与先前的单视点 DVC 的编码结果相比,率失真性能明显提高,与 H.264 帧间编码的 PSNR 性能的差距缩小到约 1 dB,而计算复杂度则大大低于 H.264 帧间编码。

由于用左右视点产生边信息不适用于场景被遮挡的情况,而用时域上相邻的帧产生边信息又不适用于大运动量的场合,而使用两种边信息融合的方法<sup>[12-13,21-22]</sup>则能取得更好的压缩效果。文献[12]构建了基于小波变换和 Turbo 码的多视点 DVC 编码的框架,多视点 DVC 编码结构的特殊性造就了灵活多样的边信息产生方法,文献[12]中,WZ 帧的边信

息就是通过构造二进制掩模(0 表示由时域相关性估计的像素,1 表示由空域相关性估计的像素)由左右视点和前后帧联合产生的,结果证明,这种方法编码视频的 PSNR 值比仅用时间上的前后帧预测高 1.5 dB。文献[13]对此加以改进,用分级树集合划分(SPIHT)的方法组织小波系数,并通过灵活选择时间和视点相关性预测技术,证明了视点间的相关性预测在多视点 DVC 编码中的作用比在传统的多视点视频编码中更明显。文献[21]在构造多视点 DVC 结构时,还定义了每一个视点上的相机都是 WZ 相机,并通过实验证明了时域运动内插映射法得到的掩模能产生最高质量的边信息。文献[22]通过运动补偿时域内插(MCTI)和单应性补偿视点间内插(HCII)获得了 MCTI,HCII-left,HCII-right 和 HCII-avg 等 4 种边信息,对多视点视频序列“Breakdancing”的测试结果(表 1)表明,HCII 与 MCTI 的融合能产生较高质量的边信息帧。文献[23]和文献[24]则基于 PRISM 框架,其中心处理器首先使用时域上取得的边信息,而解码失败后则执行视差搜索或视点分析获取边信息,结果证明,视差搜索和视点分析二者的性能是一致的,同时显示了视点间相关性的利用,能够使 PRISM 编码器应用于多视点场合具有更强的抗误码性能。

表 1 “Breakdancing”用不同方法产生边信息平均 PSNR 比较

Tab. 1 Average PSNR of side information for Breakdancing: comparison of different methods

	HCII-left	HCII-right	HCII-avg	HCII	MCTI	HCII-MCTI
PSNR 值 [dB]	24.37	24.38	26.25	29.95	26.11	32.54

## 5 差错控制和码率控制

### 5.1 信道码

信道编码与分布式信源编码有着密切的关系,而当前出现的 Slepian-Wolf 编码方法则大都是基于信道编码的。Pradhan 和 Ramchandran 提出的 DISCUS(distributed source coding using syndromes)方案<sup>[6]</sup>具有重要的意义,很多其他的分布式信源编码的方案都是建立在 DISCUS 方案的基础上的。DISCUS 方案的出现引发了人们利用信道码进行分布式信源编码的研究。

由 Aaron 和 Girod 等人提出的 Wyner-Ziv 视频

编码器<sup>[3]</sup>是以 WZ 帧与边信息帧之间的关系模型为 Laplacian 模型,并用 Turbo 码纠正这个“相关信道”产生的误码,由于这个模型能否精确地模拟“信道噪声”关系到编码效率的高低,因此,DVC 研究领域展开了对虚拟信道噪声模型的研究。由于“噪声”不是一成不变的,故文献[25]研究了编码端噪声模型参数在序列级、帧级和块级分别离线设置,以及解码端噪声模型参数分帧级、块级和像素级在线估计对编码性能的影响。

在纠错编码方法的选择问题上,DISCUS 方案、Turbo 码和低密度奇偶校验码(LDPC)是 DVC 视频编码中普遍使用的纠错编码方法,其中 DISCUS 方案是其他纠错方法的基础,Turbo 码和 LDPC 码都是目前最先进的信道码,许多对 DVC 编码中纠错编码方法的改进方案都是建立在这些基本方案的基础上的。文献[26]分析了基于 Turbo 码的 DVC 视频编码器编解码的复杂度和率失真性能,指出其编码性能随着 Turbo 码记忆长度的加长而提高,但是解码复杂度却呈指数增长。文献[27]针对删余 LDPC 码的性能,设计了自适应码率的 LDPC 码,实验证明,在中高码率下,其显示了优异的性能。文献[28]分析了 LDPC 码与 Turbo 码相比的优势,并用 LDPC 码代替了先前研究中所用的 Turbo 码,随后的 DVC 研究大都是以 LDPC 码来进行纠错的。

## 5.2 抗误码编码

虽然 DVC 视频编码系统中应用“纠错”编码模块的目的主要是为了去除视频信号的相关性,但是“纠错”码固有的抗误码特性,使得 DVC 视频编码比传统的基于预测的混合编码系统具有更强的抗误码能力。然而这里的“纠错”码还不足以抵御信道噪声的干扰和满足应用需求,抗误码编码仍然是 DVC 视频编码中重要的研究课题之一。

文献[29]展示了基于 DVC 的视频编码器超强的纠错能力,编码端由一个 MPEG 编码器和一个 WZ 编码器组成,视频信号在传输过程中受到噪声干扰后,用 MPEG 中的纠错方法进行纠错后的信息难免还包含误码。为了纠正这些误码,可对同样的视频信号粗量化后再进行 Turbo 编码,即 WZ 编码,解码端则将 MPEG 纠错输出的信息作为 Wyner-Ziv 解码的边信息,于是边信息中的误码就可以由 Wyner-Ziv 解码的过程纠正。文献[30]把 Wyner-Ziv 视频编码的思想引入到 H. 264“冗余条带”的编码中,即编码端利用 Reed-Solomon 编码器产生“校验

条带”并传输,而解码端则将从解码得到的“主条带”中提取信息作为边信息解出“冗余条带”,这样不仅降低了传输“冗余条带”的比特率,而且能提供解码视频较好的视觉效果。文献[31]针对视频信号在有噪无线信道中传输的情况,在基本的 DVC 编码框架中加入基于多人多出(MIMO)的发射分集技术,以减轻无线信道衰落的副作用,结果证明,在信道信噪比(SNR)较低的情况下,率失真性能优于 H. 264。

Wyner-Ziv 理论还可以用于解决预测编码丢包时,解码端与编码端预测不匹配的问题<sup>[28]</sup>,结果显示了与传统的纠错编码方法(如 FEC(forward error correction))相比的明显优势。

此外,通过可分级编码也可以提高 DVC 编码传输的抗误码能力,对于资源有限并易受噪声干扰的无线视频传感网络来说,视频信号的分级编码特性显得尤为重要<sup>[11]</sup>。文献[32]给出了无噪信道和有噪信道上,DVC 视频编码的信噪比可分级的实现方案,即把 Wyner-Ziv 信源编码的思想引入到视频信号的可分级编码,基本层用标准的视频编码方法进行编码,而增强层则用 WZ 编码,Wyner-Ziv 编码器由 DCT 变换(或条件 KLT(cKLT))、嵌套标量量化(NSQ)以及基于 LDPC 的 Slepian-Wolf 编码器(SWC)组成。NSQ 先将输入的 DCT 系数分成不同的陪集,然后只输出陪集的索引,而且通过 NSQ 只保留中间的比特面,最后用由多级 LDPC 编码器构成的 SWC 编码器来对这些比特面进行编码,以形成增强层码流。解码时,先用 H. 26L 解码器对基本层码流进行解码来得到基本层,然后以该基本层作为边信息,再利用 Wyner-Ziv 解码器对增强层码流进行解码来得到效果更好的视频流,这样就实现了无噪信道上视频信号的分级编码。通常信号在有噪信道上传输时,需设计以下两种信道码:Wyner-Ziv 编码中的信道码和传输时的纠错码。为了简化设计,文献[32]增加了不规则重复累积码(irregular repeat accumulate, IRA)的设计,其不仅挖掘了信源与边信息之间的相关性,同时提供了信号在有噪信道上传输的误码保护功能。如图 3 所示,假设信号经相关信道和噪声信道两个信道传输,并将信源经相关信道传输形成的边信息  $Y$  作为 IRA 码的系统部分;将 Wyner-Ziv 编码比特流作为 IRA 码的校验部分,经有噪信道传输。在接收端,边信息  $Y$  和受噪声干扰的 Wyner-Ziv 比特流  $U$  执行联合信源-信道解码,即

IRA 码的系统部分和校验部分都用于信道解码。实验证明,基于 IRA 码的 Wyner-Ziv 分级编码器不仅

保持了分级编码的特性,同时提供了对增强层码流的有效保护。

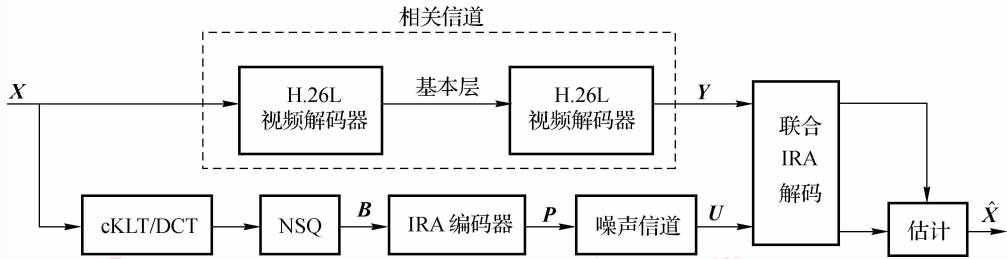


图 3 分层 Wyner-Ziv 视频编解码器结构

Fig. 3 Layered Wyner-Ziv video codec architecture

除了信噪比可分级以外,还可以在空间域和时间域实现视频信号的分级编码,文献[33]还把 Wyner-Ziv 方法用于增强层的编码,即在视频帧编解码过程中,通过改变空间分辨率实现了空间域可分级编码。

### 5.3 码率控制

DVC 的编码码率依赖于边信息和待编码数据间的“误差”,码率控制可以在编码端进行,也可以在解码端进行。解码端进行的码率控制是由当前帧和边信息的统计相关特性来决定的,而反馈信道中的码率则随当前帧和边信息的统计特性的变化而变化,因此,反馈信道在解码端的码率控制中起着至关重要的作用,文献[34]和文献[35]为了设计有效的码率控制方案,分别研究了像素域和变换域反馈信道的相关统计特性。虽然在解码端进行码率控制,明显降低了编码端的负担,但是这种方法有以下两个比较明显的缺点:首先需要有一个反馈信道,这会造成较长的延迟和复杂的编解码系统,且每收到一次请求比特数,都要重新进行解码;其次统计特性估计和解码过程都是在线执行,这就要求解码端有很强的计算能力。在这些条件的限制下,人们把码率控制的实现转向在编码端进行。

文献[36]先在编码端估计当前帧与相邻帧间的均方差、错误概率、要传输的比特数等参数,然后向解码端传输校验比特,最后解码端通过估计重建帧与信源的残差错误概率即可把失真限制在一定的范围内。该算法简化了 DVC 结构,但这是以率失真性能的下降为代价的。文献[37]通过在编码端建立率失真分析模型,在失真最小的前提下确定量化参数,结果显示,尽管模型参数估计不够准确,但所建立的率失真模型比传统的 DVC 编码方法的失真

更接近目标的失真度。此外,Ramchandran 等人提出的 PRISM 结构也是一种无反馈信道的 DVC 方案<sup>[5]</sup>。对于多视点 DVC,除了存在同一视点内的关键帧和 WZ 帧的码率分配问题外,不同视点间的码率分配也是多视点视频编码面临的问题。

## 6 结 语

本文首先介绍了分布式视频编码的理论和应用,然后回顾了从单视点 DVC 到多视点 DVC 的研究概况。尽管 DVC 编码端的复杂度和传统的帧内编码相当,而解码复杂度却大大增加,但是这并不妨碍 DVC 在多视点视频编码中的应用。

尽管经过了几年的发展,DVC 的率失真性能与传统的混合编码器相比,仍然存在差距,其原因主要在于:

- (1)“噪声”模型的选择上,Slepian-Wolf 理论和 Wyner-Ziv 理论都是以高斯模型为前提的,而在 DVC 视频编码中,这个“噪声”更趋近于拉普拉斯分布;
- (2)由于“纠错”所产生的校验信息几乎是一个等概率的无记忆信源,无法用熵编码方法对其进行编码;
- (3)边信息帧的准确性有待提高。

目前,单视点 DVC 编码正向传统视频编码器的性能迈进,现已逐步实现并完善了传统视频编码器的各项功能,而多视点 DVC 则由于受单视点 DVC 编码性能不高的影响,还未能投入实际应用。两大基于分布式信源编码理论的视频编码系统 WZ 和 PRISM,正演绎着各自的精彩,尽管 PRISM 以其卓越的抗误码性能而著称,但 80% 的系数帧内编码也

许使得其编码效率受到影响,在文献[23]和[24]中都没有提及这个问题。由于 DVC 是一种符合多视点视频编码特点的编码方式,它的巨大优势推动了多视点视频编码的研究进程,因此在视频编码领域具有非常广阔的应用前景。

### 参考文献 (References)

- Slepian J, Wolf J. Noiseless coding of correlated information sources [J]. IEEE Transactions on Information Theory, 1973, **19** (4): 471-480.
- Wyner A, Ziv J. The rate-distortion function for source coding with side information at the decoder [J]. IEEE Transactions on Information Theory, 1976, **22**(1):1-10.
- Aaron A, Zhang R, Girod B. Wyner-Ziv coding of motion video [A]. In: Proceedings of Asilomar Conference on Signals and Systems [C], Pacific Grove, CA, USA, 2002: 240-244.
- Aaron A, Rane S, Setton E, *et al.* Transform-domain Wyner-Ziv codec for video [A]. In: Proceedings of SPIE Visual Communications and Image Processing [C], San Jose, CA, USA, 2004: 520-528.
- Puri R, Ramchandran K. PRISM: A new robust video coding architecture based on distributed compression principles [A]. In: Proceedings of Allerton Conference on Communication, Control and Computing [C], Allerton, IL, USA, 2002:211-219.
- Pradhan S S, Kusuma J, Ramchandran K. Distributed compression in a dense micro-sensor network [J]. IEEE Signal Processing Magazine, 2002, **19**(3): 51-60.
- Xiong Z, Liveris A D, Cheng S. Distributed source coding for sensor networks [J]. IEEE Signal Processing Magazine, 2004, **21** (9): 80-94.
- Areia J D, Brites C, Pereira F, *et al.* Wyner-Ziv stereo video coding using a side information fusion approach [A]. In: Proceedings of the IEEE 9th Workshop on Multimedia Signal Processing [C], Chania Crete, Greece, 2007: 453-456.
- Tagliasacchi M, Trapanese A, Tubaro S, *et al.* Intra mode decision based on spatio-temporal cues in pixel domain Wyner-Ziv video coding [A]. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing [C], Toulouse, France, 2006, **2**: 57-60.
- Bernardini R, Rinaldo R, Zontone P, *et al.* Wavelet domain distributed coding for video [A]. In: IEEE International Conference on Image Processing [C], Atlanta, GA, USA, 2006:245-248.
- Kang L W, Lu C S. Low-complexity Power-scalable Multi-view Distributed Video Encoder [EB/OL]. <http://www.eurasip.org/Proceedings/Ext/PCS2007/defevent/papers/cr1176.pdf>.
- Guo X, Lu Y, Wu F, *et al.* Distributed multi-view video coding [A]. In: Proceedings of SPIE Visual Communications and Image Processing [C], San Jose, California, USA, 2006, **6077**: 290-297.
- Guo X, Lu Y, Wu F, *et al.* Wyner-Ziv based multiview video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, **18**(6): 713-724.
- Pereira F, Ascenso J, Brites C. Studying the GOP size impact on the performance of a feedback channel-based Wyner-Ziv video codec [A]. In: Proceedings of IEEE Pacific-Rim Symposium on Image and Video Technology-PSIVT [C], Santiago, Chile, 2007: 801-815.
- Ascenso J, Brites C, Pereira F. Content adaptive Wyner-Ziv video coding driven by motion activity [A]. In: Proceedings of IEEE International Conference on Image Processing [C], Atlanta, GA, USA, 2006: 605-608.
- Liu L, Li Z, Delp E J. Backward channel aware Wyner-Ziv video coding [A]. In: Proceedings of the IEEE International Conference on Image Processing [C], Atlanta, GA, USA, 2006: 1677-1680.
- Weerakkody W A R J, Fernando W A C, Martinez J, *et al.* An iterative refinement technique for side information generation in DVC [A]. In: Proceedings of IEEE International Conference on Multimedia and Expo [C], Beijing, China, 2007: 164-167.
- Varodayan D, Chen D, Flierl M, *et al.* Wyner-Ziv Coding of Video with Unsupervised Motion Vector Learning [EB/OL]. <http://www.stanford.edu/~bgirod/pdfs/VarodayanImageCommunication2008.pdf>.
- Song B, Bursalioglu O, Roy-Chowdhury A K, *et al.* Towards a multi-terminal video compression algorithm using epipolar geometry [A]. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing [C], Toulouse, France, 2006, **2**: 49-52.
- Flierl M, Girod B. Coding of multi-view image sequences with video sensors [A]. In: Proceedings of IEEE International Conference on Image Processing [C], Atlanta, GA, USA, 2006: 609-612.
- Pierre F, Dimitris A, David B. Fusion methods for side information generation in multi-view distributed video coding systems [A]. In: Proceedings of IEEE International Conference on Image Processing [C], San Antonio, Texas, USA, 2007, **6**: 409-412.
- Dufaux F, Ouaret M, Ebrahimi T. Recent advances in multi-view distributed video coding [A]. In: Proceedings of SPIE Mobile Multimedia/Image Processing for Military and Security Applications [C], Orlando, Florida, USA, 2007, **6579**: 1-11.
- Yeo C, Ramchandran K. Distributed Video Compression for Wireless Camera Networks [EB/OL]. <http://www.eecs.berkeley.edu/~zuohao/documents/vcip2007-summary.pdf>.
- Yeo C, Wang J, Ramchandran K. View synthesis for robust distributed video compression in wireless camera networks [A]. In: Proceedings of the IEEE International Conference on Image Processing [C], San Antonio, Texas, USA, 2007, **3**: 21-24.
- Brites C, Ascenso J, Pereira F. Studying temporal correlation noise modeling for pixel based Wyner-Ziv video coding [A]. In: Proceedings of the IEEE International Conference on Image Processing [C], Atlanta, GA, USA, 2006: 273-276.
- Belkoura Z M. Analysis and Application of Turbo Coder Based Distributed Video Coding [D]. Technische University Berlin, Berlin, German, 2007.
- Varodayan D, Aaron A, Girod B. Rate-adaptive distributed source coding using low-density parity-check codes [A]. In: Proceedings of Asilomar Conference on Signals, Systems, and Computers [C],

- Pacific Grove, CA, USA, 2005: 1203-1207.
- 28 Sehgal A, Jagmohan A, Ahuja N. Wyner-Ziv coding of video: an error-resilient compression framework [J]. *IEEE Transactions on Multimedia*, 2004, **6**(4): 249-258.
- 29 Aaron A, Rane S, Zhang R, Girod B. Wyner-Ziv coding for video: Applications to compression and error resilience[A]. In: *Proceedings of the IEEE Data Compression Conference [C]*, Snowbird, UT, USA, 2003: 93-102.
- 30 Rane S, Girod B. Systematic lossy error protection of video based on H.264/AVC redundant slices [A]. In: *Proceedings of SPIE Conference on Visual Communications and Image Processing (VCIP'06) [C]*, San Jose, CA, USA, 2006, **6077**: 1-9.
- 31 Muraleetharan K, Fernando W A C, Weerakkody W A R J, *et al.* MIMO wireless channel for video communications using distributed video coding [A]. In: *Proceedings of Canadian Conference on Electrical and Computer Engineering [C]*, Vancouver, BC, Canada, 2007: 1610-1613.
- 32 Xu Q, Stankovic V, Xiong Z. Layered Wyner-Ziv video coding for transmission over unreliable channels[J]. *Signal Processing: Special Section on Distributed Source Coding*, 2006, **86**(11): 3212-3225.
- 33 Mukherjee D, Macchiavello B, Queiroz R L de. A simple reversed-complexity Wyner-Ziv video coding mode based on a spatial reduction framework [A]. In: *Proceedings of SPIE conference on Visual Communications and Image Processing [C]*, San Jose, CA, USA, 2007, **6508**: 1-12.
- 34 Brites C, Ascenso J, Pereira F. Feedback Channel in Pixel Domain Wyner-Ziv Video Coding: Myths and Realities [EB/OL]. [http://amalia.img.lx.it.pt/~cicb/papers/EUSIPCO'06\\_FeedbackChannel.pdf](http://amalia.img.lx.it.pt/~cicb/papers/EUSIPCO'06_FeedbackChannel.pdf).
- 35 Pedro J, Brites C, Ascenso J, *et al.* Studying the Feedback Channel in Transform Domain Wyner-Ziv Video Coding [EB/OL]. <http://amalia.img.lx.it.pt/~cicb/papers/Conftele07.pdf>.
- 36 Morb'ee M, Prades-Nebot J, Pi'zurica A, *et al.* Rate allocation algorithm for pixel-domain distributed video coding without feedback channel [A]. In: *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing [C]*, Hawaii, USA, 2007: 521-524.
- 37 Roca T, Morb'ee M, Prades J, *et al.* A distortion Control Algorithm for Pixel-domain Wyner-Ziv Video Coding [EB/OL]. <http://ltswww.epfl.ch/ltsftp/PCS2007/defevent/papers/cr1182.pdf>.