

## 摘要

随着无线传感器网络快速发展和广泛应用,人们对感知信息精细化、丰富化的要求越来越迫切,具有感知多媒体信息的无线多媒体传感器网络(WMSNs)在此背景下孕育而生。由于引入了多媒体信息,使得传统的视频编解码因编码复杂、解码简单的特性,与节点能量受限的 WMSNs 不相适宜。具有编码简单、解码复杂特点的分布式视频编码(DVC)能够有效地解决 WMSNs 节点在资源受限条件下进行信息处理压缩的问题,但目前对分布式视频编码技术的研究还处于起步阶段。因此,为了解决 WMSNs 发展与应用中的“瓶颈”、实现感知信息多媒体化,对 DVC 的关键技术展开研究具有重要意义。

论文以 WMSNs 作为研究背景,首先概述分布式视频编码的理论基础及特点,分析分布式视频编码的典型系统方案和关键技术。对边信息在分布式视频系统中的重要性以及典型边信息生成算法进行研究,针对运动估计产生的预测重叠问题,提出一种块分类的双向估计加权边信息生成算法,采用块分类思想对不同运动变化的像素块进行分类处理,通过加权判决提高算法效率。考虑到运动剧烈、细节丰富的视频将影响解码质量的问题,提出一种相关性分类的解码检测优化算法。将原始视频帧之间的相关性信息作为检测解码帧质量的标准,动态调整检测阈值,对低质量的解码像素块,采用重建边信息算法进行优化处理,以提高系统解码帧质量。最后针对 WMSNs 结构特点和应用需求,结合边信息生成算法和解码检测优化算法,提出一种面向 WMSNs 的分布式视频编解码系统方案,并对各个功能模块进行分析。

论文分别对边信息生成算法、解码检测优化算法以及面向 WMSNs 的分布式视频编解码系统方案进行了性能仿真测试。测试结果表明所提算法能有效提高边信息质量,实现对解码帧进行检测及优化的作用。通过与传统的帧内编码和帧间编码方式进行系统编解码测试比较,验证了所提系统方案的有效性和可行性。

**关键词:** 分布式视频编码, 无线多媒体传感器网络, 边信息, 解码检测

## ABSTRACT

With the rapid development and wide application of wireless sensor networks, people demand perception information to be refined and rich more and more urgently and forcefully. In this social background, the wireless multimedia sensor networks (WMSNs) was born which could perceive multimedia information. Since the multimedia information was introduced into WMSNs, the traditional Video Codec has become unsuitable for the WMSNs with restricted node resources due to its complex encoding and simple decoding. The Distributed Video Coding (DVC) which has the feature of simple encoding and complex decoding can solve the problems that occur when WMSNs nodes process and compress information in the resource-constrained conditions. But the research on DVC techniques is still in the initial stage at present. Therefore, in order to solve the bottleneck problem in the development and application of WMSNs and realize the multimedia perception information, it has an important significance to carry out a research on the key technology of DVC.

Based on the research background of WMSNs, this paper reviews the theoretical basis and features of DVC and analyzes typical systems and key technologies of DVC at first. Then, researches about the importance of side information in DVC and typical generation algorithms of side information are made. Afterwards, a block-classified side information generation algorithm with bidirectional motion estimation and weight decision is presented to settle the problem of overlapped prediction resulting from motion estimation. By using the block-classified thought, the pixel blocks are classified into two parts according to the difference of motion variation, and the efficiency of the algorithm is improved by the weight judgment. Taking into account the issue that intense movement and details will affect the decoding quality in video system, this article proposes an optimizing algorithm for decoding detection classified by correlation. The correlation between the original video frames is used as the standard for

detecting the quality of decoding frames, and the detecting threshold is adjusted dynamically. For the low-quality decoding pixel blocks, it adopts the algorithm of rebuilding side information to optimize for high quality of decoding frames in system is adopted. Finally, combined with the generation algorithm of side information and optimizing algorithm for decoding, a DVC system for WMSNs is proposed to meet the structural feature and application requirements of WMSNs, and then each functional module of the system is analyzed.

This paper performs some simulations to test function of the proposed algorithms, which are the generation algorithm of side information, the optimizing algorithm for decoding detection and the DVC system for WMSNs. Simulation results show that the proposed algorithms can produce improvement in the quality of side information, and realized the function that the decoding frame can be automatically detected and optimized. The effectiveness and feasibility of the proposed DVC system is verified by comparing with the traditional encoding system.

**Key words:** Distributed Video Coding, Wireless Multimedia Sensor Networks, side information, decoding detection

# 第一章 绪论

## 1.1 研究背景及意义

被誉为全球未来三大高科技产业之一的无线传感器网络(Wireless Sensor Networks, WSNs)是 20 世纪末发展起来的一种无线自组织网络<sup>[1]</sup>。随着科学技术的发展和监测环境的复杂化, WSNs 获取的简单数据信息(如温度、湿度、光强度、振动等)已经不能满足人们对事物进行全面监测的应用需求。为实现信息细粒化、精准化、丰富化, 人们对监控活动中加入图像、音频和视频等多媒体信息的需求越来越迫切。

无线多媒体传感器网络(Wireless Multimedia Sensor Networks, WMSNs)在这种应用需求背景下孕育而生。作为一种新兴的具有感知多媒体信息的无线网络, WMSNs 综合了传感器技术、多媒体技术、嵌入式计算技术、无线通信技术等, 实现了丰富数据和准确信息的监测, 它将人类的感知视野从单一的数据信息扩展到丰富的多媒体信息, 为实现细粒度、精准信息的监测提供了技术保障<sup>[2-3]</sup>。WMSNs 在 WSNs 的基础上引入了多媒体信息, 它不仅继承了传统 WSNs 的节点数量大、成本低、无线多跳路由、自组织、节点资源受限、能耗敏感、无人值守等特点, 还具有以下显著特点<sup>[4]</sup>:

### (1) 能耗呈“均匀”分布

传统的 WSNs 的能耗绝大部分“聚集”在无线收发过程中, 但在 WMSNs 中, 由于采集和传输的对象是大数据量的多媒体信息, 所需能耗增加。实验表明, 数据采集和处理的能耗与无线收发的能耗呈“均匀”分布状态<sup>[5]</sup>。

### (2) 感知信息丰富

WMSNs 中各种传感器节点对监测区域进行不同种类信息的采集, 不仅包含 WSNs 所具有的简单数值信息, 还有蕴含丰富信息量的多媒体信息。这些丰富数据信息通过数据融合、网络传输, 最终送至控制中心, 并由控制中心根据用户需求进行定位、重构与处理, 实现更加全面而准确的监测。

### (3) 数据处理复杂

在 WMSNs 中, 大量的多媒体传感器节点、普通节点和汇聚节点被投放在监测区域中, 采用无线通信方式进行各节点间的数据通信以提高网络灵活性。为了降低网络传输能耗, 保证传输质量, 必须将采集到的多媒体信息进行数据压缩、融合等复杂运算。而这些数据量的处理一般达到 Mb 数量级, 导致 WMSNs 相比 WSNs, 处理任务复杂度提高, 运算量增大。

#### (4) 网络带宽要求高

在 WMSNs 中传输的数据除了简单信息以外, 还存在大数据量的多媒体信息, 这使得 WMSNs 对传输网络的带宽提出较高要求。对于 WSNs 而言, Kb 量级带宽的传输网络能够满足应用需求, 而具有图像、音频和视频等多媒体信息采集和传输功能的 WMSNs, 网络需要达到 Mb 量级<sup>[6]</sup>。

针对 WMSNs 所具有的特点, 传统的视频编码由于其编码端复杂且能耗较大、解码端相对简单的特性变得不再适合。如何在传感器节点能量、网络传输、计算能力等资源受限的情况下, 高效地实现以多媒体数据信息为中心的编码压缩与传输, 成为 WMSNs 由理论转向应用发展中一个急需解决的问题。

分布式视频编码(Distributed Video Coding, DVC)<sup>[7]</sup>是一种全新的编解码方式, 与传统的视频编解码技术不同, 它根据分布式编码理论, 采用对多个信源进行独立编码、联合解码的方式, 将系统的复杂度从编码端搬移至解码端, 实现与传统联合编解码相同的压缩编码效率。独立编码的方式不仅降低了编码端运算、内存、能耗等资源的消耗, 而且使信源相互之间的依赖性减小, 适用于相关信源之间的相互通信少的应用, 联合解码方式是对信源相关性的预测, 能够有效地提高解码效率。DVC 的这些特性, 与 WMSNs 中传感节点因计算能力和内存容量有限, 不能进行复杂的联合编码的要求相符合, 能够解决 WMSNs 应用中的“瓶颈”问题。

WMSNs 中采用分布式视频编码技术, 使多媒体传感节点在资源受限的情况下对感兴趣事件进行信息数据的采集和压缩, 通过无线方式传送至控制中心。由于控制中心处于能耗不受限的环境中, 能够对数据进行较复杂的联合解码, 并对还原数据进行分析处理。独立编码时充分探索相关性并去除冗余, 实现对信息的

高效压缩,减小网络传输的数据量,降低网络负载与能耗,延长网络生存时间。解码时对多信息进行联合解码,增加了应对网络传输错误的容错性,通过进行运动估计、运动补偿等复杂运算降低误码率,从而提高网络监测性能。

随着科学技术的发展,采用分布式视频编解码技术的 WMSNs 将实现具有细粒度、精准而丰富的多媒体信息监控,具有重要的意义和较高的应用价值,可广泛应用于军事战场监控、环境监控、安全监控、交通监控、智能家居监控、医疗卫生、敏感区域监控、公共安全等领域,具有很好的发展前景。本文将以分布式视频编码技术为研究重点,对应用于 WMSNs 的分布式视频编码技术展开研究。

## 1.2 国内外研究现状

早在上世纪 70 年代分布式视频编码理论由 Slepian 和 Wolf、Wyner 和 Ziv 相继提出,但随着无线网络技术的发展与新应用需求的推动,直到最近十年才使得广大研究人员逐渐重视对分布式视频编码技术的研究与实践,并取得了一定的成果。2002 年,斯坦福大学的 Bernd Girod 等人提出了基于像素域的 Wyner-Ziv 分布式视频编码方案,获得的编码性能优于传统帧内编码,但与传统帧间编码相比,仍存在一定差距<sup>[8]</sup>;在此基础上,Bernd Girod 研究小组将视频编码扩展到变换域,提出了基于变换域的 Wyner-Ziv 视频编码方案<sup>[9]</sup>;加州大学 Berkeley 分校的 Ramchandran 等人在充分研究分布式视频编码后提出了 PRISM 视频编码方案<sup>[10-11]</sup>;欧盟 ISTFET 的 DISCOVER 小组提出基于 LDPCA 的变换域的 DISCOVER 编码框架,其编码性能相比 H.264 视频编码标准有了一定的提升<sup>[12]</sup>;为了减少因系统预测检测匹配码所引入的大量通信数据,Sehgal 等提出“state-free”分布式视频编码方案<sup>[13]</sup>;2005 年,Javier Garcia-Frias 等人在基于 Turbo 编码的 DVC 方案基础上,对剩余 Turbo 码进行压缩,取得了较好的效果<sup>[14]</sup>;Zixiang Xiong 等证明了采用低密度奇偶校验码(Low Density Parity Check Code, LDPC)<sup>[15]</sup>相比 Turbo 码可以获得较好的压缩性能,在此基础上提出了分级的 Wyner-Ziv 视频编码方案<sup>[16-17]</sup>;2006 年,Bernd Girod 所在研究小组通过对 LDPC 码的研究,提出了基于 LDPC 的变换域分布式视频编码方案<sup>[18]</sup>;Fowler 等提出了基于小波编码的分布式视频编码方案<sup>[19]</sup>;2007 年 Adikari 等提出了递归解码分布式视频编码方案<sup>[20]</sup>;

Xiong 等人提出了一种压缩性能较高的分布式视频编码方案, 通过在编码端进行运动搜索来得到视差信息, 为解码过程提供可靠信息<sup>[21]</sup>; 2007 年 Varodayan 等提出了多视角图像的分布式编码方案<sup>[22]</sup>; Catarina Brites 等提出了一种基于变换域分布式视频编码的解码码率控制算法, 能够有效地预测校验码码率<sup>[23]</sup>。

国内院校也展开了对分布式视频编码技术的研究, 其中中科院研究小组为了将编码速率接近于带有条件熵的分布式系统, 提出了一种最优不均匀量化方案<sup>[24]</sup>; 南京邮电大学的干宗良等提出了一种空间域 Wyner-Ziv 视频编码系统的性能改进算法<sup>[25]</sup>; 蒋刚毅等人对分布式视频编码的高效压缩和边信息插值两大关键技术进行了讨论<sup>[26]</sup>; 北京交通大学的王安红对基于 LVQ(Lattice Vector Quantization) 的 DVC 高效性能方案设计以及鲁棒性能等展开了深入研究<sup>[27]</sup>; 四川大学的卿粼波等人将 H.264 解码算法引入像素域视频编码系统, 提出一种采用并行编码、迭代解码的混合视频编码框架<sup>[28]</sup>。

总体来说, 分布式视频编码技术已成为一大研究热点, 但针对 WMSNs 的特点的实际应用依然处于初步探索和实验阶段, 实际应用中还存在很多问题需要解决。

### 1.3 研究的主要内容

论文以无线多媒体传感器网络为应用背景, 通过深入研究分布式视频编解码技术和边信息生成算法, 提出块分类的双向估计加权边信息生成算法和相关性分类的解码监测优化算法, 完成面向 WMSNs 的分布式视频编解码方案的设计。全文组织结构如下:

第一章阐述了课题的研究背景以及分布式视频编码技术的研究现状, 引出课题的研究目的和意义。

第二章详细论述了分布式视频编码的两大理论基础, 分析了分布式视频编码的特点, 并对典型的分布式视频编码系统方案进行阐述, 在此基础上讨论了分布式视频编码的关键技术。

第三章首先阐述了边信息的概念、在分布式视频编码系统中的重要性以及典型边信息生成算法, 通过对 WMSNs 应用以及当前几种改进算法的分析, 提出一

种块分类的双向估计加权边信息生成算法，并与典型算法进行仿真测试比较。

第四章针对解码失真严重情况下，系统解码检测以及提高解码质量问题，对几种边信息优化算法进行分析，在此基础上提出一种相关性分类的解码检测优化算法，并对算法进行了详细描述，最后给出仿真测试结果与分析。

第五章针对面向 WMSNs 的应用，分析网络结构特点和应用要求，结合论文所提边信息生成算法和解码检测优化算法，提出一种 WMSNs 的分布式视频编解码系统方案，在详细描述系统功能模块之后，对整个系统进行仿真测试并对结果进行分析。

第六章对全文的工作进行一个整体总结，并对下一步研究工作进行展望。

## 第二章 分布式视频编码技术

分布式视频编码(Distributed Video Coding, DVC)是分布式信源编码理论(Distributed Source Coding, DSC)在视频编码中的应用。本章将在简述分布式视频编码理论及特点的基础上, 详细分析典型分布式视频编码系统方案及关键技术。

### 2.1 分布式视频编码原理

分布式信源编码指分别对两个或多个相关且独立的随机序列进行独立编码, 各编码器编码生成的比特流通过信道传输, 发送到同一个解码器。解码器根据随机序列间的统计相关特性, 对输入的所有比特流进行联合解码, 以去除信源之间的冗余信息。

在分布式信源编码的理论基础上, 利用信道编码技术的分布式视频编码是一种全新发展起来的视频编码方案, 它将系统的计算复杂度由编码端搬移至解码端。分布式信源编码的理论主要分为: Slepian-Wolf 分布式无损编码理论和 Wyner-Ziv 分布式有损编码理论。Slepian 和 Wolf 于上世纪 70 年代率先提出了分布式无损编码理论, 论证了对独立同分布的随机序列, 进行独立编码、联合解码可以达到与联合编码、联合解码相同的压缩效果<sup>[29]</sup>。Wyner 和 Ziv 在此基础上进行研究, 提出了利用辅助边信息在有率失真限制的情况下进行联合解码的分布式有损编码理论<sup>[30]</sup>。

#### 2.1.1 Slepian-Wolf 分布式无损编码理论

由信息论中的香农采样定理可知<sup>[31]</sup>, 对于两个统计相关的独立同分布随机序列 X 和 Y 来说, 进行独立编码、独立解码时, 为了保证无差错解码, 各自压缩码率必须满足计算式(2-1)。

$$\begin{aligned} R_x &\geq H(X) \\ R_y &\geq H(Y) \end{aligned} \quad (2-1)$$

其中  $R_x$  和  $R_y$  分别是随机序列 X 和 Y 的码率,  $H(X)$  和  $H(Y)$  分别是随机序列 X 和 Y 的信息熵。

对两个统计相关的独立同分布随机序列 X、Y 进行联合编码、联合解码时, 如图 2.1 所示, 要使解码端无失真地恢复出 X 和 Y, 则随机序列必须满足计算式

(2-2), 其中  $H(X, Y)$  为联合信息熵。

$$R_x + R_y \geq H(X, Y) \quad (2-2)$$

由于联合编码时利用了随机序列之间的统计相关性, 使得对序列  $X$ 、 $Y$  进行联合编码时可以获得比独立编码更好的效果。

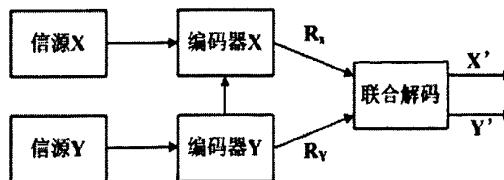


图 2.1 联合编码、联合解码

假定两个互相关的离散无记忆信源  $X$  和  $Y$ , 对  $X$  进行无损压缩时利用  $Y$  作为参考信息。如果在编码端和解码端均能得到参考信息  $Y$ , 由香农定理可知, 已知  $Y$  的情况下的  $X$  的条件熵是  $X$  的理论无损压缩极限, 即  $H(X|Y)$ 。1973 年, Slepian 和 Wolf 两位学者对两个统计相关的独立同分布随机序列进行独立编码、联合解码, 如图 2.2 所示, 通过对这种编解码方式的研究, 提出了分布式视频无损压缩编码的码率边界。

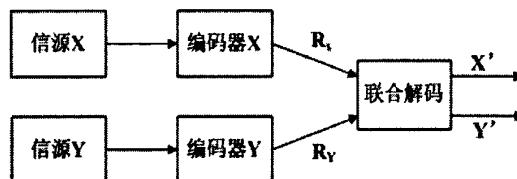


图 2.2 独立编码、联合解码

若仅在解码端得到参考信息  $Y$ , 而编码器在没有参考信息  $Y$  的情况下只需知道  $X$  和  $Y$  的联合概率分布, 同样能够取得与编码端已知参考信息一样的编码效率, 这就是 Slepian-Wolf 编码理论。根据这一理论, 当满足计算式(2-3)时, 互相关的两个信源  $X$ ,  $Y$  能够以各自的码率进行独立编码, 而在解码端将以任意小的概率进行联合解码。

$$\begin{aligned} R_x &\geq H(X|Y) \\ R_y &\geq H(Y|X) \\ R_x + R_y &\geq H(X, Y) \end{aligned} \quad (2-3)$$

其中,  $R_x$  和  $R_y$  分别为序列  $X$ 、 $Y$  编码后的码率,  $H(X|Y)$  是已知  $Y$  情况下  $X$

的条件熵,  $H(Y|X)$ 是已知  $X$  情况下  $Y$  的条件熵,  $H(X, Y)$ 为联合信息熵。

Slepian 和 Wolf 证明了对两个统计相关的独立同分布随机序列进行编码时, 如果允许存在一个任意小(但通常不为零)的解码错误概率, 则对  $X$  和  $Y$  采用独立编码、联合解码的编解码方式。虽然分别对  $X$  和  $Y$  进行独立编码, 在解码端仅利用  $X$  和  $Y$  之间的相关性进行联合解码, 但是总的码率  $R_x + R_y$  可以达到与联合熵  $H(X, Y)$ 相同, 与采用联合编码、联合解码的编解码方式的信息传输率相同。因此, 分布式视频编码方案的独立编码和传统的视频编码方案的联合编码相比较, 两者压缩效率在理论上相同, 压缩并未引入失真。

Slepian-Wolf 编码理论所取得的码率可以用图 2.3 来描述, 横轴表示信源  $X$  的码率, 纵轴表示信源  $Y$  的码率,  $H(X)$ 和  $H(Y)$ 分别是  $X$  和  $Y$  的信息熵。图中的两个转折点 A、B 为分布式无损编码需要达到的界限, 只要  $R_x$  和  $R_y$  的取值范围处在阴影区域中, 即位于直线  $AB$  右上部分的区域, 则在解码端可以以任意小的误差概率进行联合解码, 达到无损压缩的要求。

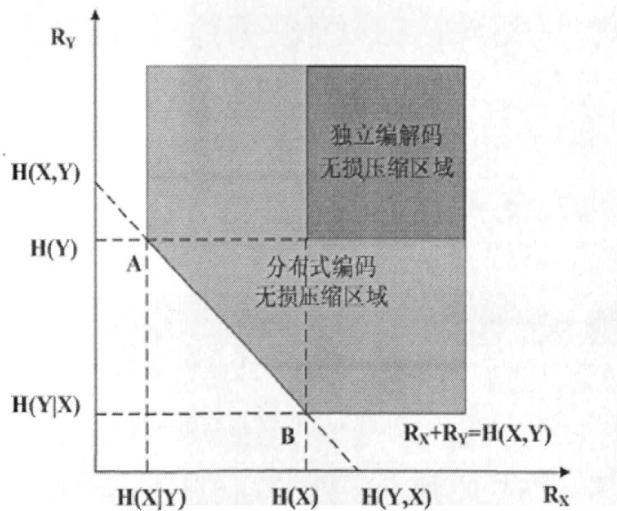


图 2.3 Slepian-Wolf 理论中的容错码率区域

在解码端利用辅助边信息是分布式编码的一个特例如图 2.4 所示, 原始信源  $X$  对应图 2.3 中 A 点,  $X$  和  $Y$  在编码端进行独立编码, 在解码端对  $X$  进行解码时将  $Y$  作为辅助边信息。能对  $X$ 、 $Y$  进行无损编码的条件分别满足计算式 (2-4)(2-5), 得到总码率  $R_x + R_y$  满足计算式(2-6)。

$$R_y \geq H(Y) \quad (2-4)$$

$$R_x \geq H(X|Y) \quad (2-5)$$

$$R_x + R_y \geq H(Y) + H(X|Y) \quad (2-6)$$

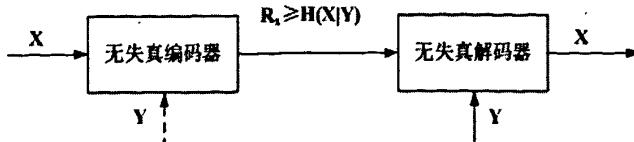


图 2.4 使用边信息的无损分布式编解码

### 2.1.2 Wyner-Ziv 分布式有损编码理论

1975 年 Wyner 和 Ziv 两位学者在 Slepian-Wolf 分布式无损编码理论的基础上, 将分布式编码扩展到了有损压缩编码领域, 建立了解码端使用辅助边信息的分布式有损编码理论。

设  $X$  和  $Y$  为两个统计相关的独立同分布随机序列, 分别对应原始信源  $X$  和边信息  $Y$ , 如图 2.5 所示。在编码端对  $X$  和  $Y$  分别进行独立编码, 在解码端先解码边信息  $Y$  并将其作为辅助边信息, 利用  $X$  与  $Y$  之间的相关性对信源  $X$  进行重构, 得到信赖度较高的重构信息  $X'$ , 其中解码失真度为  $D = E[d(X, X')]$ , 表示能接受的容错范围。

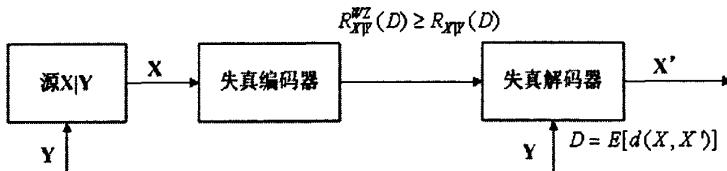


图 2.5 具有统计边信息的有损分布式信源编码

当仅在解码端知晓边信息时, 定义  $R_{X|Y}^WZ(D)$  为率失真函数, 表示某一失真度  $D$  编码器能够达到的输出码率下界; 在编码端和解码端都已知边信息的情况下, 定义  $R_{X|Y}(D)$  为率失真函数, 表示解码端边信息有效时对编码端的请求码率。Wyner 和 Ziv 证明了在编码端无法获悉边信息  $Y$  的情况下, 信源序列  $X$  的最低码率  $R_{X|Y}^WZ(D)$  满足计算式(2-7)。

$$R_{X|Y}^WZ(D) - R_{X|Y}(D) \geq 0 \quad (2-7)$$

当信源是高斯无记忆信源且失真函数采用均方误差时,  $R_{X|Y}^WZ(D)$  和  $R_{X|Y}(D)$  相等, 即满足计算式(2-8)。

$$R_{X|Y}^{WZ}(D) = R_{X|Y}(D) \quad (2-8)$$

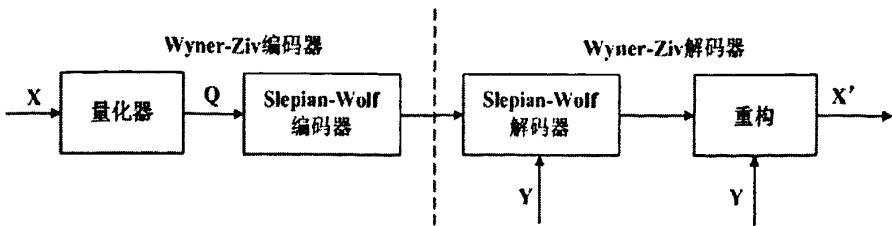


图 2.6 典型的 Wyner-Ziv 编解码器

目前实际中大多数信号是连续信号源,需要在一个设定的失真度前提下进行有损压缩,因此现有大多数分布式视频编码方案是基于 Wyner-Ziv 有损分布式编码理论。一个典型的 Wyner-Ziv 编解码系统中,编码端主要由量化器、Slepian-Wolf 编码器组成,而解码端由 Slepian-Wolf 解码器、反量化器组成对称结构,如图 2.6 所示。

## 2.2 分布式视频编码特点

无损情况下的 Slepian-Wolf 理论和有损情况下的 Wyner-Ziv 理论的共同特点是,突破传统视频编码理论仅在编码端利用信源统计特性,创新性地提出将信源统计特性转移至解码端,利用信源统计特性实现高效编码压缩。归纳起来, DVC 主要有以下几个特点:

- (1) 编码低复杂度、解码高复杂度。DVC 将运算量大、处理复杂的数据处理、运动估计等由编码端搬移至解码端,满足编码简单、解码复杂的应用需求。
- (2) 具有较好的鲁棒性。各采集点通过分布式方式将采集的信息进行传输,当处于传输质量较差的网络环境时,可以采用从相邻数据进行恢复的方式应对生成的误码信息,增加容错性。
- (3) 较高压缩效率。采用分流方式进行视频帧处理,一部分采用 H.26X 等标准压缩技术,确保多媒体信息的编解码质量;一部分采用分布式压缩技术,最大程度地进行压缩处理。
- (4) 容易形成立级编码码流。为了在出现严重误码时,能够通过相关联码流来进行纠错, DVC 将传输数据进行划分,在保证独立性的同时,又使其存在一定的相关性。

分布式视频编码技术将传统的编码方式编码端的复杂度搬移至解码端,降低编码端的负担。由于具有编码简单、解码复杂的特点,使得它适用于计算能力差、内存容量小、能耗等因素受限的无线通信终端的应用,如移动视频电话、无线监控系统、无线传感器网络等。

## 2.3 典型的分布式视频编码系统方案

在 Slepian-Wolf 无损编码理论和 Wyner-Ziv 有损编码理论的基础上,1999 年,Pradhan 和 Ramchandran 提出了第一个分布式视频编码系统方案<sup>[32]</sup>。随着应用需求的推动,各种分布式视频编码方案由众多的研究者相继提出。其中较典型的有基于像素域的 Wyner-Ziv 视频编码系统、基于变换域的 Wyner-Ziv 视频编码系统和 PRISM 视频编码系统等。

### 2.3.1 基于像素域的 Wyner-Ziv 视频编码系统

基于像素域的 Wyner-Ziv 视频编码系统由 Bernd Girod 研究小组在 2002 年提出,系统结构如图 2.7 所示。系统方案基于独立编码、联合解码思想,在编码端将视频帧分为相互独立的 Wyner-Ziv 帧 W 与关键帧 K。对于关键帧 K,采用传统的帧内编解码方式进行压缩,对于 Wyner-Ziv 帧,其编解码算法主要通过以下几步:

编码端:采用均匀量化器对 W 帧进行  $2^M$  均匀量化处理,使所有系数量化后都处于  $2^M$  个区间中,量化系数用 M 个比特表示,将所有系数按相同位的比特集中,形成 M 个位平面。重要的量化系数被依次送入 Turbo 编码器进行独立编码,得到的编码校验码放入缓存中,等待解码时的发送请求,再传输给解码端使用。

解码端:利用已解码关键帧 K 和 W 帧生成辅助边信息  $\tilde{W}$ ,用于对当前 W 帧的解码过程,将边信息进行量化,生成的系数带送入 Turbo 解码器。解码器结合边信息信息,通过反馈信道向编码端请求校验码的方式进行解码。当解码误码率满足特定的值后解码完成,若解码失败,则需要请求更多的校验码重复解码,直至达到解码误差标准并获得解码图像的量化系数。最后进行反量化与帧重构,得到 Wyner-Ziv 帧的解码视频帧 W'。

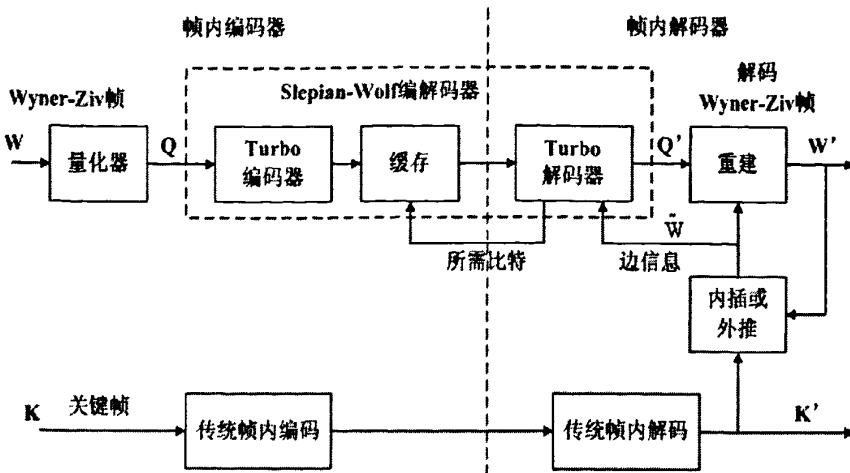


图 2.7 基于像素域的 Wyner-Ziv 视频编解码系统

基于像素域的 Wyner-Ziv 视频编码系统是最简单的 DVC 系统，编码端对输入视频无需进行离散余弦变换，直接进行量化，大运算的运动估计和运动补偿转移至解码端，保证了整个系统的复杂度较低。实验结果表明，基于像素域的 Wyner-Ziv 视频编码系统性能优于传统帧内编码，但与帧间编码性能存在差距。

### 2.3.2 基于变换域的 Wyner-Ziv 视频编码系统

基于像素域的 Wyner-Ziv 视频编码系统能够达到去除帧间时间相关性的效果，但未考虑到 Wyner-Ziv 帧的空间相关性。为了提高压缩效率，2004 年 Bernd Girod 研究小组在基于像素域的 Wyner-Ziv 视频编码系统的基础上，提出了基于变换域的 Wyner-Ziv 视频编码系统，系统框图如图 2.8 所示<sup>[33-34]</sup>。

系统中视频帧被分为关键帧 K 和 Wyner-Ziv 帧 W，对关键帧 K 采用传统的帧内编解码方式，如 H.264 等。对 Wyner-Ziv 帧 W 采用帧内编码、帧间解码的方式。系统对 W 帧的编解码流程如下：

为了降低像素的空间冗余，编码端首先对 W 帧进行离散余弦变换得到 DCT 系数，并针对系数所处位置进行频段划分，组合成 DCT 系数子带；对各 DCT 系数子带进行量化，将量化符号  $q_k$  中具有相同特性的比特构成比特面；采用 Turbo 码对每个比特面进行独立编码，编码所产生的奇偶校验码暂时存放到缓存中，等待需求信息再传送到解码端。

在解码端首先利用相邻已解码帧 K 与当前帧的相关信息，通过内插或外推

算法生成边信息  $W$ ，并进行 DCT 变换和频段划分；Turbo 码解码器利用校验码和边信息，依次对各频率子带系数进行纠错解码，若解码失败，则通过反馈信道向编码端请求更多的校验码进行，直至解码错误率低于设定阈值，解码码流经过位平面重组得到符号流  $q_k'$ ；重构模块根据边信息和符号流进行频段系数的重构；最后通过 IDCT 变换得到解码后的 WZ 帧  $W'$ 。

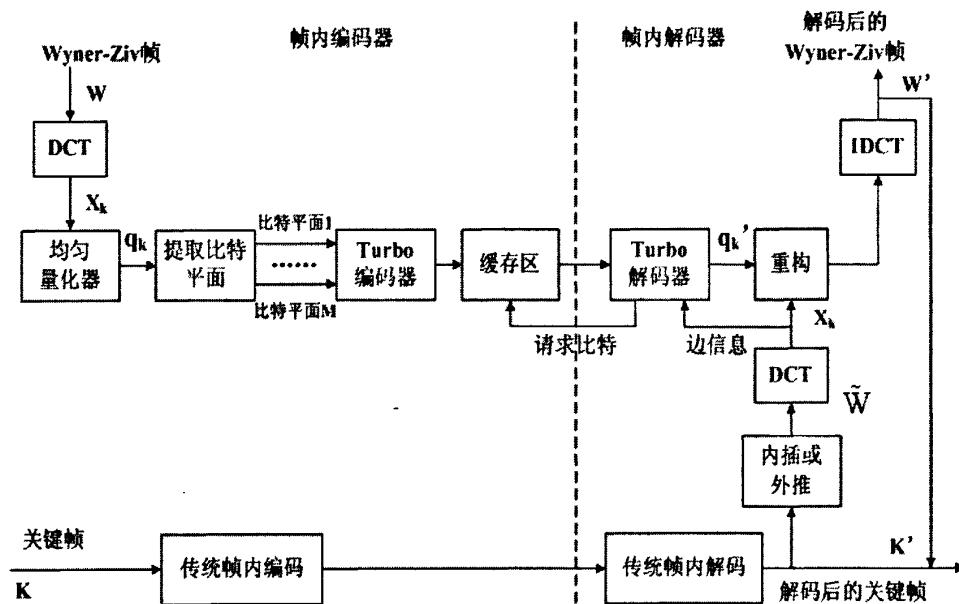


图 2.8 基于变换域的 Wyner-Ziv 视频编解码系统

变换域 Wyner-Ziv 视频编解码系统与像素域 Wyner-Ziv 视频编解码系统相比最大的区别是 WZ 帧处理中引入了 DCT 变换，视频帧由像素域变换到频率域，使得编码复杂度有一定的增加，但采用离散余弦变换有效压缩了视频信号空间冗余信息。实验结果表明，基于变换域的 Wyner-Ziv 视频系统能够获得比基于像素域的 Wyner-Ziv 视频系统更好的压缩性能与图像质量<sup>[35]</sup>。

### 2.3.3 PRISM 视频编码系统

加州大学 Berkeley 分校的 Ramchandran 研究小组与 Bernd Girod 研究小组同时对分布式视频编解码展开了研究，并于 2002 年提出并实现了基于宏块的 PRISM 视频编码系统<sup>[10-11]</sup>，其系统结构如图 2.9 所示。系统框架的具体工作流程如下：

编码端：将视频帧划分成互不重叠的像素块，根据当前块与对应预测块的相

关性进行模式划分，分别为 Skip 模式、传统编码模式和综合编码模式。属于 skip 模式的块无需进行编码，传统编码模式的块，采用传统的帧内编解码方式，综合编码模式的各像素块进行离散余弦变化得到 DCT 系数，并对 DCT 系数进行量化，量化步长的设定与相关噪声方差成正比。量化后得到的量化系数进行混合编码，对于量化后具有低频特征的比特位，采用信道编码并在解码端通过边信息估计得到，只对高频特征的比特位进行熵编码，量化各个宏块的同时，生成一个循环冗余校验作为 Hash 码，进一步降低解码错误率。

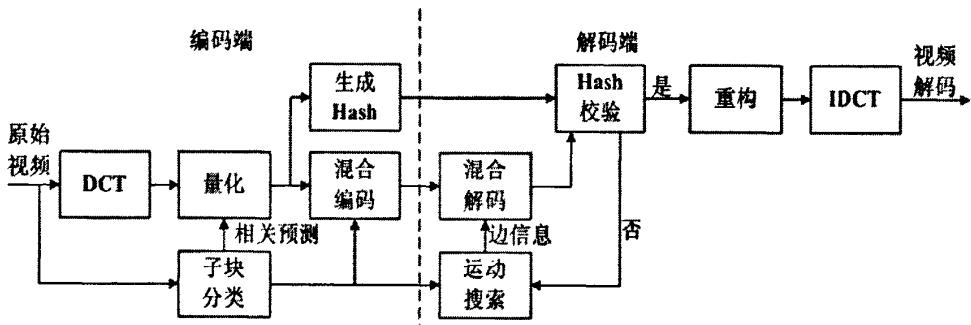


图 2.9 PRISM 视频编解码系统

解码端：与编码端对应，对不同模式的宏块采用对应的解码方式。skip 模式的块，利用已解码帧前一帧中对应块直接作为当前块，对于综合编码模式的块，解码端对高频部分采用熵解码方式，对低频部分通过运动估计生成边信息，并与接收码流一起进行解码。利用 Hash 校验码对当前解码进行检测，校验正确则解码成功，否则重新生成边信息，直至校验正确。将低频和高频各系数组合得到量化系数，与边信息一起进行 IDCT 变换，最终得到解码帧。

PRISM 编解码系统与 Wyner-Ziv 系统一样，将复杂度从编码端搬移至解码端，此外它可以根据编码端的情况，灵活地调整编码端和解码端的运算复杂度，保持整个系统的编解码性能，具有较高的灵活性。由于引入了 Hash 校验码，使得系统的鲁棒性较高，能够抵抗传输误码。实验结果表明，在无丢帧的情况下，PRISM 的性能介于 H.263 帧内编码与帧间编码之间，而在有丢帧发生情况下，PRISM 的具有比传统视频系统更高的容错性，且其编码端的运算复杂度低于传统编码系统。

## 2.4 分布式视频编码关键技术

### 2.4.1 边信息生成

在分布式视频编解码系统中,边信息是通过利用待解码的 WZ 帧与相邻关键帧的相关特性,对当前 WZ 帧进行估计产生的预测信息,用于对当前 WZ 帧的解码提供辅助信息。与传统的视频编解码方式不同,分布式视频系统解码是在解码端未知当前帧任何信息的前提下,通过帧间预测来进行最佳匹配块的搜索,生成辅助边信息。生成的边信息与当前帧的相似度越大,表明边信息的质量越好,随之降低解码过程中的错误概率,减少校验位的请求,最终提高分布式视频编解码的效率。

分布式视频系统的解码依靠边信息获得视频帧之间的相关信息。边信息不仅参与解码过程,还与经过解码器产生的量化系数一起对当前帧进行重构,由此可见,边信息不仅在很大程度上影响着分布式视频编解码系统的复杂度,而且在提高系统性能方面有着十分重要的作用。因此,如何保证边信息与原始帧差值尽量小是分布式视频编解码的关键技术之一。边信息的生成以及优化方法将在之后的章节重点介绍。

### 2.4.2 编解码

在分布式视频编码系统中,编解码方式以及对应的编解码器是一个重要组成部分。编码器根据信源生成对应码字,只传输校验信息,实现对采集信息的压缩。生成对应校验信息将存放在缓存中,待接收到来自解码端的请求,再将校验信息传输至解码端。解码器解码时需要向缓存发送请求校验信息,用于对解码信息进行正确验证。可见,一个典型的 Slepian-Wolf 编解码器主要由编码器、缓存、解码器构成。

目前,在分布式视频系统中,主要采用基于 Turbo 码<sup>[36]</sup>和低密度校验码(Low Density Parity Check Codes, LDPC)的编解码方式。由于 Turbo 码较好地运用了香农信道编码定理中的随机性编解码条件,能够逼近香农理论中的极限值,很长一段时间内,分布式编码系统方案均采用 Turbo 码。但是由于 Turbo 码采用迭代编解码方式,使得解码过程存在较大时延,不适用于对实时性要求较高的应用,其

次 Turbo 码存在的容错性较低, 解码复杂的缺点。近年来, 由 Mackay 等提出的 LDPC 码以其越来越受到研究者的关注。由于没有采用交织器, 使得 LDPC 码的时延性较低, 差错平底特性有利于并行操作的实现, 解码复杂度相比 Turbo 码明显降低, 且在解码过程中能够实现错误检测。由于 LDPC 能以较低的复杂度获得与 Turbo 码相同的性能, 使得目前采用 LDPC 码的分布式视频系统更为普及。

#### 2.4.3 相关噪声模型

相关噪声统计指当前帧与边信息之间的相关性, 即边信息可看成是当前帧通过一个虚拟信道传输得到。对相关噪声的估计包括建模和参数估计, 一般通过虚拟信道来实现。好的相关噪声模型不仅能够很好的降低当前帧与边信息间的时间空间冗余, 还可以有效地提高分布式视频编解码系统性能。

早期的 DVC 系统采用加性高斯白噪声信道作为信道模型<sup>[37]</sup>; 现有的分布式编码方案通常假设当前帧与边信息的相关噪声服从拉普拉斯分布, 一般使用 DCT 变换来降低空间冗余。Trapanese 等人采用混合指数分布, 将当前帧与边信息的相关噪声模型看成满足拉普拉斯分布<sup>[38]</sup>。Brites 等人在解码端对基于帧层的相关噪声相关性进行估计, 利用前、后向运动补偿帧的差值对相关噪声的方差进行估计<sup>[39]</sup>。考虑到实际信号的非平稳性, Brites 等人采用频带级、像素系数级获变换系数级的相关噪声建模取得较好的解码性能<sup>[40]</sup>。当前解码帧与边信息的噪声模型影响着解码帧的重建, 因此有待进一步研究更有效的信道模型或动态改变模型参数等方法, 提高 DVC 系统的效率。

#### 2.4.4 码率控制

在分布式视频编解码系统中, 系统的码率由当前帧和边信息的统计相关特性以及它们的码率来决定。

- (1) 关键帧和 WZ 帧的比例。关键帧在编码端采用传统的帧内编码方式, 因此关键帧的比例过大将导致码率的过高。
- (2) 关键帧的码率。由关键帧的量化步长来决定。
- (3) WZ 帧的码率。由 Wyner-Ziv 编码器中的量化器的量化步长以及传输校验位的数量来决定。

借鉴传统视频编码方案可以实现对前两个影响因素的控制,而对于 WZ 帧的码率控制则是分布式视频编码中一个全新的研究内容。目前完全依赖于解码端的反馈信息是进行 WZ 帧码率控制的方法之一,由于码率控制伴随着编解码同时进行,该方法只适用于实时视频传输的应用。Brites 等提出了一种由编码端进行码率控制的算法,适用于基于变换域的 Wyner-Ziv 视频编码系统<sup>[41]</sup>。在编码端通过简单的时域相关性估计来确定编码码率是另一种方法码率控制方法,在 PRISM 编解码系统中得到了运用。

## 2.5 本章小结

本章由四个部分组成,第一部分首先阐述了分布式视频编码的两个理论基础: Slepian-Wolf 分布式无损编码理论和 Wyner-Ziv 分布式有损编码理论,为论文提供理论基础。第二部分概述了分布式视频编解码特点。第三部分较详尽的讨论了三种典型的分布式视频编码系统,并对视频编码框架和特点进行分析。最后对分布式视频编码的关键技术进行介绍:边信息生成、编解码、相关噪声模型以及码率控制,分析了研究现状及发展趋势。本章的研究工作在整个研究过程中起着奠基的作用,是研究工作的一个基础环节。

### 第三章 块分类的双向估计加权边信息生成算法设计

与传统的视频编码系统所采用的编解码方式不同,分布式视频编码系统一般采用非对称方式,即在解码端根据相邻已解码帧与当前帧之间的相关信息来生成辅助信息——边信息(Side Information, SI),用于系统的解码过程。边信息是对当前帧的估计,用于进行帧间预测解码,边信息质量的优劣对提高分布式视频编解码系统性能是一个至关重要的因素。本章针对分布式视频编解码中运动估计产生的预测重叠问题,提出一种块分类的双向估计加权边信息生成算法,并对算法进行详细描述。

#### 3.1 典型边信息生成算法

边信息的生成在分布式视频编解码系统中扮演着一个关键的角色——为解码过程提供解码辅助信息,边信息的质量将直接影响最终解码视频信息的质量。为了得到高质量的视频解码信息,如何获得一个失真度较小的边信息,是分布式视频编码的主要研究课题之一。

边信息主要利用视频序列中视频帧之间高度相关的时间、空间相关特性,在解码端利用运动估计、运动补偿等技术进行重构得到。典型的边信息生成算法有前帧替代算法、像素平均算法、运动补偿内插算法、运动补偿外推算法。

##### (1) 前帧替代算法

前帧替代算法是边信息生成算法中最简单的方法,其基本思想是利用相邻视频帧之间的较高相关性,假设前一帧能够无失真的解码,忽略相邻帧之间的差异,将前一解码帧( $X_{2k-1}$ 帧)的图像信息直接作为当前帧( $X_{2k}$ 帧)边信息的图像信息,如计算式(3-1)所示。

$$Y_{2k} = X_{2k-1} \quad (3-1)$$

由于采用直接替换处理,省略了对当前帧的运动估计过程,因此运算十分简单,对解码器运算复杂度的要求很低,解码端只需存储前一解码帧信息,使得系统所需的缓存空间开销小。生成边信息时只需要前向相邻解码帧,无需等待其他视频帧的解码完成,实时性较高,适用实时性较高的应用场合。

在运动变化相对剧烈的环境下,视频序列中相邻帧之间的运动变化量较大,算法简单地将前解码帧作为边信息帧的替代,得到的边信息与实际情况相差较大。解码时需要向编码端发送更多的校验比特码请求,从而降低了系统解码速度,导致系统性能较低。

### (2) 像素平均算法

在前帧替代边信息生成算法的基础上,为提高边信息生成质量,Aaron等人加入了简单的运算,提出了像素平均边信息生成算法,即将当前帧( $X_{2k}$ 帧)的前后两个相邻已解码帧( $X_{2k-1}$ 帧)中相同坐标位置的像素值进行平均计算,得到的平均值作为边信息对应坐标位置的像素值,如计算式(3-2)所示。

$$Y_{2k}(i, j) = \frac{1}{2} [X_{2k-1}(i, j) + X_{2k+1}(i, j)] \quad (3-2)$$

像素平均算法虽然引入了均值计算,但没有进行运动补偿,复杂度较小。相比前帧替代算法,算法利用了前后相邻关键帧中对应坐标位置像素值之间的相关性,一定程度地提高了生成的边信息可靠性。

与前帧替代算法一样,算法没有深入考虑视频序列的运动特性,当画面运动变化较大时,所得边信息与当前帧的相似度将降低,导致解码出现错误的概率较大。由于解码时需要前后两相邻已解码帧,使得解码过程存在一定时延。

### (3) 运动补偿外推算法

运动补偿外推算法(Motion Compensated Extrapolation, MCE)又称推演法。MCE 算法假设相邻像素帧的运动特性前后保持一致,先计算当前帧的前两个时刻已解码帧  $X_{2k-2}$  与  $X_{2k-1}$  间的运动矢量  $MV_{2k-1}$ ,利用线性外推思想,根据计算式(3-3)得到当前帧  $X_{2k}$  与已解码帧  $X_{2k-1}$  间的运动矢量  $MV_{2k}$ ,如图 3.1 所示。根据运动矢量  $MV_{2k}$  和已解码帧  $X_{2k-1}$  进行运动外推计算,得到当前帧的对应边信息,如计算式(3-4)所示。

$$MV_{2k} = -MV_{2k-1} \quad (3-3)$$

$$X_{2k}(i, j) = X_{2k-1}(i - MV_{2k,x}, j - MV_{2k,y}) \quad (3-4)$$

运动补偿外推算法不需要利用当前时刻以后的已解码帧,能够按时间顺序进行解码,保证了解码的实时性。利用相邻帧之间的运动矢量特性来进行运动估计

操作, 对视频序列中相邻帧之间的相关性进行了一定程度的挖掘。

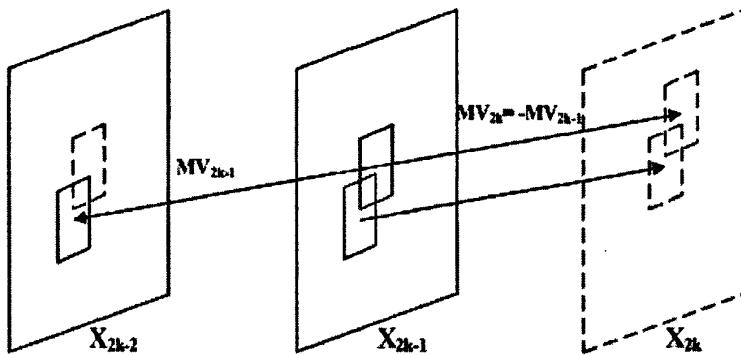


图 3.1 运动补偿外推算法

运动补偿外推算法需要额外的存储空间来保存两个前向已解码帧信息, 增大了系统解码端的存储开销。由于算法假设相邻帧前后运动特性一致, 将前一时刻的运动矢量直接作为当前时刻的运动矢量, 对于视频画面中出现曲线运动轨迹等情况, 将导致估计矢量与实际情况出现较大失真。算法仅利用前一时刻已解码帧的数据信息, 未考虑后一时刻与当前帧之间的相关性, 对于某些与后向帧相关性较大的视频序列(如开门视频中, 与之前图像相关性小, 而与之后的图像相关性大), 将丢失相关性较大的数据信息, 导致边信息与原始帧存在较大失真。

#### (4) 运动补偿内插算法

运动补偿内插算法(Motion Compensated Interpolation, MCI)利用类似传统视频编码中 B 帧的运动补偿内插方法来实现<sup>[42]</sup>。算法假设前后两相邻关键帧  $X_{2k\pm 1}$  能够无失真传输至解码端, 根据内插帧的前向运动矢量  $MV_{f2k}$  和后向运动矢量  $MV_{b2k}$  具有方向相反、模值相等的特性, 利用前后两相邻关键帧  $X_{2k\pm 1}$  进行运动内插估计。

采用块匹配算法, 在关键帧  $X_{2k-1}$  中搜索对应匹配块, 得到对应运动矢量  $MV_{2k+1}$ 。由运动矢量  $MV_{2k+1}$  得到插值帧中相同位置的像素块相对于  $X_{2k-1}$  帧的后向运动矢量  $MV_b$ , 如图 3.2 所示。根据运动矢量特性, 运动矢量  $MV_b$  与运动矢量  $MV_{2k+1}$  存在倍数关系, 如计算式(3-5)所示。前向运动矢量的获得方法与后向运动矢量的获得方法相同, 得到前向运动矢量  $MV_f$  与运动矢量  $MV_{2k-1}$  存在倍数关系, 如计算式(3-6)所示。

$$MV_b = MV_{2k+1} / 2 \quad (3-5)$$

$$MV_f = MV_{2k-1} / 2 \quad (3-6)$$

针对前后已解码帧中各像素块，利用对应的前向和后向运动矢量，通过运动补偿得到对应插值帧  $X_b$  和  $X_f$ 。最后对前后插值帧  $X_f$  和  $X_b$  取平均值，得到插值帧  $Y_{2k}$ ，如计算式(3-7)所示。

$$Y_{2k} = (X_f + X_b) / 2 \quad (3-7)$$

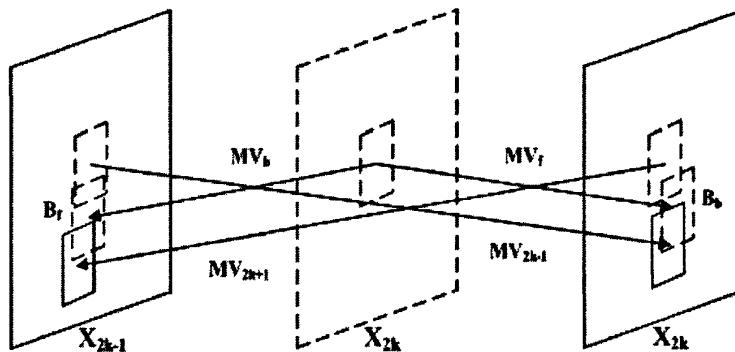


图 3.2 运动补偿内插法

运动补偿内插算法需要利用与其相邻的前后已解码帧，考虑了当前帧与前后帧的相关性对内插帧的影响，利用运动矢量对称特性，提高了内插帧的生成质量，算法能够增大生成边信息与待解码 WZ 帧间的相似度，降低解码时的错误概率。

运动补偿内插算法需要利用前后已解码帧的数据信息，使得系统需要额外的存储空间，增大了解码端的存储空间开销。由于需要计算各个像素块对应的运动矢量，计算量较大，增加了解码端的复杂度，算法需要等待前后两帧的解码完成，才能对当前帧进行解码，增加了解码的延时性。由于内插时的像素块与划分块对应位置存在随机性，使得内插帧中存在像素空洞或像素重叠现象。

前帧替代算法和运动补偿外推算法，都仅利用前一时刻的已解码帧作为参考，忽略掉后一时刻的已解码帧与当前帧的相关性，因此在当前帧与前后相关性相等或与后一时刻相关性更大的情况下，这两种算法的预测精度比像素平均算法和运动补偿内插算法要低。运动补偿内插算法相比像素平均算法而言，其边信息的生成细化到了像素块，像素平均算法假设相邻帧中各像素块的位移按相同线性移动为基础，仅简单地将先后两相邻关键帧进行平均，导致某些像素块的运动矢

量失真较大。运动补偿内插算法细化了对各个像素块的运动矢量的计算，相比得到的边信息质量最高。

### 3.2 算法设计思想

#### 3.2.1 问题分析

围绕如何生成高质量边信息这一问题，在典型边信息生成算法的基础上，许多学者提出了自己的算法思想。Kubasov and Guillemot 提出了基于 mesh 的运动补偿帧内插算法，每个视频帧中的 mesh 包含了部分量化的 DCT 系数，使得所需要的数据位相比原始数据减少，降低了边信息生成的复杂度<sup>[43]</sup>。Ascenso and Pereira 提出了基于 Hash 的运动补偿内插算法，将非关键帧的部分信息转换为 Hash 校验码传输至解码端，利用 Hash 码字对参考帧进行运动补偿来进行边信息生成<sup>[44]</sup>。文献[45]提出了基于分层变块大小运动估计的边信息生成算法，算法中采用前向运动估计(Forward Motion Estimation, FME)和双向运动估计(Bidirectional Motion Estimation, BME)，取得了质量较好的边信息。文献[46]提出了一种有效的自适应边信息帧内插方法，根据当前帧与前后两相邻帧的时间相关性，采用不同的加权系数进行运动补偿帧内插，从而提高边信息的质量。在基于整像素运动补偿内插的边信息重建的基础上，文献[47]提出了基于 1/2 像素和基于 1/4 像素的亚像素 MCI 边信息重建算法。这两种算法分别对原始帧进行 1/2 和 1/4 像素插值，利用运动估计得到运动矢量，然后进行亚像素双向补偿内插，提高了边信息的重建质量。文献[48]设计了一种复杂的边信息生成模块并提出了一种新的边信息估计算法。该模块由两部分组成，一部分采用运动补偿内插(MCI)，而另一部分采用新提出的运动补偿质量加强算法(MCQE)。以上算法思想都是对整帧所有像素块进行运动估计，将其运动变化程度同一化，忽略了像素帧中局部变化的差异，使预测得到的边信息帧与实际情况存在较大的运动矢量差异。采用整帧的思想在预测过程中将不可避免地形成“重叠”问题，重叠预测的出现不仅提高了解码端的复杂度和计算量，同时导致边信息质量降低。

#### 3.2.2 算法介绍

针对上述问题，论文提出一种块分类的双向估计加权边信息生成算法。考虑

到分布式视频系统编码端能耗受限, 算法的实现步骤主要集中在解码端。为了有助于提高后续产生的运动矢量的可靠性, 分别对前后两相邻已解码关键帧  $X_{2k+1}$  进行低通滤波。利用分块思想将前后相邻关键帧进行分块并编号, 针对相同编号的像素块进行绝对差值和计算, 得到相邻对应像素块之间的相关性信息。根据前后相邻解码关键帧对应块的相关性程度, 将像素块分成两类: 保留块类和运动块类。对保留块类进行前向像素块替代处理; 对运动块类中的像素块, 采用改进的双向运动补偿算法找到最佳匹配块, 通过加权判决模块生成边信息像素块。最后按照块编号顺序进行排列, 组合生成边信息帧。块分类的双向估计加权边信息生成算法框图如图 3.3 所示。

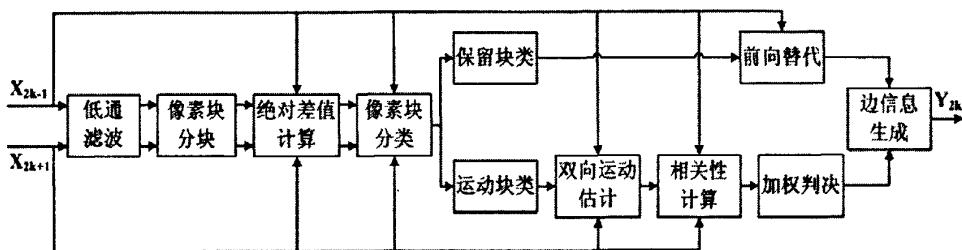


图 3.3 块分类的双向估计加权边信息生成算法框图

算法目的是尽量以较小计算量处理得到高质量的边信息。系统有针对性地对不同运动变化程度的像素块进行运动估计, 有效地提高系统效率。采用像素块分类处理的思想, 对于前后运动变化较小的像素块, 采用前像素块直接替代的方法, 在不影响解码质量的基础上减少解码端的运算量; 对于前后运动变化较大的像素块, 采用改进的双向运动估计算法得到双向运动矢量, 最后利用加权系数提高边信息的生成质量。算法思想使系统对变化较小的像素块处理过程, 能够省略不必要的运动估计计算, 并将这部分节约的资源和能耗, 转移至对复杂运动变化的像素块处理上, 实现能耗的合理分配与有效利用, 从而达到优化系统的整体编解码性能的目的。

### 3.3 算法实现

#### 3.3.1 像素块的分类

将视频中的连续像素帧依次编号, 按照编号的奇偶性将像素帧分为关键帧和非关键帧, 其中奇数编号对应关键帧, 而偶数编号对应非关键帧。

对相邻的前后关键帧  $X_{2k\pm 1}$  分别进行分块处理。由于在对图像帧分块过程中, 图像划分块的大小影响着图像精确度的高低, 划分块大, 则图像的精确度低, 如  $16 \times 16$  块适合于纹理简单且平坦的区域, 但是在运动幅度大的序列中会引入估计不准确的问题; 划分块小, 则图像的精确度高, 如  $4 \times 4$  块适合于细节较多、纹理较丰富的区域, 但在运动幅度小或者纹理简单的序列中, 不能很好地估计出运动矢量<sup>[49]</sup>。为了在保证运动估计准确度的同时降低算法计算量和复杂度, 本文采用  $8 \times 8$  大小的像素块划分格式, 同时将每个像素块进行编号, 如图 3.4 所示(为了说明简单, 图中只将像素帧划分成 16 个像素块)。将前后两相邻关键帧的对应像素块分别编号为  $B_{2k-1,x}$  与  $B_{2k+1,x}$ , 边信息帧的对应像素块编号为  $B_{2k,x}$ 。其中, 下标  $x$  表示边信息帧的像素块与相邻两关键帧的像素块的对应位置。由此可知, 相邻两关键帧和边信息帧中存在下标编号相同的像素块, 根据下标编号是否相同, 可以判断对应的像素块在各自帧中的位置是否一致。

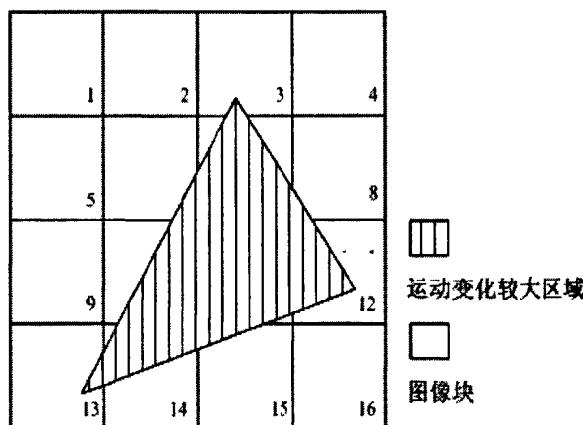


图 3.4 图像帧的分块

对前后两相邻关键帧  $X_{2k\pm 1}$  对应像素块  $B_{2k-1,x}$  和  $B_{2k+1,x}$  进行相关性分析。为了减少计算量, 采用最小绝对差和函数 SAD 作为相关性判决准则, 再根据所得 SAD 值进行相关性判断, 绝对差和值 SAD 的计算公式如计算式(3-8)。

$$SAD_x = \sum_{i=1}^n \sum_{j=1}^m |B_{2k-1,x}(i,j) - B_{2k+1,x}(i,j)| \quad (3-8)$$

其中,  $B_{2k-1,x}$  和  $B_{2k+1,x}$  分别表示关键帧  $X_{2k-1}$  和  $X_{2k+1}$  中对应的像素块,  $B(i,j)$  为像素块的灰度值,  $n$  和  $m$  分别为像素块的长度和宽度, SAD 下标  $x$  与关键帧对应。由上述可知, 当 SAD 越小, 说明其相邻对应像素块的相关性越强, 反之

亦然。

根据视频序列特性, 每组对应像素块的相关程度不同, 有些像素块(比如属于背景部分的像素块), 其前后帧之间的运动变化不大, 因而对运动估计的影响不大; 而有些像素块(比如属于物体运动范围的像素块), 其前后帧之间存在较大的运动变化或运动位移, 其变化程度对之后边信息帧的运动估计十分重要。针对上述分析, 论文提出像素块分类思想:

- (1) 根据具体视频序列的运动特性, 对计算得到的绝对误差和值  $SAD_x$  进行分析与判断, 设定一个阈值  $T_1$ 。
- (2) 将各对应像素块的  $SAD_x$  值分别与阈值  $T_1$  进行比较。根据比较结果, 将像素块分为两类: 保留块类和运动块类, 如图 3.5 所示。

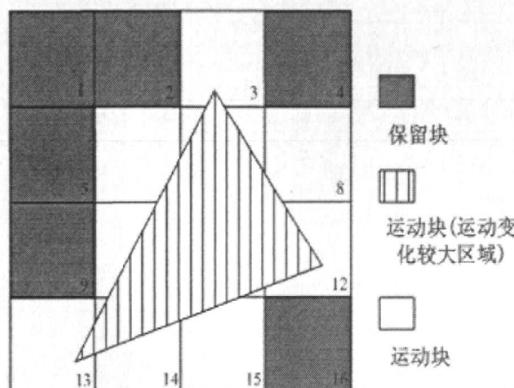


图 3.5 图像帧划分块类型

若  $SAD_x$  值大于阈值  $T_1$ , 判定对应像素块对  $B_{2k-1,x}$  和  $B_{2k+1,x}$  的运动变化程度较大, 将其像素块对归于运动块类; 若  $SAD_x$  值小于阈值  $T_1$ , 判定对应像素块对  $B_{2k-1,x}$  和  $B_{2k+1,x}$  的运动变化程度较小, 将其像素块对归于保留块类。其判决公式如计算式(3-9)。

$$SAD_x = \begin{cases} SAD_x, & |SAD_x| \geq T_1 \\ 0, & |SAD_x| < T_1 \end{cases} \quad (3-9)$$

(3) 对于属于保留块类的像素块, 由于其运动变化程度较小, 因而其运动特性对运动估计的影响不大, 为了简化解码过程, 忽略其相邻像素块之间的变化。对属于该类的边信息帧像素块  $B_{2k,x}$  采用前帧替代算法, 即由前向关键帧中对应像素块  $B_{2k-1,x}$  直接替代来生成对应边信息像素块  $B_{2k,x}$ , 如计算式(3-10)所示。

$$B_{2k,x} = B_{2k-1,x} \quad (3-10)$$

(4) 对于属于运动块类的像素块, 由于前后对应像素块之间的运动变化程度较大, 一定程度上将影响之后进行的运动估计。因此对属于该类的边信息像素块  $B_{2k,x}$ , 采用改进的双向运动补偿插值算法得到, 改进的双向运动补偿差值算法将在 3.3.2 小节介绍。

(5) 为了衡量运动块类中像素块的运动变化程度, 在运动补偿块类中, 将各对应像素块的  $SAD_x$  值进行累加, 得到绝对差值总和  $SAD_{sum}$ , 计算方法如计算式(3-11)所示。

$$SAD_{sum} = \sum_{x=1}^n SAD_x \quad (3-11)$$

当运动块类中各像素块对之间的运动变化较小时, 计算得到  $SAD_x$  的值相应较小, 使得累加得到的绝对误差总和值  $SAD_{sum}$  也随之较小; 当运动块类中各像素块对的运动变化较大时, 计算得到的  $SAD_x$  值相应较大, 导致累加得到  $SAD_{sum}$  值也随之增大。

(6) 由于不同视频序列的运动变化程度不同, 使得各像素块对的绝对差总和值存在差异。当绝对差值总和  $SAD_{sum}$  达到一定上限或超过这一上限时, 说明运动补偿块类中存在较大运动变化, 划分块格式较大, 这将影响之后的运动估计, 需要对属于该类的像素块对的大小进行再次细分。对步骤(4)计算得到的  $SAD_{sum}$  值进行分析, 并设置一个阈值  $T_2$ 。当  $SAD_{sum}$  值大于或等于阈值  $T_2$  时, 说明像素块之间的运动变化较剧烈, 将该类中每一个像素块对的大小再均匀细分为 4 个子块, 同时将边信息帧中下标编号相同的对应像素块进行相应的再划分; 当  $SAD_{sum}$  值小于阈值  $T_2$  时, 则将各像素块的划分大小保留不变。

### 3.3.2 改进的双向运动补偿插值算法

改进的双向运动补偿差值算法描述如下:

(1) 将前向关键帧  $X_{2k-1}$  作为参考帧, 后向关键帧  $X_{2k+1}$  作为当前帧, 像素块  $B_{2k+1,x}$  作为当前块, 采用自适应十字形模式搜索算法(Adaptive Root Pattern Search, ARPS)<sup>[50]</sup>对  $X_{2k-1}$  整帧进行匹配块搜索。同时采用绝对误差值和函数  $SAD$

作为匹配准则, 找到最佳匹配块并命名为  $B_{2k-1,p1}$ , 根据搜索得到匹配块, 计算得到前向运动矢量  $MV_{f2k,x}$ 。

(2) 设边信息帧像素块  $B_{2k,x}$  相对于参考前关键帧  $X_{2k-1}$  的前向运动矢量为  $MV_{f2k,x}$ , 利用上述步骤(1)计算得到的前向运动矢量  $MV_{f2k,x}$ , 通过计算式(3-12)可得到  $MV_{ff2k,x}$  的值。同时根据前向运动矢量  $MV_{ff2k,x}$  的特点, 由计算式(3-12)得到前向关键帧  $X_{2k-1}$  相对于边信息帧的对应后向运动矢量  $MV_{fb2k,x}$ 。

$$MV_{ff2k,x} = -MV_{fb2k,x} = \frac{MV_{f2k,x}}{2} \quad (3-12)$$

由计算式(3-12)可知, 前向运动矢量  $MV_{ff2k,x}$  与  $MV_{f2k,x}$  方向相同, 模值大小为两倍关系; 而与后向运动矢量  $MV_{fb2k,x}$  具有方向相反, 模值相等的特点。

(3) 将边信息帧中像素块  $B_{2k,x}$  作为参考, 利用上述计算得到的前向运动矢量  $MV_{ff2k,x}$ , 在前向关键帧  $X_{2k-1}$  中找到最佳匹配块并命名为  $B_{2k-1,p2}$ , 如图 3.6 所示。

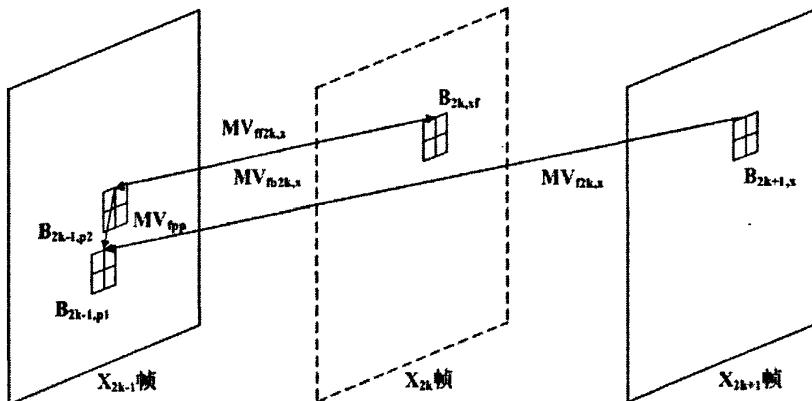


图 3.6 改进的前向运动补偿内插

(4) 以匹配像素块  $B_{2k-1,p2}$  为起点, 匹配像素块  $B_{2k-1,p1}$  为终点, 根据计算式(3-13)计算得到位移向量  $MV_{fpp}$ 。

$$MV_{fpp}(i, j) = (i_{2k-1,p2} - i_{2k-1,p1}, j_{2k-1,p2} - j_{2k-1,p1}) \quad (3-13)$$

(5) 分别以匹配块  $B_{2k-1,p2}$  和当前像素块  $B_{2k+1,x}$  为起点, 以像素块  $B_{2k-1,p1}$  为对应起点, 利用前向运动矢量  $MV_{f2k,x}$  和位移向量  $MV_{fpp}$ , 对边信息帧  $X_{2k}$  进行运动估计内插计算, 得到对应匹配像素块并命名为  $B_{2k,xf}$ , 计算方法如计算式(3-14),

其中,  $x$  表示其下标位移向量的横向分量,  $y$  表示其下标位移向量的纵向分量。

$$B_{2k,x,f}(i, j) = [B_{2k-1,p2}(i - \frac{x_{MV_{f2k,x}}}{2}, j - \frac{y_{MV_{f2k,x}}}{2}) \\ + B_{2k+1,x}(i + \frac{x_{MV_{f2k,x}}}{2} - \frac{x_{MV_{fp}}}{2}, j + \frac{y_{MV_{f2k,x}}}{2} - \frac{y_{MV_{fp}}}{2})] / 2 \quad (3-14)$$

(6) 将后向关键帧  $X_{2k+1}$  作为参考帧, 前向关键帧  $X_{2k-1}$  作为当前帧, 像素块  $B_{2k-1,x}$  为当前块, 利用 ARPS 搜索算法对  $X_{2k+1}$  整帧进行匹配块搜索, 找到最佳匹配块并命名为  $B_{2k+1,p1}$ , 根据所得匹配块, 计算得到后向运动矢量  $MV_{b2k,x}$ 。

(7) 设边信息帧对应像素块  $B_{2k,x}$  相对于参考后关键帧  $X_{2k+1}$  的后向运动矢量为  $MV_{bb2k,x}$ , 利用步骤(6)得到的后向运动矢量  $MV_{b2k,x}$ , 通过计算式(3-15)可得到  $MV_{bb2k,x}$  的值。同时根据后向运动矢量  $MV_{bb2k,x}$  的特点, 由计算式(3-15)可以得到后向关键帧  $X_{2k+1}$  相对于边信息帧的对应后向运动矢量  $MV_{bf2k,x}$ 。

$$MV_{bb2k,x} = -MV_{bf2k,x} = \frac{MV_{b2k,x}}{2} \quad (3-15)$$

从计算式(3-15)可知, 后向运动矢量  $MV_{bb2k,x}$  与  $MV_{b2k,x}$  方向相同, 模值为两倍关系; 而与前向运动矢量  $MV_{bf2k,x}$  具有方向相反, 模值相等的特点。

(8) 以边信息帧像素块  $B_{2k,x}$  为参考, 利用上述计算得到的后向运动矢量  $MV_{bb2k,x}$ , 在后关键帧  $X_{2k+1}$  中找到最佳匹配块并命名为  $B_{2k+1,p2}$ , 如图 3.7 所示。

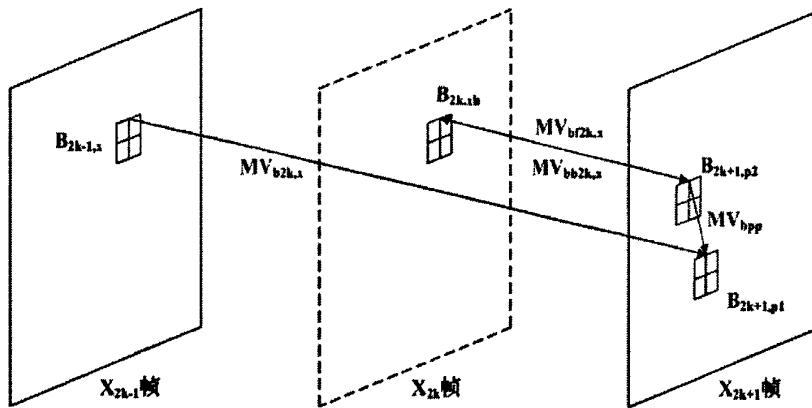


图 3.7 改进的后向运动补偿内插

(9) 以匹配像素块  $B_{2k+1,p2}$  为起点, 匹配像素块  $B_{2k+1,p1}$  为终点, 根据计算式(3-16)计算得到位移向量  $MV_{bpp}$ 。

$$MV_{b_{pp}}(i, j) = (i_{2k+1, p2} - i_{2k+1, p1}, j_{2k+1, p2} - j_{2k+1, p1}) \quad (3-16)$$

(10) 分别以匹配块  $B_{2k+1, p2}$  和当前像素块  $B_{2k-1, x}$  为起点, 以像素块  $B_{2k+1, p1}$  为对应起点, 利用后向运动矢量  $MV_{b_{2k, x}}$  和位移向量  $MV_{b_{pp}}$ , 对边信息帧  $X_{2k}$  进行运动估计内插计算, 得到对应匹配块并命名为  $B_{2k, xb}$ , 计算方法如计算式(3-17)。

$$B_{2k, xb}(i, j) = [B_{2k-1, x}(i + \frac{x_{MV_{b_{2k, x}}}}{2} - \frac{x_{MV_{b_{pp}}}}{2}, j + \frac{y_{MV_{b_{2k, x}}}}{2} - \frac{y_{MV_{b_{pp}}}}{2}) \\ + B_{2k+1, p2}(i - \frac{x_{MV_{b_{2k, x}}}}{2}, j - \frac{y_{MV_{b_{2k, x}}}}{2})] / 2 \quad (3-17)$$

### 3.3.3 加权判决

(1) 将 3.3.2 小节步骤(5)生成的像素块  $B_{2k, xf}$  与匹配块  $B_{2k-1, p2}$  采用绝对差值和函数 SAD 来进行相关性判决, 得到  $SAD_{f2k, x}$ 。

(2) 将 3.3.2 小节步骤(10)生成的像素块  $B_{2k, xb}$  与匹配块  $B_{2k+1, p2}$  采用绝对差值和函数值 SAD 来进行相关性判决, 得到  $SAD_{b2k, x}$ 。

(3) 利用上述求得的绝对差值  $SAD_{f2k, x}$  和  $SAD_{b2k, x}$ , 计算得到最终边信息帧中的内插估计像素块  $B_{2k, x}$ , 计算方法如计算式(3-18)。

$$B_{2k, x} = \frac{SAD_{b2k, x}^2}{SAD_{f2k, x}^2 + SAD_{b2k, x}^2} B_{2k, xf} + \frac{SAD_{f2k, x}^2}{SAD_{f2k, x}^2 + SAD_{b2k, x}^2} B_{2k, xb} \quad (3-18)$$

### 3.3.4 边信息帧生成

(1) 根据前后相邻关键帧中属于运动块类的各像素块对, 在边信息帧  $X_{2k}$  中找到对应下标编号相同的像素块, 将这些像素块按照算法步骤 3.3.1 至步骤 3.3.3 的描述, 分别对其进行计算, 最终得到边信息帧中对应的运动类像素块  $B_{2k, 1} \sim B_{2k, n}$ 。

(2) 根据前后相邻关键帧中属于保留块类的各像素块对, 在边信息帧  $X_{2k}$  中找到对应下标编号相同的像素块, 按照算法步骤 3.3.1 的描述, 将像素块  $B_{2k, x}$  与其对应下标编号的前像素块  $B_{2k-1, x}$  保持一致。

(3) 通过步骤 3.3.4(1)得到的属于运动块类的各像素块和步骤 3.3.4(2)中得到保留块类的像素块, 其下标编号与其对应的前后关键帧的像素块相同且唯一, 按照其下标编号进行顺序排列组合, 最终得到边信息帧  $Y_{2k}$ 。

### 3.4 仿真测试与分析

#### 3.4.1 仿真条件设置

为了验证所提边信息生成算法的性能, 以 Matlab 作为仿真平台, 分别对典型视频测试序列 Foreman、Container 的前 100 帧进行了仿真测试, 采用的仿真测试条件如下:

- (1) 设置奇数帧为关键帧 K, 偶数帧为 WZ 帧, 且假设关键帧 K 能不失真地传输到解码端;
- (2) 序列格式均为 QCIF( $176 \times 144$ ), 速率为 30fps, 实验中只计算 WZ 帧, 因而 WZ 帧的帧率为 15fps;
- (3) 设置 GOP 的长度设置为 2;
- (4) 采用与基于变换域的 WZ 视频编码相同的系统设置: K 帧采用传统的 H.264 Intra 方式<sup>[51]</sup>进行编解码;
- (5) 实验中对于运动剧烈的 Foreman 序列, 设定像素块分类阈值  $T1=2000$ , 像素块再分判决阈值  $T2=8000$ ; 对于运动变化缓慢的 Container 序列, 设定像素块分类阈值  $T1=200$ , 像素块再分判定阈值  $T2=800$ 。

为了更明确的显示算法的有效性, 在相同的关键帧恢复质量下生成边信息图像, 使用 WZ 帧的亮度信号平均峰值信噪比 PSNR(单位: dB)作为衡量边信息质量的测度。PSNR 计算公式如计算式(3-19)。其中, MSE 为原始帧与边信息帧之间的均方误差函数, 如计算式(3-20)所示。

$$PSNR = 10 \times \log \frac{255^2}{MSE(i, j)} \quad (3-19)$$

$$MSE = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} |f(x, y) - f'(x, y)|^2}{M \times N} \quad (3-20)$$

#### 3.4.2 测试与分析

将论文所提的边信息生成算法与传统运动补偿内插算法(MCI)、运动补偿外推算法(MCE)、基于图像纹理信息的边信息改进算法<sup>[52]</sup>进行仿真比较, 测试结果

如图 3.8-3.9 所示

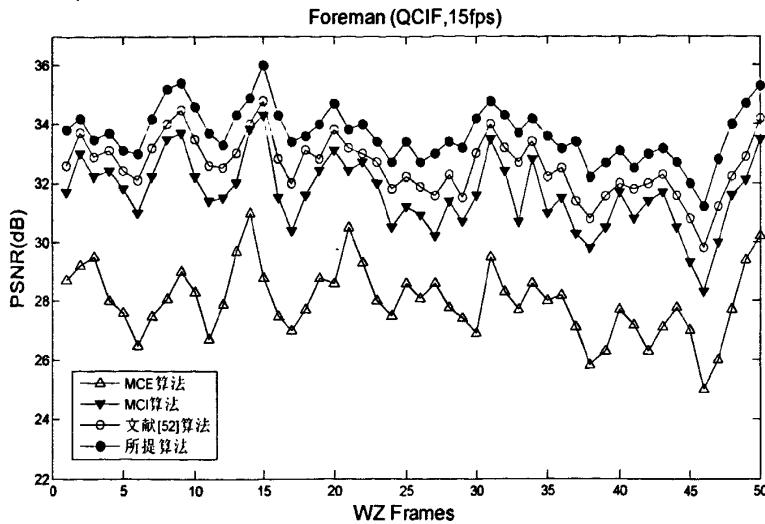


图 3.8 Foreman 序列边信息帧 PSNR 对比图

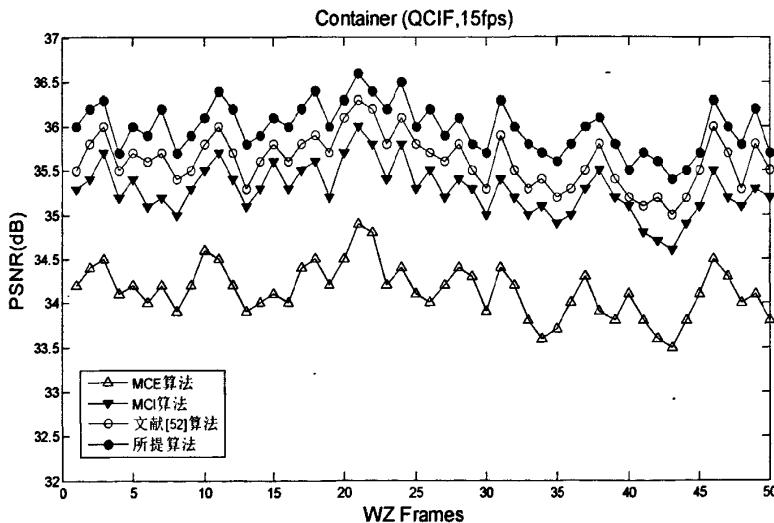


图 3.9 Container 序列边信息帧 PSNR 对比图

从图 3.8 和图 3.9 可以看出, 由于 MCE 算法仅根据前一时刻的运动特性来预测当前时刻的运动情况, 忽略了当前时刻与后一时刻的相关特性, 使得 MCE 算法所得边信息的 PSNR 平均值最低。MCI 算法、基于图像纹理信息的边信息改进算法和本文所提出算法所得边信息的 PSNR 值较为接近, 这是由于 MCI 算法根据前后两相邻时刻的运动特性对当前运动矢量进行匀速线性运动假设, 相比 MCE 算法获得性能更佳, 但对于局部细节变化, 预测帧质量较为粗糙, 与真实帧仍然存在差异。图 3.8 中的 PSNR 曲线波动相比图 3.9 更加严重, 这是由于

Foreman 视频序列中, 像素块的运动变化程度更为剧烈, 包括人物的移动和面部表情的不均匀变化, 使得预测帧与前后两相邻关键帧的差异较明显。由于 Container 序列中只存在船只和旗帜的缓慢移动, 帧间变化差距较小, 使得图 3.9 中的 PSNR 相比图 3.8 整体有一定的提高。

对于 Foreman 序列和 Container 序列, 所提算法得到的边信息 PSNR 值, 相比传统的 MCI 算法平均提高了约 2.01dB 和 1.41dB, 相比基于图像纹理信息的边信息改进算法平均提高了约 1.64dB 和 1.08dB。分析性能提高原因, 所提算法采用对像素块进行分类的思想, 对变化程度不大的像素块进行保留处理, 简化复杂度, 对变化程度较大的像素块采用改进的双向运动补偿内插以及加权判决, 深入挖掘相关信息, 使得边信息的质量提高。

由于对保留块采用直接替代处理, 忽略了部分影响较小的运动变化, 对运动块进行较复杂的运动估计, 获得较真实的运动矢量, 导致对于运动变化缓慢的 Container 序列, 所提算法得到的边信息 PSNR 值与采用其他算法得到的边信息 PSNR 值较为接近, 对于运动变化较剧烈的 Foreman 序列, 边信息 PSNR 值的提升则更为明显。由此验证, 所提的块分类的双向估计加权边信息生成算法对运动变化较大的视频序列的性能提升更大。

### 3.5 本章小结

本章首先介绍了边信息在分布式视频编码系统中的作用, 通过对典型边信息生成算法以及几种改进算法的分析, 针对边信息预测带来的预测重叠问题, 提出了一种块分类的双向运动估计加权边信息生成算法。在对提出算法实现步骤进行了详细描述后, 在 Matlab 平台上对测试视频序列进行仿真, 并与典型算法进行对比与分析, 验证所提算法能有效提高生成的边信息质量。

## 第四章 相关性分类的解码检测优化算法设计

在运动复杂和细节变化丰富的情况下,生成的边信息将与原始视频帧存在较大差距,导致得到的解码帧质量降低。针对如何实现自动检测解码帧质量,并对低质量的解码帧进行优化以提高系统性能的问题,提出一种相关性分类的解码检测优化算法。

### 4.1 算法设计思想

#### 4.1.1 问题分析

在物体运动剧烈变化时,解码端得到的边信息失真严重,采用向编码端请求更多校验码方式进行反复解码,能够一定程度减小解码失真。但是当系统获得的解码效果常与原始信息存在较大差异,甚至出现局部像素块失真严重的情况时,采用反馈请求方式将严重影响系统编解码速率,解码质量优化效果不高。而对于人工检测方式,一般需要用户在视频编解码系统完成解码流程之后,通过对输出的解码结果进行阅览与校对,才能对系统解码效果进行判定。这种检测方式需要用户对解码结果进行实时检测,并对解码错误进行及时纠错操作,不仅提高了系统错误概率,同时增加了系统纠错的时延,导致系统的错误反馈机制效率较低。如何实现系统对解码质量高效检测,并对失真度较高的误码进行优化,成为当前一个急需解决的问题。

在分布式视频编码系统中,解码帧的失真度很大程度上取决于边信息的生成质量和运动估计算法。为了提高系统解码质量,针对失真度较高的解码帧,对已生成的边信息质量进行优化成为目前分布式视频编码技术的研究热点之一。文献[53]通过在 DC 域中进行运动估计来对生成边信息实现进一步修正;文献[54]提出在解码端采用连续 Turbo 解码方式,根据时间和空间相关性对边信息进行重建,从而减少解码所需的校验码;文献[55]提出在解码端对生成边信息的性能进行运动补偿时域内插分析,根据得到的运动矢量对 WZ 帧进行重复解码以提高边信息准确性;在较高位比特面解码后,文献[56]对边信息修正的方法进行设计,利用解码值纠正出错系数,从而增强边信息与信源的相关性;以去噪声为目的,

文献[57]提出基于多假设运动补偿预测的方法对边信息质量进行优化,对传统方式生成的原始边信息,进行多假设运动补偿产生补偿块,并采用线性组合方式生成质量较高的边信息。文献[58]充分利用初始边信息所包含的丰富信息,提出一种利用图像颜色及纹理信息的边信息修正重建方案,根据运动信息、颜色和纹理,将初始边信息进行划分,并对可信度低的区域采用三维递归进行修正,从而实现优化边信息质量的效果。由于在生成边信息帧中,各个像素块的失真度不一致,因此可以针对不同失真情况进行对应处理,对失真较小的区域可以采取保留操作,降低复杂度;对失真度大的像素块进行重建优化处理,改善边信息整体质量,优化系统解码性能。

#### 4.1.2 算法介绍

在分布式视频编码中,针对系统解码帧质量检测方式,以及在解码帧部分像素块与原始视频帧存在较大失真时,如何进行解码帧优化以提高分布式编码系统效率的问题,提出了一种相关性分类的解码检测优化算法。

在编码端和解码端,分别对原始帧、已解码帧以及对应前后两相邻关键帧进行绝对差值计算,得到原始相关性信息与解码相关性信息,根据获得的相关性数据进行置信度计算,得到解码帧质量的判定阈值。对已解码帧中各像素块进行生成质量判定并将像素块分类:将失真较小的像素块进行保留;对质量较差的像素块,采用边信息重构算法对解码帧进行优化处理。在进行重构边信息像素块时,充分利用解码端计算能力强、资源不受限的优势,采用复杂计算对像素块进行较准确的运动估计,得到对应像素块的优化值。最后根据下标值进行优化像素块与保留像素块的排列组合,最终得到重建后的边信息,并利用重建边信息对当前帧进行再次解码,获得质量优化后的解码帧。相关性分类的解码检测优化算法框图如图 4.1 所示。

算法充分利用了编码端和解码端的不同特性,将编码前的原始 WZ 帧( $X_{2k}$ 帧)与前后相邻关键帧( $X_{2k\pm1}$ 帧)之间的相关度,同解码帧与其相邻已解码关键帧之间的相关度进行差值比较,从而获得检测解码帧质量的参考信息。为了提高重构质量,对于需要优化的像素块进行较复杂的边信息重构处理,对于质量较高的

像素块，采用保留方式，有效节约能耗的使用，提高解码端能耗的使用效率，将资源进行更有效的分配使用。

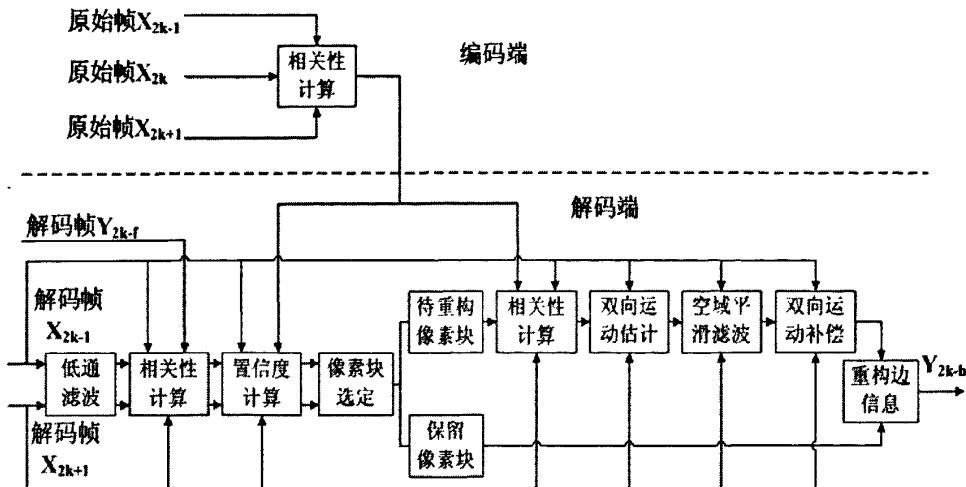


图 4.1 相关性分类的解码检测优化算法框图

## 4.2 算法实现

### 4.2.1 置信度计算

置信度计算需要在编码端和解码端分别进行相关性计算，并将编解码两端分别计算得到的相关值进行比较，得到解码信息与原始信息的差值。根据这一差值计算置信度，并对解码端的解码质量进行判断，具体步骤如下。

在编码端：

- (1) 在编码之前，将视频序列进行奇偶分类，其中奇数帧为关键帧( $X_{2k\pm 1}$ 帧)，而偶数帧为 WZ 帧( $X_{2k}$ 帧)。
- (2) 为了均衡算法的精确性和复杂度，采用  $8 \times 8$  的格式大小将前后相邻关键帧分为若干大小相同、互不重叠的像素块。分块后将各像素块进行编号，如  $B_{2k,x}$  即为第  $2k$  个视频序列帧中编号为  $x$  的像素块。根据编号，可以判断相邻像素帧的对应像素块的位置是否正确。
- (3) 分别计算每一个 WZ 帧( $X_{2k}$ 帧)像素块与其前一相邻关键帧  $X_{2k-1}$  中对应像素块的前向相关值  $SAD_{fl,x}$ ，以及与其后一相邻关键帧  $X_{2k+1}$  中对应像素块的后向相关值为  $SAD_{bl,x}$ ，计算方法如计算式(4-1)和(4-2)。

$$SAD_{f1,x} = \sum_{i=1}^n \sum_{j=1}^m |B_{2k,x}(i,j) - B_{2k-1,x}(i,j)| \quad (4-1)$$

$$SAD_{b1,x} = \sum_{i=1}^n \sum_{j=1}^m |B_{2k+1,x}(i,j) - B_{2k,x}(i,j)| \quad (4-2)$$

其中,  $n$  和  $m$  分别表示像素块的长度和宽度,  $B(i,j)$  为像素块的灰度值, 相关值下标 1 表明该相关值由编码端计算得到, 与之后在解码端计算得到的相关值相区别。

(4) 将计算得到的相关值  $SAD_{f1,x}$  和  $SAD_{b1,x}$  进行量化编码后发送到解码端。解码端进行解码时, 将这些数据将作为边信息质量检测与优化的辅助信息。

在解码端:

(1) 将解码帧标记为  $X_{2k-f}$ , 采用下标-f 表示是检测优化处理前得到的解码帧。采用分块思想, 将解码帧与前后两相邻已解码关键帧, 均采用  $8 \times 8$  的分块处理, 与编码端一样, 采用下标 x 表明对应的坐标位置。

(2) 在前后两相邻已解码关键帧中, 根据下标编号, 分别找到与解码帧像素块  $B_{2k,x}$  相对应的像素块  $B_{2k-1,x}$  和  $B_{2k+1,x}$ 。按计算式(4-3)(4-4)分别进行相关值计算, 得到解码帧  $X_{2k-f}$  与前解码关键帧  $X_{2k-1}$  对应像素块的前向相关值  $SAD_{f2,x}$ , 以及与后解码关键帧  $X_{2k+1}$  对应像素块的后向相关值  $SAD_{b2,x}$ 。其中下标 2 表明该相关值是由解码端计算得到, 与之前由编码端计算得到的相关值相区别。

$$SAD_{f2,x} = \sum_{i=1}^n \sum_{j=1}^m |B_{2k,x}(i,j) - B_{2k-1,x}(i,j)| \quad (4-3)$$

$$SAD_{b2,x} = \sum_{i=1}^n \sum_{j=1}^m |B_{2k+1,x}(i,j) - B_{2k,x}(i,j)| \quad (4-4)$$

(3) 将编码端得到的前向相关值  $SAD_{f1,x}$  与解码端得到的前向相关值  $SAD_{f2,x}$ , 按计算式(4-5)进行差值计算, 将其差值定义为前向置信度  $S_1$ ; 将编码端得到的后向相关值  $SAD_{b1,x}$  与解码端得到的后向相关值  $SAD_{b2,x}$ , 按计算式(4-6)进行差值计算, 将其差值定义为后向置信度  $S_2$ 。

对于每个像素块来说, 通过计算式(4-5)(4-6)得到的差值越小, 说明得到的解

码帧与原始帧间的失真越小，对应的边信息像素块质量越高；反之亦然。

$$S1_x = \sum_{x=1}^N |SAD_{f1,x} - SAD_{f2,x}| \quad (4-5)$$

$$S2_x = \sum_{x=1}^N |SAD_{b1,x} - SAD_{b2,x}| \quad (4-6)$$

(4) 将前向信息置信度  $S1$  和后向信息置信度  $S2$  进行平均计算，得到解码帧中对应像素块的置信度  $S_x$ ，如计算式(4-7)。

$$S_x = \frac{(S1_x + S2_x)}{2} \quad (4-7)$$

如果解码帧  $X_{2k-f}$  与原始帧  $X_{2k}$  相差较大，得到的前向相关值和后向相关值都将随之较大，从而导致最后得到的边信息置信度增大；反之亦然。

#### 4.2.2 像素块的选定

为了判定解码帧中各像素块的解码质量，需要定义一个阈值  $T$ ，将它与 4.2.1 小节计算得到的置信度  $S_x$  进行比较的。阈值  $T$  的选择是所提算法的一个关键，如果将阈值  $T$  设置过低，则像素块质量超过  $T$  的概率将过大，使得需要进行边信息重建的像素块个数增多，增加了系统复杂度；如果将阈值  $T$  设置过高，则像素块质量超过  $T$  的概率将过低，失去了检测解码质量的作用性。为了均衡系统复杂度与性能，算法根据解码帧的整体质量自动地调整阈值。阈值  $T$  的确定通过计算各像素块置信度的均值而获得，计算方法如计算式(4-8)。

$$T = \sum_{x=1}^N S_x / N \quad (4-8)$$

其中， $n$  为整个解码帧中像素块的数目。由上式可知，如果解码帧的平均质量越差，则计算得到的置信度值越高，从而导致阈值的设定也随之提高。对于解码帧  $X_{2k-f}$  的各像素块，通过计算式(4-7)可以得到对应置信度值  $S_x$ ，将其分别与阈值  $T$  进行比较判决：

- (1) 当  $S_x < T$  时，表明像素块  $B_{2k,x}$  解码质量较好，与原始像素块的失真较小，对其进行保留操作，不再进行解码优化操作；
- (2) 当  $S_x \geq T$  时，表明像素块  $B_{2k,x}$  解码质量较差，与原始像素块的失真较

大。由于较大失真的存在，需要采用以下描述的边信息重构算法对其像素块进行优化，以降低其失真度，提高解码质量。

#### 4.2.3 前/后相关性判定

通常运动估计将运动变化简化成在较短时间内直线运动，但是很多时候，物体的运动变化与估计的直线轨迹不相符，且当前变化与前后的相关性不同，这是边信息存在失真的原因之一。为了减小边信息失真度，算法引入了解码帧与前后相邻关键帧的相关性判定方法，通过比较前后相关值，计算出前后相关性加权值，从而达到优化边信息质量的目的。

由 4.2.1 小节描述可知，由编码端计算得到的前向相关值  $SAD_{f1,x}$  和后向相关值  $SAD_{b1,x}$ ，通过信道传输至解码端。当解码端进行边信息重构时，根据相关值  $SAD_{f1,x}$  和  $SAD_{b1,x}$ ，设前向运动估计加权值  $\alpha$  和后向运动估计加权值  $\beta$ ，并分别按以下情况进行简单计算得到对应的加权值。

(1) 当  $\left| \frac{SAD_{f1,x}}{SAD_{b1,x}} \right| = 1$  时，表明原始帧中像素块  $B_{2k,x}$  分别与前后对应像素块的

相关程度相同，因此取前向加权值  $\alpha$  和后向加权值  $\beta$  为  $1/2$ ；

(2) 当  $1/4 \leq \left| \frac{SAD_{f1,x}}{SAD_{b1,x}} \right| \leq 1$  时，表明前向相关值小于或等于后向相关值，即

原始帧像素块  $B_{2k,x}$  与对应前向像素块  $B_{2k-1,x}$  的差值较小，与前向像素块  $B_{2k-1,x}$  的相关程度更大。前向加权值  $\alpha$  与后向加权值  $\beta$  分别按计算式(4-9)和(4-10)取得。

$$\alpha = \frac{SAD_{f1,x}}{SAD_{f1,x} + SAD_{b1,x}} \quad (4-9)$$

$$\beta = 1 - \alpha \quad (4-10)$$

(3) 当  $1 \leq \left| \frac{SAD_{f1,x}}{SAD_{b1,x}} \right| \leq 4$  时，表明前向相关值大于或等于后向相关值，即原

始帧像素块  $B_{2k,x}$  与对应后向像素块  $B_{2k+1,x}$  的差值较小，与后向像素块  $B_{2k+1,x}$  的相关程度更大。前向加权值  $\alpha$  与后向加权值  $\beta$  分别按计算式(4-9)和(4-10)计算得到。

(4) 当  $0 \leq \left| \frac{SAD_{f1,x}}{SAD_{b1,x}} \right| \leq 1/4$  时, 表明原始帧的像素块  $B_{2k,x}$  与对应前向像素块  $B_{2k-1,x}$  的差值相比对应后向像素块  $B_{2k+1,x}$  的差值要小得多, 与对应前向像素块  $B_{2k-1,x}$  的相关性很大, 将与对应后向像素块  $B_{2k+1,x}$  的相关性忽略, 前向加权值  $\alpha$  取 1, 后向加权值  $\beta$  取 0。

(5) 当  $4 \leq \left| \frac{SAD_{f1,x}}{SAD_{b1,x}} \right|$  时, 表明原始帧像素块  $B_{2k,x}$  与对应后向像素块  $B_{2k+1,x}$  的差值相比对应前向像素块  $B_{2k-1,x}$  的差值要小得多, 与对应后向像素块  $B_{2k+1,x}$  的相关性很大, 将与对应前向像素块  $B_{2k+1,x}$  的相关性略掉, 前向加权值  $\alpha$  取 0, 后向加权值  $\beta$  取 1。

#### 4.2.4 双向加权运动估计

采用双向运动估计算法分别在前后两个相邻关键帧  $X_{2k\pm 1}$  中进行匹配块搜索。为了后续算法进行优化像素块的质量估计, 将匹配块的搜索框大小从  $8 \times 8$  扩展至  $12 \times 12$ , 使得相临近的匹配块出现重叠区域。

(1) 将后关键帧像素块  $B_{2k+1,x}$  作为当前块, 对前关键帧  $X_{2k-1}$  进行前向匹配块搜索, 找到的最优匹配块定义为  $B_{12k-1,x}$ 。匹配算法采用改进的匹配判定准则作为检测标准, 其计算式如(4-11)和(4-12)所示。

$$MAD(x, y) = \frac{1}{N \times M} \sum_{i=1}^N \sum_{j=1}^M |f_{2k+1}(i, j) - f_{2k-1}(i+x, j+y)| \quad (4-11)$$

$$CF(x, y) = \sqrt{x^2 + y^2} \times MAD(x, y) \quad (4-12)$$

计算式(4-11)为平均绝对误差函数, 其中  $(x, y)$  是位移矢量,  $f_{k+1}$  和  $f_{k-1}$  分别是前后相邻已解码关键帧的亮度值,  $N$  与  $M$  分别是像素块的长度与宽度。根据各个点的坐标值计算得到的 MAD 值, 以及相应的运动矢量 MV。

利用平均绝对误差函数值, 以及匹配的运动矢量 MV, 定义搜索像素块的代价函数, 如计算式(4-12)所示。如果在像素帧  $B_{2k-1,x}$  中某点坐标位置  $(i_n, j_n)$  使得代价函数  $CF(x_n, y_n)$  值最小, 则以该点为坐标中心的像素块即为最优匹配块, 对应的前向运动矢量为  $MV_f = (x_n, y_n)$ 。

(2) 将前关键帧  $X_{2k+1}$  中的像素块  $B_{2k+1,x}$  作为当前块, 对后关键帧  $X_{2k-1}$  进

行后向匹配块搜索, 找到的最优匹配块定义为  $B1_{2k+1,x}$ 。同样采用计算式(4-11)和(4-12)进行匹配块判定, 得到匹配的后向运动矢量  $MV_b$ , 其双向运动估计如图 4.2 所示。

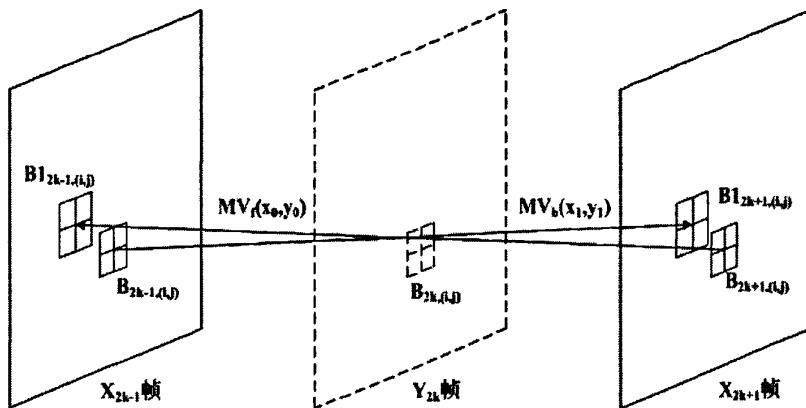


图 4.2 双向运动估计

(3) 将上述计算获得的匹配块  $B1_{2k-1,x}$  与  $B1_{2k+1,x}$ , 利用 4.2.3 小节计算得到的前向加权系数  $\alpha$  和后向加权系数  $\beta$ , 按计算式(4-13)进行加权计算, 得到重建边信息帧中对应的预测像素块  $B1_{2k,x}$ 。

$$B1_{2k,x} = \sum_{i=1}^n \sum_{j=1}^m (i + \alpha i_{2k-1,x} + \beta i_{2k+1,x}, j + \alpha j_{2k-1,x} + \beta j_{2k+1,x}) \quad (4-13)$$

利用上述得到的前向运动矢量  $MV_f$ 、后向运动矢量  $MV_b$  以及加权系数, 按计算式(4-14)计算获得最优运动矢量  $MV$ 。

$$MV = \alpha MV_f + \beta MV_b \quad (4-14)$$

通过采用前向加权系数与后向加权系数, 在对像素块以及对应运动矢量进行估计的过程中, 加入了与之相邻的前向相关值与后向相关值的影响因素, 使得估计像素值更加接近于真实情况, 起到提高重建边信息帧质量的作用。

#### 4.2.5 空域平滑滤波

采用双向运动估计时, 由于受图像噪声、基于分块的运动估计等因素的影响, 得到的运动矢量场容易陷入局部最优, 导致空域上出现不连贯的现象, 部分像素块的预测运动矢量与实际存在偏差, 甚至会出现与相邻像素块的运动矢量不一致的情况, 如运动矢量的方向不同、运动矢量的大小相差很大等, 造成边信息中块

效应的出现, 如图 4.3 所示。

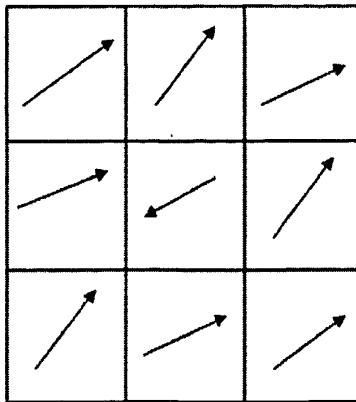


图 4.3 运动矢量不一致性

算法采用自适应加权矢量中值滤波器(Adaptively weighted vector-median filters)<sup>[59]</sup>对待优化像素块的对应运动矢量进行平滑处理。根据像素块和运动矢量的局部空间相关性, 对偏差较大的运动矢量进行滤波以减少错误数, 使运动矢量场变得平滑。为了控制平滑滤波的幅度, 根据当前像素块及周围八个相邻像素块的运动矢量残差值来确定加权系数。

根据空域平滑算法的中值滤波器的定义, 分别对当前像素块的运动矢量  $MV_1$  以及相邻像素块对应运动矢量  $MV_i$  按计算式(4-15)进行运动矢量加权和值计算。

$$\sum_{i=1}^N W_i \|MV_j - MV_i\| \quad (4-15)$$

其中,  $W_1, \dots, W_N$  是对应的自适应加权系数, 选择运动矢量加权和值最小的运动矢量  $MV_j$  作为当前像素块的运动矢量  $MV_{best}$ 。加权系数的计算方法如计算式(4-16)。

$$W_i = \frac{MSE(MV_1, B)}{MSE(MV_i, B)} \quad (4-16)$$

其中,  $MV_1$  表示当前像素块对应的运动矢量, 均方误差  $MSE(MV_i, B)$  表示前后关键帧中, 采用运动矢量  $MV_i$  进行运动补偿得到的匹配块与对应块之间的绝对差值。当  $MSE$  值越大, 说明运动矢量的双向运动矢量差异程度越大, 得到的加权系数将相应地降低, 如此对运动矢量滤波输出的影响随之减小, 反之亦然。

#### 4.2.6 加权双向运动补偿

在进行了前/后相关性判定、双向加权运动估计以及空域平滑滤波后，获得了最终的运动矢量场，重建边信息帧中待优化像素块都得到一个对应的运动矢量。由于双向运动估计内插是以像素块中心进行的，因此像素块中心位置的内插值相比块边界部分更加准确。为了提高块边界内插值质量，减小重叠块运动补偿带来的边界衰弱问题，利用最终运动矢量场和加权双向运动补偿，对 4.2.4 小节生成的各像素块初值进行优化处理，即针对不同重叠区域内的像素值计算出相应的加权系数。

由于对像素块的匹配搜索框进行了扩大处理，导致边信息帧中各相邻内插像素块产生重叠区域，如图 4.4 所示。根据相邻像素块的重叠情况，可将运动补偿的像素块划分为 A、B、C 三类区域。对于属于这三类不同区域中的像素点，分别采用以下算法进行计算。

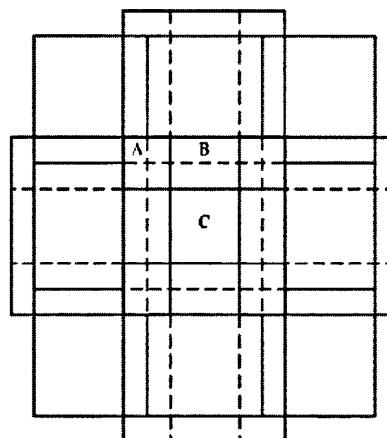


图 4.4 运动估计内插的像素块重叠

(1) A 类区域由 4 个相邻像素块互相重叠组成，对于处于 A 类区域的像素点，按计算式(4-17)进行计算。

$$B_{2k}(p) = \sum_{i=1}^4 w_{ai} [\alpha B_{2k-1}(p - mv_i) + \beta B_{2k+1}(p + mv_i)] / 4 \quad (4-17)$$

(2) B 类区域由 2 个相邻块互相重叠组成，对于处于 B 类区域的像素点，按计算式(4-18)进行计算。

$$B_{2k}(p) = \sum_{i=1}^2 w_{bi} [\alpha B_{2k-1}(p - mv_i) + \beta B_{2k+1}(p + mv_i)] / 2 \quad (4-18)$$

(3) C 类区域内像素块不存在重叠, 只存在所在相同中心坐标的像素块。对于处于 C 类区域的像素点, 按计算式(4-19)进行计算。

$$B_{2k}(p) = \alpha B_{2k-1}(p - mv_i) + \beta B_{2k+1}(p + mv_i) \quad (4-19)$$

计算式(4-17)、(4-18)和(4-19)中, 权值 $\alpha$ 和 $\beta$ 分别采用由 4.2.3 小节中求得的权值。计算式(4-17)和(4-18)中的加权系数  $w_i$  的取值方法如下: 对于处于当前内插像素块的重叠像素点, 包含的信息量对当前运动补偿影响较大, 设置  $w_i$  值设置较高; 处于当前内插像素块的相邻块的重叠像素点, 包含的信息量对当前运动补偿影响相对较小, 设置  $w_i$  值相对较低, 且加权系数值应满足计算式(4-20)。算法加权系数取值如下:  $w_{a1}=w_{a2}=w_{a3}=2/9$ ,  $w_{a4}=1/3$ ;  $w_{b1}=1/3$ ,  $w_{b2}=2/3$ 。

$$\sum_{i=1}^N w_i = 1 \quad (4-20)$$

由于当前像素块与前后相邻关键帧对应像素块的相关性不同, 算法在传统双向运动补偿内插算法的基础上, 加入了前向相关加权值与后向相关加权值, 深入挖掘出相邻像素块之间的相关性, 使得像素值更接近于原始值。

#### 4.2.7 解码优化

通过以上算法步骤, 得到重建边信息帧中待优化的各像素块的估计值  $B_{2k,x}$ 。重建边信息帧存在两类像素块, 一类为保留像素块, 因生成质量较高进行保留处理; 另一类为待优化像素块, 采用上述算法进行像素块重建, 以提高生成质量。每个像素块  $B_{2k,x}$  的下标编号表示对应所在的位置, 根据两类像素块的下标编号, 进行像素块的排列组合, 得到的像素帧即为重建后的边信息帧。

将重建边信息替代原始边信息送入系统解码器, 参与系统的解码过程, 获得优化解码帧  $X_{2k-b}$ , 采用下标-b 表示是通过检测优化处理后得到的解码帧。通过对与原始帧偏差较大的像素块进行优化处理, 提高了边信息质量, 解码帧失真度随之得到优化。

### 4.3 仿真测试与分析

#### 4.3.1 仿真条件设置

为了验证所提算法性能, 在 Matlab 仿真平台上分别对典型视频测试序列 Foreman、Container 进行仿真测试, 实验采用了以下测试条件:

- (1) 序列格式均为 QCIF( $176 \times 144$ ), 速率为 30fps;
- (2) 设置奇数帧为关键帧(K 帧), 偶数帧为 WZ 帧, 且假设关键帧 K 帧能不失真地传输到解码端;
- (3) 实验中只计算 WZ 帧, WZ 帧的帧率为 15fps, GOP 的长度为 2;
- (4) 采用与变换域 WZ 视频编码一样的系统设置: K 帧采用传统的 H.264 Intra 编码方式进行编码;
- (5) 为了明确显示算法有效性, 采用 WZ 帧的亮度信号的平均峰值信噪比 PSNR 作为衡量边信息质量的测度。

#### 4.3.2 测试与分析

本章所提算法是针对初始边信息下获得的解码帧进行质量检测与优化, 因此将本章提出的相关性分类解码检测优化算法与第三章提出的边信息生成算法进行测试比较, 仿真结果如图 4.5-4.6 所示。

从图 4.5 可以看出, 对于运动变化较剧烈的 Foreman 序列, 采用解码检测算法得到的重建边信息与第三章所提算法得到的初始边信息相比, 优化边信息的 PSNR 性能平均提高 1.45dB。分析原因, 首先是利用原始帧以及解码帧的相关信息, 通过进行解码帧的置信度计算, 得到初始边信息像素块与原始像素块的差值, 并对解码帧的失真度进行判断。对于失真较大的像素块, 采用边信息优化算法进行前后相关性计算得到双向运动估计的加权值, 利用空域平滑滤波减小运动矢量场的失真, 最后通过加权双向运动补偿得到优化的边信息像素块, 并利用重建边信息进行再次解码, 以提高解码帧质量。

从图 4.6 可以看出, 对于运动变化较小的 Container 序列, 采用解码检测算法得到的重建边信息与第三章所提算法相比, 优化边信息的 PSNR 性能平均提高 0.49dB, 提升程度较小。由于初始边信息生成算法对运动变化不大的视频序列,

能够充分利用其前后相关性，使其得到的初始边信息与原始帧之间的差异不大，获得的边信息质量较好。因此，相比运动剧烈的序列需要进行失真优化的像素块数量减少，使得优化效果没有运动变化剧烈的视频序列显著，PSNR 性能提升幅度有限。

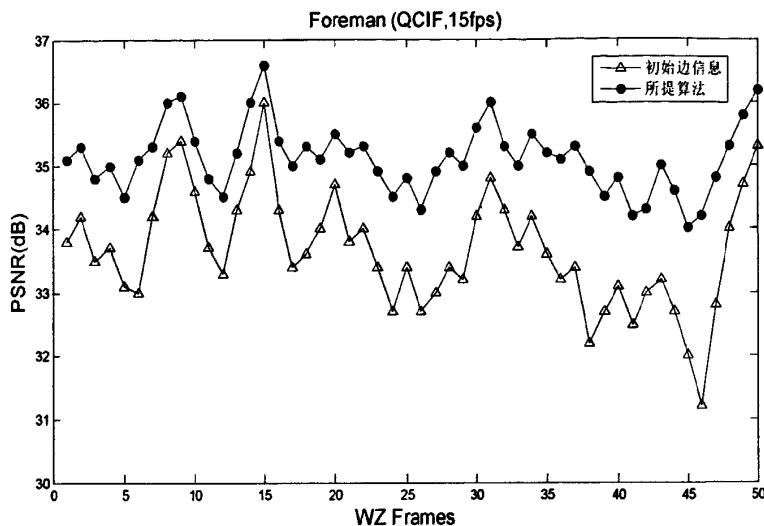


图 4.5 Foreman 序列边信息帧 PSNR 对比图

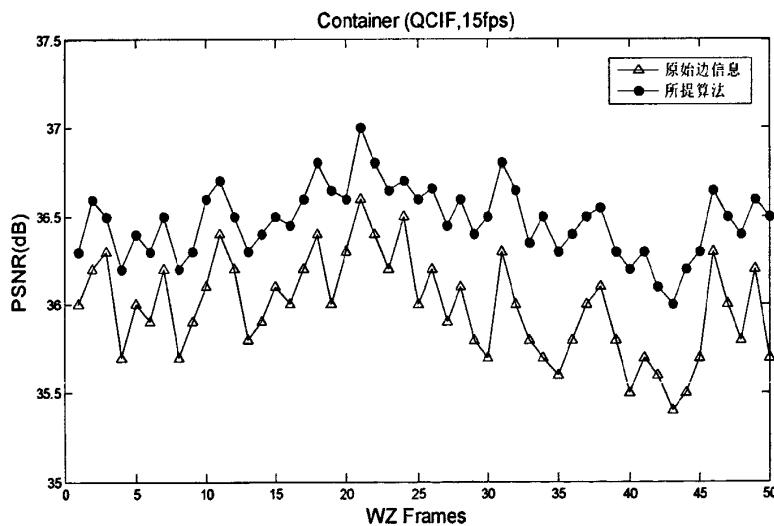


图 4.6 Container 序列边信息帧 PSNR 对比图

为了更直观对解码检测优化算法进行分析，分别对 Foreman 测试视频序列前 100 帧进行测试，以下测试结果中第一列图像为原始帧图像，第二列图像是采用块分类的双向估计加权初始边信息算法所得的解码帧图像，第三列图像为经过前后相关性分类解码检测及边信息优化算法得到的解码图像，如图 4.7-4.9 所示。

从三组测试解码帧结果对比可以看出,解码帧质量与运动变化剧烈程度相关,运动变化较小时,解码帧质量较好;运动变化剧烈时,解码帧质量随之降低,与原始帧存在较大失真。仅采用块分类的双向估计加权边信息得到的系统解码帧,在运动变化小的情况下,对于相邻帧间较小的变化,能够进行很好的帧间估计,得到质量较好的解码帧;而随着运动变化加剧,其解码质量变差,对于人物细节变化的解码存在失真,尤其在运动剧烈时,人物眼睛的转动、说话时的嘴部闭合以及帽子的位移等局部变化,不能进行有效的解码还原,存在较大失真。

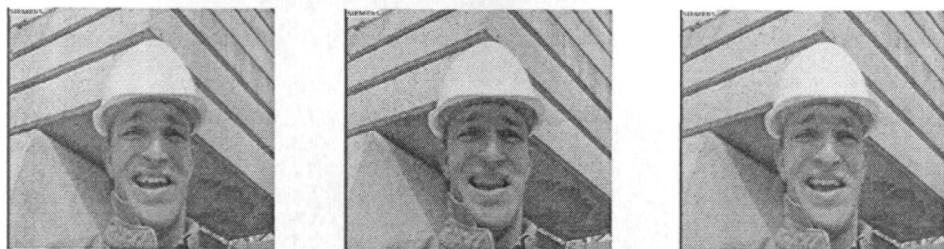


图 4.7 运动变化较小的解码帧结果



图 4.8 运动变化中等的解码帧结果



图 4.9 运动变化较大的解码帧结果

在采用初始边信息解码帧的基础上,采用前后相关性分类估计边信息优化算法将解码帧与原始帧进行对比,将解码质量较好的像素块保留,针对失真较大的像素块进行优化处理。经过本文所提边信息优化后,对于运动变化小的视频,由于仅对较少部分进行优化,改进效果不显著,而对于运动剧烈的视频,主要针对解码错误较大的部分进行纠正,使得失真明显的人物表情、图像边缘位移部分的

解码得到较大的改善。

#### 4.4 本章小结

本章对生成边信息中部分低质量的像素块的优化,以及系统检测解码质量问题进行了分析,提出一种相关性分类的解码检测优化算法。着重对算法的实现步骤进行了详细描述,对运动变化剧烈和缓慢的视频序列分别进行测试,并将测试结果与第三章测试结果进行比较与分析。仿真测试结果表明所提算法对运动剧烈的视频有较明显的检测与优化作用。

## 第五章 面向 WMSNs 的分布式视频编解码系统

本章针对无线多媒体传感器网络的结构特点以及应用要求,在结合上两章所提边信息生成算法和解码检测优化算法的基础上,提出一种面向无线多媒体传感器网络的分布式视频编解码系统实现方案,并对工作流程和各功能模块进行详细描述,最后通过仿真验证系统性能。

### 5.1 无线多媒体传感器网络分析

#### 5.1.1 结构特点

根据无线多媒体传感器网络中采用的传感器节点类型是否相同,分为同构 WMSNs<sup>[60]</sup>和异构 WMSNs<sup>[61]</sup>。为了降低网络能耗、保证信息质量、延长网络生存时间,研究人员一般采用异构 WMSNs 的网络结构。异构 WMSNs 主要由普通传感器节点、视频采集节点、汇聚节点、传输网络、控制中心构成,其网络结构如图 5.1 所示。在 WMSNs 中,大量普通传感器节点和视频采集节点被随机布设在指定的监测区域内,通过自组织方式构建网络进行相互通信。普通传感器节点进行普通信息监测,当发现感兴趣事件时,唤醒视频采集节点,让其开始进行数据采集和数据压缩。采集处理后的数据通过普通节点,以无线多跳方式传输至汇聚节点,或由视频采集节点直接传输至汇聚节点。汇聚节点将接收到的数据信息通过无线网络进行发送,最后通过互联网送达控制中心(图 5.1 中各个汇聚节点均采用无线方式进行数据传输,为了说明简单,只对部分汇聚节点画出了无线通信方式,其他省略)。在控制中心,用户实时地对监测网络进行任务发布,同时对接收到的监测信息进行数据处理和信息管理<sup>[62-63]</sup>。

与传统的 WSNs 相比, WMSNs 的能耗“均匀”地分布在无线收发、数据采集和数据处理中。异构 WMSNs 的检测网络节点分为视频采集节点、普通传感器节点和汇聚节点,这样能够实现数据传输与数据采集功能的分离,保证了网络能耗的合理利用。

(1) 普通传感器节点: WMSNs 中,普通传感器节点的主要作用是对普通数据进行采集与传输,并对视频采集节点进行唤醒。由于简单监测信息的数据量

较小，并且普通节点进行监测所需的能耗不大，因此可以根据实际需求，将普通节点设置成具有对光敏、温度、湿度、气体等简单信息变化进行监测的功能。

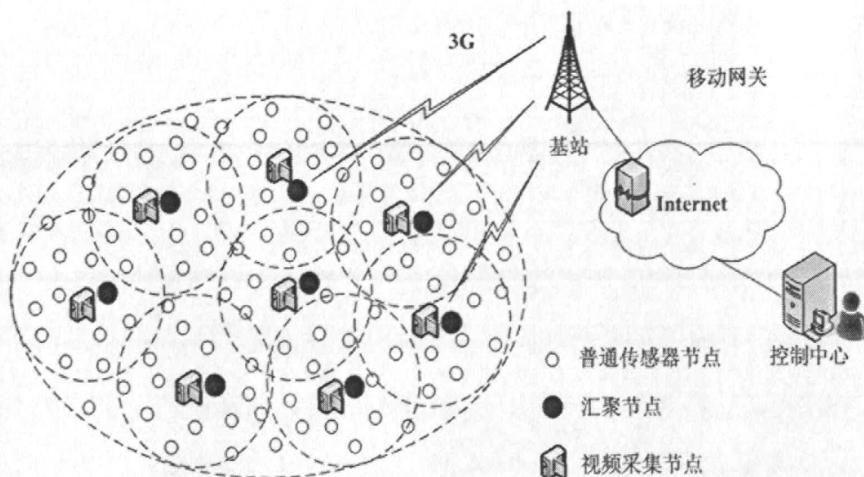


图 5.1 异构 WMSNs 网络结构图

(2) 视频采集节点：WMSNs 中，对于图像、视频等信息的采集和数据处理，由装配有 CMOS 摄像头的节点来完成。一般情况下，这些视频采集节点处于休眠状态，当有事件发生时，视频采集节点接到来自普通传感器节点的唤醒信号后，由休眠状态转为工作状态并开启摄像头，开始对事件进行采集并进行相应数据压缩。

(3) 汇聚节点：WMSNs 中，汇聚节点主要用于存储节点数据、负责与网络节点及控制中心进行数据传输。它具有存储量大的特点，接收来自普通节点和视频采集节点传输的数据，并将这些数据通过 3G 网络发送给基站，不参与 WMSNs 的信息采集和监测，保证能耗有效地运用于数据传输。

WMSNs 中的主要特点是三类节点相互协调工作，这也是区别于 WSNs 的特点之一。普通传感器节点根据设置的时间间隔，周期性的进行工作与休眠，当处于工作状态时，进行对网络区域内信息变化的检测，对普通信息进行数据采集，并将信息发送至汇聚节点。当感兴趣事件发生时，普通节点唤醒一定范围内的视频采集节点，之后转入到休眠状态。视频采集节点接收到来自普通传感器节点的唤醒信号，开始对事件进行多媒体信息的采集和压缩，经过压缩处理，原始的大数据量视频信息得到了有效精简。将处理好的数据码流直接传输给汇聚节点，为

有效保存节点能量,数据传输完毕之后关闭摄像头,转入休眠状态,以延长节点的生存时间。具有能耗不受限特性的汇聚节点,将接受的数据通过3G网络传输给基站,最终通过Internet网送至控制中心。

### 5.1.2 应用要求

鉴于以上对无线多媒体传感器网络的分析,对面向WMSNs的分布式视频编码系统方案提出了以下几点要求:

(1) 由于普通节点能耗、无线通信能力等资源有限,应用环境恶劣,大数据量的码流将严重缩短网络的生存时间,同时还将造成无线传输链路质量降低,码字错误率提高或码字丢失等问题。因此,WMSNs要求视频采集节点中采用的编码方式应该具有高效压缩性,以尽量低的能耗充分挖掘关键视频信息。

(2) 在WMSNs中,视频解码在控制中心进行。应充分利用控制中心能耗、计算、存储等资源不受限制的特点,将视频编解码的复杂度从编码移至解码。如此提高了在解码端对视频信息进行准确还原的要求,可以采用大数据量的运动估计、运动补偿等处理算法,以计算处理的复杂性换取视频解码信息的可靠性。

(3) WMSNs中采用的传输方式为无线传输,使得传输的可靠性降低。这一特性要求采用分布式视频编码方式在解码时必须充分利用视频帧之间的相关信息,尽可能准确地进行原始帧预测与还原,以保证解码后视频序列的质量。

## 5.2 WMSNs的分布式编解码系统

为了优化系统的编解码效果,提高辅助边信息的生成质量,本文在典型的基于变换域的分布式视频编解码系统的基础上进行了改进与创新,提出一种面向WMSNs的分布式编解码系统方案,该方案主要由编码端和解码端两部分组成,其系统结构框图如图5.2所示。为了优化对运动剧烈视频的解码质量,论文提出了先进行编解码、后进行检测判断与优化的编解码方式。在编码端引入了相关性提取模块,用来判定原始帧与前后帧之间的相关性,然后将计算获得的相关系数与当前帧的相关信息码流,经过编码压缩一起传送至解码端。当解码端接收到编码码流后,利用前后相邻已解码关键帧采用改进的边信息生成算法得到辅助边信息,并采用辅助边信息完成对当前帧的解码还原。解码后的像素帧需要与前后两

相邻关键帧再次进行相关性检测，对于达到检测标准阈值的像素块，将其视为合格的像素块还原，对未达到检测标准阈值的像素块，采用优化算法再次生成边信息像素块，根据重构后生成的辅助边信息，再次进行原始视频帧的解码，从而达到对解码帧的检测与边信息纠错优化的目的，以提高视频序列的解码质量。

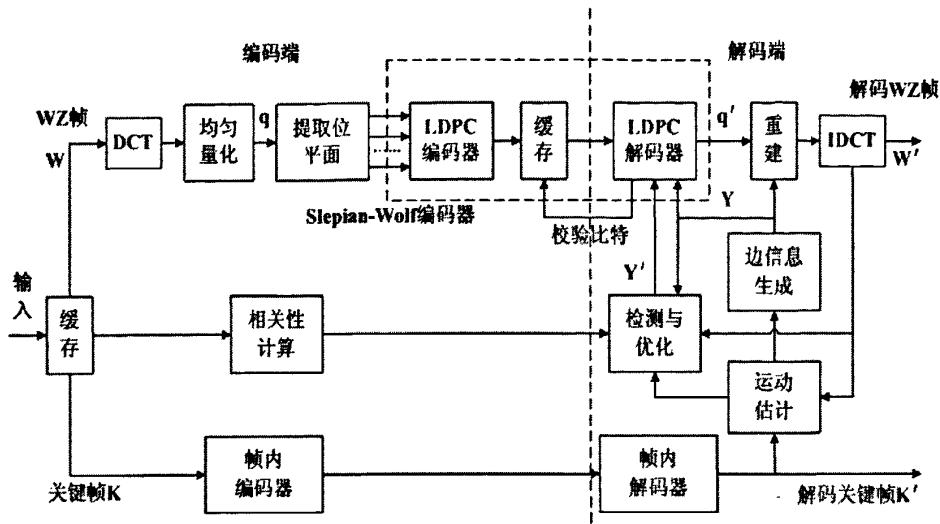


图 5.2 WMSNs 的分布式视频编码系统结构框图

在传统的分布式视频编解码系统结构中，视频帧主要分为两类：关键帧与 WZ 帧。对关键帧进行编解码时，采用传统的帧内编码、帧内解码的方式，对 WZ 帧采用帧内编码、帧间解码的方式。由于所提 WMSNs 的分布式视频编解码系统采用以上两种编解码方式的同时，还引入了编码端相关性判决模块，因此系统中的视频码流的传输将分成三条：传统的视频编解码传输、采用辅助边信息的视频编解码传输、相关性系数传输。前两类视频码流与传统的视频编解码方式保持一致，而相关性系数采用简单编码，在不增加系统复杂度的基础上以二进制码流形式传送至解码端。通过上述对 WMSNs 分布式视频编解码系统的描述可以看出，论文提出的分类编解码方式，使得系统的复杂度主要集中在解码端，与应用环境 WMSNs 的特点及应用需求相适应。

### 5.3 系统模块分析

论文提出的 WMSNs 分布式视频编解码系统主要包含变换、量化、LDPC 编码、初始边信息生成、检测与优化、LDPC 解码及重构等模块，下面将对各功能

模块进行详细介绍。

视频序列按奇偶顺序被分为关键帧和 WZ 帧：将奇数帧设为关键帧 K，偶数帧设为 WZ 帧。关键帧 K 采用传统帧内编码器进行压缩编码，编码后的码流传输至解码端，经由传统帧内解码器解码得到关键帧 K'。编码端对 WZ 帧进行变换和量化处理，经过 LDPC 编码器生成码流，传输至解码端，通过对前后两相邻关键解码帧 K' 进行运动估计等复杂运算，对原始 WZ 帧进行还原解码。

### 5.3.1 变换

视频帧由大部分变化缓慢区域和小部分细节或突变区域构成，为了去除空间相关性，需要采用离散余弦变换处理(DCT)，将高频系数、直流和低频系数的比重分别转化为少数和多数，以便后续操作步骤。

系统先把 WZ 帧划分成互不重叠的像素块，采用基于  $8 \times 8$  块的离散余弦变换，并对所有互不重叠的像素块按 4:2:0 格式对亮度块和色度块进行变换，将像素值转换为独立的 DCT 变换系数。将变换后所有的 DCT 变换系数放在同一个系数带  $X_k$  中，其中  $X_1$  为 DC(直流)系数，其余为 AC(交流)系数，将这些系数带分别进行量化与编码，并传输至解码端。在解码端，对当前 WZ 帧的预测边信息作相同的 DCT 变换，用于对 WZ 帧的解码与重构。

### 5.3.2 量化

为了对数据信息进行压缩编码，需要对每个系数带进行量化处理。根据不同频率系数带包含的图像信息量不同，采用非均匀的量化方式，以每个系数带的动态范围为依据动态调整量化步长。将系数带  $X_k$  均匀量化成  $2^M$  个区间，并用 M 位比特来表示对应的量化符号  $q_k$ ，M 的取值根据系数带  $X_k$  的不同而改变。如果 M 为零，即  $2^M=0$ ，表示对应系数带将不进行编码。DCT 变换后的系数矩阵由 DC 系数和 AC 系数组成，对 DC 系数带和 AC 系数带采用不同的量化方法。

#### (1) DC 系数带量化

DC 系数数值为正数，表示变换子块的平均能量，对其采用关于零点非对称的方式进行均匀量化。DC 系数带的量化步长  $D_l$  以及量化符号  $q_l$  对应的量化区间  $I_q$  分别通过计算式(5-1)与(5-2)得到，其中  $V_{max}$  为 DC 系数带内最大系数值， $2^M$

为量化区间数。

$$D_l = \frac{V_{\max}}{2^{M_l}} \quad (5-1)$$

$$I_q = [qD_l, (q+1)D_l] \quad (5-2)$$

## (2) AC 系数带量化

为了减少方块效应, 对同时存在正数和负数的 AC 系数带, 采用关于零点对称的量化区间进行均匀量化。量化步长  $Q_k$  以及量化符号  $q_k$  对应的量化区间  $I_q$  分别根据计算式(5-3)与(5-4)得到, 其中  $|V_k|_{\max}$  为 AC 系数带  $X_k$  的系数最大绝对值。

$$Q_k = \frac{2|V_k|_{\max}}{2^M - 1} \quad (5-3)$$

$$I_q = [(q - 0.5)V, (q + 0.5)V] \quad (5-4)$$

系数带  $X_k$  内所有系数经过量化后, 相同位的比特被放在一起, 按高比特至低比特顺序形成  $M$  个位平面。这些位平面按照由高位面到低位面的顺序, 分别送入 LDPC 编码器进行编码。当重构出 WZ 帧所有系数带后, 进行 IDCT 变换, 获得重构图像。

### 5.3.3 LDPC 编码

从最高位平面开始, 每个位平面被输入到 Slepian-Wolf 编码器中。从系统结构图 5.2 可知, Slepian-Wolf 编码器主要由信道编码和缓冲器两部分组成, 采用基于 LDPC 码的编码器模块。

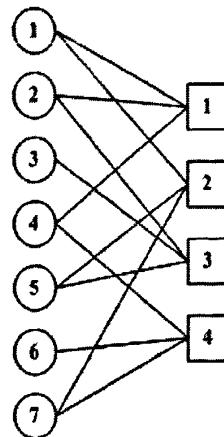


图 5.3 LDPC 码因子图

LDPC 码是一种线性分组码, 可以采用校验矩阵  $H$  定义, 同时也可以由变量节点与校验节点组成的因子图定义, 两种定义方式存在对应关系<sup>[64]</sup>。对于  $m \times n$  的校验矩阵, 其列数  $n$  对应变量节点数, 行数  $m$  对应校验节点数。变量节点表示需要传输的码字比特, 而检验节点表示对应的校验式, 如图 5.3 所示。

图 5.3 中, 左侧为变量节点, 右侧为校验节点, 若变量节点与校验校验节点间存在一条边, 则在校验矩阵中的对应元素值为 1, 以上因子图对应的校验矩阵如计算式(5-5)所示。

$$H = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \quad (5-5)$$

从计算式(5-5)可以看出校验矩阵  $H$  中, 零元素个数远大于非零元素个数, 是一个稀疏矩阵, LDPC 码正是利用稀疏矩阵实现了低复杂度、高压缩率的编码性能。校验矩阵  $H$  中, 若各行各列的非零元素个数均相同, 则称为 LDPC 规则码, 否则称为 LDPC 非规则码, 由于非规则码可以自由调整两种节点的比例, 因此性能优于规则码<sup>[65]</sup>。LDPC 编码器由一个传统编码器和一个累加器构成, 结构如图 5.4 所示。

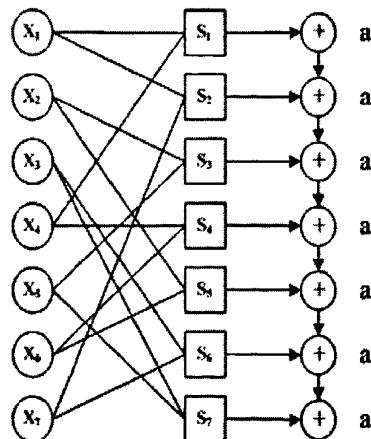


图 5.4 LDPC 编码器结构

传统编码器针对输入信源  $x$  生成对应的校验位  $s$  并送入累加器, 累加器对校验位进行模 2 相加运算, 得到累加校验位  $a$  并进行均匀间隔分组。顺序连续的累加校验位将分到不同的分组中, 最后将分组送入缓存器中, 等待解码端发送的校

验位请求，再依次将校验位传输至解码端。在 LDPC 编码器中，对信源的划分、变量节点和校验节点的个数确定以及累加校验位的分组需要根据实际情况，综合码率精确度与复杂度进行确定。

### 5.3.4 相关性模块

为了提取出原始帧的前后两相关性差别，并利用这一差值对初始解码帧质量进行判定，同时为边信息优化提供相关性判定信息，系统在编码端引入了相关性模块。在系统进行编码之前，该模块对当前帧  $X_{2k}$  和前后两相邻帧  $X_{2k\pm1}$  进行相关度计算，采用绝对差值和函数 SAD 得到前相邻相关度  $SAD_{f1,x}$  和后相邻相关度  $SAD_{b1,x}$ ，并通过比较前相邻相关度  $SAD_{f1,x}$  和后相邻相关度  $SAD_{b1,x}$  的大小来判定当前帧与前/后相关性。编码端将其所得结果进行简单编码，通过信道传输到解码端，使得原始帧间的相关度信息为解码帧质量判定以及边信息优化提供相关信息。

### 5.3.5 初始边信息生成

解码端的初始边信息生成模块采用第三章介绍的边信息生成算法思想，利用与原始帧  $X_{2k}$  相邻的两个关键帧  $X_{2k\pm1}$  的解码帧信息，进行运动估计生成初始边信息  $Y_{2k-f}$ 。先将像素帧划分成互不重叠的像素块，根据前后相邻关键帧的相关性，将像素块分成保留块类与运动块类。对保留块类直接做替代处理，对运动块类对应的相邻关键帧，采用改进的双向运动补偿算法找到对应的匹配块，通过加权判决生成初始边信息。

初始边信息一方面传输至解码器，辅助对当前 WZ 帧的解码，另一方面传输至检测优化模块，为低质量的边信息像素块的优化提供相关信息。

### 5.3.6 检测与优化

为了检测解码帧质量，系统采用检测与优化模块对当前帧的解码帧  $WZ'$  进行质量判定，并对判定质量较低的边信息像素块进行优化处理。该模块采用第四章提出的算法思想，先根据编码端相关性模块计算得到的前后相关度，在解码端计算得到解码检测阈值，利用这一检测阈值来判定解码帧中各像素块的解码质量。如果质量较高，则说明预测得到的像素块与原始像素块失真在可容误差范围

内,保留初始边信息的像素块作为对应位置的优化边信息像素块;如果质量较差,则说明解码得到的像素块与原始像素块失真较大,需要采用优化算法对初始边信息像素块进行优化。将两类像素块进行边信息重建,并将重建后的边信息进行再次解码以减小解码帧失真。

### 5.3.7 LDPC 解码

在提出的分布式视频编解码系统中,LDPC解码器是与LDPC编码器相对应的另一个重要模块。解码器接收来自缓存器的累加校验码,并将其与之前已经接收的累加校验码组合,进行与编码端对应的连续模2相减运算。接着检验节点从相邻变量节点获取码字信息,结合收到的边信息数据,根据运算后的累加校验消息对其进行校验处理,并将结果发送给变量节点,变量节点在进行了变量消息处理后,对解码结果的正确性进行检测。如果解码不成功,需要向编码端发送反馈信息,请求更多的累加校验码,并重复一轮以上迭代解码过程,直到正确解码为止。LDPC解码过程如图5.5所示。

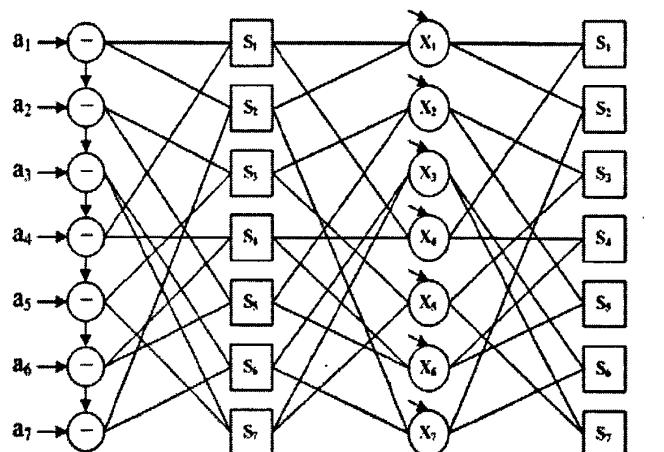


图 5.5 LDPC 解码

### 5.3.8 重构

重构模块是系统的最后一个功能模块。在解码之后生成M个位平面,其与WZ帧 $X_{2k}$ 的解码符号流相对应。通过对这M个位平面进行位平面分解的逆操作——位平面重组,得到解码量化符号流 $q'$ 。为了减少误差,当接收到解码符号流 $q'$ 与边信息后,重构模块进行像素重建,得到重构视频帧 $X_{2k}'$ 。重构函数如计算式(5-6)所示。

$$X'_{2k} = E(X_{2k} | q', S) \quad (5-6)$$

为了有效的消除尖锐噪声，提高视频解码质量，在得到解码量化符号流后，重构模块需要将对应位置的边信息像素值与该符号流确定的取值范围进行比较判断，一般分为三种情况。

(1) 如果在量化符号  $q'$  的电平范围之间，那么用边信息像素值来替换该位置的重构像素值。

(2) 如果超出解码量化符号  $q'$  的取值范围上限，那么将解码量化符号流  $q'$  中的最大值设为重构像素值。

(3) 如果低于解码量化符号  $q'$  的取值范围下限，那么解码量化符号流  $q'$  中的最小值设为重构像素值。

通过上述对 WMSNs 分布式视频编解码系统的工作原理以及各功能模块进行详细介绍，可以看出，系统的复杂度主要集中在解码端。系统的编解码流程如图 5.6 所示。

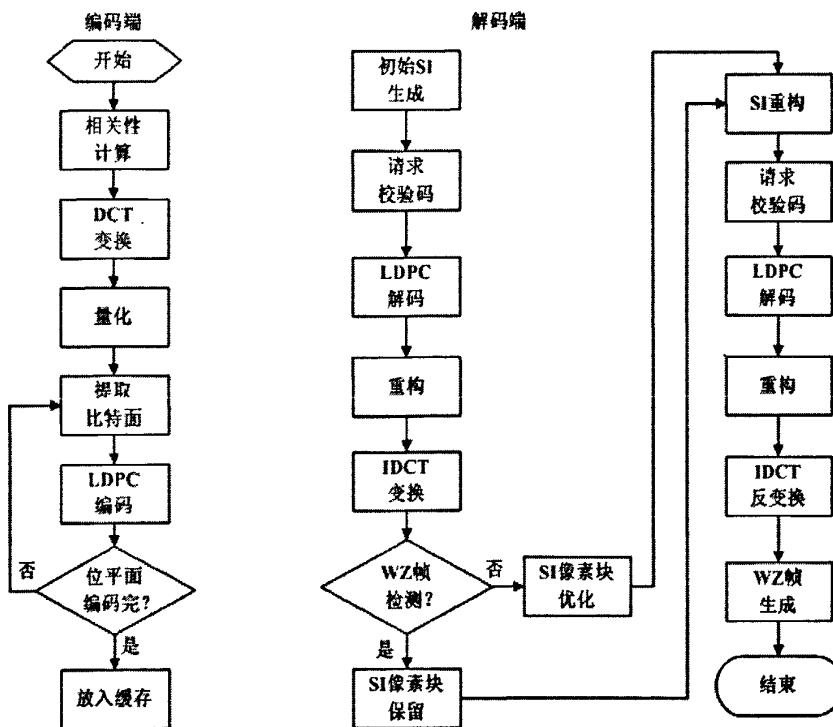


图 5.6 系统编码与解码流程图

## 5.4 系统性能仿真与分析

### 5.4.1 仿真条件设置

仿真在 matlab 平台上进行, 仿真的模块包括系统中的量化器、编码器、初始边信息生成模块、检测与优化模块、解码器和重构模块, 仿真条件设置如下。

- (1) 采用量化间隔为  $2^M \in \{2, 4, 8, 16\}$  的步长可变均匀量化器, 如此可以获取系统不同的率失真(Rate Distortion, RD)点。
- (2) 编码器。采用 LDPC 编码器作为 Slepian-Wolf 编码器<sup>[65]</sup>, K 帧采用传统的 H.264 Intra 编码<sup>[51]</sup>。
- (3) 测试序列格式均为 QCIF( $176 \times 144$ )。
- (4) 测试只计算 WZ 帧, 且 WZ 帧的帧率为 15fps, GOP 长度设置为 2。

### 5.4.2 测试与分析

分别对典型视频序列 Foreman、Carphone、Container 中的前 100 帧进行了编解码测试。为了更明确的显示算法的有效性, 测试中只计算视频帧的亮度信号的平均码率及 PSNR 值。

将论文所提分布式视频编码系统性能测试结果与以下三种视频编码系统进行编码性能比较: H.264 Intra 编码、H.264 Inter 编码、ISIG 编码算法<sup>[66]</sup>, 如图 5.7-5.9 所示。

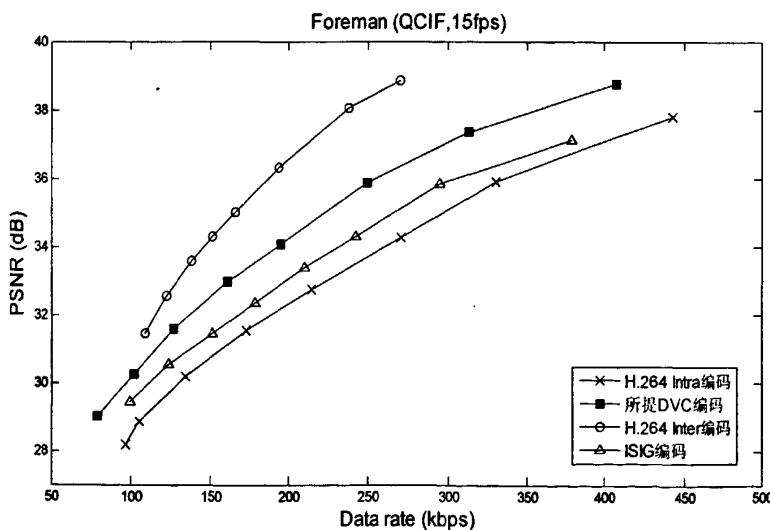


图 5.7 Foreman 序列的 RD 性能比较

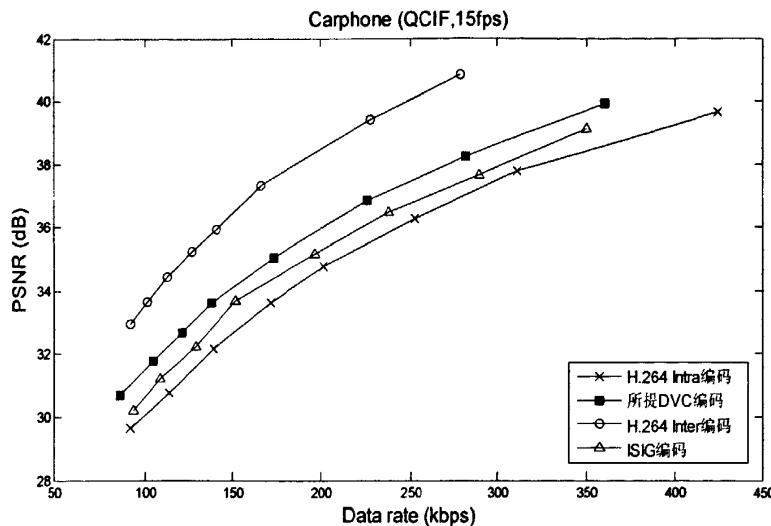


图 5.8 Carphone 序列的 RD 性能比较

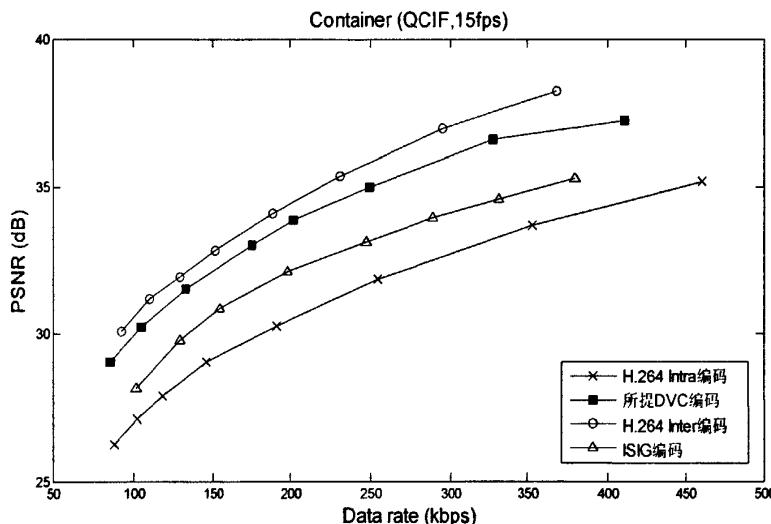


图 5.9 Container 序列的 RD 性能比较

分布式视频编解码系统方案的特点是帧内编码、帧间解码，其性能应介于帧间编解码和帧内编解码之间。从图 5.7-5.9 中可以看到，H.264 Inter 编解码方式的性能最好，而 H.264 Intra 编解码方式的性能最差，符合理论推测。提出的分布式视频编码系统方案只需要对序列进行帧内编码，其编码复杂度较 H.264 Intra 编解码方式相当，但在 RD 性能上平均提高了 1-3dB，相比 ISIG 编码方式平均提高了 0.5-1.5dB，但相比 H.264 Inter 编解码方式还存在一定的差距。

从图 5.8 中可以看出，在 Carphone 序列的 RD 性能比较中，所提出的分布式视频编码系统的 RD 性能与传统的 H.264 Inter 帧间编码差距最大，说明对于运动

剧烈的视频序列, 说明所提编码方式效率降低。这是因为 Carphone 序列中帧间运动量较大, 该序列中包含有大量的面部表情变化和背景快速移动, 使得帧间运动变化量增加, 由此边信息生成中保留块类的像素块数量减少, 从而增大了算法计算量和编码复杂度, 导致编码效率降低。

将图 5.7 和图 5.9 进行对比可以发现, 在 Foreman 序列和 Container 序列的 RD 性能对比中, 所提出编码方式相比传统的 H.264 Inter 帧间编码, 其 RD 性能平均提高较明显, 尤其在 Container 序列中。这是由于 Foreman 序列属于运动量较大的视频序列, 运动变化程相比 Container 较剧烈, 而 Container 序列中的运动变化主要是船只的移动, 背景的伸缩变化较小。在所提分布式视频编码系统方案中, 采用边信息生成模块, 对像素块进行分类处理, 将变化程度不大的像素块保留, 简化了编码复杂度, 同时采用改进的双向运动补偿内插和加权判决, 使得初始边信息的质量提高。利用初始边信息得到解码帧后, 根据原始帧信息自动对解码帧质量进行检测, 对于未到达检测阈值的像素块采用前后相关性判定, 进行双向加权运动估计、空域平滑滤波以及运动补偿, 对初始边信息进行质量优化。由于采用了边信息生成模块和监测与优化模块, 使得系统对运动变化程度较小的视频序列, 编码性能提高明显。

## 5.5 本章小结

本章在基于像素域的 DVC 系统方案基础上, 针对 WMSNs 这一应用背景, 结合之前提出的边信息生成算法和优化算法, 对面向 WMSNs 应用的分布式视频编码系统方案进行了设计。首先分析了 WMSNs 网络结构和应用要求, 对面向 WMSNs 的分布式视频编解码系统结构方案进行了整体设计, 并对各个模块进行了详细描述。最后给出实验仿真结果, 并与传统的视频编码进行比较和理论分析。

## 第六章 总结和展望

### 6.1 总结

随着科学技术的发展，在无线传感器网络(WSNs)基础上，具有多媒体信息感知能力的无线多媒体传感器网络(WMSNs)受到前所未有的关注。它在继承 WSNs 特点的同时，引入了图像、音频和视频等多媒体信息，满足了人们对精细信息的迫切需求，为实现远程监控信息的准确化、丰富化、直观化提供了条件，具有广阔的发展前景。

由于 WMSNs 中传感节点能耗、存储、计算等资源有限，如何进行大数据量的多媒体信息处理压缩成为了 WMSNs 发展的“瓶颈”。传统的视频编码方式需要在编码端进行复杂的运动估计和运动补偿，使得其不再适用于 WMSNs 的应用，因此人们将目光投向了一种新型的视频编码方式——分布式视频编码。

分布式视频编码基于分布式信源编码理论基础，采用帧内编码、帧间解码的方式，具有编码简单、解码复杂的特点。它不仅能满足 WMSNs 中传感节点在资源受限条件下进行压缩编码的应用要求，同时对无线网络的不可靠传输具有较高的鲁棒性。与传统的编解码方式不同，DVC 利用相邻关键帧的相关性信息，在完全不知道原始 Wyner-Ziv 帧的情况下对其进行预测解码。虽然解码端依靠边信息进行辅助解码，但是对于运动变化剧烈的视频序列，容易出现边信息与原始帧失真较大，导致解码出现错误。本文针对这一研究课题，提出了一种 WMSNs 的分布式视频编码方案，采用边信息生成算法和解码检测优化算法，以提高系统的编解码性能。论文的主要工作包括以下几个方面：

(1) 论文以无线多媒体传感器网络为应用对象，分析了课题的发展现状与应用前景，系统研究了分布式视频编码理论基础：Slepian-Wolf 分布式无损编码理论和 Wyner-Ziv 分布式有损编码理论、分布式视频编码特点、典型分布式视频编码系统方案和关键技术。

(2) 对分布式视频编码中的边信息生成算法进行深入研究，提出了一种块分类的双向估计加权边信息算法。采用分类思想根据帧间相关性将像素块分类，分别进行保留处理和改进的双向运动估计，最后通过加权判决生成边信息。实验

结果表明，在不增加编码端复杂度的情况下算法能有效提高了边信息质量。

(3) 对解码检测及边信息优化算法进行了研究，针对视频序列中运动剧烈导致解码质量低的问题，提出了一种相关性分类的解码检测优化算法。利用编码端原始帧的相关性，进行解码质量及前后相关性判定，对需要优化的像素块进行优化处理，最后将保留块与优化块进行边信息重建。实验结果表明，所提优化算法相比初始解码帧，PSNR 性能平均获得 0.5~1.5dB 的提高。

(4) 对无线多媒体传感器网络结构特点和应用要求进行了分析，提出了面向 WMSNs 的分布式视频编码系统方案，并对其各功能模块进行了详细描述。通过将仿真结果与其他视频编码进行性能对比，验证了该方案的可行性和有效性。

## 6.2 展望

本文对边信息生成和解码检测优化算法进行了深入研究，并结合这两种算法提出了一种面向 WMSNs 的分布式视频编码编解码方案，通过仿真实验证明其性能有了一定的提高。但本文在以下几个方面仍有待于继续研究和完善：

(1) 运动估计算法。与大多数运动估计算法一样，论文采用的运动估计是基于运动位移为线性轨迹的假设，当视频序列中变化剧烈或运动速度较快时，物体运动将与假设的线性轨迹偏移，导致估计的运动矢量与现实运动出现较大差异。怎样对当前运动变化进行有效而准确的运动矢量估计，以提高边信息生成质量，有待进一步的研究。

(2) 边信息算法。边信息的准确程度直接影响着分布式视频编解码系统的整体性能。论文所提出的边信息生成算法使得边信息质量有所提高，虽然采用了分类思想对减少了部分计算量，但是仅利用了前后两关键帧的相关信息，且计算规模较大。如何有效挖掘相邻的一系列关键帧的相关信息，生成一个质量高且计算量较小的边信息，具有很高的研究价值。

(3) 编解码器的选择。在分析关键技术时已指出，目前分布式视频编码系统中主要采用的是 Turbo 码和 LDPC 码，两者都是较耗时的迭代码，且进行解码时需要较复杂的计算量，这不能满足实现实时、高效的要求。对编解码的进一步深入研究，找到一种实时、高效和可靠的编解码方式是一个有待解决的问题。

(4) 码率控制技术。码率控制是影响分布式视频编码系统性能的因素之一,其自身受到三个因素的影响:关键帧与 WZ 帧的比例、关键帧的码率和 WZ 帧的码率。如何根据视频序列的运动特性,合理分配关键帧与 WZ 帧的比例,调节关键帧和 WZ 帧的码率,实现降低总码率,并保证重建帧质量,是今后研究的一个方向。

## 参考文献

- [1] 闫朝升,张承红,马英.传感器网络研究综述[J].信息技术,2006,30(3):117-121.
- [2] 罗武胜,翟永平,鲁琴.无线多媒体传感器网络研究[J].电子与信息学报,2008,30(6):1511-1516.
- [3] R. Cucchiara. Multimedia surveillance systems[C]. Proceedings of the third ACM international workshop on Video surveillance & sensor networks, 2005:3-10.
- [4] 马华东,陶丹.多媒体传感器网络及其研究进展[J].软件学报,2006,17(9):2013-2028.
- [5] Margi.B, Petkov V, Obraczka K, Manduchi R. Characterizing energy consumption in a visual sensor network testbed[C]. Testbeds and Research Infrastructures for the Development of Networks and Communities, 2006:332-339.
- [6] 刘海艳,视频无线传感器网络信息协同处理技术研究[D].长春理工大学.2010.
- [7] B.Girod, A.Aaron, S.Rane, D. Rebollo-Monedero. Distributed Video Coding[J]. Proceedings of the IEEE, 2005, 93(1):71-83.
- [8] Aaron A, Zhang R, Girod B. Wyner-Ziv coding of motion video[C]. Proc. Asilomar Conference on Signals, Systems and Computers, 2002, (1):240-244.
- [9] Aaron A, Girod B. Compression with Side Information Using Turbo Codes[C], Proc. IEEE Data Compression Conference Proceedings.2002, 252-261.
- [10] Puri R, Ramchandran K. PRISM: A new robust video coding architecture based on distributed compression principle[C]. Proc. Allerton Conference on Communication Control and Computing,2002.
- [11] Puri R, Ramchandran K. PRISM: A reversed multimedia coding paradigm[C].IEEE International Conference on Image Processing,2003,September,(1):617-20.
- [12] X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov, M. Ouaret. The DISCOVER Codec: Architecture, Techniques and Evaluation. Picture Coding Symposium.2007:1-4.
- [13] A. Sehgal, A. Jagmohan, N. Ahuja. A state-free causal video encoding paradigm[C]. IEEE International Conference on Image Processing, 2003(1):605-608.
- [14] Y. Zhao, J. Garcia-Frias. Joint Estimation and Compression of Correlated Nonbinary Sources Using Punctured Turbo Codes[J]. IEEE Transactions on communications,2005,53(3):385-390.
- [15] Mackay D.J.C, Neal R.M. Near Shannon limit performance of low density parity check codes[J]. IEEE Electronics Letters,1996,32(18):1645-1646.
- [16] Zixiang Xiong, Angelos D. Livers, Samuel Cheng, Zhixin Liu. Nested Quantization and Slepian-Wolf Coding: A Wyner-Ziv Coding Paradigm for i.i.d. Sources[C]. IEEE Workshop on Statistical Signal Processing,2003:399-402.
- [17] Xu Qian, Zixiang Xiong. Layered Wyner-Ziv video coding[J]. IEEE Transactions on Image Processing, 2006,15(12):3791-3803.
- [18] D. Rebollo-Monedero, S. Rane, A. Aaron, B. Girod. High-Rate Quantization and Transform Coding with Side Information at the Decoder[J]. EURASIP Signal Processing, 2006, 86(11): 3160-3179.
- [19] Fowler JE, Marco Tagliasacchi, B'eatrice PesquetPopescu. Wavelet-based distributed source coding of video [EB/OL]. <http://www.ece.msstate.edu/Publications/Papers/FTP2005.pdf>, 2005.

- [20] A.B.B.Adikari, W.A.C., Weerakkoy. W.A.R.J. Iterative Wyner-Ziv decoding for unidirectional distributed video coding[J]. IEEE Electronics Letters,2007,43(2):93-95.
- [21] Yang Y, Stankovic V, Xiong Z, Zhao W. Two-terminal video coding[J]. IEEE Transactions on Image Processing, 2009,18(3):534-551.
- [22] Lin Y, Varodayan D, Girod B. Image Authentication based on Distributed Source Coding[C]. IEEE International Conference on Image Processing, 2007.III-5-III-8.
- [23] Brites C, Pereira F. An Efficient Encoder Rate Control Solution for Transform Domain Wyner-Ziv Video Coding[C], IEEE Transaction on Circuits and System for Video Technology, 2011,21(9):1278-1292.
- [24] Bo Wu, Xun Guo, Debin Zhao, Wen Gao, Feng Wu. An optimal non-uniform scalar quantizer for distributed video Coding[C], IEEE International Conference on Multimedia and Expo,2006,165-168.
- [25] 千宗良,齐丽娜,朱秀昌.一种空间域 Wyner-Ziv 视频编码系统的性能改进算法[J].电子学报 2007,35(10):2014-2018.
- [26] 蒋刚毅,金智鹏,郁梅.分布式视频编解码技术的研究[J].中国图形图像学报,2008,13(3): 386-393.
- [27] 王安红.分布式视频编码研究[D].北京交通大学,2009.
- [28] 卿粼波,吕瑞,欧先锋,郑敏,何小海.基于迭代译码算法的分级分布式视频编码[J].山东大学学报(工学版),2011,41(2):58-61.
- [29] Slepian J, Wolf J. Noiseless Coding of Correlated Information Sources[J]. IEEE Transactions on Information Theory,1973,19(4):471-480.
- [30] 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.
- [31] A. D. Wyner. Recent result in the Shannon theory[J]. IEEE Trans Inf Theory,1974,20(1):2-10.
- [32] Pradhan S.S, Ramchandran K. Distributed source coding using syndromes(DISCUS):Design and construction[C]. Proc.IEEE Data Compression Conference DCC'99, 1999, 158-167.
- [33] A. Aaron, S. Rane, B. Girod. Transform-Domain Wyner-Ziv Codec for Video [C].Proceedings of SPIE Visual Communications and Image Processing Conference, USA, 2004.
- [34] Aaron A, Rane S, Girod B. Wyner-Ziv Video Coding with Hash-Based Motion Compensation at the Receiver[C].Proc.IEEE International Conference on Image Processing, 2004, 24-27.
- [35] 王海芳.分布式视频边码边信息研究[D].太原科技大学,2011.
- [36] Berrou C, Glavieux A, Thitimajshima P. Near Shannon Limit error-correcting coding and decoding: Turbo-codes[C].Proceeding of ICC'93,1993:1064-1070.
- [37] P. Ishwar, V. Prabhakaran, K. Ramehandran.Towards a theory for video coding using distributed compression principles. Proceeding of the IEEE International Conference on Image Processing, Barcelona, Spain, IEEE Computer Society Press, 2003,687-690.
- [38] Trapanese A, Tagliasacchi M, Tubaro S. Improved correlation noise statistics modeling in frame-based pixel domain Wyner-Ziv video coding[C]. Proceedings of the International Workshop on Very Low Bitrate Video Coding, 2005,1-4.
- [39] Brites C, Ascenso J, Pereira F. Modeling correlation noise statistics at decoder for pixel based Wyner-Ziv video coding[C].Proceedings of the International Picture Coding, 2006b, 1-6.
- [40] Brites C, Pereira F. Correlation Noise Modeling for Efficient Pixel and Transform Domain Wyner-Ziv Video Coding [J]. IEEE Trans.Circuits and Systems for Video Technology, 2008,

- 18(9):1177-1190.
- [41] Brites C, Pereira F. Encoder rate control for transform domain Wyner-Ziv video coding [C]. Proceedings of the IEEE International Conference on Image Processing, USA, 2007, 5-8.
- [42] Ascenso J, Brites C and Pereira F. Improving frame interpolation with spatial motion smoothing for pixel domain distributed video coding [C]. Proceedings of the 5th EURASIP Conference on Speech and Image Processing, Multimedia Communications and Services, 2005.
- [43] Kubasov D, Guillemot C. Mesh-based motion-compensated interpolation for side information extraction in distributed video coding [C]. Proceedings of the IEEE International Conference on Image Processing, Atlanta, 2006, 261-264.
- [44] Ascenso J and Pereira F. Adaptive hash-based side information exploitation for efficient Wyner-Ziv video coding [C]. Proceedings of the IEEE International Conference on Image Processing, San Antonio, 2007, (3):29-32.
- [45] LIU R, YUE Z, CHEN C W. Side Information Generation Based on Hierarchical Motion Estimation in Distributed Video Coding[J]. Chinese Journal of Aeronautics, 2009, 22(2): 167-173.
- [46] 王凤琴,樊养余,张政伟,等.一种有效的自适应边信息帧内插算法[J].系统仿真学报,2009, 21(22):7215-7224.
- [47] 卿器渡,何小海,吕瑞.DVC 中基于亚像素 MCI 的边信息重建[J].计算机工程,2010,36(3): 224-226.
- [48] ASCENSO J, BRITES C, PEREIRA F. A flexible side information generation framework for distributed video coding[J]. Multimedia Tools and Applications, 2010, 48(3):381-409.
- [49] 薛冲冲.视频压缩中运动估计算法及预测搜索起始点的研究.西南大学,2011
- [50] Yao Nie, Kai-Kuang Ma. Adaptive Rood Pattern Search for Fast Block-Matching Motion Estimation[J]. IEEE Trans. Image Processing, 2002, 11(12):1442-1448.
- [51] KARSTEN S. H.264/AVC Software Coordination[EB/OL].<http://iphome.hhi.de/suehring/tm1>.
- [52] Chen Jian, Zhang Tao, Cui Wei, Wu Wei. Reducing Quantization Loss and Improving Side Information in Distributed Video Coding[J]. Electronic Letters, 2011, (47):30.
- [53] Badem M B, Mrak M, Fernando W A C. Side Information Refinement Using Motion Estimation in DC Domain for Transform-based Distributed Video Coding[J]. IEEE Electronics Letters, 2008, 44(16): 965-966.
- [54] Ko B, Shim H J, Jeon B. Wyner-Ziv Coding with Spatio-temporal Refinement Based on Successive Turbo Decoding[C]. Proc. of IEEE International Conference on Multimedia and Expo. [S. l.]:IEEE Press, 2008, 785-788.
- [55] YE S, OUARET M, DUFAUX F. Improved side information generation with iterative decoding and frame interpolation for distributed video coding[A]. Proceedings of the 15th International Conference on Image Processing. 2008, 2228-2231.
- [56] 张金荣,基于分布式编码的图像视频编码关键技术研究[D],中国科学技术大学,2009.
- [57] 马力,苏卓涵.基于多假设运动补偿预测的边信息改进方法[J],2011,37(12): 248-250
- [58] Mengjie Hou, Aidong Men, Bo Yang. An improved side information refinement using color-texture features for distributed video eoding, 2010 International Conference on Network Infrastructure and Digital Content,2010,230-234.
- [59] L. Alparone, M. Barni, F. Bartolini, V. Cappellini. Adaptively weighted vector-median filters for motion fields smoothing. IEEE ICASSP,1996,4:2267-2270

- [60] Holman R, Stanley J, OzkanHaller T. Applying video sensor networks to nearshore environment monitoring[J]. IEEE Transactions on Pervasive Computing, 2003,2(4):14-21.
- [61] V.Mhatre, C.Rosenberg. Homogeneous vs heterogeneous clustered sensor networks:a comparative study[C]. Proc.of the IEEE International Conference on Communications, 2004,6:3646-3651.
- [62] 沙超,孙力娟,王汝传,黄海平.无线多媒体传感器网络中能量高效的采样与传输方法[J].通信学报,2011,32(2):1-10.
- [63] 鲁琴,杜列波.无线多媒体传感网节点能耗问题评述[J].传感器与微系统,2008,27(12): 1-3
- [64] 夏洪涛,基于 LDPC 码的分布式视频编码研究[D],华南理工大学,2010.
- [65] VARODAYAN D, AARON A, GIROD B. Rate-adaptive for distributed source coding[J]. Signal Processing, 2006, 86(11):3123-3130.
- [66] SHUI M Y, MOURAD O, FREDERIC D. Improved side information generation for Distributed Video Coding by exploiting spatial and temporal correlations[J]. EURASIP Journal on Image and Video Processing, 2009, 2009(1):1-15

## 致 谢

转眼间，三年的读研生活即将过去，回想起这三年的学习和生活，心里总会呈现出一些难忘的时刻。正是由于老师和同学的帮助，我才得以一步步的成才，对此，心里充满无限的感激。

首先，我要谢谢我的导师陈祖爵教授，在这三年的读研生活中，对我学习上的指导。特别是这篇论文，从选题、构思到资料收集以及修改阶段都得到了导师的精心指导。在此论文完稿之际，特向导师致以最诚挚的感谢！

感谢闫述老师、朱娜老师、周莲英老师、王洪金老师等众多老师在论文开题和中期汇报阶段给我提出了许多宝贵意见。

感谢实验室里单凯丽、郭秋月、施展、王君、黄福春等同学，在我读研期间的学习和生活上提供的帮助。

感谢在我成长过程中给予关心和支持的所有老师、同学和朋友。

最后，我要深深地感谢我的家人和男友，感谢他们一直以来对我的支持、鼓励和关心。

## 攻读硕士期间参与的科研项目和发表的文章

- [1] 陈祖爵, 陈媛, 董亮, 江晓明, 李杰. 《块分类的双向估计加权边信息生成算法》, 计算机应用研究, 2012 年 1 月。
- [2] Zu-jue Chena, Yuan Chenb, Yong-fu Zhao. “A Block-unoverlapped Algorithm of Side Information Generation in DVC”, IEEE World Conference on Science and Engineering 2011.
- [3] 陈祖爵, 麻勰光, 陈媛. 《能量均衡的动态间隔分层路由协议》, 计算机应用研究, 2011 年 1 月。