

上海大学出版社

2005年上海大学博士学位论文 113



混和时域 / 空域 / SNR 的 MC + FGS 视频编码 及其相关技术研究

● 作者：江 涛

● 专业：通信与信息系统

● 导师：张兆扬



9 643/158

上海大学出版社

001280760

2005年上海大学博士学位论文 113



混和时域 / 空域 / SNR 的 MC + FGS 视频编码 及其相关技术研究

- 作者：江涛
- 专业：通信与信息系统
- 导师：张兆扬



贵阳学院图书馆



图书在版编目(CIP)数据

2005 年上海大学博士学位论文. 第 2 辑/博士论文编辑部编. —上海: 上海大学出版社, 2009. 6

ISBN 978-7-81118-367-2

I. 2… II. 博… III. 博士—学位论文—汇编—上海市—2005 IV. G643. 8

中国版本图书馆 CIP 数据核字(2008)第 180878 号

2005 年上海大学博士学位论文

——第 2 辑

上海大学出版社出版发行

(上海市上大路 99 号 邮政编码 200444)

(<http://www.shangdapress.com> 发行热线 66135110)

出版人: 姚铁军

*

南京展望文化发展有限公司排版

上海华业装潢印刷厂印刷 各地新华书店经销

开本 890×1240 1/32 印张 274.25 字数 7641 千

2009 年 6 月第 1 版 2009 年 6 月第 1 次印刷

印数: 1~400

ISBN 978-7-81118-367-2/G·490 定价: 980.00 元(49 册)

Shanghai University Doctoral Dissertation (2005)

Hybrid Spatio-Temporal-SNR MC+FGS Video Coding and Research of Relative Technology

Candidate: Jiang Tao

Major: Communication and Information System

Supervisor: Zhang Zhaoyang

Shanghai University Press

• Shanghai •

上海大学

本论文经答辩委员会全体委员审查,确认符合上海大学博士学位论文质量要求。

答辩委员会名单:

主任:	余松煜	教授,上海交通大学	200240
委员:	王朔中	教授,上海大学	200072
	王汝笠	研究员,中科院技术物理研究所	730000
	严壮志	教授,上海大学	200072
	李 先	高级工程师,新上广	200001
导师:	张兆杨	教授,上海大学通信学院	200072

学大商上

合共人商，查审员委科全会员委粮答强文信本
。本要量通文信学士制学大商上

评阅人名单：

- | | | |
|------------|------------------|--------|
| 余松煜 | 教授，上海交通大学图象通信研究所 | 200240 |
| 王翔中 | 教授，上海大学通信学院 | 200072 |
| 周源华 | 教授，上海交通大学 | 200240 |

江涛博士论文答辩评语

江涛的博士论文研究时空域可分级的精细粒度可伸缩视频编码。精细粒度可伸缩编码能够精细匹配网络的带宽波动,且具备较好的抗差错能力,是 Internet 流媒体应用中先进的视频编码解决方案。本研究有重要的学术意义和良好的应用价值。

论文以提高编码效率和提供时域/空域/SNR 可分级为目标,主要创新成果有:

1. 提出了两种 FGS 视频编码方案-双环 MC+FGS 和单环 MC+FGS,使编码效率得到提高。同时提出了一套用来确定 FGS 增强层中参考位平面个数的算法,较好地解决了预测漂移问题。

2. 提出了 MC+FGST 视频编码方案,其编码效率优于 MPEG-4 的 FGST 方案。

3. 提出了 MC+FGSS 方案,实现了将可分级和空域 SNR 的精细粒度可分级相结合的混合方案。

4. 提出了一种时域/空域可分级和 SNR 精细粒度可分级相结合的 FGSST 方案,并将其与 MC+FGS 相结合,得到一种 MC+FGSST 的编码方案,可同时支持三种可分级,且有较高的编码效率。

论文立论正确,分析详实,实验结果可信;条理性较好,层次分明,逻辑性强;在答辩时回答问题正确,表明作者有坚实的理论基础和系统的专业知识,分析问题解决问题的能力强。

答辩委员会表决结果

经论文答辩委员会讨论,并投票表决,一致同意通过江涛同学的毕业论文答辩,并建议授予工学博士学位。

答辩委员会主席: **余松煜**

2005年6月24日

摘 要

近年来随着计算机与网络技术的飞速发展,人们通过 PC 或者非 PC 设备接入互联网或无线网络进行流媒体点播服务成为现实。流媒体技术彻底改变了传统的多媒体 Web 服务,它允许用户无需等到整个多媒体文件被完全下载,便可以享受在线的实时的视频点播服务。因此,流媒体技术有着非常光明的应用前景。

网络的异构性、带宽的波动性和各类不同用户的多重服务要求对当前的视频编码技术提出了新的挑战。视频编码后得出的码流必须能够适应网络带宽不断波动的变化,具备一定的抗差错能力,以及能同时提供图像质量、时域、空域及解码复杂度的可分级能力。

可分级视频编码方案是解决 Internet 流媒体应用中带宽波动的一种有效方法。但是,传统的可分级编码仅能提供粗糙的可分级能力,无法精细地匹配网络带宽的变化。MPEG-4 标准中采纳的 FGS 编码方案可有效地解决精细匹配网络带宽波动问题。不仅如此,采用 FGS 方案编码后的码流还获得较好的差错复原能力和解码复杂度可分级的能力。但是,FGS 获得的所有这些特性都是以牺牲编码效率为代价的。FGS 方法的编码效率较低,因为 FGS 编码中的增强层未使用任何运动补偿的措施来去除原始视频序列在增强层上的时域冗余。

针对 FGS 编码效率较低的缺点,我们提出了两种在增强层

也使用运动补偿的视频编码方案:双环 MC+FGS 结构编码方案和单环 MC+FGS 结构编码方案。双环 MC+FGS 和单环 MC+FGS 的编码方案虽然在实现结构上借鉴了 Mihaela van der Schaar 的思想,即使用图像质量更好的增强层图像作参考,通过提高运动补偿的效率来达到提高编码效率的目的,但是本文对此作了重要的改进。在所提出的方案中引入了接收端驱动(receiver-driven)的思想使得能根据网络的可用带宽动态地调整用于重建高质量参考帧所使用位平面的个数,这样不仅可妥善解决上述两种编码结构在低比特率情况下会产生预测漂移的问题,而且还进一步提高了在高比特率情况下的编码效率。对于在无法引入接收端驱动思想的情况下,也对两种结构的 FGS 增强层中究竟使用多少个位平面来重建高质量的参考提出了解决方案。通过这些重要改进后,使得这两种结构的编码性能产生了质的提高,不但预测漂移的问题得以解决,而且编码效率还得到进一步的改善。

针对时域可分级的 SNR FGS, MPEG-4 标准中引入了 FGST(时域 FGS)方案,该方案实现了混合时域/SNR 的精细粒度可分级。但是,FGST 方案继承了 FGS 方案编码效率较低的缺陷。因此,我们把 MC+FGS 与 FGST 相结合,得到一种 MC+FGST 结构的编码方案,提高了 FGST 结构方案的编码效率。在传输过程中,我们提出了一种有效的速率分配策略,用来管理 SNR FGS 比特流数据和 FGST 比特流数据在传送过程中的分配。

MPEG-4 标准中的 FGS 方案并未提供对空域可分级的 SNR FGS 支持。因此,针对 MPEG-4 中的这一空白,本文提

出一种 FGSS(空域 FGS)方案,实现了混合空域/SNR 的精细粒度可分级能力。但由 FGS 扩展得到的 FGSS 方案同样也继承了 FGS 方案编码效率较低的缺陷,为此把 MC+FGS 与 FGSS 相结合,得到一种 MC+FGSS 结构的编码方案,提高了 FGSS 结构方案的编码效率。

针对混合时空域可分级的 SNR FGS,本文进而又提出了 FGSST(时空域 FGS)方案。FGSST 方案实现了混合时域/空域/SNR 的精细粒度可分级能力。然后,我们再对 FGSST 方案的编码效率较低的缺陷进行改进,把 MC+FGS 与 FGSST 相结合,得到一种 MC+FGSST 结构的编码方案,它既实现了混合时域/空域/SNR 的精细可分级粒度可分级能力,又提高了 FGSS 结构方案的编码效率。

关键词 精细粒度可分级,编码效率,时域/SNR 精细粒度可分级,空域/SNR 精细粒度可分级,时域/空域/SNR 精细粒度可分级,可用带宽

本论文研究受到上海市青年科学基金项目(03AQ86)的资助。

Abstract

With the high development of computer and network technology recent years, it is becoming true for people to enjoy VOD (video-on-demand) service by accessing network using PC. Streaming media technology changes the traditional multimedia service based on the Web. It allows clients to enjoy VOD service real-time, without waiting for whole multimedia file to be downloaded on local computer. Therefore, there is great commercial future for multimedia stream.

However, to realize such VOD service, it demands great challenges for current video compression technology. Firstly, current network of Internet is heterogeneity. To cope with network's available bandwidth varying, bit-streams after encoded should provide scalable ability to adapt bandwidth's varying. Secondly, current best-effort Internet does not offer any quality of service (QoS) guarantees to streaming media over the Internet. Thus, bit-streams after encoded should provide robust error resilience. Finally, clients hope to obtain different services according to their available resource, such as different quality service, different frame-rate service and different resolution service. So, bit-streams after encoded should provide SNR, temporal, spatial and decoding complexity scalabilities.

Scalable coding has already been recognized as an effective solution to resolve the network's bandwidth varying. Though traditional scalable video coding schemes provide scalability, the scalability is coarse and need expensive cost. To obtain fine granularity scalability, MPEG-4 standard adopts FGS coding scheme. FGS scheme permits bit-streams after encoded be truncated and transmitted arbitrarily, which matches bandwidth varying well. FGS scheme also provides robust packet-loss resilience and decoding complexity scalability. However, all the advantages of FGS scheme are obtained at sacrifice of coding efficiency. This is mainly due to lack of motion compensation on FGS enhancement-layer.

To improve coding efficiency, a two-loop MC+FGS and a single-loop MC+FGS coding schemes are presented. The key idea of both MC+FGS coding schemes is using higher quality picture of enhancement-layer as reference. By improving the efficiency of motion compensation, both MC+FGS coding schemes obtain higher coding efficiency than FGS. In both presented MC+FGS coding schemes, an effective algorithm is presented to determine how many bit-planes on FGS enhancement-layer should be used to reconstruct reference frame. The algorithm not only solves the problem of prediction drift at low bit-rate, but also further improves coding efficiency at high bit-rate.

To solve temporal scalability, MPEG-4 standard introduces FGST scheme, which provides temporal-SNR scalability. However, the coding efficiency of FGST scheme

is as poor as FGS. Thus, we combine FGST and MC + FGS and develop a MC + FGST coding scheme. MC + FGST scheme not only provides temporal-SNR scalability, but also obtains considered coding efficiency. Then, we present a rate-allocation strategy to manage how to transmit SNR FGS bit-stream and FGST bit-stream.

To solve spatial scalability, we present a FGSS scheme for MPEG - 4 standard. FGSS scheme expend MPEG - 4 FGS to spatial domain, which provides spatio-SNR scalability. Owing to deriving from FGS, FGSS scheme also inherits the defect of poor coding efficiency of FGS. Thus, we combine FGSS and MC + FGS and develop a MC + FGST coding scheme. MC + FGSS scheme not only provides spatio-SNR scalability, but also obtains considered coding efficiency.

To solve spatio-temporal scalability, we present a FGSST scheme, which provides spatio-temporal-SNR scalability. Furthermore, we combine FGSST and MC + FGS and develop a MC + FGST coding scheme to improve FGSST's coding efficiency. MC + FGSST scheme not only provides spatio-SNR scalability, but also obtains considered coding efficiency.

Key words fine-granular-scalability, coding efficiency, temporal-SNR scalability, spatial-SNR scalability, spatio-temporal-SNR scalability, available bandwidth

目 录

第一章 绪论	1
1.0 引言	1
1.1 流媒体应用在传输上面临的挑战	4
1.2 视频压缩编码技术	6
1.2.1 视频图像压缩编码标准及其应用场合	7
1.2.2 视频图像压缩编码的主要方法	8
1.2.3 适用于流视频传输的可分级视频编码技术	10
1.3 本文的结构与创新	15
第二章 精细粒度可分级视频编码的分析研究	18
2.0 引言	18
2.1 精细粒度可分级的 FGS 编码	19
2.1.1 DCT 系数的位平面编码方案	21
2.1.2 原始 FGS 编解码器的结构	25
2.1.3 FGS 方案的编码性能	28
2.1.4 FGS 方案的高级功能	32
2.1.5 FGS 编码方案的性能总结	35
2.2 渐进精细粒度可分级编码方案	36
2.2.1 基本的 PFGS 编码的基本思路	36
2.2.2 简化的 PFGS 结构	38
2.2.3 改进的 PFGS 结构	40
2.2.4 PFGS 结构的性能总结	46
2.3 本章小结	46

第三章 基于增强层运动补偿的 MC+FGS 视频编码方案	48
3.0 引言	48
3.1 双环 MC+FGS 编码方案	50
3.1.1 双环 MC+FGS 编码方案的结构	50
3.1.2 双环 MC+FGS 编码方案的编解码器	52
3.1.3 如何确定双环 MC+FGS 编码方案中用于 参考的位平面个数	56
3.1.4 双环 MC+FGS 编码方案的优缺点	61
3.2 单环 MC+FGS 编码方案	62
3.2.1 单环 MC+FGS 编码方案的结构	62
3.2.2 单环 MC+FGS 编码方案的编解码器	63
3.2.3 如何确定单环 MC+FGS 编码方案中用于 参考的位平面个数	67
3.2.4 单环 MC+FGS 编码方案的优缺点	70
3.3 两种 MC+FGS 结构编码方案的性能比较	70
3.3.1 两种结构编码效率的比较	71
3.3.2 两种结构编码方案的复杂度比较	73
3.4 从两种结构中选取 MC+FGS 的标准结构	74
3.5 仿真实验结果及分析	74
3.6 本章小结	76
第四章 时空域可分级的 MC+FGS 视频编码方案	78
4.0 引言	78
4.1 时域可分级的 MC+FGS 视频编码方案	79
4.1.1 FGST 结构的编码方案	80
4.1.2 MC+FGST 结构的编码方案	82
4.1.3 实验结果及分析	91
4.2 空域可分级的 MC+FGSS 视频编码方案	93
4.2.1 FGSS 结构的编码方案	94

4.2.2	MC+FGSS 结构的编码方案	100
4.2.3	空域下抽样及上抽样策略	103
4.2.4	实验结果及分析	105
4.3	时空域可分级的 MC+FGSST 视频编码方案	107
4.3.1	FGSST 结构的编码方案	108
4.3.2	MC+FGSST 结构的编码方案	112
4.3.3	实验结果及分析	115
4.4	本章小结	116
第五章 总结与展望		118
参考文献		121
附录一 网络可用带宽的估计		135
附录二 文中部分缩写词汇的中英文含义表		150
攻读博士学位期间发表的学术论文		151
致谢		152

们的工作和生活方式。Internet 的发展也使人们对该网络寄予了越来越多的期望。数字信息处理技术和骨干网络传输带宽(SDH)的不断提高以及网络接入技术(如 DSL)的不断进步,使得在 Internet 网络上传输多媒体信息成为可能。人们已经成功实现了在 Internet 上传输音频信息,也不断尝试在该网络上传输视频信息。Internet 作为尽力而为的网络,不能保证类似在 ATM 网络上提供的多媒体通信业务的服务质量。于是研究者开始思考如何在本网络提高文件运行速度的方式,考虑在播放端设置缓冲区来解决很多而这些问题就成了流媒体技术发展的指导思想之一。

流媒体(Streaming Media)技术是当前非常活跃的多媒体技术,其基础就是多媒体通信技术^[1]。流媒体一般是指通过 IP 网络传输多媒体(如视频、音频)的技术总称。一般来说,流包含两种含义,广义上的流是使音频和视频形成稳定和连续的数据流和回放流的一系