中国科学院研究生院

博士学位论文

计算耗费可控的率失真性能联合优化视频编码技术研究

苏 荔

指导教师 高 文 教授

中国科学院研究生院

申请学位级别 工学博士 学科专业名称 计算机应用

论文提交日期 2009 年 3 月 论文答辩日期 2009 年 5 月

培养单位 中国科学院研究生院

学位授予单位 中国科学院研究生院

答辩委员会主席 王鼎兴教授

学位论文中文摘要

由于终端设备计算性能和电力供应的约束，设备的计算耗费（即，处理器载荷）成为制约移动视频通信发展的瓶颈。基于传统率失真理论的视频编码优化的研究重点，在于寻找“码率”与“失真”二者之间的均衡关系，没有将节省编码计算耗费作为优化目标。因此，如何针对无线视频通信的特点建立一个高效、鲁棒、低耗的信源编解码系统，存在着巨大的商业前景和重要的学术价值。本文研究的计算耗费定量可控的率失真性能联合优化的视频编码系统，就是在这种背景下提出的。本文的主要研究成果包括：

1. 提出了一套计算资源受限条件下率失真性能联合优化的视频编码方法。

该方法可用于在终端设备电力供应或计算性能受限的应用环境中实现计算耗费精细可调的视频编码器。首先将设备的功耗或计算限制转化为由缩放的计算单元（CU）数目定量衡量的编码计算量耗费，然后重点解决有限的计算资源分配和利用的问题。本方法调整精度高，实时性强，在计算资源受限的约束下可尽量减小对率失真性能的损耗。

2. 提出了一种编码计算资源自适应优化分配的模型。

该模型用以解决有限的计算资源优化分配问题。提出了虚拟计算资源缓冲区（VCB）概念，可参考已完成编码的视频帧历史参数信息，结合编码器当前可用的有限计算资源，在开始编码当前帧之前，自适应地对每帧所需的计算资源进行预估计和预分配。该方法摆脱了传统的内容自适应分配算法需要对视频内容进行预识别和分类的思路，因此实时性强、控制精度高、率失真性能平稳。

3. 提出了一套计算量耗费自适应、定量可调的编码方案。

该方案可以充分利用所分配的有限计算资源优化编码率失真性能。针对混合视频编码器实现中耗费计算资源最多的“运动估计”和“模式选择”两大模块，设计了计算量耗费分层调整算法：帧级提出了计算量耗费可调的运动估计算法（CAME），宏块级提出了计算量耗费可调的模式选择算法（CAMD），通过参考已经完成编码的前一帧或周围宏块信息，适当利用视频序列在时域、空域上的相关特性，自适应地完成计算耗费的调整。

4. 提出了一套应用于解码器端的视频流实时错误隐藏方法。

该方法用以改进由于信道传输错误带来的接收端视频质量下降的问题。该方法包含简单有效的场景信息检测和运动活跃性检测模块，可快速、自适应选择视频内容在时域和空域的相关性特征进行丢失信息的恢复。同时由于采用了分层执行策略，避免了同类算法附加复杂度高的缺点，实验表明，在主观和客观的错误隐藏效果上均明显优于现有的 H.264/AVC 非标准解码器。

5. 实现了一个基于新一代编码标准的、具有实用价值的视频通信验证系统 。

该系统以新一代视频编码标准 H.264 和 AVS 标准为基础，参照国际通行的测试规范，实现了一个完整的无线视频通信实例：从可自适应定量调节计算耗费的编码器、到模拟丢包环境的差错传输信道、再到支持实时容错算法的解码器，为实现率失真性能优化的编码计算耗费控制和容错控制提供了一个具有实用价值的参考方案。

学位论文英文摘要

The computational costs (i.e. processor workloads) to encode the video content become the bottleneck of wireless video communication, mainly due to the constraint of computational capability and power supply for portable devices. The encoders that based on the traditional rate-distortion theory focus on looking for the balance between "rate" and "distortion". However, its optimization target does not consider saving of power and computational costs. Therefore, there are great commercial prospects and important academic value to build a source codec system with efficiency, robust and low computational consumption for wireless video communication. This dissertation studies the computational costs adjustable joint rate-distortion optimal video coding system. Main research achievements are listed as follows:

1. Proposes a joint optimal video coding scheme considering both rate-distortion performance and computational costs under the constraint of computational resource.

The scheme can be used to design the computational costs adjustable encoder under computational capability constrained environments. Firstly, the algorithm transforms the constraint of power or computational resources to encoding computational costs that quantitatively measured by the number of scalable Computation Unit (CU). And then it focuses on solving the problem of the distribution and the utilization of limited computing resources. The algorithm performs high adjusting precision of computational costs and low additional complexity; at the same time it tries to reduce the loss of rate-distortion performance with limited computational resources.

2. Proposes an adaptive distribution model to solve the optimized distribution problem of limited computational resources.

We propose a concept termed as Virtual Computation Buffer (VCB), which can adaptively carry out pre-estimation and pre-distribution for each frame before encoding, with combination of knowledge of the historical parameters of finishing coded frames and the limited computing resources. The proposed model avoids the pre-identification and pre-classification for video content required by conventional approach. Therefore, it has the advantages of high real-time, high control precision, small fluctuation of rate-distortion performance.

3. Proposes a solution of real-time encoder design strategy with adaptive and quantitative adjusting of the computational costs.

The solution can fully use the limited computing resources to achieve optimal rate-distortion performance. Considering two modules consuming most computational costs in a typical hybrid video coding framework -- motion estimation module and mode decision module, we propose the computational costs adjustable algorithms hierarchically. On the frame level, we propose a Computational costs Adjustable Motion Estimation (CAME) algorithm. On the macro block level, we propose a Computational costs Adjustable Mode Decision (CAMD) algorism. These algorithms adaptively adjust the computational costs hierarchically, considering the encoded information of the previous frame and neighboring macroblocks, i.e. the relevant characteristics of video sequences on time and space field.

4. Proposes a set of approaches for real-time error concealment of video stream in decoder.

The proposed method can solve the decline of the video quality cased by channel transmission errors. The proposed method includes the scene change detection and motion activity detection, which can carry out effective error concealment by adaptively select the relevant information of time-field or space-field. Based on the hierarchical scheme, it has low additional complexity. Simulation experiments show that the improved algorithm outperforms the H.264/AVC non-normative decoder on both subjective and objective effects.

5. Accomplishes a valuable wireless video communication system based on the state-of-the-art video coding standards.

The system includes the computational costs adjustable encoder that can adaptively select joint rate-distortion optimal coding parameters, the transmission channel that can simulate package-losing environment, and the decoder that uses high-performance error concealment algorithm. The experimental platform based on the H.264 and AVS standards, providing a valuable solution for the computational costs adjustable joint rate-distortion optimal video coding system.