摘　要

无线网络和因特网等传输信道的不可靠性向高效的视频编码和传输提出了挑战，其原因在于，在有噪信道中传输压缩视频流容易产生严重的误差积累，进而影响视觉质量。目前绝大多数网络本身提供了差错控制机制，但应用于视频传输时存在过保护或欠保护的问题，有时还需要反馈信道的支持。针对视频的特点，实现容错的视频编码与传输可以获得更好的服务质量。

本文主要研究容错视频编码中的率失真优化问题，首先对传输系统的特性进行准确的理论分析，结合信源统计特性、信道特性和解码端的错误隐藏方法推导端到端失真估计函数。在此基础上，研究联合信源信道率失真模型和影响端到端率失真特性的各项因素，以及如何优化控制编码参数以提高端到端的传输效率。具体而言，本文的主要研究成果包括以下几个部分：

(1) 提出了高效低复杂度的端到端失真估计算法。端到端失真的估计是实现容错编码的关键，传统的方法通常假定端到端失真为信源失真和信道失真之和，在信道错误率较大时这一假设并不准确。本文基于理论推导，将端到端失真划分为信源失真、错误积累失真和错误隐藏失真，并进一步划分为若干个失真项，每个失真项都具有明确的物理意义，可以直接或递归计算。所推导的模型可以跟踪来自先前帧中的错误积累，即使信道错误率较大仍然保持估计的准确性，并且可以消除子像素运动补偿、环内滤波等技术对理想模型的影响。大量的实验结果证实所提出的端到端失真估计算法在帧级别上和宏块级别上都相当准确。理论分析也可以证实所提出的方法具有很低的计算复杂性。

(2) 提出了联合信源-信道率失真优化及编码控制策略。本文通过扩展传统率失真函数的定义，在率失真优化中引入误码因素，根据当前宏块的量化系数、运动矢量、可恢复性和编码所用的比特数确定当前编码模式下的编码代价，在同时考虑编码效率和容错性能的前提下实现总体优化编码。基于提出的端到端失真模型，首次精确推导出端到端失真与码率之间的关系，并得到丢包环境下的拉格朗日乘子。基于端到端失真模型与导出的拉格朗日乘子，提出联合信源-信道率失真优化算法的一般框架。实验结果与分析表明，所提出的方法比其它同类方法效率更高，且计算复杂性大大降低。

(3) 提出了基于率失真优化的混合容错编码方法。在视频编码中，多种不同的容错工具可以分别实现某种程度的差错控制，但其联合实现往往达不到预期的目的。本文通过对常用容错工具的性能分析，提出不同容错编码工具的联合优化算法，其核心在于将提出的端到端失真模型扩展到混合容错编码框架中，采用基于大数定理的方法估计与错误隐藏有关的失真项。针对灵活宏块排序和数据划分两种容错工具，进一步提出结合率失真优化的混合容错编码方案，在端到端率失真优化的框架下，将编码参数选择与不同的容错编码工具结合使用，可以设计更为高效的混合容错编码方案，达到了更好的容错

性能。

(4) 提出了面向无线信道的自适应打包方法。本文首先考察视频流打包机制对无线视频传输的端到端失真以及失真方差的影响。理论与统计分析结果表明，对于某些视频图像帧，视频打包机制对包丢失时错误隐藏失真的影响远大于对包正确接收时期望失真的影响，并且严重地影响了失真方差。由于所提出的端到端失真模型已包含了独立表示的错误隐藏失真项，因而更易于控制比特流的打包。有鉴于此，本文提出一种自适应的打包策略，在编码每帧前，根据该帧的错误隐藏失真决定划片与打包的机制。本文所提出的打包策略可以有效地改善传输效率，对接收端的质量波动改善尤其明显。

关键词：视频压缩，容错编码，视频传输，率失真优化，端到端失真估计

Study on Rate-distortion Optimization in Error Resilient Video coding

The transmission of compressed video over wireless networks as well as the Internet presents many new challenges due to the problems caused by channel errors. Most of today's video coding standards are based on a hybrid coding framework. In the error-prone environment, transmitting the hybrid-coded video may suffer from the error propagation and thereby lead to error drifting. While the transmission errors can be controlled in the network layer, the error-resilient source coding can also reduce the network burden and becomes even more efficient, and hence achieve better Quality of Services (QoS).

This dissertation targets at the rate-distortion optimized error-resilient video coding. The typical video transmission system is first studied, including the theoretical analysis and modeling of the transmission errors. The end-to-end distortion function is then derived in terms of the source and channel characteristics as well as the employed error concealment method. Based on the end-to-end distortion model, the joint source-channel rate-distortion optimized error-resilient coding is proposed, including the selection of coding control parameters. The contributions of this dissertation are summarized as below.

Firstly, a low-cost yet efficient end-to-end distortion estimation approach is proposed. The end-to-end distortion estimation plays a key role in error-resilient coding. The traditional end-to-end distortion models usually simply add the source and channel distortions, which are inaccurate especially when the error rate is large. Distinctively, the proposed distortion model associates all distortion items (including source, error-propagation and error-concealment distortions) with the error rate in theory. Each distortion item is further separated into several small distortion items that can be calculated either directly or recursively. In particular, the recursive calculation scheme is employed to trace the error propagation from all previous frames. Since the overall distortion in the proposed model is taken as the sum of several distortion items, it can suppress the approximation errors from sub-pixel motion-compensated prediction and in-loop filter. The proposed model can be easily extended to the block-level implementation. Experimental results show that the proposed end-to-end distortion model has very high precision with either block-level or pixel-level implementation, while its computational cost is also very low.

Secondly, a joint source-channel rate-distortion optimization method is proposed. The traditional rate-distortion function is extended in terms of the above end-to-end distortion model, in which the coding parameters such as quantization stepsize and

motion vectors are jointly considered for the overall rate-distortion optimization. The Lagrange method is employed due to its wide adoption. Intuitively, the Lagrange multiplier should be related to the channel conditions. Thanks to the proposed distortion model, for the first time, the Lagrange multiplier in packet-loss environment is derived in theory. Based on the derived end-to-end distortion and the new Lagrange multiplier, a generalized joint source-channel rate-distortion estimation framework is developed. Experimental results show that the proposed method can achieve high efficiency while it has much lower computational cost.

Thirdly, a rate-distortion optimized hybrid error-resilient video coding approach is proposed, in which the various error-resilient coding tools are jointly optimized. The typical error-resilient tools are statistically analyzed in terms of the relationship between the source of redundancies and the final quality at the decoder side. The statistics show that it is less efficient in the control of error propagation by separately optimizing these coding tools. Accordingly, jointly using these coding tools leads to a further optimization problem. Thanks to generalized framework of the proposed end-to-end distortion model, it is feasible to consider the influence of error concealment at the encoder side. In particular, the flexible macroblock order (FMO) and data partitioning tools are jointly optimized with the help of the extended end-to-end distortion model.

Finally, an adaptive packetization method is proposed, targeting at the video transimission over the wireless network. Above all, the mean and the variance of overall end-to-end distortion are studied. The statistics show that the error-concealment distortion and its variance are more sensitive to the packetization scheme rather than the overall distortion. Thanks to the proposed end-to-end distortion model that represents the error-concealment distortion separately, it is readily feasible to have the adaptive packetization mechanism. In particular, the slicing and packetization parameters are decided according to the calculated error-concealment distortion prior to the coding of the current frame. The proposed approach significantly improves the overall transmission efficiency. In particular, the fluctuation of reconstruction quality at the decoder is reduced.

Keywords: video compression, error resilient coding, video transmission, rate-distortion optimization, end-to-end distortion estimation