摘　要

随着视频编码技术和标准化的进步与发展，以及网络基础架构和各种 PC、移动设备计算能力的提高，与数字视频相关的各种应用越来越深入人们的日常生活，例如从多媒体消息、视频电话、视频会议，因特网和无线网络上的流媒体服务，到标清和高清分辨率的电视广播等。可伸缩视频编码 SVC 具有时域、空域和质量可分级能力，更加适应这些新型视频应用的发展。具有质量可伸缩性质的视频码流尤其适合在带宽不稳定、误码率较高的因特网和无线网络传输。现在有越来越多的视频通讯应用为了保证通话参与者能获得满意的视频交互感，既要求高质量的视频图象更要求这种服务的低延时性。因此，本文将研究如何提高具有低延时特性的质量可伸缩视频编码的编码效率和错误鲁棒性能。此外，由于隔行扫描技术和隔行视频源仍被广泛地使用，本文也将结合 SVC 和隔行视频序列的特点，研究如何改进隔行序列的质量可伸缩编码性能。本文的主要内容和创新点如下：

通过对视频编码标准 MPEG-4 和 SVC 中现有 FGS 编码技术的学习，提出了一个增强的质量可伸缩视频编码方法 CFGS（Cycle-based FGS）。类似于 SVC 中的 AR-FGS，具有低延时性的 CFGS 在进行质量增强层中的运动补偿时使用自适应参考帧。不过由于 AR-FGS 没能很好地解决错误累积传递问题，其编码性能的提高受到了限制。理论分析和实验结果表明质量增强层中的高频数据对提高编码效率的作用有限，并且更容易在网络传输时被迫丢弃，导致严重的错误累积传递，影响最终解码出的图像质量。考虑到质量增强层中各部分数据的优先级高低，以及其对预测图象质量的贡献大小，应该将它们分别进行重构之后为相对低（高）频数据赋予较高（低）加权系数（不大于 1），然后再和基本层数据有选择地组成自适应参考帧；这样可以很好地平衡编码效率和错误恢复能力两者之间貌似矛盾的关系。因此 CFGS 可以为因特网和无线网络中要求低延时的流媒体服务提供更好的编码性能和更强的错误鲁棒性。

通过对隔行视频序列的特性和 SVC 的预测/分解结构的综合分析，提出了在 SVC 编码框架下编码隔行视频序列的更高效的 FGS 预测编码结构。SVC 特有的 MCTF 分解架构/等级 B 图象预测结构可以有效地消除非关键场图象之间的时域冗余，但是关键场图象之间的，尤其是在同一帧内的顶场和底场之间的时间相关性却不能很有效地消除。为了解决这个问题，本文提出在编码质量基本层的关键场图象时使用高质量参考帧，即采用开环预测结构，来取代原有的低质量参考帧闭环预测结构，由此编码效率可以明显地提高。不过为了在编码效率和错误鲁棒性能之间取得合理的平衡，关键场图象应该交叉地隔一帧进行一次开环预测，剩

余另一半关键场则采用闭环预测来防止可能出现的错误累积传递。改进后的隔行序列FGS编码方法不仅可以在保证错误恢复能力的前提下获得更高的编码效率，还可以提供增强的时域可分级能力，并且不增加任何编解码复杂度。

综上，本文的研究内容是基于新一代的可伸缩视频编码（SVC）平台，针对两种不同的应用背景，提出相应的更加有效的质量可伸缩编码方案，使之能够更加灵活地适应网络环境的变化，提供更高的视频图象质量和错误鲁棒性能。

关键词：FGS；SVC；质量可伸缩；错误鲁棒性；低延时；部份/完全重构；隔行编码

Study on SNR Scalable Video Coding

Yanyan Zheng (Computer Application)
Directed By Debin Zhao

The advances in video coding technology and standardization along with the improvements of network infrastructures and the increasing growth of computing power are bringing digital video into our daily lives. Application areas today range from multimedia messaging, video telephony, and video conferencing over mobile TV, wireless and Internet video streaming to standard- and high-definition TV broadcasting. An attractive solution to these challenging modern video applications is scalable video coding (SVC) with temporal, spatial, and quality scalability. SNR scalability is particularly desirable for video transmission over Internet and wireless networks with unpredictable bandwidth variations and relatively high packet loss rate. Nowadays, there are more and more requirements for real time video communications applications that demand low delay to ensure good interactions among the participants and viewers with satisfying video quality. Besides, the interlaced scanning technology has persisted in many camera and display designs, and the interlaced video contents are still widely used in digital TV broadcast and storage applications. Therefore, in this thesis, we study how to enhance SNR scalability with high coding performance and strong error resilience for low-delay applications and interlaced video coding. The main contributions of this thesis are as follows:

Based on the in-depth study for the current FGS coding techniques in MPEG-4 and SVC, an enhanced SNR scalable coding scheme (CFGS, Cycle-based FGS) is developed. For low-dealy applications, as in AR-FGS, an adaptive reference is used in CFGS to perform motion compensation in the SNR enhancement layer. However, the inherent error drifting problem of AR-FGS inhibits its further improvement of coding performance. Theoretical analysis and experimental results show that the contribution of high-frequency data to coding efficiency is limited, and the data with high cycle index is more inclined to get truncated and result in error drifting and low coding performance. Considering the priorities among those cycles of the SNR enhancement layer and their contributions to the quality of the generated prediction picture, high (low) leaky parameters (not larger than 1) are set for relatively low (high) cycles to selectively form the adaptive reference with base layer data. In this way, CFGS can

flexibly balances between the compression efficiency and error robustness. With better coding performance and stronger error resilience, CFGS is more suitable for Internet and wireless video streaming applications with low-delay demand.

Based on a detailed analysis on the features of interlaced video sequences and the prediction/decomposition scheme in SVC, an improved FGS coding scheme for interlaced SVC is realized. Although the decomposition architecture of MCTF or the prediction structure of hierarchical B in SVC can efficiently reduce the temporal redundancy among the non-key fields as done for progressive sequences; the temporal correlations between the key fields in each GOP, especially those between the top and bottom fields in the same frame, can not be exploited efficiently. To solve this, an improved interlaced FGS coding scheme is proposed, which perform open-loop prediction with higher-quality references when coding the base layer of the key fields, instead of close-loop prediction with lower-quality references. In this way, the coding efficiency can be significantly improved; but in order to keep the error resilience, only half key fields predict in open loop, while the other half in close loop to prevent the possible error drifting. Moreover, besides achieving higher coding efficiency without adding any encoding/decoding complexity, this proposed scheme can also provide additional temporal scalability.

To sum up, in this paper, based on the new developed Scalable Video Coding scheme, two high efficient FGS coding techniques are proposed to suit appropriate important application scenarios. The enhanced FGS coding methods can more flexibly adapt to the changing network, and provide better coding quality and stronger error robustness.

Keywords: FGS, SVC, SNR scalability, error robustness, low delay, partial/complete reconstruction, interlaced coding