摘　要

多视点视频由多个摄像机从不同位置和角度对同一场景或物体进行拍摄得到，相对于单视点视频，多视点视频可以提供同一场景或物体多角度的信息，使人们获得更为丰富的视觉体验，因此多视点视频技术的研究受到越来越多的关注。多视点视频数据量庞大，传输时需要更多的带宽。为提高压缩效率，减少传输信号所需带宽，往往需要利用多个视点之间的冗余，这需要各个视点的编码器之间进行大量的数据交换。然而，在实际的系统中，各个视点的编码器之间大规模的数据交换十分困难。此外，许多应用需要在无线分布式的环境下获取多视点视频数据并进行传输，编码方法的计算复杂度和容错能力都是必须重点考虑的问题。这些都对传统的视频编码方法提出了挑战。

分布式视频编码可以很好地解决多视点视频技术中存在的以上问题。分布式视频编码的理论基础是 Slepian-Wolf 理论和 Wyner-Ziv 理论。Slepian-Wolf 理论证明了在无损压缩的情况下，如果两个相关信源在编码端独立编码而在解码端联合解码，那么它能够达到与联合编码相同的编码效率。Wyner 和 Ziv 将这一理论扩展到了有损压缩的情况，并且认为这一理论可以通过信道编码的方式实现，编码器对信源进行信道编码，编码后可以只传送纠错位，解码器借助边信息进行解码，可以逼近联合编码的码率-失真性能。分布式视频编码引起了国内外众多研究者的广泛关注。本文分析了分布式多视点视频编码的特点，对其中存在的问题进行了深入的研究，并提出了相应的解决方案。

本文的主要工作及取得的成果如下：

1. 提出了基于运动模型融合的边信息生成方法

多视点视频中，各个视点存在很强的相关性，在运动信息方面表现为各个视点的运动矢量具有很强的相似性。在基于运动模型融合的边信息生成方法中，我们利用视间运动模型描述各个视点运动矢量的相关性，并通过它估计当前视点的运动矢量，从而生成当前视点的边信息，这种方法能够很好的生成运动强度较高区域的边信息；同时，我们利用线性运动模型对线性运动区域的边信息进行增强，进而提高边信息的整体质量。

2. 提出了基于区域的相关性模型

分布式视频编码中，相关性模型描述原始帧与边信息的统计关系。从信道编码的角度看，相关性模型描述边信息中噪声分布的统计特性，为信道解码提供软输入信息。在基于区域的相关性模型中，我们根据预测的相似性，对边信息中噪声信号的分布进行估计，并且通过独立建模的方式，体现不同区域噪声信号分布的差异, 从而为信道解码提供更为准确的软输入信息，提高解码效率。

3. 提出了基于多假设的解码技术

在多视点视频中，通过利用时域相关性以及视间相关性可以在解码端得到多个边信息。分布式视频编码中采用信道编码作为熵编码器，因此充分利用边信息与原始帧之间的统计相关性是提高熵编码效率的关键。在基于多假设的解码技术中，我们通过充分挖掘多边信息之间的相关性以及它们与原始帧之间的统计相关性提高解码效率以及重构图像质量。

4. 提出了基于人眼视觉特性的分布式视频编码技术

去处视觉冗余是提高视频编码效率的重要手段。在基于人眼视觉特性的分布式视频编码中，我们在编码端通过低复杂度的插值对边信息进行近似的估计，并通过 JND(Just noticeable distortion)模型评估边信息中噪声信号的视觉特性。根据噪声信号的视觉特性，对不同区域采用不同的编码策略；对于包含人眼可察觉噪声的区域，对其进行实际的 Wyner-Ziv 编码；其它区域则不进行实际的编码，其原始信号由边信息直接恢复。该方法可以有效地去除视频信号中的视觉冗余，从而在保持主观视觉质量的前提下，降低编码码率。

.

关键词：分布式视频编码；多视点视频编码；边信息生成；相关性模型；基于视觉的编码；分布式多视点视频编码

Abstract

Yongpeng Li (Computer Application)

Supervised By Prof. Wen Gao

Multi-view video, which is obtained from multiple cameras capturing the same scene from different positions and viewpoints, can provide more detailed and vivid information regarding to the scene being captured. For this reason, multi-view video is receiving increasing attentions. Usually, more bandwidth is required to transmit multi-view video due to the large amount data involved. Therefore, it is important to improve the coding efficiency of multi-view video coding. The redundancy among the neighbor views is usually exploited to achieve this. However, this requires the inter-view camera communication, which is hard to perform in practice. Besides, in many applications multi-view video has to be captured and transmitted in distributed wireless environment, so the encoding complexity and robustness of the encoding scheme has to be considered. All these pose great challenges to the traditional hybrid video coding framework.

Distributed video coding is a promising solution to the problems mentioned above. Distributed video coding is built on Slepian-Wolf theory and Wyner-Ziv theory. Slepian and Wolf first proved that although two statistical sources are independently encoded, the similar performance can be achieved as long as the joint decoding of them is allowed for the lossless coding. Wyner and Ziv extended the distributed source coding theory to the lossy coding, and proved that the Slepian-Wolf theory can be realized by channel coding. The encoder encodes the original signals with channel codes and only transmits the pairty bits to the decoder. The decoder recovers the original signals by the transmitted parity bits and side information. Recently, distributed video coding is receiving more and more attentions. This paper analyzes distributed multi-view video coding and develops several algorithms to improve its coding efficiency.

Main research achievements are listed as follows:

1. A fused motion model based side information generation method is proposed

In multi-view video, the motions of each view are highly correlated with each other. We model such correlation with inter-view motion model, and use it to generate side information. It works well for side information generation of the regions with high motions. Besides, we use linear motion model to enhance the side information generation in the region with low motion, which usually obeys the linear motion rule. In this way, the side information quality can be improved greatly.

2. A region based correlation channel model is proposed

In distributed video coding, the statistical relation between the original frame and

side information is modeled by the correlation channel model. From the viewpoint of channel coding, the correlation channel model is used to describe the distribution of the noises in side information. The proposed correlation channel model divides the current frame into two regions based on the prediction accuracy. The noise distributions of two regions are assumed to be distinguished, and they are modeled separately. This will provide more accurate soft input information in channel decoding, thus the decoding efficiency can be improved.

3. A multi-hypothesis decoding algorithm is proposed

For multi-view video, multiple side information can be generated by using temporal correlation and inter-view correlation. In distributed video coding, channel coding is used as entropy coding, so coding efficiency can be improved by exploiting the statistical redundancy between side information and the original frame. The proposed algorithm can improve the decoding efficiency and improve reconstruction by exploiting the correlation between the multiple side information and the statistical redundancy between them and the original frame.

4. A distributed video coding scheme based on the human visual system is proposed

The coding efficiency can be improved by removing the visual redundancy. In the proposed scheme, we use a low-complexity interpolation method to approach the side information at the encoder, and use the JND model to evaluate the visual characteristics of the noises in side information. Practical Wyner-Ziv coding scheme is only applied on the region where the noises can be perceived by human eyes, and the original signals of the rest region are recovered via their corresponding side information directly. By removing the visual redundancy in video signals, the proposed scheme can save the bits-rates significantly without degrading the subjective quality of the reconstructed frames.

Keywords: Distributed video coding; multi-view video coding; side information generation; correlation channel model; perceptual coding; multi-view distributed video coding.