摘要

互联网与数字多媒体技术的发展，推动了超高清视频的普及。超高清视频在极大丰富了人们的物质生活和精神享受的同时，也导致了数据量的爆炸式增长，视频的存储与传输面临着严峻的挑战，迫切需要更高效的视频编码方法。在视频编码的混合编码框架中，对预测残差的编码一般由变换、量化、熵编码等几个模块组成，其目的是通过降低残差数据之间的相关性来减少冗余信息。探索更加高效的残差编码方法有助于提升视频质量和编码效率，降低编码开销。本文面向新一代视频编码标准 AVS3（The third generation of Audio Video coding Standard），统计分析残差在空域和频域的不同分布特性，研究更加紧凑的表示方法，从而提高编码性能。本文主要创新点如下：

第一，提出一种基于帧内残差频域分布特性的隐式选择变换方法，提升了帧内预测残差的编码效率。为了解决 DCT-II 变换无法适应残差的多样分布特性的问题，本文引入了更适合帧内残差的 DST-VII 变换核作为变换候选类型。但是引入新变换后会导致额外标识开销的新问题，本文通过统计分析变换系数的分布特性，提出了标识代价更低的隐式选择方法以消除标识开销，并设计了编码端最优的系数调整策略以达到降低内容失真的效果。同时，结合隐式选择和显式标识两种表示方法各自的优势，实现了对两种表示方法的自适应选择。在 AVS3 参考软件平台上进行了测试，实验结果表明该方法可以在帧内（AI）和随机访问（RA）配置下分别获得平均 1.76% 和 0.76% 的性能增益，同时解码复杂度没有增加。所提方法已被 AVS3 标准采纳。

第二，提出一种基于帧内残差空域分布特性的变换类型预测方法，进一步提升了帧内预测残差的编码效率。为了解决单一变换核无法适应帧内残差的多样分布的问题，本文引入了基于广义图拉普拉斯矩阵计算得到的适应残差分布的新变换核，提出了增强二次变换方法。但是在加入更多变换核后，编码变换类型标志位会严重影响变换性能，本文进一步设计了解码端预测变换类型与残差符号的方法。该方法利用周围重构像素与当前块边界像素之间的一致性，降低了编码类型标志位的码率。在 AVS3 参考软件平台上进行了测试，实验结果表明该方法可以较为准确地对残差正确项进行预测，降低编码标志位的代价，在 AI 和 RA 配置下分别可以获得 0.86% 和 0.44% 的性能。所提出的增强二次变换方法已被 AVS3 标准采纳。

第三，提出一种基于屏幕内容残差空域分布特性的自适应变换跳过方法，提升了屏幕内容残差的编码效率。为了解决变换导致的屏幕内容残差能量分散的问题，本文通过分析屏幕内容残差的能量分布特性，引入了更适合屏幕内容残差分布的变换跳过方式作为变换候选类型，并设计了帧级自适应的隐式选择变换跳过方法。同时，基于

残差的空域分布特性，针对普通帧内残差，设计了风车形重排方法；针对帧间和帧内块复制预测的残差，设计了部分变换跳过方法。在此基础上，对三种子方法进行组合提升，提出了自适应变换跳过方法。在 AVS3 参考软件平台上进行了测试，实验结果表明该方法可以在 AI、低延迟（LD）和 RA 配置下，分别获得 12.79%、11.00% 和 12.04% 的性能增益，同时可以降低编解码复杂度。所提方法已经被 AVS3 标准采纳。

综上所述，本文针对残差的不同分布特性，研究了不同类型残差的空域分布和频域分布的特性。面向新一代视频编码标准 AVS3，提出了三种高效编码方法，分别为基于残差频域分布特性的隐式选择变换方法、基于残差空域特性的变换类型与残差符号预测方法以及基于屏幕内容残差空域分布特性的自适应变换跳过方法。所提出的方法均被 AVS3 标准采纳，取得了显著的性能增益，为高效编码提供了理论依据和实验证明。

关键词：视频编码，变换编码，变换跳过，二次变换，AVS3

Research on High Efficiency Video Coding based on the Residual Distribution Characteristics

Yuhuai Zhang (Computer Applications)
Directed by Prof. Wen Gao
Prof. Siwei Ma

ABSTRACT

The development of the Internet and digital multimedia technology has promoted the popularization of ultra-high-definition video. While ultra-high-definition video greatly enriches people's material life and spiritual enjoyment, it also leads to the explosive growth of data volume. Video storage and transmission face severe challenges, and more efficient video coding methods are urgently needed. In the hybrid coding framework of video coding, the coding of prediction residuals generally consists of several modules such as transformation, quantization, and entropy coding, to reduce redundant information in residual data. Exploring more efficient residual coding methods is beneficial to improving video quality and coding efficiency, and reducing overhead bits. In this thesis, the different distribution characteristics of residuals in the spatial and frequency domains are statistically analyzed, and a more compact representation method is researched to improve the coding performance for the new generation of video coding standard AVS3 (The third generation of Audio Video coding Standard). The main innovations of this thesis are as follows:

First, this thesis proposes an implicitly selected transform method based on the frequency domain distribution characteristics of intra-coded residuals to improve the coding efficiency. The widely used DCT-II transform cannot adapt to the diverse distribution characteristics of intra-coded residuals. In order to solve this problem, this thesis introduces the DST-VII transform kernel that is more suitable for intra-residuals as a transform candidate. However, introducing a new transform core will bring additional identification overhead. By statistically analyzing the distribution of the transform coefficients, this thesis proposes an implicit selection method to eliminate overhead bits and designs the optimal coefficient adjustment strategy at the decoder to reduce the content distortion. In addition, the adaptive selection of the two representation methods is realized by combining the respective advantages of the two representation methods, which further improves the coding performance. We validated the

proposed method on the AVS3 reference software platform, the experimental results show that the method can achieve an average performance gain of 1.76% and 0.76% in All-Intra (AI) and Random Access (RA) configurations, respectively, with negligible decoding time variations. The proposed method is the first to hide the transform type into the coefficients and has been adopted by the AVS3 standard.

Second, this thesis proposes a transform type prediction method based on the spatial distribution characteristics of intra-coded residuals to further improve the coding efficiency. To adapt to the diverse distribution characteristics of intra-coded residuals, this thesis introduces new transform kernels based on the generalized graph Laplacian matrix and proposes an enhanced secondary transform method. However, after introducing more transform cores, signaling flag bits to indicate transform types will seriously affect the transform performance. This thesis further designs a method to predict transform type at the decoder side. This method utilizes the consistency between the surrounding reconstructed pixels and the current block boundary pixels to reduce the bit rate of the indication flag bits. We verified the proposed method on the AVS3 reference software platform, and the experimental results show that the method can accurately predict the correct residuals, reduce the indication overhead bits, and achieve 0.86% and 0.44% performance under AI and RA configurations, respectively. The proposed enhanced secondary transform method has been adopted by the AVS3 standard.

Third, this thesis proposes an adaptive transform skip method based on the distribution characteristics of screen content residuals to improve the screen content coding efficiency. Transform will disperse the energy of the screen content residual, which is not conducive to the subsequent quantization and entropy encoding process. In order to solve this problem, this thesis introduces a transform skip method that is more suitable for screen content residuals as a transform candidate type by analyzing the spatial distribution characteristics of screen content residuals. A frame-level adaptive implicitly selected transform skip method is designed to save the indication overhead bits. In addition, based on the distribution characteristics of residuals, a windmill-shaped rearrangement method is designed for intra-coded residuals, and a partial transform skip method is designed for inter-coded and intra-block-copy-coded residuals. Based on the three sub-methods, an adaptive transform skip method is proposed. We validated the proposed method on the AVS3 reference software platform, and the experimental results show that the method can achieve performance gains of 12.79%, 11.00%, and 12.04% under AI, low latency (LD), and RA configurations, respectively, and the encoding and decoding time is reduced. The proposed method is the first to combine the coefficient-hiding method and content-adaptive method and has been adopted by the AVS3 standard.

In summary, this thesis researches the spatial distribution and frequency distribution characteristics of different types of residuals according to the diverse distribution characteristics of residuals. For the new generation video coding standard AVS3, an implicitly selected transform method based on the residual distribution characteristics in the frequency domain, a transform type prediction method based on the spatial distribution characteristics of intra-coded residuals, and an adaptive transform skip method based on the screen content residual distribution characteristics are proposed. The proposed methods are adopted by the AVS3 standard and achieve good performance gains. Moreover, these methods provide a theoretical basis and experimental proof for high-efficiency coding.

KEY WORDS: Video coding, Transform coding, Transform skip, Secondary transform, AVS3