摘要

随着多媒体、人工智能时代的来临，图像和视频成为越来越主流的信息交流方式。除了传统多媒体时代中广播电视、网络视频、视频通讯、监控视频等应用之外，人工智能领域的检测、检索、行为识别等也离不开图像和视频。同时，视频数据的产生量与日剧增，人类已经进入了大数据时代，从而对视频信息的压缩效率提出了更高的要求。在此背景下，在过去几十年的视频编码发展过程中，视频编码标准已经经历了若干代的更替、演化、发展，使得视频压缩效率不断提高，以应对时代的需求。然而，视频编码技术如今在应用当中仍面临很多问题，其中最重要的问题就是如何用更低的计算复杂度尽可能提升编码性能。现存的一些解决方案仍缺乏鲁棒性，即在不同的应用场景下难以有稳定的性能，其根本原因在于这些算法在设计过程中缺乏理论基础。基于此研究背景，本文从视频编码理论出发，通过理论分析，对如何设计出鲁棒的提升编码性能、降低编码复杂度的编码算法做出了研究，提出了率失真模型在视频编码中的零块检测技术、多路码率分配技术以及基于感知的块级码率分配技术中的应用。其中本文的主要创新点如下：

第一，根据信息论与率失真理论，对视频编码率失真过程进行分析，阐述了率失真优化过程背后的理论依据，并由此导出了视频编码中拉格朗日系数 λ、量化系数 QP、码率、失真四个变量之间的关系，阐述了现有视频编码中率失真模型与预估的算法以及其核心理论依据，为了解决快速零块检测、多路码率分配以及基于感知的码率分配算法的现有问题，推导了一种多路视频率失真模型，根据本文的感知失真提出了一种码率-感知失真模型，提出了一种更好的模型参数预估的方法，以及提出了适用于 HEVC 的码率失真预估算法。

第二，分析了现有快速零块检测技术中存在的两难问题和不足之处，提出了混合快速零块检测算法，分别在多个编码阶段分别进行检测，其中多阶段包括整像素运动估计之后、分像素运动估计之后、离散余弦变换之后的三个阶段的检测，并且第三阶段变换之后进行的伪零块检测算法基于率失真预估原理推导得到。本文所提出的检测算法复杂度很低，和每帧像素个数成线性关系。该算法在降低 47.68%总编码时间的基础上，RD 性能损失为 0.67%。相比之下，现有最优的相关算法则在降低 35.45%总编码时间的基础上，RD 性能损失为 0.51%。上述实验结果说明本文提出的混合快速零块检测算法和现有算法在具有几乎一样的 RD 性能的情况下，进一步提升了加速算法的性能，解决了两难问题。

第三，对于最小化多路失真方差的多路视频码率分配问题，现有的基于预分析的多路视频码率分配算法存在技术瓶颈。为了突破瓶颈，本文首先推导了多路视频率失

真模型，提出了基于多路率失真模型的多路视频码率分配算法。然后，分别利用一次率失真模型和幂次率失真模型推导了两种不同的分配方法，大大提高了分配性能。其中，利用幂次率失真模型的分配结果较现有算法提升了 75.29%的性能。根据分析，所提出的基于一次模型和幂次模型的分配算法的计算复杂度很低，只和序列个数成线性关系。

第四，对于基于感知的码率分配技术，分析了目前现有算法的两个缺点：第一，没有提出合理的优化目标函数，码率分配目标不明确，缺乏理论分析；第二，无法同时兼容现有宏块级码率控制算法。为了解决这些问题，本文首先提出了一种低复杂度的适用于视频编码的感知模型，再以该模型为基础，推导出一种基于码率-感知失真模型的块级码率分配算法。该分配方式有很好的主观效果，从码率误差的均值来看，打开宏块级码控误差为 1.27%，加入本文算法后误差为 1.20%，而文献[52]和文献[55]的误差分别为 1.82%和 1.90%，可见本文提出算法可以很好的与现有码率控制算法融合，并且码率误差几乎不受影响。根据分析，所提出的基于码率-感知失真模型的块级码率分配算法计算复杂度很低，只和宏块个数成线性关系，并且该码率分配过程适用于任意感知模型。

综上所述，本文从视频编码理论出发，通过理论分析，对如何设计出鲁棒的、提升编码性能、降低编码复杂度的编码算法做出了研究。提出了混合快速零块检测算法，解决了现有零块检测算法中的两难问题。同时提出了基于率失真模型的多路视频码率分配算法，在几乎不增加算法复杂度的情况下，大大提高了现有多路视频码率分配算法的性能。最后提出了基于码率-感知失真模型的块级码率分配算法，提升了视频的主观效果，更是可以很好的与现有码率控制算法融合。

关键词：视频编码，率失真模型，零块检测，多路码率分配，感知码率分配

Study on Rate-Distortion Model and Its Applications on Video Coding

Hongfei Fan (Computer Application Technology)
Directed by Xiaodong Xie

ABSTRACT

With the approach of the times of multimedia and artificial intelligence, images and videos become a more and more popular way for communication. Besides traditional applications like broadcasting, network video, video conference, surveillance videos and so on, new artificial intelligence applications including detection, retrieval, behavior recognition and so on also depend on videos and images. At the same time, the number of video data are increasing dramatically which lead to the requirements of more and more efficient video coding algorithms. Human beings have entered the era of big data. Under such backgrounds, video coding standards have gone through the replacement, evolution, and development of several generations in the past few decades to improve the coding efficiency and meet the demands of applications. However, there is still a lot of challenge on the applications of video coding techniques nowadays. How to develop a high efficiency codec with low computational complexity is the most important research topic. Some solutions are not robust under different situations mainly because of the lack of theoretical analysis. Based on the mentioned backgrounds, in this research we focus on how to design high efficiency algorithm with robust coding performance or fast algorithm to reduce the coding complexity by theoretical analysis of video coding. We solve this problem by using rate-distortion modeling and estimation on zero block detection, joint rate allocation, and perceptual based rate allocation on macro blocks. The main contribution of this paper includes,

Firstly, this paper analyze the the process of rate distortion optimization of video coding by information theory and rate-distortion theory. The theoretical basis of rate distortion optimization is described and the relationship among λ, QP, rate and distortion are derived. We explain the derivation of related rate-distortion modeling and estimating methods theoretically. In this paper, we derive a rate-distortion model for joint sequences and a rate-perceptual-distortion model. We also propose a better parameter estimation method for rate-distortion model. Besides, we propose a rate and distortion estimation method for

HEVC. Based on the mentioned contribution, we solve the problems of related hybrid zero blocks detection methods, rate-distortion based joint rate allocation methods, and perceptual based rate allocation on macro blocks methods.

Secondly, the dilemma of related zero blocks detection methods are analyzed. To overcome this problem, we propose a hybrid zero blocks detection method in which we detect zero blocks by three phases: detection after integer motion estimation, detection after fractional motion estimation, and detection after DCT. The third phase is derived by the mentioned rate-distortion estimation methods. Meanwhile, the proposed detection method only takes little computational costs. According to the results, the proposed hybrid zero blocks detection method can achieve 47.68% time savings of the whole encoder while a BDrate loss of 0.67%. On the comparison, the related method can achieve 35.45% time savings of the whole encoder while a BDrate loss of 0.51%. We can draw a conclusion that the proposed hybrid zero blocks detection method can achieve a much higher time saving with a similar loss and the dilemma of related zero blocks detection methods can be solved.

Thirdly, we analyze the bottle neck of related joint rate allocation methods to minimize the variance of distortion among different sequences. In order to break the bottle neck of the related methods, we derive a joint rate-distortion model and then derive a rate-distortion based joint rate allocation methods. We also use inverse function model and hyperbolic function model in our proposed scheme to derive two different allocation methods. According to the results, our proposed scheme can increase the performance by 75.29% than the related allocation scheme. The computational cost of the proposed method is also analyzed. The relationship between computational costs with the number of sequences is linear which means the computational cost can be negligible.

Fourthly, we analyze the shortcoming of related perceptual based rate allocation on macro blocks methods. There is no target function of the related methods and it is difficult to combine rate control scheme with related methods together. To solve this problem, we propose a low complexity perceptual based video quality assessment at first. Based on this proposed assessment, we derive a perceptual based rate allocation method by rate-distortion models. From the experimental results, our algorithm not only has a better quality but also works well with other rate control schemes. When the LCU rate control is turned on, the error between allocated bitrate and actual bitrate is 1.27%. Meanwhile, the error is 1.20% when the proposed allocation scheme is used, and the error is 1.82% and 1.90% when the allocation scheme given in [52] and [55] are used, respectively. Our allocation scheme can perfectly match the rate control algorithm and has nearly no effect to rate control. According

to the analysis, the proposed allocation method has a low computational complexity. The allocation method can work with different video quality assessments without the concerning of extra computational cost from the allocation part.

In summary, this paper focuses on how to design high efficiency algorithm with robust coding performance or fast algorithm to reduce the coding complexity by theoretical analysis of video coding. To overcome the dilemma of related zero blocks detection methods, we use rate-distortion estimation to derive a hybrid zero block detection method. We also derive a rate-distortion based joint rate allocation method by using rate-distortion modeling, and the efficiency can be significantly improved. At last, we propose a perceptual based rate allocation on macro blocks method and the proposed method not only has a better quality but also works well with other rate control schemes.

KEY WORDS: Video coding, Rate-distortion modeling, Zero block detection, Joint rate allocation, Perceptual based rate allocation