摘　要

近年来，随着互联网技术的发展和普遍应用，网络越来越影响人们的生产生活方式。由于视频信息具有直观性强，内容丰富等特点，因此视频传输在网络传输领域占据了越来越重要的地位。视频流式传输（也称流媒体）是其中一项迅猛发展的技术，能够满足人们迅即的学习娱乐需求。然而，在互联网上进行流媒体传输，对视频压缩，网络传输等提出了新的挑战。首先，由于不同的流媒体应用，人们对流媒体的质量、时延、交互性等要求也不同，所以需要压缩视频流具有服务质量的可适应性；其次，由于网络的异构性（Heterogeneity），服务质量 QOS（Quality of service）的非确保性（Best-effort service），流媒体传输需要面对网络传输中可能遇到的各种问题，提供拥塞控制、带宽适应性和视频平滑等等；第三，用户解码设备解码能力和播放方式的多种多样，要求压缩传输的视频码流具有相应的结构来方便设备多种能力的解码和播放。所有这些的核心是要求编码和调度相配合，在一定的网络状况下，最大限度地满足客户端多种应用和多层次的需求。可伸缩性视频编码（例如 MPEG-4 FGS 和 JVT SVC）和传输是流媒体中一项非常有希望的技术。本文从最基本的概率统计，信息论和量化理论入手，系统而且深入地研究了可伸缩性视频编码和传输的关键点：率失真函数（Rate-distortion function），设计了新的率失真函数模型，并在此基础上实现了一系列优化的视频编码和调度算法。本文主要的内容和创新点包括：

由于广义高斯分布 GGD（Generalized Gaussian distribution）的高度灵活性，理论上和实际中它都可以非常好地描述离散余弦变换（DCT）和小波变换（Wavelet transform）的编码系数（Transform coefficients）。本文通过信息论和量化理论的知识，分析了广义高斯分布 GGD 率失真函数的基本属性，数学上严格证明了：以峰值信噪比 PSNR（Peak signal noise ratio）或信噪比 SNR 为失真标准，广义高斯分布 GGD 的失真速率函数（Distortion-rate function）的导数存在，并且与分布的方差无关。这样，实际中只需研究特定编码量化模式下 GGD 失真速率函数的导数与 GGD 形状参数的关系和规律，就可以重建相应编码模式下的失真速率函数。

基于位平面的精细可伸缩性视频编码 FGS 是 MPEG-4 标准支持流媒体传输的关键技术。为了能够在不同的传输码率下平滑重建 FGS 目标视频流，基于帧的 FGS 增强层率失真函数的获取是一个非常关键的问题。本文首先通过广义高斯分布 GGD 描述了 FGS 增强层 DCT 系数的分布。然后分析了 FGS 编码的量化模式，并把它应用到 GGD 分布上，得到了 GGD 分布率失真函数的变化规律，从而获得实际 FGS 编码的率失真函数的变化规律，即 FGS 编码失真速率函数的导数先随着码率的增加而逐渐减小，然后在码率比较高的时候开始随码率的增加而缓慢地增加。在以上观察和分析的基础上，本文提出了一个灵活有效的率失真函数模型去近似实际的 FGS 率失真函数。大量的实验验证了模型的

有效性。

最近，以运动补偿时域滤波 MCTF（Motion compensated temporal filter）和 H.264 编码为基础的可伸缩性视频编码 SVC 成为研究的热点，并由 JVT（Joint video team）进行 SVC 相关的标准化工作。本文同样通过广义高斯分布 GGD 对 SVC 质量增强层整数变换后的系数进行统计分析，并将分析得到的 SVC 量化模式应用到广义高斯分布 GGD 上，借以得到实际 SVC 编码率失真函数的变化规律，即 SVC 编码失真速率函数的导数一般随着编码码率的增加而逐渐减小，在少数情况下，当 SVC 增强层变换系数的 GGD 估计的形状参数都很大时，SVC 编码失真速率函数的导数随着编码码率的增加而缓慢增加。因为 FGS 率失真函数模型能同时模拟这两种情况，所以我们采用同样的率失真函数模型去近似实际 SVC 编码的率失真函数。最新的实验验证了模型的有效性和准确性。

利用已有的率失真函数模型，本文提出了一种简单而且快速的恒定质量重建调度算法。算法首先计算平均分配目标码率时所有调度单元（以视频帧或视频场景(Scene)为单位）的平均视频质量，然后利用此平均质量根据每个调度单元的率失真函数得到调度单元的初始分配码率，接着根据目标码率和初始分配码率的误差，以及调度单元率失真函数的斜率最终得到每一调度单元的码率分配。算法不需要平滑算法通用的搜索运算，计算复杂度为 O（N），大大降低了服务器计算量。算法在 FGS 编码视频和 SVC 编码视频的平滑实验上都取得了很好的效果。

为了进一步地降低视频服务器实时调度的计算量，本文还提出了 FGS 两层平滑算法。首先我们设计实现了 FGS 基本层平滑编码算法，使得视频传输中 FGS 基本层有恒定的质量，然后根据 FGS 增强层率失真函数的特点计算每一帧压缩图像的调度权值，传输时根据这个离线计算的权值快速有效地为 FGS 增强层每一帧图像分配传输码率，实现传输视频的快速平滑重建。

关键词：统计模型；率失真函数；可伸缩性编码；FGS；SVC；视频平滑

Research of key problems on video coding and streaming

Jun Sun （Computer Application）
Directed By Wen Gao

With the development of network technology, Internet has transformed our lives and the way we communicate, how we learn, how we work and spend free time, in essence – it has more or less changed every aspect of human society one can think of. Since video is one of the most compelling ways to present information, video steaming is expected to dominate the future network traffic, satisfying people’s immediate requirements of study and entertainment. However, video streaming on Internet have some new challenges on video compression and network transmission. First, video compression for streaming should be adaptive on quality of service to satisfy people’s different requirements on quality, delay and interaction. Second, video streaming over Internet should provide congestion control, bandwidth adaptation and video smoothing. Third, video compression should support different decoder capacity and different requirements of play back. All these factors require the cooperation of video compression and network scheduling to satisfy different needs of streaming applications. Scalable video coding (such as MPEG-4 FGS and JVT SVC) and streaming is one of the most promising technologies. In this thesis, we investigate the key problem of scalable video coding and streaming: rate-distortion analysis and approximation, with which we also design some simple and effective smooth quality reconstruction algorithms of video streaming. The main contributions of this thesis are as follows:

Generalized Gaussian distribution (GGD) is a nice model for the DCT coefficients and wavelet coefficients in theory and practice. With information and quantization theory, we mathematically prove that the derivative of GGD distortion-rate function exists and is not related to the variance of the distribution. So if we gain the relation between the the derivative of GGD distortion-rate function and the GGD shape parameter, then we can reconstruct the rate-distortion function.

The bitplane-based FGS of MPEG-4 has been accepted as a standard coding tool for video streaming applications. To support smooth quality reconstruction under multiple rate constraints of transmission, it’s important to obtain the actual rate-distortion (R-D) function or distortion-rate function of each enhancement layer (EL) frame. In this thesis, firstly, we introduce generalized Gaussian distributions to model the distributions of DCT coefficients of an EL frame. Then, according to quantization theory, we analyze the distortion-rate function of the generalized Gaussian model, with which we conclude that for the actual FGS coding, the derivative of actual distortion-rate function usually decreases continuously as the rate increases, and then begins to increase slowly at a comparatively high bit rate. Finally, guided by the above observations, an effective and flexible R-D model is proposed to approximate the actual rate-distortion function. Extensive experiments show the effective and accuracy of our proposed R-D model.

Recently, scalable video coding (SVC) with motion compensated temporal filtering (MCTF) has been a focus of research and application. In this thesis, we also utilize generalized Gaussian distributions to model the transform coefficients of the quality enhancement layer (EL) in SVC coding. Then according to the quantization scheme of SVC quality EL, we analyze the GGD distortion-rate function, with which we conclude that for the actual SVC coding, the derivative of actual distortion-rate function usually decreases continuously as the rate increases, and on less occasion, that the derivative of actual distortion-rate function increases as the rate increases when the GGD-estimated shape parameters of the transform coefficients are all large. Since the above FGS R-D model can simulate all these two properties, we also use it to model the rate-distortion function of SVC coding. Recent experiments show the effective and accuracy of the R-D model.

Based on our R-D model, we proposed a simple and fast algorithm for constant quality reconstruction of scalable video. The algorithm first calculates the average distortion $\bar{D}$ with uniform bit rate allocation $\bar{R}$, then calculates the initial bit rate allocation using the rate-distortion function $R_i(\bar{D})$, at last gets the actual transmitting bit rate though tuning the difference between the target bit rate $\bar{R}$ and the initial allocated bit rate $\frac{1}{N}\sum_i R_i(\bar{D})$. The algorithm doesn't need any search operation. The computation complexity of the proposed algorithm is O(N). Extensive experiments on MPEG-4 FGS and JVT SVC videos show the efficiency and effective of our algorithm.

To further simplify the real-time schedule for constant quality streaming, we propose a FGS two-layer quality smooth algorithm. In base layer (BL), we design an accurate constant quality coding algorithm to realize the smoothing of BL quality. After that a weight is assigned to each frame of FGS enhancement layer (EL), then a simple but effective weight-based rate allocation of the EL is proposed for constant quality streaming. Experiments show the efficiency of our simple method.

Keywords: Statistical model, Rate-distortion function, scalable video coding, FGS, SVC, Smooth algorithm