



北京大学

博士研究生学位论文

题目： 基于视觉特征的
图像视频编码

姓 名： 张 翔

学 号： 1301111273

院 系： 信息科学技术学院

专 业： 计算机应用技术

研究方向： 视频编码技术

导 师： 高文 教授

二零一八年十二月

摘要

大数据时代的到来对图像视频以及新媒体的压缩、处理和分析提出了更高的要求 and 新的挑战。然而，视觉特征作为重要的视觉信号处理和分析工具，在传统图像视频压缩算法中没有被有效利用，限制了编码效率的进一步提升，同时也限制了其在视觉分析等应用场景中的使用。本文以视觉特征为核心，结合编码率失真优化理论，深入研究了局部描述特征、稀疏基元特征、视图稀疏特征对图像视频以及光场的高效压缩算法。

本文的创新点主要包括以下四个方面：

- **局部描述特征与视频的联合编码。** 本文提出了将视频和视频中提取的局部特征联合压缩的编码框架。从原始视频中提取特征并编码在很大程度上保证了特征在分析检索等视觉任务中的准确率和效率，同时编码的特征可以进一步用来指导视频本身的编码。首先，针对视频局部特征的压缩，设计了针对联合编码优化的特征选择机制、多种预测模式以及特征的率失真优化模型。因为局部特征最终是用来做匹配检索的，因此通过准确估计特征失真对于匹配性能的降低，可以提高特征的编码效率。此外，利用已经编码的局部特征，提出了基于局部特征匹配的帧间运动估计模型，通过帧间特征匹配导出更加精确的仿射或者平动模型作为当前编码块的运动矢量预测，极大提高了视频的编码效率。本文还提出了针对联合编码框架的联合率失真优化策略，根据不同应用场景自适应分配视频和特征码率，最小化联合率失真代价。最后，在对象匹配、对象定位以及视觉检索等多种视觉任务中验证了联合编码框架的有效性，提出的方法可以在实现高效的视觉任务的同时提高联合编码效率。
- **基于稀疏基元特征的视觉优化编码。** 稀疏表示将自然图像分解为若干稀疏基元特征的线性组合，目标是在稀疏约束下最小化重构图像的均方误差，但是忽略了重构图像的主观视觉质量。本文根据分歧归一化的思想，设计了一种基于稀疏基元特征的图像全参考客观质量评价方法，该方法结合了稀疏基元的结构化特性，有效利用了人眼视觉系统的亮度、对比度以及结构掩蔽效应，能够对失真图像的主观质量做出准确的估计。进一步，将该质量评价方法用于优化图像压缩的主观质量，通过分歧归一化技术，将图像块在稀疏基元特征上分解得到的系数映射到视觉均衡空间，通过自适应量化机制，将码率更多地分配给图像中对人眼敏感的区域，在相同码率下显著提高了整体重构图像的视觉质量。
- **基于稀疏基元特征的率失真优化编码。** 在基于稀疏基元特征的图像编码问题中，

另一个重要问题是模型优化过程中很少考虑到稀疏基元特征系数的编码代价。本文提出了通过计算系数方差来快速有效估计系数码率的方法，并将其引入到稀疏表示的优化模型中。通过交替方向乘子法将复杂的原问题分解为若干易于求解的子问题，迭代更新稀疏基元特征的系数，同时也将该方法应用到了字典学习中的系数更新过程中，使得学习到的字典在进行稀疏分解时，能够在图像编码率失真意义下达到最优，从而大幅度提高了编码效率。在实际的图像集压缩算法中进一步验证了算法的有效性。

- **基于视图稀疏特征的表面光场编码。**现有的光场表示与压缩算法大多都是基于多视点图像的，通过去除不同视点间的像素冗余达到压缩的目的，但是因为缺失了准确的几何信息，渲染的虚拟视点的质量通常较差，严重影响了用户体验。本文提出了基于表面光场的高效表示和压缩方法，本质上将多视图图像映射到三维点云空间中，点云中的每个点从不同角度观测形成了球面视图。进一步，将每个点对应的视图分解为B样条小波基函数下的一组表示系数，这组表示系数作为重要的视图稀疏特征可以用来重构表面光场。通过稀疏和平滑约束优化求解得到的视图表示系数可以实现高效的压缩。最后，利用三维点云的空间相关性，采用点云压缩方法编码视图稀疏特征。该方法能够实现解码端真正意义上的实时自由视点渲染，是未来虚拟现实、增强现实等应用中最具潜力的发展方向之一。

综上，本文针对局部描述特征、稀疏基元特征以及视图稀疏特征，利用编码率失真优化模型，提出了图像视频以及光场的高效编码算法，发掘了视觉特征在编码压缩以及视觉分析中的潜力，显著提高了编码效率，为学术研究和实际应用提供了新的思路。

关键词：视频编码，视觉特征，局部特征，稀疏表示，率失真优化