

摘要

随着互联网的普及与网络多媒体应用的广泛发展，视频数据作为主要的信息载体呈现出爆发式增长。近年来，视频图像在各领域的应用需求不断增加，对数据的存储和传输提出了严峻的挑战。为应对这一挑战，高效的图像编码技术成为研究热点。传统的图像编码方法如 JPEG、BPG、VVC-intra 等，通过分块、预测、量化及熵编码等实现数据压缩，但这些混合编码框架的局部优化难以确保全局最优。幸而，近年来深度学习技术蓬勃发展，基于深度学习的端到端图像压缩崭露头角，后者利用神经网络全局学习压缩和重构过程，实现了全局联合优化。端到端图像压缩最核心的问题在于如何充分利用神经网络强大的表示能力提升编码效率。针对这一挑战，前人通常从网络结构的角度入手，优化神经网络变换、神经网络熵模型等，取得了一定效果。然而，现有研究大多聚焦于神经网络架构的改进（如从 CNN 到 Transformer），未能充分挖掘“编码效率导向”的优化策略。为了解决这个问题，本文从变换设计、熵模型设计、率失真特性探索这三个端到端图像编码的核心研究方向出发，相应地在 λ 域、空域、统计域对隐变量进行联合建模，提出隐空间连续模型、隐空间稀疏模型、隐空间非均匀模型。主要创新点如下：

(1) 提出了一种基于 λ 域隐空间连续模型的编码方法。传统的可变码率 (VBR) 压缩虽然实现了单模型多码率，却缺乏对率失真特性的可解释性。为解决这一问题，本文以“变量中心”的视角理解端到端压缩范式，并通过对齐 λ 域和固定码率模型中的隐变量，实现了 λ 域的可变码率压缩。在以上设计的基础上，进一步从理论角度分析了神经网络编码的率失真模型，并将其应用到码率控制任务中。提出的 VBR 方法的率失真性能超过固定码率压缩方法 0.27%，而码率控制方法的平均码率控制误差低于 0.5%。

(2) 提出了一种基于空域隐空间稀疏模型的编码方法。已有的端到端图像压缩方法往往忽视了神经网络固有的稀疏性，对完整隐变量进行熵编码，引入了较大冗余。对此，本文提出通过神经网络分析隐空间的稀疏性，对显式稀疏区域进行选择熵编码，节省了 20% ~ 50% 的编解码时间。另外，针对目前超先验信息未被充分利用的问题，设计了解耦超先验模块与主变换模块的方法，使得超先

验信息能够更有效地作为纹理先验知识增强主变换，提升了 1.1% 的率失真性能。

(3) 提出了一种基于统计域马尔可夫模型的非均匀量化方法。传统的四舍五入量化方式未能充分利用图像中的上下文信息，引入了较大的统计冗余。对此，本文提出基于马尔可夫模型的非均匀量化方法，通过神经网络分析图像内容的上下文信息及其纹理特性，依赖非线性特征函数在线性复杂度内量化隐变量。进一步提出了中心补偿模块，该模块通过神经网络自适应地选择非均匀的量化中心，在不影响码率的情况下减小重建图像的失真。提出的量化方法能够有效减少统计冗余，提升了 1.4% 的率失真性能。

综上所述，本文基于对端到端图像压缩的问题分析，从 λ 域、空域、统计域三个角度对隐空间进行建模，并由此提出超先验回馈、稀疏性筛选、非均匀量化、可变码率、码率控制等方法，显著优化了端到端图像压缩的压缩效率，能够适应不同场景的编码需求。

关键词：深度学习，图像编码，端到端图像压缩，率失真理论