

数字图像处理报告九：图像压缩

姓名：鲁国锐

学号：17020021031

专业：电子信息科学与技术

2020 年 5 月 24 日

目录

1	题目描述	2
2	JPEG2000	2
2.1	原理简述	2
2.2	优势	2
3	基于深度学习的图像压缩	3
3.1	2019 年研究进展：CVPR 与 ICCV	3
3.2	各文献概述	3
4	总结	4

1 题目描述

简单描述 *JPEG2000* 的图像压缩原理，并与基于深度学习的图像压缩方法进行对比。

2 *JPEG2000*

2.1 原理简述

JPEG 是“使用最为普遍且广泛的连续色调静止帧压缩标准”[14]。但随着今年来人们对于多媒体图像资料需求的增长，旧的标准已难以适应新的潮流，于是促使人们又在其基础上进行扩充，得到了 *JPEG2000* 标准。该标准共有 8 个部分，其中 *PART 1* 在 2000 年 12 成为了国际标准，其余部分由于知识产权等原因暂未得到推广。故本节只简要介绍 *PART 1* 的压缩流程和原理。

JPEG2000 标准 *PART 1* 的压缩过程大致可分为以下几步 [16]：

1. 数据预处理：
 - (a) 区域划分：降低压缩所需的内存资源，在内存充足的情况下可忽略；
 - (b) 降低量级：将样本的范围从 $[0, 2^P - 1]$ 移到 $[-2^P, 2^{P-1} - 1]$ ，从而简化对数值溢出问题的处理；
 - (c) 分量变换：在图像有多个分量时，降低分量之间的相关性，提高压缩效率；
2. 离散小波变换 (*DWT*)：进一步降低数据之间的相关性，并可针对不同类型图像的不同区域采用不同的分辨率，以取得更好的压缩比，而且它还可以提供实现无损压缩的机制；
3. 量化：与 *JPEG* 量化基本相同：总体上都采用均匀量化；不同子带量化步长不同。但也有两处不同：
 - (a) 引入了 *deadzone* 概念，用于信噪比分级；
 - (b) 解码时量化索引的逆量化值可取量化器允许范围中的某个值而不是仅局限在中值点如果取，值策略正确，将有助于提高解码性能；
4. *EBCOT* 算法：优化截取内嵌码块编码 (*embedded block coding with optimized truncation*)，简称 *EBCOT*，是 *JPEG2000* 的核心，它分为两级 (*Two Tiers*)：
 - (a) 第一级：块编码，包括比特平面编码、算术编码，负责数据的具体压缩；
 - (b) 第二级：负责码流组织，主要根据压缩后率失真最优化算法 (*PCRD*) 实现的码流进行组织，并完成最后的打包工作 [17]。

2.2 优势

相比于 *JPEG*，*JPEG2000* 在很多层面上都有提高，包括但不限于以下几个方面 [14, 16]：

1. 压缩性能提高；
2. 支持有损压缩和无损压缩；
3. 在连续色调静止图像的压缩和压缩数据的访问方面提供了更大的灵活性；
4. 多解析度支持；
5. 可嵌入的码流；
6. 感兴趣区（用户可以选择对图像的某一区域进行更高质量的编解码或优先显示）；

7. 容错性;
8. 对码流的随机访问处理;
9. 灵活的文件格式;
-

3 基于深度学习的图像压缩

3.1 2019 年研究进展: CVPR 与 ICCV

关于深度学习在图像压缩方面应用的文献比较多, 本次报告粗略整理了一下 CVPR2019 和 ICCV2019 的论文:

1. CVPR2019:

- (a) *Feature distillation : Dnn – oriented jpeg compression against adversarial examples* [9];
- (b) *Hybrid scene compression for visual localization* [2];
- (c) *Learning image and video compression through spatial – temporal energy compaction* [3];
- (d) *Practical full resolution learned lossless image compression* [12];
- (e) *Dvc : An end – to – end deep video compression framework* [11];
- (f) *Machine vision guided 3d medical image compression for efficient transmission and accurate segmentation* [10];

2. ICCV2019:

- (a) *JPEG Artifacts Reduction via Deep Convolutional Sparse Coding* [6];
- (b) *Generative Adversarial Networks for Extreme Learned Image Compression* [1];
- (c) *DSIC : Deep Stereo Image Compression* [8];
- (d) *Variable Rate Deep Image Compression With a Conditional Autoencoder* [4];
- (e) *Learned Video Compression* [13];
- (f) *Neural Inter – Frame Compression for Video Coding* [5];
- (g) *Video Compression With Rate – Distortion Autoencoders* [7];
- (h) *Non – Local ConvLSTM for Video Compression Artifact Reduction* [15];

考虑到图像压缩历史也比较悠久, 在 2019 年仍能有这么多 (实际可能更多) 该领域的文献, 图像压缩的重要性可见一斑。

3.2 各文献概述

从收集的文献来看, 该领域又可以分为四个小类:

1. 图像压缩:

- (a) [3]: 采用时空能量压缩的基本思想对图像和视频进行压缩;
- (b) [12]: 主要采用了结合学习辅助表示的方法对图像分布进行建模, 而非仅仅只在 RGB 空间对图像建模;

- (c) [1]: 提出了一个基于 GAN 的图像压缩系统, 该系统可以在极低的比特率下运行;
- (d) [8]: 利用了两幅图像具有重叠视场这一事实以达到进一步压缩立体图像的目的;
- (e) [4]: 通过引入 *Lagrange multiplier* 和 *quantization bin size* 两个参数, 从而使得一个网络能够生成不同压缩率的图像;

2. 视频压缩:

- (a) [11]: 提出了第一个端到端的深度视频压缩模型 (DVC), 该模型可以联合学习运动估计、运动压缩和残差压缩;
- (b) [13]: 提出了一个新的视频编码算法, 学习端到端的低延迟模式, 其中每个帧只能依赖过去的信息;
- (c) [5]: 通过把所需信息编码到一个可以直接解码为运动和混合系数的隐层表示, 并直接在隐空间计算残差以重复利用相同的图像压缩网络, 从而同时提高解码效率和重建质量;
- (d) [7]: 提出的模型包含一个具有离散隐空间的三维自编码器和一个用于熵编码的自回归先验, 二者联合训练以减少失真。

3. 图像压缩质量提升:

- (a) [6]: 暂未看懂;
- (b) [15]: 引入了一个 *Non-local* 策略来捕捉全局运动模式并追踪视频序列中的时空相关性, 从而使得模型能够以一种快速、低成本的方式工作;

4. 图像压缩在其它领域的应用:

- (a) [9]: 基于 *JPEG* 的压缩框架来解决对抗样本问题;
- (b) [2]: 通过压缩 3D 模型来提高视觉定位的性能, 并减少对存储、带宽的要求, 使之可以应用在移动设备上;
- (c) [10]: 展示了人和机器看待压缩质量的不同, 并基于 *DNN* 设计了一个面向机器视觉的 3D 图像压缩框架, 达到了更高的分割精度或在相同分割精度下获得更好的压缩率。

上述文献的总结中, 只截取了其主要贡献与思想, 并未面面俱到。

4 总结

相比于如 *JPEG2000* 这样以 *DCT*、*WT* 为基础进行编码的传统算法, 深度学习更多依靠深度神经网络来实现编码器、解码器, 经过训练达到一个较好的效果。同时, 基于深度学习的图像很多时候是一个端到端的压缩模型, 并不像 *JPEG*、*JPEG2000* 这样每一步都有着清晰的数学解释。但在部分文献 (如 [6] 中, 也会将 *DCT* 与深度学习结合起来, 以达到更好压缩效果, 并取得更好的可解释性。

参考文献

- [1] Eirikur Agustsson, Michael Tschannen, Fabian Mentzer, Radu Timofte, and Luc Van Gool. Generative adversarial networks for extreme learned image compression. In The IEEE International Conference on Computer Vision (ICCV), October 2019. 3, 4

- [2] Federico Camposeco, Andrea Cohen, Marc Pollefeys, and Torsten Sattler. Hybrid scene compression for visual localization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 7653–7662, 2019. 3, 4
- [3] Zhengxue Cheng, Heming Sun, Masaru Takeuchi, and Jiro Katto. Learning image and video compression through spatial-temporal energy compaction. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 10071–10080, 2019. 3
- [4] Yoojin Choi, Mostafa El-Khamy, and Jungwon Lee. Variable rate deep image compression with a conditional autoencoder. In The IEEE International Conference on Computer Vision (ICCV), October 2019. 3, 4
- [5] Abdelaziz Djelouah, Joaquim Campos, Simone Schaub-Meyer, and Christopher Schroers. Neural inter-frame compression for video coding. In The IEEE International Conference on Computer Vision (ICCV), October 2019. 3, 4
- [6] Xueyang Fu, Zheng-Jun Zha, Feng Wu, Xinghao Ding, and John Paisley. Jpeg artifacts reduction via deep convolutional sparse coding. In The IEEE International Conference on Computer Vision (ICCV), October 2019. 3, 4
- [7] Amirhossein Habibian, Ties van Rozendaal, Jakub M. Tomczak, and Taco S. Cohen. Video compression with rate-distortion autoencoders. In The IEEE International Conference on Computer Vision (ICCV), October 2019. 3, 4
- [8] Jerry Liu, Shenlong Wang, and Raquel Urtasun. Dsic: Deep stereo image compression. In The IEEE International Conference on Computer Vision (ICCV), October 2019. 3, 4
- [9] Zihao Liu, Qi Liu, Tao Liu, Nuo Xu, Xue Lin, Yanzhi Wang, and Wujie Wen. Feature distillation: Dnn-oriented jpeg compression against adversarial examples. In 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 860–868. IEEE, 2019. 3, 4
- [10] Zihao Liu, Xiaowei Xu, Tao Liu, Qi Liu, Yanzhi Wang, Yiyu Shi, Wujie Wen, Meiping Huang, Haiyun Yuan, and Jian Zhuang. Machine vision guided 3d medical image compression for efficient transmission and accurate segmentation in the clouds. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 12687–12696, 2019. 3, 4
- [11] Guo Lu, Wanli Ouyang, Dong Xu, Xiaoyun Zhang, Chunlei Cai, and Zhiyong Gao. Dvc: An end-to-end deep video compression framework. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 11006–11015, 2019. 3, 4
- [12] Fabian Mentzer, Eirikur Agustsson, Michael Tschannen, Radu Timofte, and Luc Van Gool. Practical full resolution learned lossless image compression. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 10629–10638, 2019. 3
- [13] Oren Rippel, Sanjay Nair, Carissa Lew, Steve Branson, Alexander G. Anderson, and Lubomir Bourdev. Learned video compression. In The IEEE International Conference on Computer Vision (ICCV), October 2019. 3, 4
- [14] Rafael C. Gonzalez, Richard E. Woods. 数字图像处理（第三版）. 电子工业出版社, 2011.6. 2

- [15] Yi Xu, Longwen Gao, Kai Tian, Shuigeng Zhou, and Huyang Sun. Non-local convlstm for video compression artifact reduction. In The IEEE International Conference on Computer Vision (ICCV), October 2019. [3](#), [4](#)
- [16] 刘方敏, 吴永辉, 俞建新. Jpeg2000 图像压缩过程及原理概述. 计算机辅助设计与图形学学报, 14(10):905–911, 2002. [2](#)
- [17] 王伊洛. EBCOT 中 T2 编码器的 VLSI 设计. PhD thesis, 西安电子科技大学. [2](#)