

《计算机视觉》期末报告

201250068 陈骏

1. 问题描述

本篇报告选取去噪作为分析专题，自然界中的噪声是指在指定音频外由其他声源产生的噪声，例如在一场钢琴演出中，由台下低素质观众发出的叽叽喳喳声。而在计算机领域，噪声不仅仅局限于声音这一维度，在音频，图片，视频，甚至文本或数据中，都可能出现“噪声”，例如文本中的错别字，数据集中的异常点等，噪声开始普遍地指向于那些不在预期之类出现的异常属性。相较于音频问题中，噪声通常以异常的分贝或音调出现，在计算机视觉中，噪声是时域的，即某个值相对于连续的一段时间内，在同一个位置产生的不同值，即像素值，其误差大于某个阈值，则认为是噪声。但单张图像不能记录时序信息，即对于单个图片，而非是某连续视频的某一帧，如何在缺少上下文的情况下，进行噪声识别并进行去噪处理，提高图片的还原度，是计算机视觉去噪问题的根本内容。

2. 发展趋势

传统的图片噪声包括数字成像过程中的各个阶段产生的噪声，包括不限于光电效应的泊松抖动等。在图像文件的压缩，传输等过程，也有可能产生误差，或者用一个更为贴切的称呼：失真。对于单帧图像，有基于空间域，变换域，图像压缩，超分辨率等不同的去噪方法。而去噪随着自身方法体系的不断完善，其针对不同的场景，比如图片中的雨滴，雾气或者遮挡等等，也衍生出不同的应用算法。同时，最新的去噪问题也逐渐专业化，例如，医疗图像的去噪问题已经有了重大突破。

3. 代表性方法

本文选取了多个历史上较为经典并具有代表性的非深度学习去噪方法，分别为 NLM、BW3D、Wavelet Transform、MRF、K-SVD 算法。

3.1. NLM 算法

NLM (Non-Local Means) 算法是由 Buades 等人在 2005 发表的对传统邻域滤波方法的改进版本^[1]。算法中，首先定义了每个像素值的估计值为所有像素的加权平均 $NL[v](i) = \sum_{j \in I} w(i, j) v(j)$ ，而权重值依赖于两个像素之间灰度向量的相似性。而这个相似性的计算依赖于以这两个像素为中心，大小为 k 的两个窗口。NLM 算法采用了欧式距离作为比较相似性的方案。而权重定义为下列式子 $w(i, j) = \frac{1}{Z(i)} e^{-\frac{\|v(N_i) - v(N_j)\|_2^2}{h^2}}$ 。其中 $Z(i)$ 为归一化常数。NLM 算法不仅对图像中的单点进行比较，还对该像素附近的邻域进行比较，提高了滤波器的鲁棒性。在 Buades 的论文中，NLM 算法与高斯滤波，各向异性滤波，总变分，邻域滤波的对比结果如下表。可以看到 NLM 算法的均方误差远小于其他滤波器的结果。

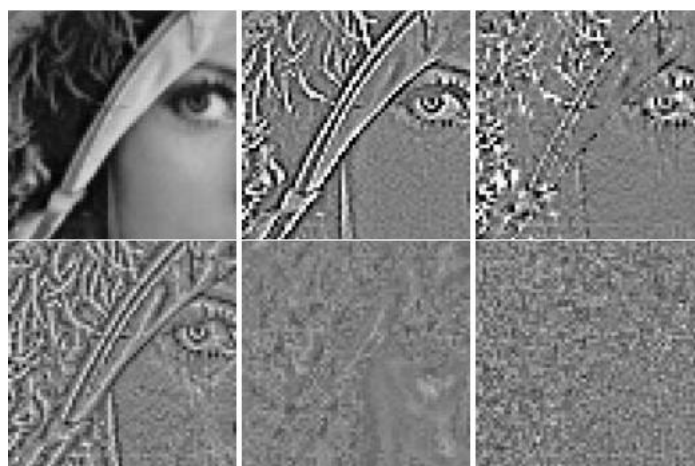


图 1 NLM 算法滤波结果比较

表 1 NLM 算法均方误差比较

| Image | GF | AF | TVF | YNF | NL |
|--------|-----|-----|-----|-----|-----|
| Lena | 120 | 114 | 110 | 129 | 68 |
| Baboon | 507 | 418 | 365 | 381 | 292 |

3.2. BW3D 算法

BW3D (Block-matching three dimension) 算法是由 Dabov 等人与 2007 年, 在 NLM 方法上, 进行改进的衍生算法, 是基于超像素的代表性方法^[2]。算法分为两级联合滤波和每级的三个步骤, 分组匹配堆叠, 协同过滤和聚合。其中分组可以通过多种方案, 例如 K-Means 聚类等, 而在 BW3D 中, 作者采用了块匹配的方案, 这和 NLM 算法中, 以窗口计算相似性的方法属于一脉相承。而在对每个块进行分组以及 Wiener 协同过滤之后, 进行聚合, 采用加权平均值来获得真实图像的估计。BM3D 也有许多衍生和变体, 例如 CBM3D 和 BM4D 等, 都沿袭了其块处理的思想, 这跟图像本身的属性也有关, 图片的特性不是靠单一像素表达的, 而是通过大片像素的块来体现。

表 2 BM3D 算法 PSNR 结果比较

| Approach to color-image denoising | Lena | House | Peppers |
|-----------------------------------|------|-------|---------|
| Lena | 120 | 114 | 110 |
| Baboon | 507 | 418 | 365 |

3.3. Wavelet Transform 算法

WT (小波过滤算法) 是 Donoho 提出的基于变换域的去噪方法^[4], 是将图片进行某种变换, 转换到指定的变换域 (类似的还有傅里叶变换等等), 在变化域上过滤完之后再返回原来的空间域。2008 年 Benyahia 提出了基于 Stein' s Principle 的小波变化去噪方法^[3], 论文中通过斯坦原理在线性变换域中, 对假定为高斯噪声的加性噪声进行去噪, 对单分量和多分量图像都进行了应用。文章首先提出了观察模型, 指出在冗余变化之下, 在 transform

域中的最优准则具有次优性， MSE_{im} ，即在空间域上的均方误差应该为优化目标。作者通过对于高斯加性噪声在空间域和变换域的雅可比矩阵，推导了 MSE_{tr} 的风险，接着通过 Stein 准则，进行二次风险的计算，再通过一个新的策略：SUREVECT 来进行优化模型，避免由于缺少无噪图片的过拟合问题。文章中涉及到了很多变换域与空间域在进行了变换之后，空间域的参数如何对应到变换域中的复杂运算，以及在冗余框架中的去噪实现。

3.4. MRF 算法

MRF（马尔可夫场）是用来表示链状概率转移关系，在 Zhang 等人于 2009 年发表的论文中^[5]，在 Wavelet 变换的基础上，再通过 Markov Random Fields 以及 Gibbs 展开，最小化能量，通过其结果进行边缘检测，再将变换域回到空间域，优化了在进行去噪过程中边缘的丢失问题。其算法实现难点和突破主要是 MRF 的建立，以及小波变换的过程参数的传递。



图 3 MRF 算法去噪结果比较

表 3 MRF 结果比较

| | Noise | Soft | Hard | Our |
|-----------|---------|---------|---------|---------|
| PSNR (dB) | 23.1168 | 22.8851 | 22.0229 | 22.9460 |
| SSIM | 0.4974 | 0.6849 | 0.7101 | 0.7187 |

3.5. K-SVD 算法

K-SVD(K-Means 奇异值矩阵分解)是基于图像矩阵稀疏表示压缩的去噪算法，由 Aharon 等人于 2006 年发表^[6]，稀疏表达通过较少的信号的线性组合来表示原有图像的大部分内容，这就是数据的稀疏性，例如一个原有的图像，由多个有限的有意义的“原子”组成，而噪声图像中，大量的噪声是无意义的不可稀疏的。在 K-SVD 算法中，将图片表示进行分解，表示为一个颜色矩阵 D 和一个系数矩阵 X 的乘积，SVD 就是不断的用 X 来优化 D ，再用 D 优化 X ，其中 Aharon 的论文通过 SVD 奇异值分解方法来求解了用 D 优化 X 的过程。后续基于稀疏矩阵的算法还有 LSSC，NCSR 等。

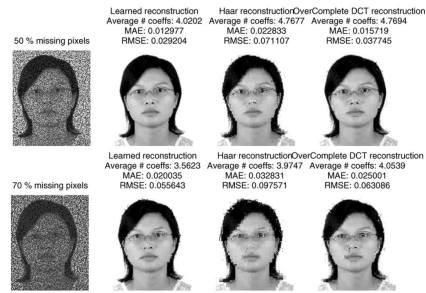


图 3 K-SVD 算法去噪结果比较

3.6. 思考

在对这五个方法的研究中,我总结到了他们的出发点大多是从图片的构成角度来进行去噪的,然后将特征提取之后的噪声检测,用了一些很经典的机器学习方法,例如使用欧式距离来进行高维数据的相似性计算以及 K-Means 算法来进行聚类。而引用的论文中的方法,大多并不是完全的创新,很多都是建立在前人的基础上的,例如基于 NLM 算法的 BW3D 算法,一些去噪方法也是属于移花接木,例如马尔可夫场并不是首先应用于去噪场景的,同理 SVD 算法也并不是,都是在其他场景有了较好的效果之后运用到去噪问题的。

不同噪声之间的去除方法也有一定的关联,如音频噪声的 EMD 及其衍生算法,在图像去噪问题中也有较好的效果。

4. 未来发展趋势

2021 年 CVPR 上关于图像噪声的论文大多是以深度学习进行,很多都是构建非常复杂的神经网络,例如 CNN 卷积神经网络对于去噪问题效果的天然的优秀,但可解释性不如传统的算法。同时去噪也有了较多的子专题,例如进行图像去雨,图像去雾, CVPR 上都有较多的论文发表,以及首个利用 Transformer 进行医学图像去噪的,这些子专题的应用场景更加具体并贴切实际,例如图像去雨对于雨天驾车行驶的帮助,类似的可能有雨天监控的优化,这些实际应用场景可能会将与去噪技术相互促进发展。此外,基于 ChatGPT 的出现,是否会出现一个基于大量高质量无噪声图片数据集的 DenoisingGPT,可以做到去噪处理,目前 Photoshop 已经实现自动拓展图片功能,且效果惊人,或许一个集成了去噪,抠图,自动补全等功能的图片工具指日可待。

致谢

感谢 Deep1 以及有道翻译对于英文论文的翻译帮助,感谢 ChatGPT 对于去噪算法相应论文的检索帮助。

参考文献

- [1] A. Buades, B. Coll and J. . -M. Morel, "A non-local algorithm for image denoising," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, pp. 60-65 vol. 2, doi: 10.1109/CVPR.2005.38.
- [2] K. Dabov, A. Foi, V. Katkovnik and K. Egiazarian, "Image Denoising by Sparse 3-D Transform-Domain

Collaborative Filtering," in IEEE Transactions on Image Processing, vol. 16, no. 8, pp. 2080-2095, Aug. 2007, doi: 10.1109/TIP.2007.901238.

- [3] A. Benazza-Benyahia, J. . -C. Pesquet and C. Chaux, "Image Denoising in the Wavelet Transform Domain Based on Stein's Principle," 2008 First Workshops on Image Processing Theory, Tools and Applications, Sousse, Tunisia, 2008, pp. 1-9, doi: 10.1109/IPTA.2008.4743802.
- [4] David L Donoho , Iain M Johnstone, Ideal spatial adaptation by wavelet shrinkage, Biometrika, Volume 81, Issue 3, September 1994, Pages 425–455, <https://doi.org/10.1093/biomet/81.3.425>
- [5] Y. Zhang, F. Duan and R. Cui, "An Edge-Preserving Wavelet Denoising Method Based on MRF," 2009 Fifth International Conference on Image and Graphics, Xi'an, China, 2009, pp. 67-71, doi: 10.1109/ICIG.2009.109.
- [6] M. Aharon, M. Elad and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," in IEEE Transactions on Signal Processing, vol. 54, no. 11, pp. 4311-4322, Nov. 2006, doi: 10.1109/TSP.2006.881199.