Image Inpainting Based on Sparse Representation

陶重阳, *1501111344*
北京大学计算机科学技术研究所

2016年06月19日

1 引言

在文献 [1–4]中稀疏表示模型被引入到图像去噪和图像修复问题中。这类问题的基本想法是用一组过完备变换的稀疏组合来表示图像，从而缺失的像素可以通过自适应更新稀疏表示来推断出来。Guleryuz [5]基于*DCT*字典对不同的受损图像进行了修复。Elad [4]等将稀疏表示理论引入了信号和图像处理领域，采用K-SVD字典训练算法来解决图像修复问题并取得了较好地修复效果。

文献 [6, 7]将图像修复(*Image Inpainting*)问题建模为非负矩阵填充(*NMC*)问题。

$$\begin{aligned} &\min_{\mathbf{X},\mathbf{e}} \|\mathbf{X}\|_* + \frac{1}{2\mu}\|\mathbf{e}\|_2^2 \\ &s.t. \quad \mathbf{b} = \mathcal{P}_\Omega(\mathbf{x}) + \mathbf{e}, \mathbf{X} \geq 0 \end{aligned} \tag{1}$$

上式中$\mathbf{b}$ 为观测数据， $\mathbf{e}$ 为噪声矩阵，$\mathcal{P}_\Omega$ 是针对位置集合 Ω 的线性映射，该优化问题可以通过 APG [8]，ADM/LADM [6, 9]，LADMPSAP [7]等算法来进行求解。

本文实现了基于 *DCT* 字典的图像修复算法和基于 *K-SVD* 字典训练的图像修复算法，此外测试了插值和基于矩阵填充算法的效果。

2 基于稀疏表示的图像修复

令$\mathbf{X} \in \mathbb{R}^N$ 为大小 $\sqrt{N} \times \sqrt{N}$ 的原始图像，$\mathbf{M}$ 为图像的掩码矩

阵*，$\mathbf{x}_{ij} \in \mathbb{R}^n$表示该图像在 (i, j) 处的大小为 $\sqrt{n} \times \sqrt{n}$ 图像块的向量表达†，可以表示如下：

$$\mathbf{x}_{ij} = \mathbf{R}_{ij}\mathbf{X} \tag{2}$$

上式中，$\mathbf{R}_{ij} \in \mathbb{R}^{n \times N}$表示从图像 $\mathbf{X}$ 中提取出块 $\mathbf{x}_{ij}$ 的矩阵算子。

由于用到的图像块 $\mathbf{x}_{ij}$ 彼此都是互相重叠的，因此这种基于块的表示是高度冗余的，这种互相重叠的技巧以及高度冗余的表示对于取得高质量的恢复效果至关重要。由公式 (2)，利用这些彼此重叠的块 $\mathbf{x}_{ij}$ 来恢复 $\mathbf{X}$ ，可以通过最小二乘算法表达如下[43]：

$$\mathbf{X} = \Big(\sum_{ij} \mathbf{R}_{ij}^T \mathbf{R}_{ij}\Big)^{-1} \sum_{ij} \Big(\mathbf{R}_{ij}^T \mathbf{x}_{ij}\Big) \tag{3}$$

上式的实质就是把这些块放到原图相应的位置然后取平均。

给定一个字典 $\mathbf{D} \in \mathbb{R}^{n \times M}$， 每一个块 $\mathbf{x}_{ij}$ 关于字典 $\mathbf{D}$ 进行稀疏编码 (*sparse coding*) 的过程是找一个稀疏的向量 $\boldsymbol{\alpha}_{ij}$， 即 $\boldsymbol{\alpha}_{ij}$ 中大部分的元素为零或者接近零， 使得 $\mathbf{x}_{ij} \simeq \mathbf{D}\boldsymbol{\alpha}_{ij}$ 。因此，整幅图可以被一组稀疏码字(*sparse code*) $\boldsymbol{\alpha}_{ij}$ 来稀疏地表示。

对于给定字典 $\mathbf{D}$， 求解每一个块 $\mathbf{x}_{ij}$ 的稀疏编码问题可以表示为：

$$\{\hat{\boldsymbol{\alpha}}_{ij}\} = \arg\min_{\boldsymbol{\alpha}_{ij}} \|\mathbf{x}_{ij} - \mathbf{M}_{ij} \otimes (\mathbf{D}\boldsymbol{\alpha}_{ij})\| + \|\boldsymbol{\alpha}_{ij}\|_0 \tag{4}$$

上式中 $\mathbf{M}_{ij}$ 表示图像 块$\mathbf{x}_{ij}$ 的掩码矩阵。通常情况下，由于 L_0 范数优化问题是非凸的并且是 NP 难问题，可以采用凸松弛得到原始问题的凸近似，然后使用一些凸优化工具(如：CVX)来进行求解；此外也可以使用贪心算法的思想得到次优解，比如：正交追踪匹配(Orthogonal matching pursuit, OMP)算法 [10]。

OMP算法的基本思想是：根据贪婪迭代的方法选择字典 $\mathbf{D}$ 的列，使得在每次迭代的过程中所选择的列与当前冗余向量最大程度的相关，从原始信号向量中减去相关部分并反复迭代，只到迭代次数达到一定的稀疏度 k，停止迭代。核心算法步骤如下：

*掩码矩阵的作用是“屏蔽”图像上的部分像素，即图像上已知的像素设置为1，缺失的像素设置为0

†本实验中采用 “Sliding” 方法选取图像块，因此其个数为 $(\sqrt{N} - \sqrt{n} + 1)^2$

参数：字典矩阵 $\mathbf{D}$，向量(图像块展开) $\mathbf{b}$ 以及误差阈值 ϵ_0

1. 初始化: $k = 0$, $\boldsymbol{\alpha}^0 = \mathbf{0}$, 残差 $\mathbf{r}^0 = \mathbf{b} - \mathbf{M} \otimes (\mathbf{D}\boldsymbol{\alpha}^0) = \mathbf{b}$, 支撑集 $S^0 = Support(\mathbf{x}^0) = \phi$

2. Repeat:

   (a) 选择：计算误差 $\epsilon(j) = \min_{z_j} \|\mathbf{a}_j z_j - \mathbf{r}^{k-1}\|_2^2$, 找到最小的$j_0$, 使得 $\epsilon(j_0) \leq \epsilon(j), \quad \forall j \notin S^{k-1}$，更新支撑集 $S^k = S^{k-1} \cup \{j_0\}$

   (b) 更新稀疏表示系数：对于给定的支撑集 S^k，求解近似信号的最优稀疏表示系数

   $$\boldsymbol{\alpha}^k = \arg\min_{\boldsymbol{\alpha}} \|\mathbf{M} \otimes (\mathbf{D}\boldsymbol{\alpha}^k) - \mathbf{b}\|_2^2 \tag{5}$$

   (c) 更新残差： $\mathbf{r}^k = \mathbf{b} - \mathbf{M} \otimes (\mathbf{D}\boldsymbol{\alpha}^k)$

   (d) 迭代：如果 $\|\mathbf{r}^k\|_2 \leq \epsilon_0$，停止迭代更新；否则，返回（a）

3. 输出: 最终的稀疏表示解 $\boldsymbol{\alpha}^k$

字典 $\mathbf{D}$ 的选择稀疏表示模型的重点，一般情况下我们可以选择 *Wavelet*，*DCT* 等固定字典，但这些固定字典缺乏自适应性，在很多情况下不能够获得很好的稀疏性。现有很多研究工作致力于如何从一组训练图像块中学习的得到冗余字典。具体来讲，给定一组图像块 $S = [s_1, s_2, \cdots, s_J]$，其中 J 是训练图像块的个数。字典学习的目标就是联合优化字典 $\mathbf{D}$ 和稀疏表示系数 $\boldsymbol{\alpha}$ ，使得 $\mathbf{s}_k = \mathbf{D}\boldsymbol{\alpha}_{ij}$ 且 $\|\boldsymbol{\alpha}_{ij}\|_0 \leq L$ 。具体优化问题可以表示如下：

$$\{\hat{\boldsymbol{\alpha}}, \hat{\mathbf{D}}\} = \arg\min_{\mathbf{D},\boldsymbol{\alpha}} \sum_{ij} \|\mathbf{R}_{ij}\mathbf{X} - (\mathbf{R}_{ij}\mathbf{M}) \otimes (\mathbf{D}\boldsymbol{\alpha}_{ij})\| + \mu_{ij} \sum_{ij} \|\boldsymbol{\alpha}_{ij}\|_0 \tag{6}$$

上述优化问题可以通过 MOD [8]和 K-SVD [8]来进行求解，本文中我们选用K-SVD来进行实验，该算法交替优化字典 $\mathbf{D}$ 和 稀疏系数 $\boldsymbol{\alpha}$。设原始图像为 $\mathbf{X}$，初始化字典 $\hat{\mathbf{D}}$ 为离散余弦变换 (*DCT*) 字典，然后进行迭代更新，具体迭代算法如下：

1. **稀疏编码**：固定字典 $\hat{\mathbf{D}}$, 利用 *OMP* 算法求解每一个*patch*的系数编码

   $\forall ij, \hat{\boldsymbol{\alpha}}_{ij} = \arg\min_{\boldsymbol{\alpha}_{ij}} \|\boldsymbol{\alpha}_{ij}\|_0, \quad \|\mathbf{R}_{ij}\mathbf{X} - (\mathbf{R}_{ij}\mathbf{M}) \otimes (\hat{\mathbf{D}}\boldsymbol{\alpha}_{ij})\|_2^2 \leq n(C\sigma)^2$

其中C为噪声因子，与图像块的维数 n 相关，n为图像块大小(字典维数)

2. **字典更新：** 固定系数矩阵 $\hat{\boldsymbol{\alpha}}$, 依次对字典 $\hat{\mathbf{D}}$ 中每个原子 $\mathbf{d}_l(l \in 1,2\cdots,k)$ 进行更新

   (a) 找出所有满足 $w_l=\{[i,j]|\hat{\boldsymbol{\alpha}}_{ij}(l)\neq 0\}$ 的图像块集合

   (b) 对每个图像块 $[i,j]\in w_l$, 计算其残差： $\mathbf{e}_{ij}^l=(\mathbf{R}_{ij}\mathbf{M})\otimes\Big(\mathbf{R}_{ij}\mathbf{X}-\sum_{m\neq l}\hat{\mathbf{d}}_l\hat{\boldsymbol{\alpha}}_{ij}(m)\Big)$，得到残差矩阵 $\mathbf{E}_l$

   (c) 更新 $\hat{\mathbf{d}}_l$ 和 $\hat{\boldsymbol{\alpha}}_{ij}(l)$ 通过最小化下式：

$$(\hat{\mathbf{d}}_l,\hat{\boldsymbol{\alpha}}^l)=\arg\min_{\hat{\boldsymbol{\alpha}}^l,\|\hat{\mathbf{d}}_l\|_2=1}\|(\mathbf{R}_{ij}\mathbf{M})\otimes(\mathbf{E}_l-\mathbf{d}_l\boldsymbol{\alpha}^l)\|_2^2 \tag{7}$$

   上式可以通过奇异值分解(SVD) 来进行求解.

通过迭代 (1) 和 (2)，直到满足一定的迭代次数 J 或误差阈值 ϵ_0。得到最终更新后的字典 $\hat{\mathbf{D}}$ 和稀疏表示矩阵 $\hat{\boldsymbol{\alpha}}$，考虑到图像块的重叠性，需要对Inpainting之后的图像块取平均，即：

$$\hat{\mathbf{X}}=\Big(\sum_{ij}\mathbf{R}_{ij}^T\mathbf{R}_{ij}\Big)^{-1}\sum_{ij}\Big(\mathbf{R}_{ij}^T\mathbf{D}\hat{\boldsymbol{\alpha}}_{ij}\Big) \tag{8}$$

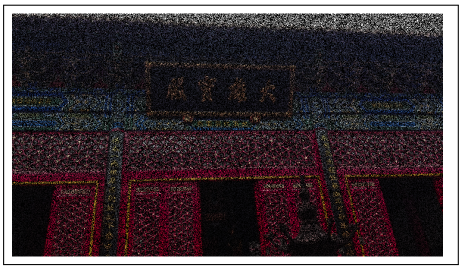
3 实验

本文中选取的图像块大小为 8×8，即字典 $\mathbf{D}\in\mathbb{R}^{64\times 256}$。实验所给测试图像为彩色图像，本文测试了以下两种方案。(在K-SVD 和 OMP 重建中采用第二种方案)

1) 三个通道分别进行稀疏重建

2) 将原始图像转换至 *YCbCr* 颜色通道上，对CbCr通道进行插值重建，Y 通道进行稀疏重建。

如图1(b)是基于德劳内三角化的插值结果；如图1(c)是采用ADM进行矩阵填充算法得到的结果；1(e) 是采用 *DCT* 固定字典进行稀疏重建的结果，可以看出采用固定字典重建的图像中有一些异常 *patch*；1(e) 是基于K-SVD算法进行重建的结果。



(a) 原始图像

(b) Delaunay triangulation based interpolation

(c) NMC,ADM,$rank = 50$

(d) DCT,OMP,$\sigma = 0.01$

(e) KSVD,$J = 25$

图 1: 实验结果对比

注：最终提供的实验结果为基于 K-SVD 重建的图像。

参考文献

[1] Mingyuan Zhou, Haojun Chen, Lu Ren, Guillermo Sapiro, Lawrence Carin, and John W Paisley. Non-parametric bayesian dictionary learning for sparse image representations. In Advances in neural information processing systems, pages 2295–2303, 2009.

[2] Jian Zhang, Debin Zhao, and Wen Gao. Group-based sparse representation for image restoration. Image Processing, IEEE Transactions on, 23(8):3336–

3351, 2014.

[3] Julien Mairal, Michael Elad, and Guillermo Sapiro. Sparse representation for color image restoration. Image Processing, IEEE Transactions on, 17(1):53–69, 2008.

[4] Michael Elad and Michal Aharon. Image denoising via learned dictionaries and sparse representation. In Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on, volume 1, pages 895–900. IEEE, 2006.

[5] Onur G Guleryuz. Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-part ii: adaptive algorithms. IEEE Transactions on Image Processing, 15(3):555–571, 2006.

[6] Yangyang Xu, Wotao Yin, Zaiwen Wen, and Yin Zhang. An alternating direction algorithm for matrix completion with nonnegative factors. Frontiers of Mathematics in China, 7(2):365–384, 2012.

[7] Risheng Liu, Zhouchen Lin, and Zhixun Su. Linearized alternating direction method with parallel splitting and adaptive penalty for separable convex programs in machine learning. In ACML, pages 116–132, 2013.

[8] Kim-Chuan Toh and Sangwoon Yun. An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems. Pacific Journal of Optimization, 6(615-640):15, 2010.

[9] Junfeng Yang and Xiaoming Yuan. Linearized augmented lagrangian and alternating direction methods for nuclear norm minimization. Mathematics of Computation, 82(281):301–329, 2013.

[10] Joel A Tropp and Anna C Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. Information Theory, IEEE Transactions on, 53(12):4655–4666, 2007.