

# 基于主成分分析与深度神经网络的 快速噪声水平估计算法

徐少平, 李崇禧, 林官喜, 唐祎玲, 胡凌燕

(南昌大学信息工程学院, 江西南昌 330031)

**摘要:** 鉴于从噪声图像分解获得的原生图块集合的协方差矩阵前若干个特征值(按照升序排序)与图像噪声水平值具有强相关性,提出了一种基于主成分分析和深度神经网络的快速噪声水平估计算法.该算法首先选用原生图块集合协方差矩阵前若干个特征值构成刻画图像噪声水平高低的特征矢量,然后在大量有代表性且已标定噪声水平值的噪声图像集合上利用深度神经网络训练预测模型以实现将特征矢量直接映射为噪声水平值,最后为获得更高的预测准确性,采用粗精预测模型相结合的两步预测方式实现.实验表明:文中算法在各个噪声级别上都具有稳定的预测准确性,且执行效率非常高,作为降噪算法的前置预处理模块具有更好的综合优势.

**关键词:** 图像降噪; 噪声水平估计; 主成分分析; 深度神经网络; 粗精结合策略

**中图分类号:** TP391 **文献标识码:** A **文章编号:** 0372-2112 (2019)02-0274-08

**电子学报 URL:** <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2019.02.003

## Fast Image Noise Level Estimation Algorithm Based on Principal Component Analysis and Deep Neural Network

XU Shao-ping, LI Chong-xi, LIN Guan-xi, TANG Yi-ling, HU Ling-yan

(School of Information Engineering, Nanchang University, Nanchang, Jiangxi 330031, China)

**Abstract:** Considering the fact that there exists the strong correlation between the first several eigenvalues (in ascending order) of the covariance matrix of the raw patches extracted from a noisy image and its noise level, we proposed a novel fast multiple image-based noise level estimation (FMNLE) algorithm using the principal component analysis (PCA) and the deep neural network (DNN). Specifically, we selected the first several eigenvalues of the raw patches to form a feature vector characterizing the noise level of an image. Then, we employed deep neural network to train an estimation model on a large number of representative natural images corrupted with known noise levels, by which the feature vector can be directly mapped into the corresponding noise level. To obtain higher estimation accuracy, a two-step estimation strategy was adopted. Extensive experiments show that, the estimation accuracy of the proposed algorithm is stable at each noise level with good efficiency, demonstrating a better comprehensive advantage as the pre-processing module for denoising algorithms.

**Key words:** image denoising; noise level estimation; principal component analysis; deep neural network; coarse-to-fine strategy

### 1 引言

大多数图像降噪(Image Denoising)算法的性能在不同程度上依赖于对噪声图像中噪声水平值的准确估计<sup>[1-7]</sup>,因此噪声水平估计(Noise Level Estimation, NLE)算法一直以来都受到了众多学者的关注<sup>[1,6,7]</sup>.早期,研究者们基于小波变换<sup>[8]</sup>、掩模邻域求和<sup>[9]</sup>、区域

分割<sup>[10]</sup>、峰度(Kurtosis)值不变性<sup>[11]</sup>、奇异值分解<sup>[12]</sup>等技术路线提出了很多具有代表性的NLE算法,这些NLE算法的预测准确率以目前发展水平来看已不能令人满意.然而,近年来以Pyatykh算法<sup>[13]</sup>和Liu算法<sup>[14]</sup>为代表的基于主成分分析(Principal Component Analysis, PCA)的NLE算法在噪声预测准确性方面取得了较大的进展.相比已提出的各类NLE算法,Pyatykh算法

和 Liu 算法在噪声估计的准确度上有了很大提升,但由于需要通过迭代执行方式从原生图块(Raw Patches)集合中筛选出同质图块(Homogenous Patches)子集,导致执行时间比较长.近期,Cui 等人<sup>[15]</sup>提出了一种基于噪声水平累积(Noise Level Accumulation, NLA)指数的 NLE 算法.该算法通过将不同亮度范围(Intensity Interval)的 NLA 指数整合起来估计噪声图像的噪声水平值,比 Pyatykh 算法在某些噪声水平段上的预测准确率要好,但是其执行时间大约是 Pyatykh 算法的两倍<sup>[15]</sup>.

现有的 NLE 算法多是基于所谓的单张噪声图像的实现策略(Single Image-based NLE, SNLE),由于仅有噪声图像本身信息可以利用,所以这些算法不得不设计相当复杂的算法流程以实现噪声水平值的准确估计,导致执行时间过长,从而严重拖累降噪主算法的执行效率.近年来,笔者基于多张图像(Multiple Image-based NLE, MNLE)的实现策略,在具有代表性的噪声图像集上提取若干噪声水平感知特征值(Noise Level-Aware Features, NLA)并构建预测模型将这些特征值映射为噪声水平值,探索从另外一个方向解决上述 SNLE 算法中存在的问题<sup>[1,7]</sup>.文献[1]和[7]的研究表明:提高 MNLE 类算法预测准确性的途径可从提取描述能力更强的 NLA 特征和构建映射能力更好的预测模型入手.

为了获得一种在预测准确性和计算效率两方面的综合性能俱佳的 MNLE 算法,本文试图利用这些原生图块集合协方差矩阵的前若干个特征值(按照升序排列)与图像中的噪声水平值强相关性,采用深度神经网络(Deep Neural Network, DNN)在具有代表性且已知噪声水平值的噪声图像集上训练预测模型,然后利用所训练的预测模型直接将原生图块集合协方差矩阵的若干特征值(构成矢量形式)映射为图像的噪声水平值,从而获得一种新的快速噪声水平估计(Fast Multiple Image-Based Noise Level Estimation, FMNLE)算法.此外,为了获得更高的预测准确率,采用粗精结合的两步预测方式完成评估过程.相对于现有算法,FMNLE 算法在预测准确性和执行效率两个方面展现出更好的综合优势,更具实用价值.

## 2 现有工作

### 2.1 基于主成分分析的 NLE 算法

对于任意给定的一张大小为  $M \times N$  的噪声图像  $I$ , 可以以滑动窗口方式逐像素点扫描各个像素点获取原生图块集合  $y = \{y_i\}_{i=1}^s \in \mathbb{R}^{r \times s}$ ,  $y$  包含  $s = (M - d + 1)(N - d + 1)$  个大小为  $r = d \times d$  的图块.为便于描述,这些图块中像素点的亮度值具有下面的关系:

$$y_i = x_i + n_i \quad (1)$$

其中,  $x_i$  为第  $i$  个无噪声(原图像)图块,  $n_i \in \mathbb{R}^{r \times 1}$  为均值为 0、方差为  $\sigma_n^2$  的高斯噪声  $N(0, \sigma_n^2)$ ,  $y_i$  为相应的噪声

图块.当  $I$  受  $N(0, \sigma_n^2)$  分布的高斯噪声干扰时,  $n_i$  则服从均值为 0、方差为  $\sigma_n^2 I$  的高斯分布  $N(0, \sigma_n^2 I)$ . 基于原生图块集合  $y$ , 它的协方差矩阵可以由下式计算:

$$\Sigma_y = \frac{1}{s} \sum_{i=1}^s (y_i - \mu)(y_i - \mu)^T \quad (2)$$

其中,  $\mu = \frac{1}{s} \sum_{i=1}^s y_i$ . 类似地, 无噪声的图块集合  $x = \{x_i\}_{i=1}^s \in \mathbb{R}^{r \times s}$  的协方差矩阵可以用下式计算:

$$\Sigma_x = \frac{1}{s} \sum_{i=1}^s (x_i - \mu)(x_i - \mu)^T \quad (3)$$

根据文献[13]和文献[14]可知:

$$\lambda_{\min}(\Sigma_y) = \lambda_{\min}(\Sigma_x) + \sigma_n^2 \quad (4)$$

其中,  $\Sigma_y$  为噪声图块集合协方差矩阵,  $\Sigma_x$  为对应的无噪声图块集合的协方差矩阵,  $\lambda_{\min}(\Sigma_y)$  和  $\lambda_{\min}(\Sigma_x)$  分别为矩阵  $\Sigma_y$  和  $\Sigma_x$  的最小特征值. 如果无噪声图块集合已知, 则根据公式(4)很容易得到噪声水平值. 但是, 实际上无噪声图块集合的协方差矩阵的最小特征值是无法获得的. 尽管如此, 找到相对平滑的区域(即所谓的同质区域), 这时无噪声图块集合  $x$  的  $\lambda_{\min}(\Sigma_x)$  近似为零, 那么公式(4)可以改写为:

$$\hat{\sigma}_n^2 = \lambda_{\min}(\Sigma_y) \quad (5)$$

其中,  $\bar{y}$  是  $y$  集合中由所谓同质图块构成的集合. 公式(5)表明同质图块集合协方差矩阵的最小特征值等于噪声水平值, 因此从原生图块集合中筛选出同质图块子集是基于 PCA 的 NLE 算法的关键步骤.

### 2.2 筛选同质图块子集

Pyatykh 算法<sup>[13]</sup>通过公式(6)所描述的约束条件从原生图块集合  $y$  筛选同质图块构成集合  $\bar{y}$ :

$$\lambda_{m+1} - \lambda_1 < \frac{T\sigma^2}{\sqrt{s}} \quad (6)$$

其中,  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_r$  是图块集合  $\bar{y}$  协方差矩阵所有的特征值(按照从小到大排序),  $T$  是阈值,  $s$  是  $\bar{y}$  集合的基数,  $\lambda_{m+1}$  是与最小特征值  $\lambda_1$  间隔  $m$  个位置的特征值. 这里  $\sigma$  是噪声水平值, 被设定为约束条件中的已知量, 而它正是所要求解的噪声水平值. 为此, Pyatykh 算法采用迭代逼近的方法获得最佳的  $\bar{y}$  图块子集.

Liu 算法<sup>[14]</sup>采用了一种基于图像梯度矩阵的弱纹理强度  $\tau$  作为筛选条件来确定图块集合  $\bar{y}$ :

$$\tau = \sigma^2 F^{-1} \left( \delta, \frac{N^2}{2}, \frac{2}{N^2} \text{tr}(\mathbf{D}_h^T \mathbf{D}_h + \mathbf{D}_v^T \mathbf{D}_v) \right) \quad (7)$$

其中,  $F^{-1}(\delta, \alpha, \beta)$  是逆 Gamma 累积分布函数,  $\alpha$  是形状参数,  $\beta$  是尺度参数,  $\delta$  是置信水平参数,  $N^2$  是图块中像素点的个数,  $\mathbf{D}_h$ ,  $\mathbf{D}_v$  是用于梯度滤波的两个矩阵,  $\sigma$  是噪声水平值. 从公式(7)可知, 选择同质图块的阈值  $\tau$  实际上是给定的置信水平  $\delta$  和噪声水平值  $\sigma$  的一个函数, 而  $\sigma$  值正是需要求解的值, 所以 Liu 算法也需采用

迭代算法求解噪声水平值.

### 2.3 存在的问题

虽然 Pyatykh 算法<sup>[13]</sup>和 Liu 算法<sup>[14]</sup>的预测准确性比较高,但他们复杂的筛选同质图块集合  $\bar{y}$  子集的过程导致他们在执行效率方面的性能都不能令人满意,从而严重影响后继降噪主算法的性能,亟待改进.

## 3 FMNLE 算法

### 3.1 特征值与噪声水平相关性

基于文献[16]中的 90 张和 10 张从互联网随机选取的图像构成测试图像集. 对这些图像分别加上 0 ~ 100, 步长为 5 的高斯噪声, 然后提取每张噪声图像原生图块集合的协方差矩阵的特征值. 参照 Liu 算法的设置, 本文提取的原生图块大小为  $8 \times 8$ , 这样原生图块集合协方差矩阵的大小为  $64 \times 64$ , 其特征值为  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{64}$ . 限于篇幅, 这里仅以最小和最大的特征值 ( $\lambda_1$  和

$\lambda_{64}$ ) 为例来说明特征值与噪声水平值之间的相关性.

由表 1, 表 2 可知: (1) 随着噪声水平的逐渐增大, 特征值也发生相应的变化, 表 1 中各行特征值与第一行真实噪声水平值之间的皮尔森线性相关系数 (Pearson Linear Correlation Coefficient, PLCC) 几乎均在 0.99 以上, 表明  $\lambda_1$  与噪声水平值之间存在高度相关性. 其主要的原因是协方差矩阵经过 PCA 变换后, 其前若干个较小的特征值主要体现的是图像中的噪声信号, 所以相关性好, 尤以  $\lambda_1$  表现得最为显著; (2) 在表 2 中,  $\lambda_{64}$  与噪声水平值的相关性很低, 都在 0.7 以下. 其主要的原因是协方差矩阵经过 PCA 变换后, 其后若干个较大的特征值主要体现的是图像内容信号, 与噪声信号的相关性低, 尤以  $\lambda_{64}$  表现得最为显著. 总之, 从噪声图像分解得到的原生图块集合的协方差矩阵前若干个特征值与噪声水平之间具有显著的相关性.

表 1 5 张噪声图像原生图块协方差矩阵的  $\lambda_1$  特征值与噪声水平值的相关性

图像编号	噪声水平								相关系数
	5	10	15	20	40	60	80	100	
1	4.81	9.08	13.21	17.72	34.86	51.72	69.73	84.90	0.9999
2	4.52	8.78	12.92	17.28	33.88	50.90	68.11	86.45	0.9998
3	4.67	8.89	13.30	17.44	34.93	52.15	67.79	85.69	0.9998
4	5.03	9.23	13.64	18.08	34.85	52.07	68.61	84.89	0.9999
5	4.87	9.17	13.24	17.91	35.05	51.72	68.07	86.17	0.9998

表 2 5 张图像原生图块协方差矩阵的  $\lambda_{64}$  特征值与噪声水平值的相关性

图像编号	噪声水平								相关系数
	5	10	15	20	40	60	80	100	
1	427.12	416.27	437.55	447.39	433.09	446.53	423.28	450.93	0.4004
2	163.14	147.36	155.46	143.37	145.53	155.05	165.89	179.73	0.6823
3	423.12	394.18	410.65	408.22	451.76	408.73	376.84	442.80	0.1001
4	822.15	824.94	821.47	782.75	846.30	805.82	803.18	793.22	0.4007
5	230.70	215.10	218.95	214.76	199.17	235.78	225.40	257.56	0.6145

### 3.2 噪声水平值映射

假设有  $n$  张用于训练的噪声图像, 第  $k$  张训练图像记为  $I^k$ , 该图像被施加已知噪声水平值为  $\sigma_k$  的噪声. 相应的特征矢量记为  $\mathbf{F}_I^k = (\lambda_1, \dots, \lambda_m)$ ,  $m \leq 32$  (文中在实现中最多考虑使用协方差矩阵的前 32 个特征值构成特征矢量, 根据下文实验数据最终确定使用前 16 个特征值). 它的映射目标, 即该噪声图像的噪声水平值记为  $\sigma_k$ . 给定训练数据  $\{(\mathbf{F}_I^1, \sigma_1), \dots, (\mathbf{F}_I^n, \sigma_n)\} \subset \mathbb{F} \times \mathbb{R}$ , 训练的目标是从所有训练数据中找到一个函数 (预测模型)  $\Psi(\mathbf{F}_I^k)$ , 在该预测模型映射下所获得的预测值与目标值  $\sigma_k$  之间的误差尽可能小.

这里, 利用若干层受限波尔曼机 (Restricted Boltzmann Machine, RBM) 和 1 层回归层 (Regression Layer) 构

建 DNN 深度神经网络. RBM 是一种无向图模型, 在输入层 (可见) 和隐藏层 (表示特征) 之间包含无向对称连接, 而层内节点相互之间则无连接. 假设输入层用  $h^{k-1}$  表示, 那么隐藏层为  $h^k$ ,  $k = 2, 3, 4, \dots$ . 这里, 输入层和隐藏层的概念是相对的. 由于 RBM 是依次堆叠的, 当前层作为前一层的隐藏层, 同时也是下一层的输入层. 在基于 RBM 构建的 DNN 深度网络中, 输入层和输出层联合配置的能量被定义为:

$$E(h^{k-1}, h^k; \theta) = - \sum_{i=1}^{H_{k-1}} \sum_{j=1}^{H_k} \omega_{ij} h_i^{k-1} h_j^k - \sum_{i=1}^{H_{k-1}} b_i h_i^{k-1} - \sum_{j=1}^{H_k} c_j h_j^k \quad (8)$$

其中,  $\theta = (w, b, c)$  为模型的参数,  $\omega_{ij}$  为层  $h^{k-1}$  中的可视

节点  $i$  和层  $h^k$  的隐藏节点  $j$  的连接权值.  $b_i$  和  $c_j$  分别为节点  $i$  和  $j$  的偏置值. 在 RBM 网络中, 输出单元在给定输入状态的条件是独立的. 当给出输入数据向量 (Data-Vector) 时, 可以获得来自后验分布的无偏样本, 可以表示为:

$$P(h|v) = \prod_i P(h_i|v) \quad (9)$$

由于  $h_i \in \{0, 1\}$ , 则条件分布概率可以定义为:

$$p(h_j^k = 1|h^{k-1}; \theta) = \varphi\left(\sum_{i=1}^{H^{k-1}} \omega_{ij} h_i^{k-1} + c_j\right) \quad (10)$$

$$p(h_i^{k-1} = 1|h^k; \theta) = \varphi\left(\sum_{j=1}^{H^k} \omega_{ij} h_j^k + b_i\right) \quad (11)$$

其中,  $\varphi(t) = (1 + e^{-t})^{-1}$ . 由公式(8~11)可知, 两层之间的权重和每层的偏置值决定能量  $E$ , 这个值在给定训练集合条件的最小化是确定网络参数的关键. DNN 深度神经网络模型的训练过程中分为两步: (1) 无监督的预训练阶段. 利用 G. Hinton 提出的对比散度 (Contrastive Divergence, CD) 方法<sup>[17]</sup>逐层训练 RBM 网络, 更新每一层模型参数  $\theta = (w, b, c)$ . (2) 有监督的微调阶段. 以预测误差为优化目标函数, 利用反向传播学习算法来调整 DNN 网络各层的模型参数. 预测误差定义为

$$\hat{\varphi} = \operatorname{argmin} \sum_{k=1}^n (\sigma_k - \varphi(g_k))^2 \quad (12)$$

其中,  $g_k$  和  $\sigma_k$  分别表示关于特征矢量  $F_i^k$  在经过多个 RBM 层映射后在最后一个隐藏层的输出值 (特征向量) 和相应的噪声水平值,  $\varphi$  是回归层的核函数. 这个阶段主要任务是利用反向传播方法对各个 RBM 层的模型参数进行微调, 使得 DNN 网络所预测的噪声水平值尽可能与真实值的误差最小.

理论上讲, 文中所提出的 DNN 网络可以通过堆叠更多的 RBM 层构成更深的网络结构从而提高预测模型的准确性. 但实验数据表明通过增加隐含层层数进

一步提升预测准确性的空间不大, 反而会导致在网络训练收敛速度和执行效率下降. 因此, 仅选用具有 3 层 RBM 的网络结构. 为了确定各层最适宜的节点个数, 首先构建具有不同隐含层节点个数的网络模型并在训练图像集合上采用文献[18]中所提出的训练算法完成训练. 然后对 50 张 BSD500 数据库<sup>[19]</sup> 图像施加不同的噪声水平值的噪声干扰, 最后利用各个已训练好不同配置的映射模型对他们的噪声水平值进行估计, 计算所有图像上不同噪声水平值下的估计值与真值的均方根误差 (Root Mean Square Error, RMSE), 列在表 3 中. 限于篇幅, 只列出部分具有代表性的隐含层节点组合情况, 最优 3 个值加粗表示. 基于表 3 中的数据, 文中最终采用隐含层各层节点个数分别设置为 (8, 4, 2) 的 3 层 RBM 构建映射模型. 需要说明的是: 上述各映射模型的网络都是通过离线形式完成训练的, 对最终算法执行时间其实影响不大.

### 3.3 特征矢量维数的确定

选择常用的 BSD500 数据库<sup>[19]</sup> 中的 50 张图像作为测试数据集, 对这些图像加入 0~100, 步长为 5 的高斯噪声, 得到 1050 张失真的图像作为测试图像集合; 另外同样在训练图像数据集合中也施加相同的噪声水平值, 然后分别采用前 4、8、16、24 和 32 个特征值构成特征矢量作为输入训练 DNN 预测网络模型, 用所训练的模型对测试图像集合中图像的噪声进行预测, 计算噪声估计值与真值之间 RMSE, 然后依据 RMSE 大小确定最优的维数. 限于篇幅, 仅给出部分具有代表性的噪声水平值下的 RMSE 数据, 列在表 4 中. 从表 4 可知, RMSE 在取前 16 个特征值的时候整体上比取其它维数要小 (虽然低于 15 和高于 80 时效果不是最好的, 但是离最优值相差不大).

表 3 网络隐含层节点个数的选取对噪声水平估计准确性的影响及其训练网络所需时间

各个隐含层节点的数	噪声水平								均值	时间(s)
	5	10	15	20	40	60	80	100		
14,7,2	18.61	18.84	17.06	14.48	11.52	13.60	10.25	0.16	13.06	329.01
13,7,2	2.06	2.51	2.95	3.78	5.31	0.64	3.14	0.15	2.57	338.35
12,6,3	0.80	0.57	0.53	0.53	1.59	2.09	4.48	0.26	1.36	312.18
11,7,4	2.07	1.67	1.05	0.71	0.99	1.07	0.50	1.23	<b>1.16</b>	370.65
11,5,2	4.55	6.39	6.84	7.51	10.28	2.47	6.52	0.03	5.57	370.02
10,7,2	4.56	7.21	8.48	9.53	13.67	15.08	7.64	7.16	9.17	363.55
10,5,3	1.51	1.30	0.87	0.61	0.90	1.31	2.10	0.50	<b>1.14</b>	348.86
9,6,2	2.29	4.99	7.61	9.49	11.21	10.13	6.27	0.38	1.45	334.94
8,5,3	1.46	1.00	0.70	0.56	1.20	3.41	3.08	0.22	6.54	354.82
8,4,2	1.72	0.94	0.57	0.64	0.52	0.55	0.81	0.92	<b>0.83</b>	346.27

表 4 采用不同维数训练的预测模型在测试图像集合上所预测结果与真值之间的均方根误差

维数	噪声水平							
	5	10	15	20	40	60	80	100
4	<b>1.08</b>	<b>0.57</b>	<b>0.51</b>	<b>0.68</b>	<b>0.96</b>	<b>0.85</b>	<b>0.74</b>	0.96
8	<b>1.47</b>	<b>0.95</b>	<b>0.57</b>	<b>0.50</b>	<b>0.76</b>	<b>0.70</b>	<b>1.10</b>	<b>0.71</b>
16	<b>1.55</b>	<b>0.72</b>	<b>0.45</b>	<b>0.48</b>	<b>0.61</b>	<b>0.68</b>	<b>0.73</b>	0.84
24	1.61	2.01	2.29	2.29	1.90	1.94	1.59	<b>0.72</b>
32	3.28	4.18	4.41	4.34	4.06	3.90	1.81	<b>0.66</b>

### 3.4 两步预测

在全局噪声水平值范围 $[0, 100]$ 内训练一个全局的预测模型,这个模型用来确定噪声水平值大致的范围.然后,在 $[0, 10)$ 、 $[10, 20)$ 、 $[20, 40)$ 、 $[40, 60)$ 、 $[60, 80)$ 和 $[80, 100]$ 六个细分范围上用相同的方法进行训练,得到六个更为精细的预测模型(根据实验数据,粗

精预测模型均使用 16 维特征值作为输入).在 50 张 BSD500 图像集<sup>[19]</sup>上,不同噪声水平段的预测结果的 RMSE 数据列在表 5 中.由表 5 可知,在各个细分噪声水平段上获得的预测误差比全局 $[0, 100]$ 范围内下降很多,这充分说明了两步预测方法的有效性和必要性.

表 5 不同噪声水平范围下 DNN 预测模型的均方根误差比较

噪声水平范围	$[0 - 100]$	$[0 - 10)$	$[10 - 20)$	$[20 - 40)$	$[40 - 60)$	$[60 - 80)$	$[80 - 100]$
均方根误差	1.34	0.64	0.58	0.85	0.85	0.85	0.81

## 4 实验及其分析

### 4.1 实验配置

实验均在 Intel (R) Core (TM) i5-6500 CPU RAM 8GB 主机上完成,软件运行环境为 Matlab R2014a.测试集包括两个图像集合:第 1 个集合由各类文献中常用的 10 张图像构成,如图 1 所示;第 2 个集合为 BSD500

数据库中 50 张图像<sup>[19]</sup>(与确定最佳特征矢量维数所用的图像不同),这个测试集图像纹理信息丰富,非常适合用来测试噪声水平估计的精确性和鲁棒性.参与比较算法包括 Chen 算法<sup>[20]</sup>、Immerkaer 算法<sup>[11]</sup>、Yang 算法<sup>[21]</sup>、Zoran 算法<sup>[11]</sup>、Liu 算法<sup>[14]</sup>和 LME 算法<sup>[2]</sup>.预测准确性和平均执行时间是衡量各个算法性能的 2 个主要评价指标.

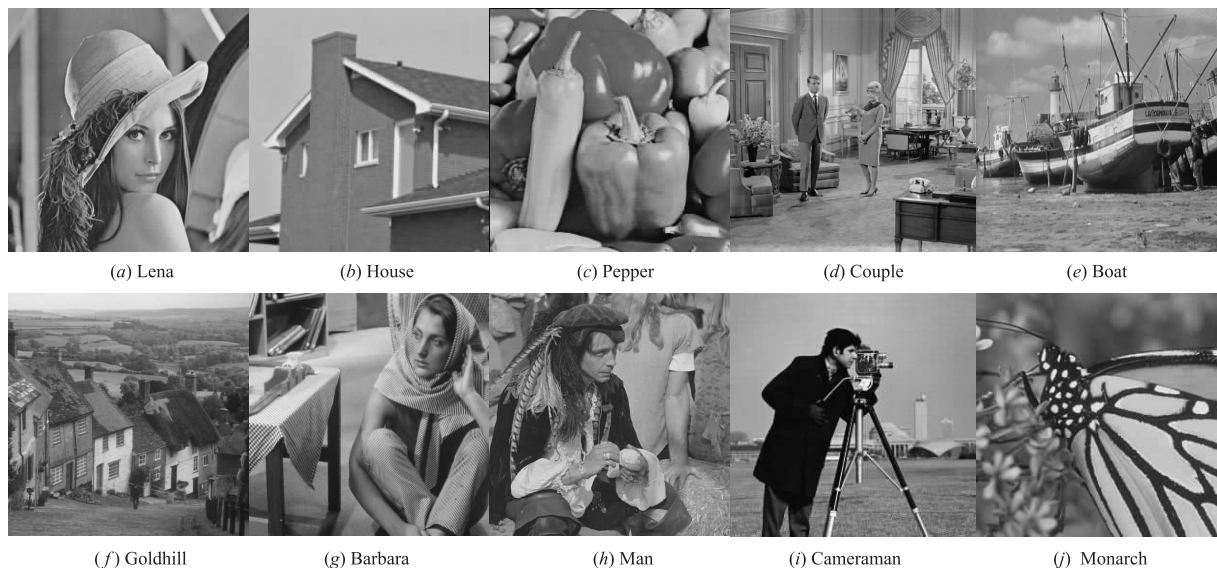


图 1 各类文献中常用 10 张的测试图像

### 4.2 预测准确性

各个 NLE 算法对第 1 个测试图像集合中所有图像预测结果列在表 6 中,使用预测噪声值与真实噪声值

之间的 RMSE 作为评价指标.从表 6 中可以看出:在预测准确性方面,FMNLE 算法在多张图像上具有很好的稳定性,它在各个噪声级别上均进入了前 3 名. Liu 算

法和 Chen 算法虽然在某些噪声水平级别上较 FMNLE 算法要好一些,但由下文表 10 中的执行时间比较实验数据可知,他们的执行时间比 FMNLE 算法要长很多,严重影响了他们的实际使用效果. FMNLE 算法利用基于训练的实现策略和两阶段预测方法在各个级别的噪声条件下实现了比较稳定的预测结果.

为了更加严格地测试各个算法的预测准确性,在

由纹理图像构成的第 2 个测试图像集合上完成了测试,数据列在表 7 中. 由表 7 可知,测试结果与在第 1 个测试图像集合上的测试情况基本类似(各个算法性能存在一定程度的下降),FMNLE 算法仍然表现出非常好的预测准确性和稳定性,这说明它的性能并不依赖图像集合.

表 6 各算法在不同级别下对 10 张测试图的预测结果与真实结果的均方根误差

算法	噪声水平							
	5	10	15	20	40	60	80	100
Chen	1.40	1.02	0.65	0.59	<b>0.68</b>	<b>0.51</b>	<b>0.49</b>	<b>0.58</b>
Immerkaer	1.25	0.94	0.70	0.58	0.93	4.42	10.95	20.46
Yang	<b>0.36</b>	<b>0.42</b>	<b>0.36</b>	<b>0.32</b>	0.82	1.25	1.39	1.84
Zoran	0.88	0.95	1.19	1.45	1.43	2.26	2.41	1.83
Liu	<b>0.07</b>	<b>0.15</b>	<b>0.19</b>	<b>0.22</b>	<b>0.63</b>	<b>0.85</b>	<b>1.32</b>	<b>1.87</b>
LME	1.95	2.97	2.72	2.84	2.35	2.54	2.90	1.33
FMNLE	<b>0.35</b>	<b>0.61</b>	<b>0.49</b>	<b>0.52</b>	<b>0.71</b>	<b>0.78</b>	<b>0.83</b>	<b>0.68</b>

表 7 各算法在不同级别下对 50 张测试图像的预测结果与真实结果的均方根误差

算法	噪声水平							
	5	10	15	20	40	60	80	100
Chen	1.47	0.86	0.66	<b>0.59</b>	<b>0.50</b>	<b>0.47</b>	<b>0.48</b>	<b>0.62</b>
Immerkaer	0.96	0.64	0.52	0.62	2.68	6.81	13.45	22.76
Yang	<b>0.28</b>	<b>0.20</b>	<b>0.22</b>	<b>0.27</b>	0.66	0.82	1.33	1.89
Zoran	1.35	1.27	1.32	1.37	1.83	1.52	1.40	1.34
Liu	<b>0.12</b>	<b>0.12</b>	<b>0.16</b>	<b>0.23</b>	<b>0.51</b>	<b>0.80</b>	<b>1.09</b>	<b>1.48</b>
LME	1.74	1.90	2.04	1.90	1.96	2.88	3.72	1.22
FMNLE	<b>0.24</b>	<b>0.64</b>	<b>0.58</b>	<b>0.85</b>	<b>0.85</b>	<b>0.85</b>	<b>0.85</b>	<b>0.81</b>

为了验证所训练的预测模型的泛化能力,对 50 张测试图像施加不同于构建训练集合时使用的噪声级别(即 7.5、17.5、32.5、47.5、57.5、77.5、92.5)进行测试,测试结果列在表 8 中. 由表 8 可知:FMNLE 算法的预测准确性总体上与在表 6~7 中的表现类似,这说明了 FMNLE 算法具有很好的泛化能力.

### 4.3 BM3D 算法<sup>[7]</sup>降噪效果

为了进一步验证 FMNLE 算法的实际使用效果,用经典 BM3D 算法对加入不同噪声水平的 Lena 图像进行

降噪. 分别使用各对比算法对各噪声级别下的 Lena 图像的噪声水平估计值以及真实噪声水平值作为 BM3D 算法的入口参数进行降噪,将降噪后图像的峰值信噪比 PSNR 值与真实降噪后图像的 PSNR 值之间的绝对差值及其平均值进行对比,数据列在表 9 中. 由表 9 可知,FMNLE 算法在不同噪声水平下对应的误差平均值最小,FMNLE 算法对图像噪声水平值的预测准确性总体上是所有参与比较的 NLE 算法中最好的.

表 8 各算法在不同级别下对 50 张测试图像的预测结果与真实结果的均方根误差

算法	噪声水平						
	7.5	17.5	32.5	47.5	57.5	77.5	92.5
Chen	1.10	0.60	<b>0.51</b>	<b>0.50</b>	<b>0.50</b>	<b>0.54</b>	<b>0.52</b>
Immerkaer	0.77	0.54	2.18	4.00	6.21	12.48	19.04
Yang	<b>0.23</b>	<b>0.25</b>	<b>0.40</b>	<b>0.60</b>	<b>0.73</b>	1.13	1.60
Zoran	1.26	1.34	1.39	1.52	1.54	1.31	1.46
Liu	<b>0.10</b>	<b>0.20</b>	<b>0.36</b>	0.62	0.79	<b>1.10</b>	<b>1.36</b>
LME	1.66	2.04	2.16	2.72	2.79	3.68	4.93
FMNLE	<b>0.32</b>	<b>0.39</b>	0.61	<b>0.46</b>	<b>0.48</b>	<b>0.50</b>	<b>0.68</b>

表 9 使用真实值和各 NLE 算法预测值作为口参数的 BM3D 算法在 PSNR 差值上的比较(单位: dB)

算法	噪声水平								平均值
	5	10	15	20	40	60	80	100	
Chen	0.07	0.03	0.03	0.09	0.16	0.12	0.04	0.18	0.09
Immerkaer	0.06	0.03	0.03	0.09	0.02	<b>0.02</b>	0.60	2.26	0.39
Yang	<b>0.02</b>	<b>0.02</b>	0.01	0.08	0.16	0.10	0.05	0.24	0.09
Zoran	0.07	0.10	0.07	<b>0.01</b>	0.10	0.07	<b>0.01</b>	0.23	0.08
Liu	<b>0.02</b>	<b>0.02</b>	0.02	0.08	<b>0.01</b>	0.11	0.03	0.22	0.06
LME	0.78	0.34	0.07	0.03	0.16	0.11	0.04	0.20	0.22
FMNLE	0.04	0.08	<b>0.00</b>	0.02	0.04	<b>0.02</b>	0.02	<b>0.01</b>	<b>0.03</b>

#### 4.4 算法执行时间的对比

为了比较各算法的执行效率,选取大小为  $512 \times 512$  图像,各算法在图像上重复执行 10 次,以平均执行时间作为实验结果进行比较,实验数据列在表 10 中.从表 10 中可以看出:FMNLE 算法的执行效率排名处于第

2 位,仅比 Immerkaer 算法要慢.考虑到 Immerkaer 算法的噪声水平值预测准确性比较差,FMNLE 算法的综合优势是最好的.故作为众多图像处理算法的前置模块,从预测的准确性和执行效率这两个方面综合考虑,FMNLE 算法相对其他算法而言更具有综合优势.

表 10 各算法在不同噪声级别下的平均执行时间对比(单位:毫秒)

对比	噪声水平							
	5	10	15	20	40	60	80	100
Chen	129.0	131.8	131.9	131.1	129.1	128.7	128.4	129.2
Immerkaer	<b>2.9</b>	<b>3.0</b>	<b>3.4</b>	<b>3.2</b>	<b>3.1</b>	<b>3.4</b>	<b>3.4</b>	<b>3.7</b>
Yang	<b>86.6</b>	<b>87.8</b>	<b>89.3</b>	<b>88.6</b>	<b>88.3</b>	<b>89.0</b>	<b>88.5</b>	<b>88.5</b>
Zoran	1589.6	1563.8	1616.5	1618.7	1637.2	1568.7	1539.2	1534.3
Liu	589.5	629.2	661.6	671.9	697.0	701.3	688.2	688.0
LME	40.3	39.5	38.3	41.5	38.4	41.7	41.3	42.1
FMNLE	<b>12.1</b>	<b>11.6</b>	<b>12.1</b>	<b>12.2</b>	<b>11.7</b>	<b>12.1</b>	<b>11.9</b>	<b>11.9</b>

## 5 总结

FMNLE 算法采用了基于训练的实现策略,没有冗长的筛选同质图块的迭代过程,效率得到了大幅度提高;同时充分利用了 DNN 神经网络强大的映射能力,在预测准确性上也有所提高.未来,可考虑在卷积神经网络(Convolutional Neural Network, CNN)<sup>[22]</sup> 框架下整体一步完成图像特征提取和噪声水平值的映射,所获得的 NLE 算法的性能有望进一步提高.

### 参考文献

- [1] Xu S P, Zeng X X, Jiang Y N, et al. A multiple image-based noise level estimation algorithm [J]. IEEE Signal Processing Letters, 2017, 24(11): 1701 - 1705.
  - [2] 徐少平, 张兴强, 姜尹楠, 等. 局部均值噪声估计的盲 3 维滤波降噪算法[J]. 中国图象图形学报, 2017, 22(4): 422 - 434.
- Xu Shao-ping, Zhang Xing-qiang, Jiang Yin-nan, et al. Noise level estimation based on local means and its appli-

cation to the blind BM3D denoising algorithm [J]. Journal of Image and Graphics, 2017, 22(4): 422 - 434. (in Chinese)

- [3] 周先春, 吴婷, 石兰芳, 等. 一种基于曲率变分正则化的小波变换图像去噪方法 [J]. 电子学报, 2018, 46(3): 621 - 628.
- Zhou Xian-chun, Wu Ting, Shi Lan-fang, et al. A kind of wavelet transform image denoising method based on curvature variation regularization [J]. Acta Electronica Sinica, 2018, 46(3): 621 - 628. (in Chinese)
- [4] Dabov K, Foi A, Katkovnik V, et al. Image denoising with block-matching and 3D filtering [A]. SPIE 6064, Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning [C]. San Jose, CA, United States, 2006, 354 - 365.
  - [5] 杨学志, 叶铭, 周芳, 等. 采用混合特征相似性的极化 SAR 图像降噪算法 [J]. 电子学报, 2016, 44(11): 2583 - 2591.
- Yang Xue-zhi, Ye Ming, Zhou Fang, et al. Speckle reduc-

- tion for polSAR images using hybrid features similarity [J]. *Acta Electronica Sinica*, 2016, 44(11): 2583 – 2591. (in Chinese)
- [6] Ghazi M M, Erdogan H. Image noise level estimation based on higher-order statistics[J]. *Multimedia Tools and Applications*, 2017, 76(2): 2379 – 2397.
- [7] Xu S P, Hu L Y, Yang X H. Quality-aware features-based noise level estimator for block matching and three-dimensional filtering algorithm [J]. *Journal of Electronic Imaging*, 2016, 25(1): Article No. 013029.
- [8] Donoho D L, Johnstone I M. Adapting to unknown smoothness via wavelet shrinkage [J]. *Journal of the American Statistical Association*, 1995, 90(432): 1200 – 1224.
- [9] Immerkaer J. Fast noise variance estimation[J]. *Computer Vision and Image Understanding*, 1996, 64(2): 300 – 302.
- [10] 张旗, 梁德群, 樊鑫. 基于小波域的图像噪声估计新方法[J]. *计算机工程*, 2004, 30(8): 37 – 39.  
Zhang Qi, Liang De-qun, Fan Xin. Estimating image noise based on region segmentation in the wavelet domain[J]. *Computer Engineering*, 2004, 30(8): 37 – 39. (in Chinese)
- [11] Zoran D, Weiss Y. Scale invariance and noise in natural images[A]. 2009 IEEE 12th International Conference on Computer Vision [C]. Kyoto, Japan: IEEE, 2009, 2209 – 2216.
- [12] Liu W, Lin W S. Additive white Gaussian noise level estimation in SVD domain for images[J]. *IEEE Transactions on Image Processing*, 2013, 22(3): 872 – 833.
- [13] Pyatykh S, Hesser J, Zheng L. Image noise level estimation by principal component analysis [J]. *IEEE Transactions on Image Processing*, 2013, 22(2): 687 – 699.
- [14] Liu X H, Tanaka M, Okutomi M. Single-image noise level estimation for blind denoising [J]. *IEEE Transactions on Image Processing*, 2013, 22(12): 5226 – 5237.
- [15] Cui G M, Feng H J, Xu Z H, et al. No-reference image noise estimation based on noise level accumulation [J]. *Optical Review*, 2016, 23(2): 208 – 219.
- [16] Zhang L, Zhang L, Bovik A C. A feature-enriched completely blind image quality evaluator [J]. *IEEE Transactions on Image Processing*, 2015, 24(8): 2579 – 2591.
- [17] Hinton G, Deng L, Dong Y, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups [J]. *IEEE Signal Processing Magazine*, 2012, 29(6): 82 – 97.
- [18] Tanaka M, Okutomi M. A novel inference of a restricted Boltzmann machine [A]. 2014 22nd International Conference on Pattern Recognition [C]. Stockholm, Sweden: IEEE, 2014, 1526 – 1531.
- [19] Abdullah S M, Tischer P, Wijewickrema S, et al. Hierarchical mutual nearest neighbour image segmentation [A]. 2016 International Conference on Digital Image Computing: Techniques & Applications (DICTA) [C]. Gold Coast, QLD, Australia: IEEE, 2016, 1 – 8.
- [20] Chen G Y, Zhu F Y, Heng P A. An efficient statistical method for image noise level estimation [A]. 2015 IEEE International Conference on Computer Vision [C]. Santiago, Chile: IEEE, 2015, 477 – 485.
- [21] Yang S M, Tai S C. Fast and reliable image-noise estimation using a hybrid approach [J]. *Journal of Electronic Imaging*, 2010, 19(3): 033007.
- [22] Zhang K, Zuo W M, Chen Y J, et al. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising [J]. *IEEE Transactions on Image Processing*, 2017, 26(7): 3142 – 3155.

#### 作者简介



徐少平 男, 1976 年生于江西九江. 博士, 南昌大学信息工程学院计算机科学与技术系教授, 博士生导师. 主要研究方向为图形图像处理、机器视觉、虚拟手术仿真等.  
E-mail: xushaoping@ncu.edu.cn



李崇禧 男, 1994 年生于江西吉安. 现为南昌大学硕士研究生, 主要研究方向为图像处理与计算机视觉.



林官喜 女, 1992 年生于江西吉安. 2018 年获南昌大学硕士学位, 主要研究方向为图像处理与计算机视觉.

唐祎玲 女, 1977 年生于浙江奉化. 现为南昌大学博士研究生, 主要研究方向为图像处理与计算机视觉.

胡凌燕(通信作者) 女, 1978 年生于江西临川. 博士, 南昌大学信息工程学院自动化系教授, 博士生导师. 主要研究方向为机器视觉、自动控制、虚拟现实等.

E-mail: hulingyan@ncu.edu.cn