

## 基于局部期望最大化注意力的图像降噪

李泽田, 雷志春

引用本文:

李泽田, 雷志春. 基于局部期望最大化注意力的图像降噪[J]. 液晶与显示, 2020, 35(4): 350–359.

LI Ze-tian, LEI Zhi-chun. Local expectation-maximization attention network for image denoising[J]. *Chinese Journal of Liquid Crystals and Displays*, 2020, 35(4): 350–359.

在线阅读 View online: <https://doi.org/10.3788/YJYXS20203504.0350>

## 您可能感兴趣的其他文章

Articles you may be interested in

### 基于局部路径特征信息神经网络的图像去噪

Image denoising based on local path feature in formation neural network

液晶与显示. 2020, 35(1): 70–79 <https://doi.org/10.3788/YJYXS20203501.0070>

### 基于注意力掩模融合的目标检测算法

Object detection algorithm based on attention mask fusion

液晶与显示. 2019, 34(8): 825–833 <https://doi.org/10.3788/YJYXS20193408.0825>

### 高斯混合模型自适应微光图像增强

Automatic low light level image enhancement using Gaussian mixture modeling

液晶与显示. 2015(2): 300–309 <https://doi.org/10.3788/YJYXS20153002.0300>

### 基于核模糊C-均值和EM混合聚类算法的遥感图像分割

Remote sensing image segmentation based on KFCM and EM hybrid clustering algorithm

液晶与显示. 2017, 32(12): 999–1005 <https://doi.org/10.3788/YJYXS20173212.0999>

### 异构无线网络多用户多业务接入中多目标优化控制方法研究

Multi-objective optimization control method for multi-user multi-service access in heterogeneous wireless networks

液晶与显示. 2019, 34(7): 708–715 <https://doi.org/10.3788/YJYXS20193407.0708>

文章编号:1007-2780(2020)04-0350-10

## 基于局部期望最大化注意力的图像降噪

李泽田, 雷志春\*

(天津大学 微电子学院, 天津 300072)

**摘要:** 图像降噪算法生成的图像往往会因丢失高频信息而造成细节模糊, 并且随着网络的加深, 噪声在特征域会不断累加。为了恢复图像降噪中生成图像的高频细节, 增强网络的拟合能力和泛化能力, 减少噪声在特征域的传播, 本文提出了一种基于局部期望最大化注意力机制的降噪网络。网络以经典的自动编码器为基本模块以达到足够的深度, 实现足够大的感受野, 并滤除不必要的低频信息, 促进网络学习高频信息。局部期望最大化注意力模块利用局部期望最大化算法迭代出紧凑的局部基, 并在该局部基上应用注意力机制, 大大降低网络复杂度。大量实验表明, 本文提出的局部期望最大化注意力网络可以恢复更多的高频细节, 在基准测试集 BSD68 上, PSNR 值比当前最先进算法 N<sup>3</sup>Net 提升约 0.13 dB, 并有更好的视觉质量。

**关键词:** 图像降噪; 局部期望最大化; 注意力; 自动编码器

**中图分类号:** TP391.1    **文献标识码:** A    **doi:** 10.3788/YJYXS20203504.0350

## Local expectation-maximization attention network for image denoising

LI Ze-tian, LEI Zhi-chun\*

(School of Microelectronics, Tianjin University, Tianjin 300072, China)

**Abstract:** Images generated by image denoising algorithms often result in missing details due to loss of high frequency information. In order to restore the high-frequency details of the reconstructed images and enhance the representation ability of the network, a noise removing network based on local expectation-maximization attention (LEMA) is proposed. The network uses the classic autoencoder as the basic module to achieve sufficient depth, so as to filter out unnecessary low-frequency information and promote network learning high-frequency information. The proposed channel and spatial attention module adaptively learns informativel features by computing the interdependence of image channels and pixel spatial positions. A series of experiments show that the LEMA can recover more high frequency details. On the benchmark BSD68, the PSNR value is about 0.13 dB higher than the state-of-the-art algorithm N<sup>3</sup>Net, and has better visual quality.

**Key words:** image denoising; local expectation-maximization; attention; autoencoder

收稿日期: 2019-09-13; 修订日期: 2019-11-04.

基金项目: 天津市科技计划项目 (No.16YFZCGX00760)

Supported by Tianjin Science and Technology Program (No.16YFZCGX00760)

\* 通信联系人, E-mail: zclei@tju.edu.cn

## 1 引言

图像降噪是计算机视觉中的基本问题,也是许多其他图像恢复任务的重要预处理步骤<sup>[1-6]</sup>。通常,图像降噪的目的是从其退化图像恢复出无噪声图像<sup>[7]</sup>。

图像降噪具有非适定性(Ill-posed)<sup>[8]</sup>。传统图像降噪方法<sup>[9-12]</sup>应用各种约束来模拟噪声特征(例如,均匀、非均匀、深度感知),并利用不同的图像先验来规范解空间。然而,在大多数方法中使用的图像先验都是手工制作的,并且消耗大量的计算资源。

最近几年,随着深度卷积神经网络(Deep Convolutional Neural Network, DCNN)的发展,图像去噪性能得到了显著提升。文献[13]通过将批量归一化集成到残差网络中,提出了一种称为DnCNN(Denoising Convolutional Neural Networks)的有效图像降噪网络,该网络优于传统的降噪算法。文献[14]基于小波分解,提出了一种基于多级小波的降噪算法,称为MWCNN(Multi-level Wavelet-CNN)。文献[15]提出了一种快速灵活的网络称为FFDNet,可以处理具有不均匀噪声的图像。为了利用DCNN中图像特征的非局部性,文献[16]在去噪网络中采用k-最近邻匹配,提出N<sup>3</sup>Net算法。

由于DCNN强大的拟合能力,上述基于DCNN的降噪算法较传统算法效果提升显著。但这些算法也有其局限性。首先,采用“编码器-解码器”级联的方式构建模型,在计算每一个解码器的状态时需要关注所有的编码器输入,计算量比较大,灵活性差,网络的泛化能力较低。其次,平等地处理不同通道、不同区域的特征,效率较低,浪费有限的计算资源。最后,在DCNN中,随着图像在网络中的传播,特征域的噪声会不断累加<sup>[17]</sup>。

为克服以上基于DCNN的降噪算法的缺点,本文设计了一个基于局部期望最大化注意力的图像降噪算法。近年来,注意力机制在自然语言处理领域取得卓越成果<sup>[18]</sup>。Nonlocal<sup>[19]</sup>第一次将注意力机制应用到计算机视觉领域。RCAN<sup>[20]</sup>提出后,注意力在低层视觉领域也受到了广泛的关注。注意力机制关注图像中突出的特征(纹理

细节等),并使每个像素充分利用不同通道的语义信息,有利于提升网络效率。然而,这些注意力机制需要生成一个巨大的注意力图,其空间复杂度和时间复杂度巨大。其瓶颈在于,每一个像素的注意力图都需要对全图计算。期望最大化注意力机制(Expectation-Maximization Attention, EMA)<sup>[21]</sup>,摒弃了在全特征图上计算注意力图的流程,首先通过期望最大化(EM)算法迭代出一组紧凑的基,然后在这组基上运行注意力机制,大幅降低了计算复杂度。然而,EMA仍然是针对全图计算期望值,这对于降噪任务不是最优方案。因为噪声属于快速变化的高频信号,本文认为,应该利用特征图的局部信息提取噪声注意力图,即采用局部期望最大化(Local Expectation-Maximization Attention, LEMA)算法替代EMA算法,通过局部期望最大化(LEM)算法迭代出一组紧凑的局部基,在这组局部基上运行注意力机制,这样不仅可以有效滤除噪声,还可以进一步降低复杂度。本文把这一机制嵌入网络中,构造出轻量且易实现的LEMA Unit,在多个数据集上取得了较好的降噪效果。本文的主要贡献如下:

(1)改进了期望最大化算法,提出了更适于降噪任务的局部期望最大化算法(LEMA),该算法对输入的局部特征进行迭代得到一组紧凑的局部基,在这组局部基上计算注意力图可以大幅降低计算复杂度。

(2)将LEMA机制引入DCNN中,结合通道注意力与空间注意力,构造出LEMA Unit,并将LEMA Unit嵌入到自动编码器网络中,提出LEMANet模型。

本文进行了大量的图像降噪实验,以验证LEMANet的有效性。结果表明,LEMANet在视觉质量和定量测量方面均实现了领先的降噪性能。

## 2 基于DCNN的图像降噪

近几年的时间里,卷积神经网络发展的一个趋势是为了获得较好的性能,不断地增加网络层数。DCNN通过堆叠多个简单的非线性单元从大量标记数据中学习高度非线性映射<sup>[22]</sup>。

自动编码器(Autoencoder)是一种无监督的学习算法,主要用于数据的降维或者特征的提取,

由编码器(Encoder)和解码器(Decoder)两部分组成,每一部分均采用堆叠 DCNN 的方式构成。编码器将输入图像(设为  $x$ )经过加权  $W$ 、映射( $\varphi$ )之后提取到图像主特征  $y$ :

$$y = f_{\theta}(x) = \varphi(W^{(L-1)} \varphi(W^{(L-2)} \cdots \varphi(W^{(1)} x) \cdots)) \quad (1)$$

其中: $L$  为编码器层数。

解码器对  $y$  反向加权  $W^T$ 、映射( $\varphi$ )得到输出  $z$ ,

$$z = g_{\theta}(y) = W^{(1)T} \varphi(W^{(2)T} \cdots \varphi(W^{(L)T} y) \cdots) \quad (2)$$

其中: $L$  为解码器层数。

反向传播过程通过反复迭代训练两组权值  $W$ 、 $W^T$ ,使得损失函数(MSE,感知损失等)最小,尽可能保证  $z$  近似于  $x$ ,即可重构输入图像  $x$ 。自动编码器结构如图 1 所示。 $f_{\theta}$  是编码器映射,由于输入数据  $x$  中存在相关性,为了提取图像表征的关键特征,在编码过程施加“瓶颈”(Bottleneck),即对输入数据进行压缩,提取得到关键特征  $y$ 。解码器负责由特征  $y$  解码恢复出重构数据  $z$ 。在文献[23-25]中,采用降噪自动编码器(Denoising Autoencoder, DAE)方法进行高斯噪声去除,并获得与 BM3D<sup>[26]</sup>、K-SVD<sup>[27]</sup>等传统算法相当的结果,显示了自动编码器方法的降噪潜力。

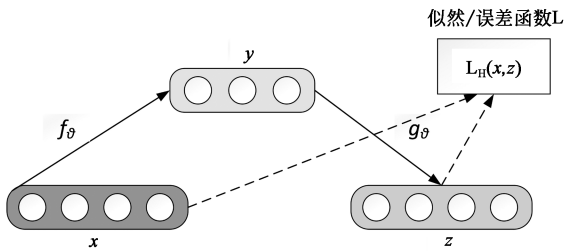


图 1 自动编码器结构图

Fig.1 Structure of autoencoder

### 3 局部期望最大化注意力机制

#### 3.1 局部期望最大化算法

期望最大化(Expectation-Maximization, EM)<sup>[28]</sup>算法旨在找到隐含变量模型的最大似然解。用  $Y$  表示观测变量的数据,  $Z$  表示隐含变量的数据,  $Y$  和  $Z$  一起构成完全数据。假设观测数据  $Y$  概率分布为  $P(Y|\theta)$ , 其中  $\theta$  是模型待估计的参数,  $Y$  和  $Z$  的联合概率分布为  $P(Y, Z|\theta)$ 。

一般情况下,  $Z$  的分布未知, 在  $\theta$  给定的情况下, 通过后验概率  $P(Z|Y, \theta)$  迭代得到。完全数据的似然函数为  $\ln P(Y, Z|\theta)$ , EM 算法通过迭代的方法求  $\ln P(Y, Z|\theta)$  的极大似然估计, 每次迭代分两步: E 步和 M 步, 不断重复 E 步和 M 步, 直到收敛。

在 E 步, 根据给定的  $\theta$  通过后验概率  $P(Z|Y, \theta)$  求得  $Z$  的分布, 然后求完全数据的似然函数的数学期望  $E_{Z|Y, \theta}$ :

$$E_{Z|Y, \theta^i} [\ln P(Y, Z|\theta^{i+1})] = \sum_Z \ln P(Y, Z|\theta^{i+1}) P(Y, Z|\theta^i) \quad (3)$$

式中:  $\theta^i$  为第  $i$  次迭代的模型参数,  $\theta^{i+1}$  为第  $i+1$  次迭代的模型参数。

在 M 步, 求使式(3)极大化的  $\theta^i$ , 确定第  $i+1$  迭代的参数估计值  $\theta^{i+1}$ ,

$$\theta^{i+1} = \arg \max_{\theta} E_{Z|Y, \theta^i} \quad (4)$$

式中  $\arg \max$  为极大值函数。

考虑到图像不同区域的噪声分布并不一定是均匀的, 所以本文提出在图像局部应用期望最大化机制, 其结构如图 2 所示。首先采用边缘检测, 区分不同的噪声区域<sup>[29]</sup>。然后在每一区域内部应用期望最大化机制, 迭代出一组紧凑的局部基, 为后续计算注意力特征图做准备。最后, 在局部基上计算注意力特征图:

$$\text{LEM}(z) = \varphi_{\text{local}}([E(z), M(z)]) \quad (5)$$

其中:  $\text{LEM}(z)$  为局部基,  $\varphi_{\text{local}}$  为局部噪声检测算子,  $[E(z), M(z)]$  为期望最大化算子。

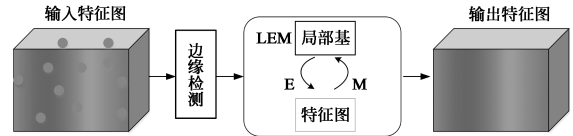


图 2 局部期望最大化注意力模块

Fig.2 Local expectation-maximization attention module

#### 3.2 通道注意力

网络的浅层特征包含丰富的高频信息, 这对降噪后图像细节的恢复非常重要, 但这些高频信息在特征通道间被平等对待, 这阻碍了网络的表达能力。为克服这一缺点, 本文将通道注意力机制引入到图像降噪中。

通道注意力主要关注输入的抽象信息<sup>[30]</sup>。输入图像各特征通道间的信息有高度相关性,

利用这一相关性,滤除通道冗余信息,使网络专注于关键信息。生成通道注意力,不仅可以提高网络性能,还可以降低网络复杂度。为了有效计算通道注意力,需要对输入特征图进行降维,即减少冗余信息,主要是低频信息。常用的降维方法是利用均值池化<sup>[31-32]</sup>,但是均值池化在滤除低频的同时,会损失一些高频信息。为此,本文同时使用了均值池化与最大池化,实验证明效果比单独使用均值池化性能优异。通道注意力满足方程:

$$M_C(F) = \sigma(\text{ReW}(\text{AvgPool}(F)) + \text{ReW}(\text{MaxPool}(F))) = \sigma(W_1(W_0(F_{\text{avg}}^C)) + W_1(W_0(F_{\text{max}}^C))) \quad (6)$$

式中: $\sigma$ 为Sigmoid函数, $\text{ReW}()$ 为特征评估函数, $\text{AvgPool}()$ 为均值池化函数, $\text{MaxPool}()$ 为最大池化函数, $F \in R^{C \times H \times W}$ 为某中间特征图, $W_0$ 、 $W_1$ 为降维后特征权值, $F_{\text{avg}}^C$ 和 $F_{\text{max}}^C$ ,分别为降维后的特征图。视觉质量对比实验如图3、4所示。通道注意力模块的结构如图5所示。

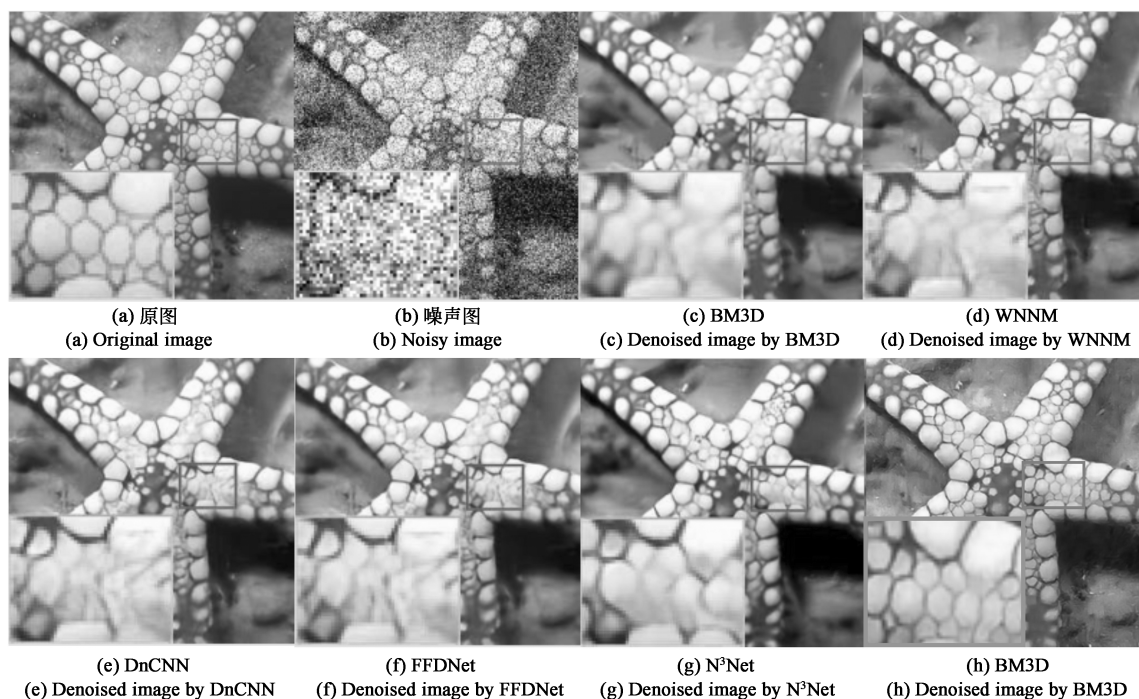
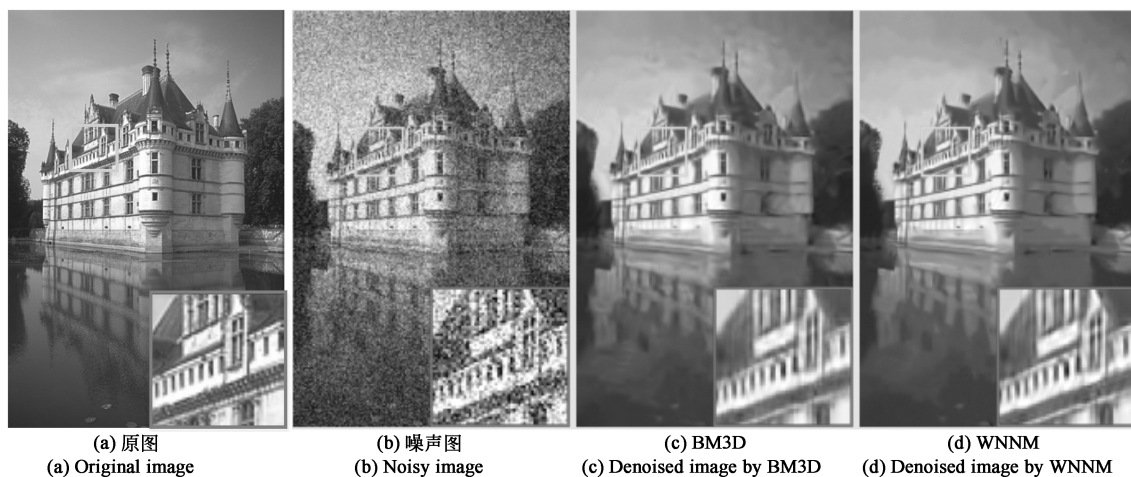


图3 视觉质量对比。图像噪声水平为25的降噪结果。

Fig.3 Visual quality comparison. Denoising results of one image from BSD68 with noise level 25.



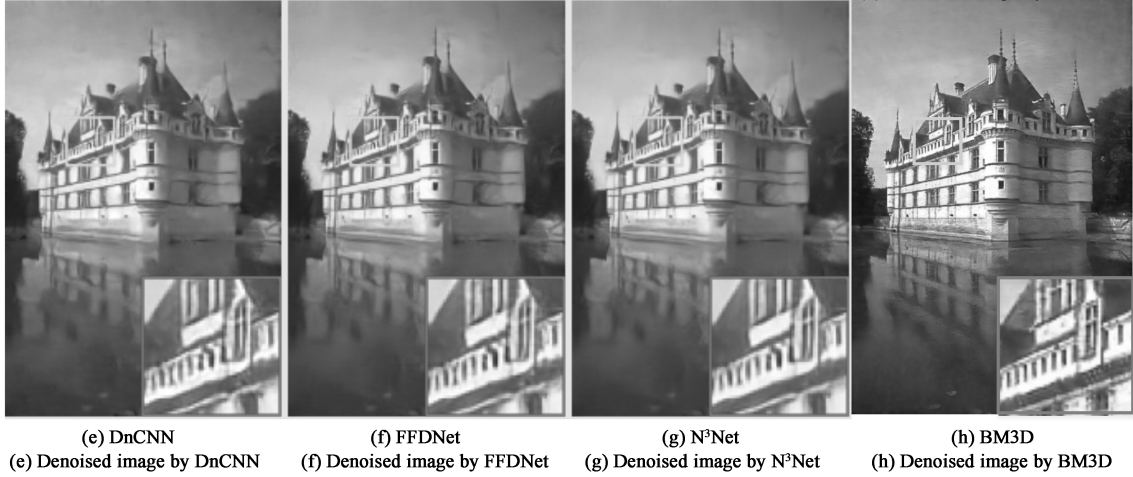


图 4 视觉质量对比。图像噪声水平为 50 的降噪结果。

Fig.4 Visual quality comparison. Denoising results of one image from BSD68 with noise level 50.

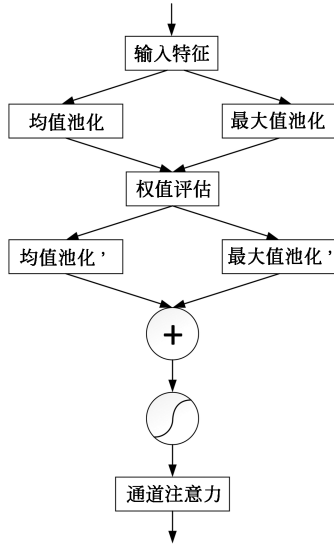


图 5 通道注意力模块结构

Fig.5 Structure of channel attention moduel

上一层特征图分别经过均值池化层和最大池化层,前者主要提取图像的背景信息,记为  $F_{avg}^C$ ,后者主要提取图像的纹理信息,记为  $F_{max}^C$ 。将二者通过权值评估模块进行参数压缩,分别得到  $F_{avg}^{C'}$  和  $F_{max}^{C'}$ ,为了减少参数量,对二者进行特征压缩,则  $F_{avg}^{C'} \in \mathbb{R}^{C/r \times 1 \times 1}$ ,  $F_{max}^{C'} \in \mathbb{R}^{C/r \times 1 \times 1}$ ,其中  $r$  是压缩率。然后进行元素级求和实现特征融合,并通过 Sigmoid 函数做归一化处理,即得到通道注意力。

### 3.3 空间注意力

空间注意力主要关注输入图像中的位置信

息<sup>[26]</sup>。首先在特征通道的维度上使用最大池化和平均池化得到两个不同的特征图:  $F_{avg}^S \in \mathbb{R}^{1 \times H \times W}$  和  $F_{max}^S \in \mathbb{R}^{1 \times H \times W}$ ,然后在通道维度上将这两个特征图连接,最后经过一个  $7 \times 7$  的卷积层得到空间注意力。空间注意力满足方程:

$$M_S(F) = \sigma(f^{7 \times 7}$$

$$([\text{AvgPool}(F); \text{MaxPool}(F)])) \quad (7)$$

式中:  $\sigma$  为 Sigmoid 函数;  $f^{7 \times 7}()$  为  $7 \times 7$  的卷积操作; AvgPool() 为均值池化函数; MaxPool() 为最大池化函数;  $[\cdot]$  表示通道维度上特征的连接操作。

空间注意力模块的结构如图 6 所示。

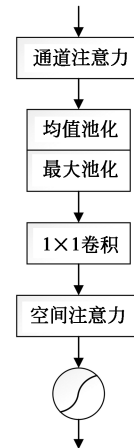


图 6 空间注意力模块结构

Fig.6 Structure of spatial attention moduel

### 3.4 全局注意力网络

在卷积神经网络中,设卷积核为  $n \times n$ ,则该

层卷积只关注邻域为  $n$  的区域,即使通过增加层数、采用扩张卷积使感受野越来越大,也难以学习全局语义信息<sup>[19]</sup>。全局注意力网络就是提取全局特征,也就是图像中任何像素在当前像素的映射的叠加。

全局注意力满足方程:

$$z_i = F(x_i, \delta(\sum_{j=1}^{N_p} \alpha_j x_j)) \quad (8)$$

式中:  $x_j$  表示第  $j$  个像素;  $\alpha_j$  为第  $j$  个像素对当前像素的映射权值,  $\sum_{j=1}^{N_p}()$  为所有像素映射在当前像素的加权和;  $\delta(\cdot)$  为特征传递函数,  $F(\cdot, \cdot)$  表示特征聚合函数。

全局注意力网络如图7所示,其结构由3部分组成:(1)注意力生成,采用  $1 \times 1$  卷积和 softmax 函数来获取注意力权重,然后与输入特征图相乘得到全局注意力特征;(2)特征传递,此部分采用  $1 \times 1$  卷积、实例归一化和非线性激活层(ReLU)级联的方式;(3)全局注意力聚合,得到全局注意力。该模块中的  $1 \times 1$  卷积可以实现特征数的压缩,从而降低计算复杂度。

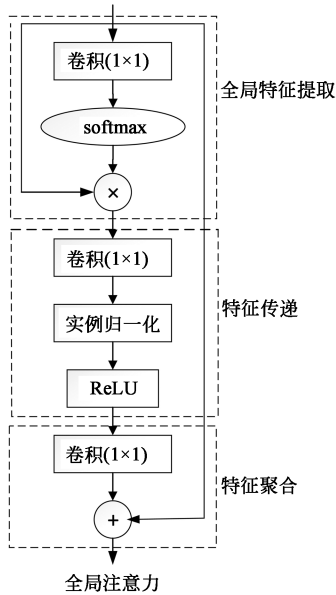


图7 全局注意力模块结构

Fig.7 Structure of global attention network

#### 4 局部期望最大化注意力网络

局部期望最大化注意力网络,利用局部期望最大化机制迭代出一组紧凑的局部基,然后在这组局部基上利用通道注意力模块提取图像的抽象

信息,即图像中的内容;利用空间注意力模块提取图像中的位置信息,即图像中各要素间的相对位置。全局注意力模块负责计算图像中所有像素点对于某一特定像素的映射,其中全局池化操作可以提高网络的拟合能力,减少网络层数,加速训练。

局部期望最大化注意力网络由5部分组成:

(1)浅层特征提取模块,由一个  $3 \times 3$  的卷积层构成,用于提取浅层特征,这些特征包含大量的高频信息;(2)深层特征提取模块,用于提取深层特征,这些特征包含抽象信息,由一个自动编码器 Autoencoder 构成,Autoencoder 采用的是 Encoder-Decoder 结构;(3)局部期望最大化模块,用于迭代出一组紧凑的局部基;(4)注意力模块,由通道注意力、空间注意力网络及全局注意力网络构成;(5)图像重建模块,由一层卷积核大小为  $1 \times 1$  的卷积层构成,因为生成的是彩色图像,所以输出通道为3。整个网络结构如图7所示。

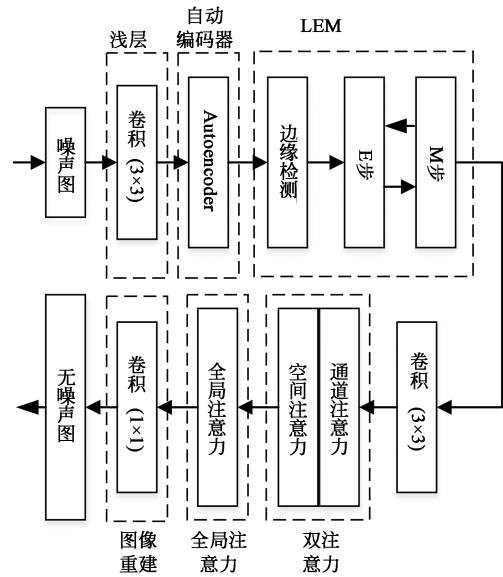


图8 全局双注意力网络结构

Fig.8 Structure of global dual attention network

通道注意力模块与空间注意力模块中采用最大池化与均值池化操作,突出主要特征,抑制次要特征,生成通道注意力与空间注意力。全局注意力模块采用全局池化提取图像全局特征,生成全局注意力。由于抑制了对降噪任务贡献较小的次要特征,因而使用注意力机制降低了特征图计算复杂度。

## 5 测量实验与结果

为了验证本文推出的局部期望最大化注意力网络的降噪性能,进行了多次实验,并与其他现有的降噪模型进行了全面对比,包括深度学习和非深度学习的方法:BM3D<sup>[26]</sup>、WNNM<sup>[12]</sup>、DnCNN<sup>[13]</sup>、FFDNet<sup>[15]</sup>、N<sup>3</sup>Net<sup>[16]</sup>。

### 5.1 实验参数与数据集

#### 5.1.1 参数设置

浅层与深层特征提取模块,卷积核尺寸为  $3 \times 3$ ,步长为 1,特征通道数为 64;全局特征提取模块,卷积核尺寸为  $1 \times 1$ ,步长为 1,特征通道数为 128;图像重建模块,卷积核尺寸为  $1 \times 1$ ,步长为 1,特征通道数为 3。

训练集原图使用 DIV2K 数据集<sup>[33]</sup>,该数据集来自 CVPR NTIRE 2017 挑战的高质量(2K 分辨率)图像数据集。DIV2K 数据集包含 800 个训练图像,100 个验证图像和 100 个测试图像。噪声图则依据文献<sup>[14]</sup>,生成 3 种噪声等级的噪声图,即  $\sigma=15, 25, 50$ 。测试集使用 3 个标准数据集: Set12<sup>[17]</sup>、BSD68<sup>[34]</sup> 和 Urban100<sup>[35]</sup>。实验结果测试指标在 YCbCr 空间的 Y 通道上用峰值信

噪比(PSNR)和结构相似性指数(SSIM)<sup>[36]</sup>评价。

#### 5.1.2 实验细节

对 800 个训练图像进行数据增强,训练图像随机旋转  $90^\circ, 180^\circ, 270^\circ$  并进行水平翻转。在每个训练批次中,提取 16 对  $64 \times 64$  的原图与噪声图作为输入。训练过程中,使用 ADAM 优化器<sup>[37]</sup>最小化损失函数。初始学习率设置为  $1E-4$ ,并随训练过程呈指数衰减,衰减系数为 0.9。

### 5.2 实验参数与数据集

与 5 个最先进算法进行了对比实验: BM3D<sup>[26]</sup>、WNNM<sup>[12]</sup>、DnCNN<sup>[13]</sup>、FFDNet<sup>[15]</sup>、N<sup>3</sup>Net<sup>[16]</sup>。测试集为 Set12<sup>[13]</sup>、BSD68<sup>[34]</sup> 和 Urban100<sup>[35]</sup>。实验结果如表 1 所示。可以看出,本文提出的 LEMA 的 PSNR 值比其他最先进算法均有显著提升。 $\sigma=25$  时,对于测试集 Set12, LEMA 的 PSNR 值比传统方法 BM3D 和 WNNM 提升  $0.36 \sim 0.67$  dB 左右,比 DnCN 提升  $0.20$  dB 左右,比 FFDNet 提升  $0.21$  dB,比 N<sup>3</sup>Net 提升  $0.06$  dB。在数据集 BSD68 与 Urban100 上, LEMA 均表现出较明显的提升。

视觉质量对比实验如图 3、图 4 所示。可以看出,相较于其他算法,LEMA 的降噪效果明显,保留更多的细节信息。

表 1 与最先进算法对比实验

Tab.3 Performance comparison with the state of the art

| 算法                 | Set12       |             |             | BSD68       |             |             | Urban100    |             |             |
|--------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
|                    | $\sigma=15$ | $\sigma=25$ | $\sigma=50$ | $\sigma=15$ | $\sigma=25$ | $\sigma=50$ | $\sigma=15$ | $\sigma=25$ | $\sigma=50$ |
| BM3D               | 32.37       | 29.97       | 26.72       | 31.08       | 28.57       | 25.60       | 32.34       | 29.70       | 25.94       |
| WNNM               | 32.70       | 30.28       | 27.05       | 31.37       | 28.83       | 25.87       | 32.97       | 30.39       | 26.83       |
| DnCNN              | 32.86       | 30.44       | 27.18       | 31.73       | 29.23       | 26.23       | 32.67       | 29.97       | 26.28       |
| FFDNet             | 32.75       | 30.43       | 27.32       | 31.63       | 29.19       | 26.29       | 32.42       | 29.92       | 26.52       |
| N <sup>3</sup> Net | —           | 30.50       | 27.43       | —           | 29.30       | 26.39       | —           | 30.19       | 26.82       |
| LEMA               | 32.88       | 30.64       | 27.65       | 31.86       | 29.43       | 26.52       | 32.91       | 30.36       | 26.94       |

## 6 结 论

本文提出了一种基于局部期望最大化注意力的图像降噪网络,该网络利用期望最大化在图像局部迭代出局部基,然后在局部基上采用注意力网络,在保证精度的同时,大大降低特征图的计算

复杂度。与最先进的降噪算法对比显示该算法的性能提升明显。 $\sigma=25$  时,对于测试集 Set12, LEMA 的 PSNR 值比传统方法 BM3D 和 WNNM 提升约  $0.36 \sim 0.67$  dB,比 DnCN 提升约  $0.20$  dB,比 FFDNet 提升  $0.21$  dB,比 N<sup>3</sup>Net 提升  $0.06$  dB。在数据集 BSD68 与 Urban100 上, LEMA 也均表现出较明显的提升。

## 参 考 文 献:

- [1] HEIDE F, STEINBERGER M, TSAI Y T, *et al.* FlexISP: a flexible camera image processing framework [J]. *ACM Transactions on Graphics*, 2014, 33(6): 231.
- [2] ZHANG K, ZUO W M, GU S H, *et al.* Learning deep CNN denoiser prior for image restoration [C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA: IEEE, 2017: 2808-2817.
- [3] 肖启阳, 李健, 吴思进, 等. 基于能量的 DSPI 相位图正余弦降噪 [J]. 光学 精密工程, 2018, 26(4): 936-943.  
XIAO Q Y, LI J, WU S J, *et al.* Denoising of DSPI phase map using sine-cosine filtering and signal energy [J]. *Optics and Precision Engineering*, 2018, 26(4): 936-943. (in Chinese)
- [4] 陈真诚 吴贤亮 赵飞骏. EEMD 结合小波阈值的光电容积脉搏波信号降噪方法及实现 [J]. 光学 精密工程, 2019, 27(6): 1327-1334.  
CHEN Z C, WU X L, ZHAO F J. Denoising and implementation of photoplethysmography signal based on EEMD and wavelet threshold [J]. *Optics and Precision Engineering*, 2019, 27(6): 1327-1334. (in Chinese)
- [5] 孙京阳, 喻春雨, 董仕佳. 非局部均值噪声预测的独立成分分析降噪研究 [J]. 光学 精密工程, 2018, 26(2): 511-516. SUN J Y, YU C Y, DONG S J. Noise reduction of independent component analysis based on NLmeans noise prediction. [J]. *Optics and Precision Engineering*, 2018, 26(2): 511-516. (in Chinese)
- [6] 丁明君, 习江涛, 李光旭, 等. 基于三波长相移轮廓术的表面测量降噪 [J]. 光学 精密工程, 2017, 25(10s): 297-303.  
DING M J, XI J T, LI G X, *et al.* Denoising method for shape measurement based on three wavelength phase shift profilometry. [J]. *Optics and Precision Engineering*, 2017, 25(10s): 297-303. (in Chinese)
- [7] BUADES A, COLL B, MOREL J M. A non-local algorithm for image denoising [C]//*Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Diego, USA: IEEE, 2005: 60-65.
- [8] ROTH S, BLACK M J. Fields of experts [J]. *International Journal of Computer Vision*, 2009, 82(2): 205-229.
- [9] CHO S, LEE S. Fast motion deblurring [J]. *ACM Transactions on Graphics*, 2009, 28(5): 145.
- [10] XU L, JIA J Y. Two-phase kernel estimation for robust motion deblurring [C]//*Proceedings of the 11th European Conference on Computer Vision*. Heraklion, Crete, Greece: Springer, 2010: 157-170.
- [11] 唐玲, 陈明举. 基于图像局部方向特性的自适应全变分去噪模型 [J]. 液晶与显示, 2016, 31(5): 477-483.  
TANG L, CHEN M J. Adaptive directional total variation denoising model based on image local direction [J]. *Chinese Journal of Liquid Crystals and Displays*, 2016, 31(5): 477-483. (in Chinese)
- [12] GU S H, ZHANG L, ZUO W M, *et al.* Weighted nuclear norm minimization with application to image denoising [C]//*Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, USA: IEEE, 2014: 2862-2869.
- [13] 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.
- [14] LIU P J, ZHANG H Z, ZHANG K, *et al.* Multi-level wavelet-CNN for image restoration [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. Salt Lake City, USA: IEEE, 2018: 88601-88609.
- [15] ZHANG K, ZUO W M, ZHANG L. FFDNet: toward a fast and flexible solution for CNN-based image denoising [J]. *IEEE Transactions on Image Processing*, 2018, 27(9): 4608-4622.
- [16] PLÖTZ T, ROTH S. Neural nearest neighbors networks [C]//*Proceedings of the 32nd Conference on Neural Information Processing Systems*. Montréal, Canada: NuruIPS, 2018: 1087-1098.
- [17] LIAO F Z, LIANG M, DONG Y P, *et al.* Defense against adversarial attacks using high-level representation guided denoiser [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA: IEEE, 2018: 1778-1787.

- [18] VASWANI A, SHAZEER N, PARMAR N, *et al.* Attention is all you need [C]//*Proceedings of 31st Conference on Neural Information Processing Systems*. Long Beach, CA, USA: NeurIPS, 2017: 5999-6009.
- [19] WANG X L, GIRSHICK R, GUPTA A, *et al.* Non-local neural networks [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA: IEEE, 2018: 7794-7803.
- [20] ZHANG Y L, LI K P, LI K, *et al.* Image super-resolution using very deep residual channel attention networks [C]//*Proceedings of European Conference on Computer Vision*. Munich, Germany: Springer, 2018: 294-310.
- [21] LI X, ZHONG Z S, WU J L, *et al.* Expectation-maximization attention networks for semantic segmentation [C]//*Proceedings of 2019 International Conference on Computer Vision*. Seoul, Korea: ICCV, 2019.
- [22] MAO X J, SHEN C H, YANG Y B. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections [C]//*Proceedings of the 30th International Conference on Neural Information Processing Systems*. Barcelona, Spain: Curran Associates Inc, 2016: 2802-2810.
- [23] BURGER H C, SCHULER C J, HARMELING S. Image denoising: can plain neural networks compete with BM3D? [C]//*Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition*. Providence, RI, USA: IEEE, 2012: 2392-2399.
- [24] XIE J Y, XU L L, CHEN E H. Image denoising and inpainting with deep neural networks [C]//*Proceedings of the 25th Neural Information Processing Systems*. Lake Tahoe, Nevada: Curran Associates Inc, 2012: 341-349.
- [25] CHO K. Simple sparsification improves sparse denoising autoencoders in denoising highly corrupted images [C]//*Proceedings of the 30th International Conference on Machine Learning*. Atlanta, USA: PMLR, 2013: 432-440.
- [26] DABOV K, FOI A, KATKOVNIK V, *et al.* Image denoising by sparse 3-D transform-domain collaborative filtering [J]. *IEEE Transactions on Image Processing*, 2007, 16(8): 2080-2095.
- [27] ELAD M, AHARON M. Image denoising via sparse and redundant representations over learned dictionaries [J]. *IEEE Transactions on Image Processing*, 2006, 15(12): 3736-3745.
- [28] DEMPSTERA P, LAIRD N M, RUBIN D B. Maximum likelihood from incomplete data via the EM algorithm [J]. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1977, 39(1): 1-22.
- [29] 于长淦,方超.基于小波变换的 ESPI 图像去噪及边缘提取[J].*液晶与显示*,2011,26(6):818-822.  
YU C S, FANG C. ESPI image denoising and edge extraction based on wavelet transform [J]. *Chinese Journal of Liquid Crystals and Displays*, 2011, 26(6): 818-822. (in Chinese)
- [30] WOO S, PARK J, LEE J Y, *et al.* CBAM: convolutional block attention module [C]//*Proceedings of the 15th European Conference on Computer Vision*. Munich, Germany: Springer, 2018: 3-19.
- [31] JOHNSON J, ALAHI A, LI F F. Perceptual losses for real-time style transfer and super-resolution [C]//*Proceedings of the 14th European Conference on Computer Vision*. Amsterdam, The Netherlands: Springer, 2016: 694-711.
- [32] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. arXiv preprint arXiv:1409.1556, 2014.
- [33] TIMOFTE R, AGUSTSSON E, VAN GOOL L, *et al.* NTIRE 2017 challenge on single image super-resolution: methods and results [C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Honolulu, USA: IEEE, 2017: 1110-1121.
- [34] MARTIN D, FOWLKES C, TAL D, *et al.* A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics [C]//*Proceedings of the 8th IEEE International Conference on Computer Vision*. Vancouver, Canada: IEEE, 2001: 416-423.
- [35] HUANG J B, SINGH A, AHUJA N. Single image super-resolution from transformed self-exemplars [C]//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston, USA: IEEE, 2015: 5197-5206.
- [36] WANG Z, BOVIK A C, SHEIKH H R, *et al.* Image quality assessment: from error visibility to structural similarity [J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.

- [37] KINGMA D P, BA J. Adam: a method for stochastic optimization [C]//*Proceedings of the 3rd International Conference for Learning Representations*. San Diego, USA: ICLR, 2015.

作者简介:



李泽田(1991—),男,河北邢台人,硕士,2015年于河北工业大学大学获得硕士学位,主要从事图像降噪,计算机视觉的研究。E-mail: lizetian527@tju.edu.cn



雷志春(1964—),男,河北迁安人,博士,教授,1998年于德国多特蒙德大学获得博士学位,主要从事高动态范围与宽色域算法研究。E-mail: zclei@tju.edu.cn