

一种基于深度学习的多聚焦图像融合算法

陈清江¹, 李毅^{1*}, 柴昱洲²

¹西安建筑科技大学理学院, 陕西 西安 710055;

²空间电子信息技术研究院, 陕西 西安 710100

摘要 针对深度学习在计算机视觉上的良好表现,提出一种基于深度学习的多聚焦图像融合算法,在原有的 AlexNet 网络模型基础上改进了卷积核大小、步长等;利用改进后的深度学习网络特有的得分机制分类了聚焦图像块与散焦图像块;使用矫正矩阵矫正了误判图像块,并细分、修复了融合后的图像聚焦与散焦分界区域,得到了融合图像;选取 6 组多聚焦图像验证了本文算法的有效性。实验结果表明:与其他算法相比,运用本文算法进行图像融合,能够保存较多的图像原始高频信息,并在互信息、边缘信息保持度、平均梯度和熵等评价指标上取得了较好的表现。

关键词 图像处理; 多聚焦图像融合; 深度学习网络; 矫正矩阵; 图像块分类; 边界修复

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/LOP55.071015

A Multi-Focus Image Fusion Algorithm Based on Depth Learning

Chen Qingjiang¹, Li Yi¹, Chai Yuzhou²

¹ College of Science, Xi'an University of Architecture and Technology, Xi'an, Shaanxi 710055, China;

² Xi'an Institute of Space Radio Technology, Xi'an, Shaanxi 710100, China

Abstract Aiming at the good performance in computer vision for depth learning, a multi-focus image fusion algorithm based on depth learning is proposed. Based on the existing AlexNet network model, the convolution kernel size and step size are improved. The focused image blocks and the defocused image block are classified by using the improved scoring mechanism of deep learning network. Then, the correction matrix is used to correct the misjudgment image blocks. The boundary zone of image focus and defocus is subdivided and repaired. Six pairs of multi-focus images are randomly selected to verify the effectiveness of the proposed method. The experimental results show that, compared with other algorithms, the fusion results obtained by this algorithm can preserve the original high-frequency information of the image, and achieve good performance on four evaluation indexes of mutual information, edge information retention, entropy and average gradient.

Key words image processing; multi-focus image fusion; deep learning network; correction matrix; image block classification; boundary restoration

OCIS codes 100.3008; 200.4260; 350.2660

1 引言

图像融合是将同一场景,由不同传感器获得的两幅或多幅图像在精确配准后融合为一幅图像的过程。多聚焦图像融合是图像融合领域的一个重要分支,由于光学镜头在景深上的限制,可见光成像系统很难将拍摄场景中距离差异较大的所有目标都清晰

成像。要获得一个全聚焦图像要将同一场景多个不同聚焦设置的图像进行融合,就需采用多聚焦图像融合法。多聚焦图像融合算法还可以用于其他图像融合任务,如可见光红外图像融合和医学图像融合等。现有的多聚焦图像融合方法^[1]大致可以分为变换域方法和空间域方法,其中最经典的变换域融合方法是基于多尺度变换(MST)理论上的融合方法,

收稿日期: 2018-01-03; 收到修改稿日期: 2018-02-07

基金项目: 国家自然科学基金(61403298)、陕西省自然科学基金(2015JM1024)、陕西省教育厅专项科研计划(2013JK0586)

作者简介: 陈清江(1966—),男,博士,教授,主要从事小波分析及其在图像处理中的应用等方面的研究。

E-mail: chen66xauat@126.com

* 通信联系人。E-mail: 314098601@qq.com

随着拉普拉斯金字塔(LP)的融合方法^[2-3]的出现,大量基于多尺度变换的图像融合方法开始应用于这个领域。代表性的例子包括以形态金字塔(MP)、离散小波变换(DWT)^[4]、双树复小波变换(DTCWT)为基础的方法,以及基于非子采样轮廓波变换(NSCT)与剪切波^[5]的方法等。近年来,变换域融合方法带来了新的冲击,其采用滑动窗口技术追求近似平移不变的融合过程,例如基于稀疏表示的方法^[6]。

早期的空间域方法通常采用基于图像块的融合策略,其中源图像分解成块,图像块大小对融合结果的质量有很大的影响。最早出现了使用手动固定尺寸的基于图像块的方法,已经提出了许多改进的版本,如基于自适应块的方法和基于轮廓的方法^[7],使用差分进化算法获得固定的最佳块尺寸;最近又引入了基于二叉树^[8]方法,其中图像可根据图像内容自适应地分成不同大小的图像块。另一种类型的空间域方法是基于分块的方法,与基于图像分割的思想相似,但是这些方法的融合质量在很大程度上依赖于分割的准确性。

文献[9]提出基于小波与自适应块的融合(DWTDB)算法,文献[10]提出基于离散余弦变换与方差(DCT+Var)算法,以及文献[11-12]提出的基于脉冲卷积神经网络和拉普拉斯金字塔变换的像融合(LP-PCNN)算法与基于多尺度加权梯度的图像融合(MWGF)算法,这些方法在提取高频信息的同时,很容易造成高频信息的损失。近年来,深度学习在图像处理领域取得了巨大的成功。结合深度学习,本文提出一种基于深度学习图像块分类的多聚焦图像融合算法。对 AlexNet 网络模型^[13]进行较大改进,将改进后的网络模型成功运用在小图像块的聚焦与散焦分类上。实验结果表明:改进后的网络模型在分类的精度上取得了极大的提高。将用 Sobel 算子预处理的后的图像块在 Caffe 框架下输

入改进后的网络模型,利用深度学习网络较好的学习能力和优秀的得分机制对聚焦清晰的图像块与聚焦模糊的图像块做分类,得到两幅互补的多聚焦图像。利用简单的矫正矩阵矫正误判图像块,并对聚焦与散焦边界区域做细分与修复。与已有的融合算法相比较,运用本文算法进行融合的图像效果较好。

2 基于深度学习的分类网络模型

经典的基于分类的网络模型有 LeNet 网络模型^[14]、AlexNet 网络模型^[13]、VGGNet 网络模型^[15]以及 GoogLeNet 网络模型^[16]等。本文的深度学习网络主要用作对聚焦清晰的图像块与聚焦模糊的图像块进行分类。由于输入的图像块大小只有 16 pixel×16 pixel,深层的网络模型会造成最后的特征信息损失。尽管 VGGNet 与 GoogLeNet 在精度上更优于 AlexNet,但由于网络层数过多,因此将 AlexNet 作为分类网络模型,并在其模型结构基础上做出改进。

2.1 AlexNet 网络模型及分类原理

AlexNet 共有 8 层,其前 5 层是卷积层,后 3 层是全连接层,最后还有一个 softmax 损失层。图像输入大小为 227 pixel×227 pixel,第一个卷积层滤波器个数为 96 个,大小为 11 pixel×11 pixel,卷积后利用修正线性单元(ReLU)激活函数进行非线性映射,再进行局部归一化,池化后的输出作为第二个卷积层的输入。第二个卷积层与第一个卷积层类似,卷积核个数为 256 个,大小为 5 pixel×5 pixel。第三个卷积层到第五个卷积层之间直接连接,这三个卷积层之间没有池化层,第五个卷积层进行 ReLU 与池化后连接到第一个全连接层,共有 4096 个节点,第一个全连接层经 ReLU 与 dropout 后输入到第二个全连接层,以此类推,连接到第三个全连接层,最后输入到 softmax 损失层,输出结果。其网络结构如图 1 所示。

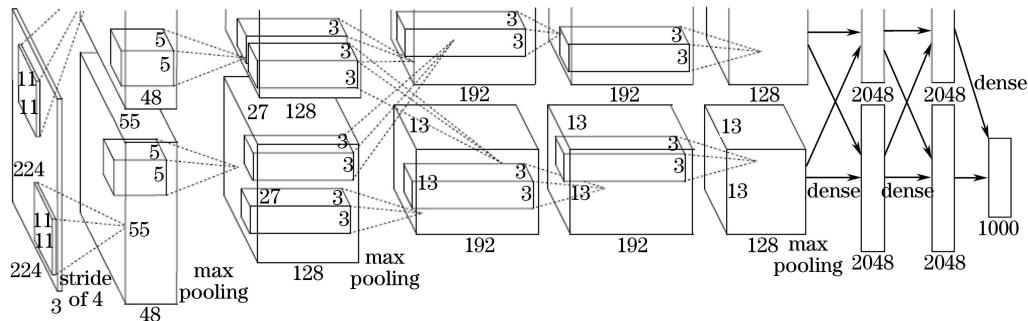


图 1 AlexNet 网络结构

Fig. 1 AlexNet network structure

该模型采用 $\text{ReLU}(x) = \max(0, x)$ 作为非线性激活函数,并对激活后的结果进行局部归一化,归一化公式为

$$b_{x,y}^i = a_{x,y}^i \cdot \left[k + \alpha \cdot \sum_{j=\max(0, i-n/2)}^{\min(N-1, i+n/2)} (a_{x,y}^j) \right]^{-2\beta}, \quad (1)$$

式中: $a_{x,y}^i$ 为在点 (x, y) 处, 计算出的神经元的激活值; $b_{x,y}^i$ 为归一化后的激活度, 然后再通过 ReLU 激活函数, 得到激活后的值; N 为层中卷积核的数量; 常数 α, β, k, n 是超参数, 一般设置为 $\alpha = 10^{-4}, \beta = 0.75, k = 2, n = 5$ 。这一策略可以降低误差率。

AlexNet 网络模型在池化层采用了最大重叠池化, 与非重叠池化相比, 重叠池化将 top1 和 top5 的误差分别降低了 0.4% 和 0.3%, 而且重叠池化不容易过拟合。模型在全连接层后加了 Dropout 层, 起到了很好的效果, 减少网络参数, 防止过拟合。

2.2 改进的分类网络模型

由于图像输入大小的特殊性, 不能直接应用 AlexNet 网络模型进行分类, 所以对其进行改进。

在 AlexNet 网络模型中, 第一个卷积层 conv1 中采用大小为 $11 \text{ pixel} \times 11 \text{ pixel} \times 3 \text{ pixel}$ 的核, 考虑到输入的图像大小只有 $16 \text{ pixel} \times 16 \text{ pixel}$, 而较大的卷积核会造成细节特征损失, 则需要对卷积核大小做出改变。参考 VGGNet 模型中对深度卷积网络的实验结果与经验可知, $3 \text{ pixel} \times 3 \text{ pixel}$ 大小的卷积核是最小的能够捕获上下左右和中心概念的尺寸。多个 $3 \text{ pixel} \times 3 \text{ pixel}$ 的卷积可以替代大尺寸滤波器, 不仅具有更少的参数, 而且具有更好的非线性, 使得判决函数更加具有判决性。

借鉴于 VGGNet 网络模型中将所有卷积核大小调节到 $3 \text{ pixel} \times 3 \text{ pixel}$ 取得的优秀表现, 在改进的 AlexNet 网络模型中将第一个卷积层 $11 \text{ pixel} \times 11 \text{ pixel}$ 的卷积核改为 $3 \text{ pixel} \times 3 \text{ pixel}$, 同时第二个卷积层中 $5 \text{ pixel} \times 5 \text{ pixel}$ 的卷积核也改为 $3 \text{ pixel} \times 3 \text{ pixel}$, 其余三个卷积层的 $3 \text{ pixel} \times 3 \text{ pixel}$ 卷积核保持不变。由于输出的目标分类较少, 只有两类, 因此不需要增加卷积核的数量, 在原网络模型中保持卷积核的数目不变。

原网络模型在第一层卷积层 conv1 与池化层 pool1 的步长分别为 4 和 2, 第二层与第五层的池化层(pool2 与 pool5)步长为 2。但由于输入图像的大小只有 $16 \text{ pixel} \times 16 \text{ pixel}$, 步长太大就会导致图像大小缩小过快, 不利于特征的提取。为了避免步长太大, 造成卷积和池化后的图像信息损失过快, 本文

将网络中所有的步长设置为 1。改进后的网络模型在训练过程中取得了较好的结果。当迭代 5000 次时, 精度达到了 0.992。

3 基于深度学习的多聚焦图像融合算法

3.1 训练数据图像获取与预处理

对训练图像选取了大量纹理清晰、多样的聚焦图像作为样本。由于网络训练同时需要聚焦清晰图像与聚焦模糊的图像做有监督训练, 聚焦清晰的图像容易得到, 而人工提取聚焦模糊的图像需要做大量提取与分割工作, 因此选用高斯滤波器对聚焦清晰的图像做平滑处理以代替自然的聚焦模糊图像。高斯滤波是一种线性平滑滤波, 可以使滤波后的图像在细节特征上与聚焦模糊的图像相似, 通过实验发现这种替代方法是可行的。选用大小为 $7 \text{ pixel} \times 7 \text{ pixel}$ 的高斯滤波器对聚焦清晰的图像进行滤波处理, 为了更加接近自然的聚焦模糊图像, 得到不同程度散焦的图像, 在高斯滤波器处理后的图像基础上, 再次使用高斯滤波进行滤波处理, 以此类推。数据选取时, 随机选取不同版本的聚集模糊图像与聚焦清晰的图像组成图像对作为数据集, 实验结果表明, 3 个不同版本的聚焦模糊图像训练得到的网络精度最高, 如表 1 所示。

表 1 数据集选择与处理的测试结果

Table 1 Test results of data set selection and processing

Filter version	1	2	3	4	5
Sobel pretreatment	0.886	0.942	0.992	0.986	0.980
No pretreatment	0.852	0.891	0.936	0.928	0.924

为了训练的网络能够更好的提取图像特征, 在之前使用 Sobel 算子对数据集进行预处理, 如表 1 所示, 测试结果表明预处理可以较为明显的提高网络精度。Sobel 算子可以较好的提取图像轮廓及纹理特征, 并主要用于边缘检测, 从而加快网络训练收敛速度。

改进后的网络模型输入大小为 $16 \text{ pixel} \times 16 \text{ pixel}$, 需要对预处理后的聚焦图像与散焦图像数据集进行分块, 选择适当的训练集与测试集, 然后再输入网络进行训练与测试。整个训练过程在 Caffe 框架下经过 5000 次的迭代达到收敛, 得到了精度为 0.992 的网络模型。

3.2 图像分类矫正

在融合实验中, 使用 Sobel 算子对两幅多聚焦图像进行滤波处理, 利用训练好的网络模型对其逐块计算聚焦得分情况, 并比较对应图像块的清晰聚

焦得分情况,分别赋予 0、1 值。对清晰聚焦得分高的记为 1,低的记为 0,得到了两个互补的 0、1 矩阵。

虽然得到的网络模型精度较高,但是得到的 0、1 矩阵还是有可能存在误判情况,对此可以分情况设置矫正矩阵以提高算法准确性。对于 0、1 矩阵判定为 0 的非边界点 $e(i, j)$,如图 2 所示,根据其临近 8 个点为 1 的个数进行判断:

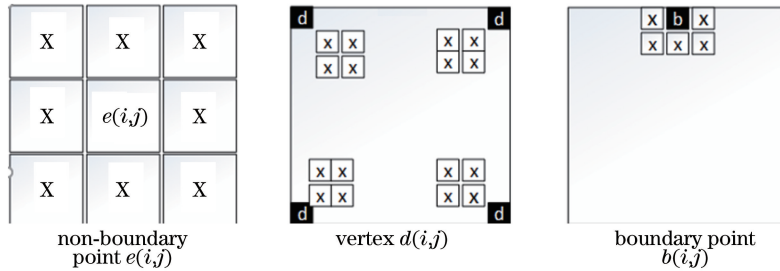


图 2 矫正矩阵

Fig. 2 Correction matrix

同样,若顶点 $d(i, j)=0$,则选取其对角线 4 个点设置矫正矩阵对其进行矫正,由图 2 可知,若矫正矩阵 4 个点中非零点个数大于 2,则判定 $d(i, j)=1$,同时判定对应的另一个 0、1 矩阵顶点值为 0,反之则相反。

对于非顶点的边界点 $b(i, j)=0$,则如图 2 所

$$\begin{cases} e(i, j) = 1, & \text{cot} > 4 \\ e(i, j) = 0, & \text{cot} \leq 4 \end{cases}, \quad (2)$$

式中 cot 是临近点为 1 的个数,若 cot 大于 4,则判断该点为 1,同时判定其对应点为 0,反之则相反。为了提高判断的精度,得到更新后的 0、1 矩阵对其进行再判断,判断两次后得到了新的 0、1 矩阵。

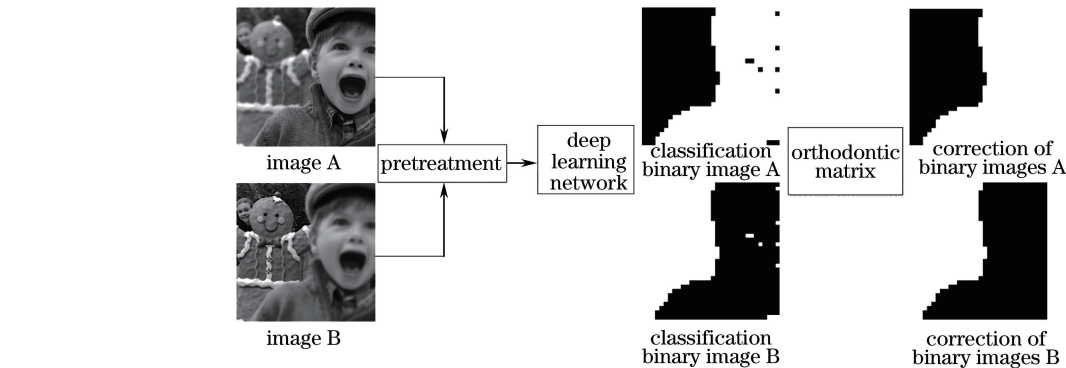


图 3 矫正后分类图像

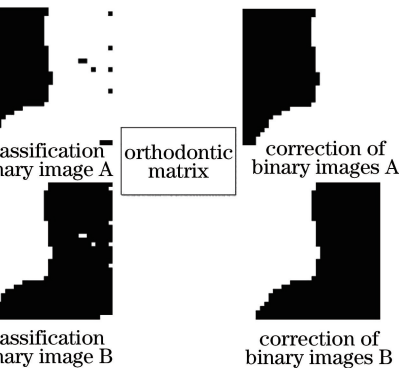
Fig. 3 Corrected classification images

3.3 图像分类边界修复

经过矫正矩阵处理,可以得到两个互补的二值图,利用原始多聚焦图像对两幅二值图进行填补,得到两幅互补的清晰聚焦图像。将其叠加后可得到一幅如图 4 所示的聚焦图像。选取的分块图像大小为 $16 \text{ pixel} \times 16 \text{ pixel}$,叠加后的图像在叠加边界出现了明显的块效应。为了消除边界的块效应,提高融合后图像的质量,可将融合前的两幅图像的聚焦与散焦分界区域部分进行修复处理。

对于所提取出来的边界区域,仍可以看作若干

个 $16 \text{ pixel} \times 16 \text{ pixel}$ 的小块进行处理。利用 Brenner 梯度函数、Tenengrad 梯度函数以及拉普拉斯能量函数等图像清晰度评价指标对小块做像素级分类。其对应数学表达式如下。



1) Brenner 梯度函数。Brenner 梯度函数是最简单的梯度评价函数,它只是简单的计算相邻两个像素灰度差的平方,该函数定义为

$$D(f) = \sum_j \sum_i |f(i+2, j) - f(i, j)|^2, \quad (3)$$

式中 $f(i, j)$ 为像素点所在位置。

2) Tenengrad 梯度函数。Tenengrad 梯度函数

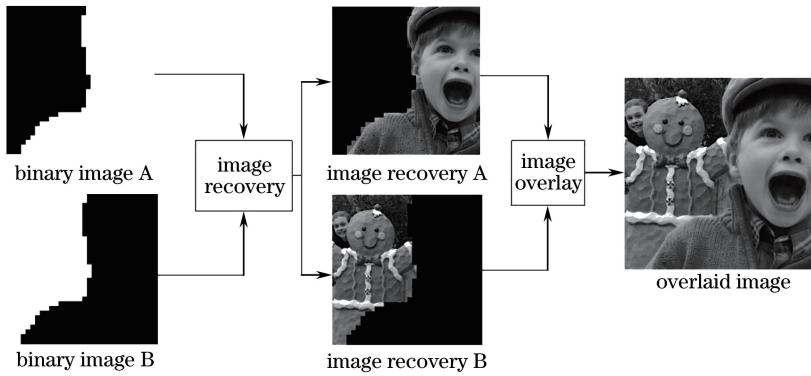


图 4 叠加后聚焦图像

Fig. 4 Focus image after overlay

采用 Sobel 算子分别提取水平和垂直方向的梯度值,基于 Tenengrad 梯度函数的图像清晰度定义为

$$D(f) = \sum_j \sum_i |G(i, j)|, \quad G(i, j) > T, \quad (4)$$

式中: $G(i, j) = \sqrt{G_i^2(i, j) + G_j^2(i, j)}$, G_i 和 G_j 分别为点 (i, j) 在水平和垂直方向的边缘检测 Sobel 算子的卷积; T 为给定的边缘检测阈值。

3) 拉普拉斯能量函数是空间域焦点检测的首选函数,其函数计算方法为

$$M_q = H \otimes (\nabla^2 I_q)^2, \quad q \in \{A, B\}, \quad (5)$$

式中 H 为平均滤波器, \otimes 、 ∇^2 分别为卷积和拉普拉斯算子。对于滤波器 H , 选用 $5 \text{ pixel} \times 5 \text{ pixel}$ 大小的高斯滤波器。拉普拉斯算子对于具体像素点 (i, j) 表示为

$$\begin{aligned} \nabla^2 I &= 123\partial^2 I / \partial i^2 + \partial^2 I / \partial j^2 = \\ &[-I(i-1, j-1) - 4I(i-1, j) - \\ &I(i-1, j+1) - 4I(i, j-1) + 20I(i, j) - \\ &4I(i, j+1) - I(i+1, j-1) - \\ &4I(i+1, j) - I(i+1, j+1)]. \end{aligned} \quad (6)$$

分别利用这三个评价指标来计算分界区域图像

块的每一个像素点的得分情况,并与其对应图像块的得分进行比较,采取三取二的标准进行分类。若其中两项或以上评价指标表明边界图像块在点 (i, j) 处计算值大于其对应图像块中的计算值,则令 (i, j) 为 1, 反之则为 0。经评价指标判断后的图像块变为二值图图像块。

所得到的决策图可能在聚焦与散焦的边界区域仍存在坏点的情况,与文献[17]处理的情况类似,本文同样采用引导滤波器来提高决策图的质量。引导滤波器是一种非常有效的边缘保留滤波器。将图像边界的结构信息导入到滤波器中,新得到的决策图就作为引导图像来指导图像滤波。引导过滤算法中有两个自由参数:窗口半径 r 和正则化参数 ϵ , 分别设定为 8 和 0.1。具体处理步骤如图 5 所示。将修复的边界与融合后的图像叠加即可得到最终的融合结果。

4 实验结果及分析

为了验证和评估本文算法的有效性,将测试 6 组多聚焦图像。并与当前主流算法: DWTDB 算法^[9]、DCT + Var 算法^[10]、LP-PCNN 算法^[11]、

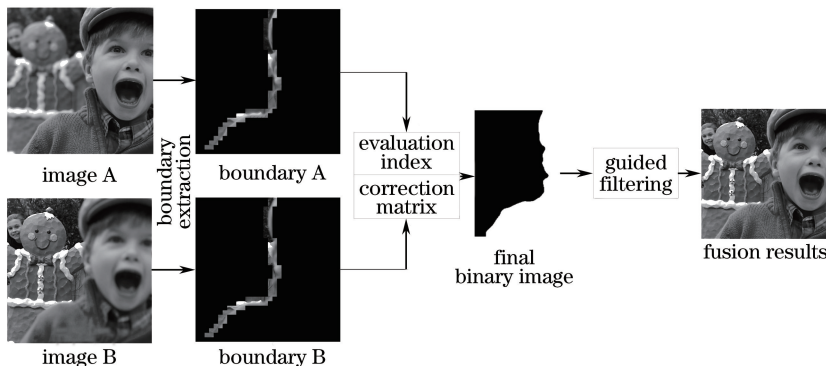


图 5 边缘修复后的融合结果

Fig. 5 Fusion results after edge reparation

MWGF 算法^[12]以及基于稀疏表示的融合(SR)算法^[17]进行比较。并使用互信息(MI)、边缘信息保持度($Q^{AB/F}$)、熵(E)与平均梯度(AG)等评价指标对融合结果进行分析对比。

4.1 实验步骤

1) 训练聚焦与散焦分类的网络模型,并对训练好的网络模型参数进行提取。

2) 对实验的多聚焦图像进行 Sobel 算子滤波预处理,将预处理后的图像输入训练好的网络模型进行逐块分类判断。

3) 利用矫正矩阵分情况对误判图像块进行矫正。

4) 对分类边界部分运用图像清晰度评价指标以三取二的机制进行像素级细分,并对分类后的边界部分进行滤波处理。

5) 将修复的边界与分类好的图像叠加即为融合的最终结果。

4.2 算法有效性分析

随机选取了 Lytro 数据集中的 6 对多聚焦图像 Diving、Boy、Statue、Tree、Baby 和 Book 进行实验,图像大小均为 $512 \text{ pixel} \times 512 \text{ pixel}$,并将其转换到灰度空间进行融合。图 6~11 分别为 6 对多聚焦图像利用不同算法的融合结果,表 2 所示为 6 对多聚焦图像在 4 种评价指标下的表现结果。



图 6 不同算法对源图像 Diving 融合结果

Fig. 6 Fusion results of the source image Diving with different algorithms

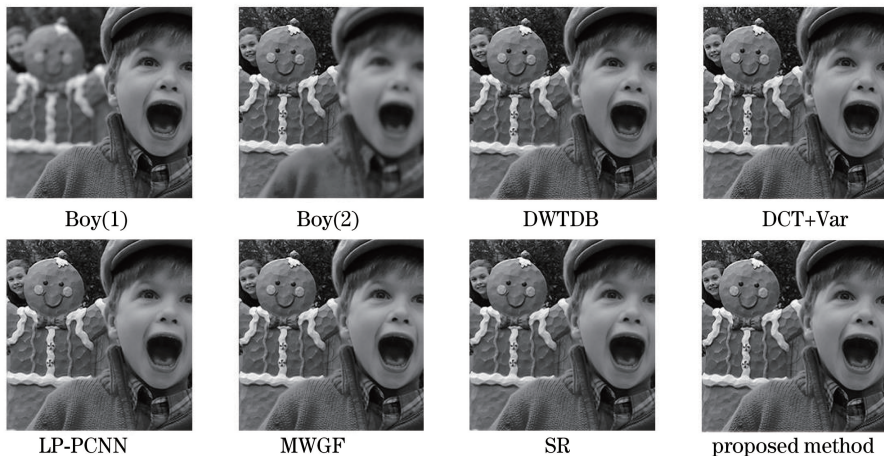


图 7 不同算法对源图像 Boy 融合结果

Fig. 7 Fusion results of the source image Boy with different algorithms

通过观察融合结果可以发现,本文算法在融合结果的主观视觉上与其他 5 种算法相比,能够较为完整的保存图像高频信息,无论在细节处理还是聚焦与散焦的分界处都取得了较好的视觉效果。

观察表 2 可以发现,本文算法在 4 项评价指标

下,得分情况基本上优于其他算法,在 4 项评价指标方面都取得了不错的结果。说明本文算法能够较好地保存图像细节信息与边缘信息,而且携带信息量较为丰富,图像层次较多,图像清晰度也较高。其他图像融合算法在提取高频信息的同时,也损失了一部

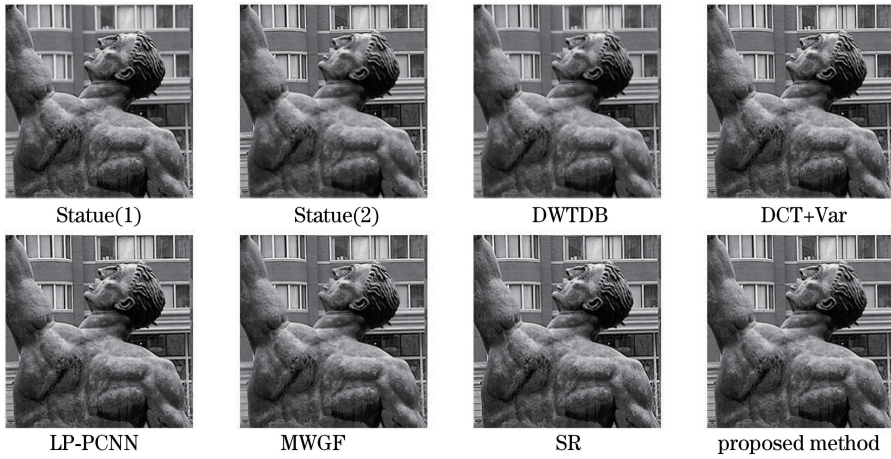


图 8 不同算法对源图像 statue 融合结果

Fig. 8 Fusion results of the source image Statue with different algorithms

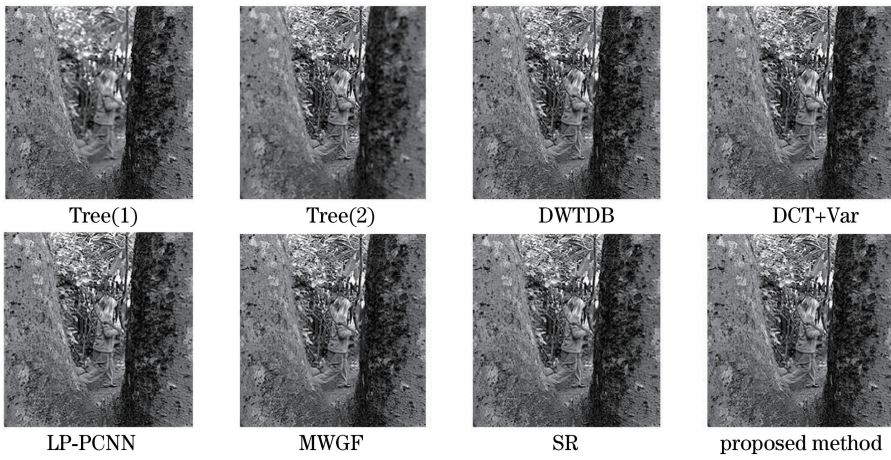


图 9 不同算法对源图像 tree 融合结果

Fig. 9 Fusion results of the source image Tree with different algorithms



图 10 不同算法对源图像 baby 融合结果

Fig. 10 Fusion results of the source image Baby with different algorithms

分高频信息。而本文算法结合深度学习可以从原始图像中较为完整地分离出高频信息,并利用图像处理

算法对融合结果进行优化,融合后的图像无论在主观视觉还是客观评价指标上都取得了不错的结果。

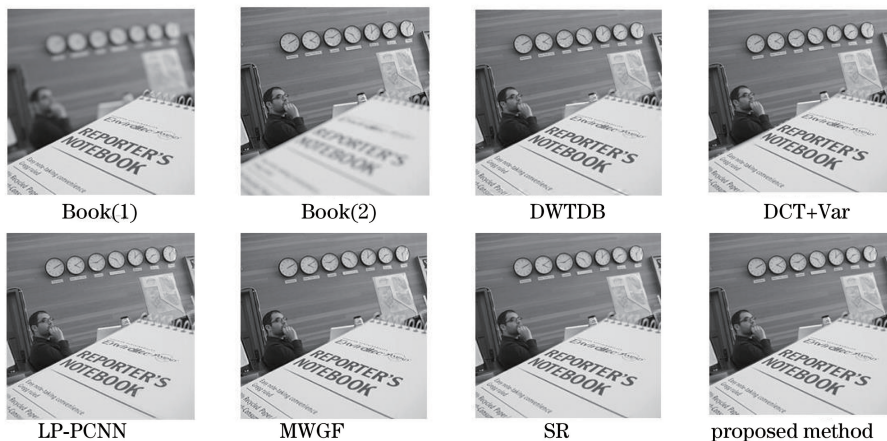


图 11 不同算法对源图像 Book 融合结果

Fig. 11 Fusion results of the source image Book with different algorithms

表 1 不同算法对应 4 种评价指标结果分析

Table 1 Four kinds of evaluation index results of different algorithms

Source image	Evaluation index	Method					
		DWTDB	DCT+Var	LP-PCNN	MWGF	SR	Proposed
Diving	MI	8.4243	8.8022	8.6994	8.6116	7.7642	8.8050
	$Q^{AB/F}$	0.7404	0.7514	0.6964	0.7520	0.7326	0.7527
	E	7.3152	7.3178	7.3097	7.3167	7.3019	7.3183
	AG	5.8150	5.9273	5.8815	5.9303	5.5156	5.9341
Boy	MI	7.5354	7.9745	7.9762	7.5593	6.9264	7.9874
	$Q^{AB/F}$	0.7127	0.7376	0.6590	0.7363	0.7169	0.7392
	E	7.5393	7.5088	7.5328	7.5406	6.9204	7.5413
	AG	5.8618	5.9145	6.0214	5.9394	5.5156	5.9434
Statue	MI	7.7094	8.3327	8.3421	7.6251	7.4576	8.3470
	$Q^{AB/F}$	0.7477	0.7723	0.7403	0.7553	0.7139	0.7765
	E	7.6887	7.7072	7.6882	7.5449	7.6032	7.6986
	AG	11.1419	11.8021	11.6657	11.6765	11.4094	11.8278
Tree	MI	7.0794	7.3241	7.3571	6.6251	6.1025	7.3654
	$Q^{AB/F}$	0.7290	0.7407	0.6919	0.7330	0.7404	0.7427
	E	7.6375	7.6324	7.6307	7.6249	7.6359	7.6426
	AG	17.1354	17.5293	17.1550	17.3998	16.8677	17.5413
Baby	MI	9.0224	6.9761	8.9361	6.9013	8.1202	9.1409
	$Q^{AB/F}$	0.7570	0.6557	0.7192	0.6447	0.7442	0.7729
	E	7.2432	7.2439	7.3249	7.3366	7.3330	7.3411
	AG	5.1372	5.1511	5.2579	4.9910	4.7349	5.2511
Book	MI	8.5496	7.3451	8.2.147	7.2413	7.3247	8.7218
	$Q^{AB/F}$	0.7421	0.6142	0.7041	0.6128	0.7547	0.7889
	E	7.0210	7.0012	6.8513	6.8129	7.0001	7.3081
	AG	5.6312	5.1511	5.2541	5.0146	4.9642	6.2336

5 结 论

所提出的基于深度学习的多聚焦图像融合算法,能够对多聚焦图像的聚焦图像块与散焦图像块做出准确的分类。结合传统融合策略进行优化与改进,取得了较好的效果。将深度学习应用于多聚焦图像融合,再次验证了深度学习在计算机视觉上的应用潜力。针对实际问题设计相应的深度学习网络可以快速地解决许多图像处理问题。然而,有待研究的问题是适当缩减网络层数,设计相应的损失函数,或者对训练样本进行预处理来提高网络运算速度,并保持相应的精度,以期得到更加简便快速的网络模型。

参 考 文 献

- [1] Stathaki T. Imagefusion: algorithms and applications [M]. London: Elsevier, 2008.
- [2] Burt P J, Adelson E H. The Laplacian pyramid as a compact image code [J]. IEEE Transactions on Communications, 1983, 31(4): 532-540.
- [3] Cheng D Q, Chen G, Gao L Z, *et al.* An adaptive region fusion rule multi-focus image fusion algorithm [J]. Laser & Optoelectronics Progress, 2017, 54(10): 101005.
程德强, 陈刚, 高凌志, 等. 一种自适应区域融合规则多聚焦图像融合算法[J]. 激光与光电子学进展, 2017, 54(10): 101005.
- [4] Zhu D R, Xu L, Wang F B, *et al.* Multi-focus image fusion algorithm based on fast finite shearlet transform and guidance filter [J]. Laser & Optoelectronics Progress, 2018, 55(1): 011001.
朱达荣, 许露, 汪方斌, 等. 基于快速有限剪切波变换与引导滤波的多聚焦图像融合算法[J]. 激光与光电子学进展, 2018, 55(1): 011001.
- [5] Ding W S, Bi D Y, He L Y, *et al.* Fusion of infrared and visible image based on shear wave transform and neighbor structure[J]. Acta Optica Sinica, 2017, 37(10): 1010002.
丁文杉, 毕笃彦, 何林远, 等. 基于剪切波变换和邻域结构特征的红外与可见光图像融合[J]. 光学学报, 2017, 37(10): 1010002.
- [6] Yang B, Li S T. Multifocus image fusion and restoration with sparse representation [J]. IEEE Transactions on Instrumentation & Measurement, 2010, 59(4): 884-892.
- [7] Cheng X J, Cheng X L, Hu M J, *et al.* Building detection and contour extraction by fusion of aerial images and LIDAR point cloud[J]. Chinese Journal of Lasers, 2016, 43(5): 0514002.
程效军, 程小龙, 胡敏捷, 等. 融合航空影像和 LIDAR 点云的建筑物探测及轮廓提取[J]. 中国激光, 2016, 43(5): 0514002.
- [8] Bai X Z, Zhang Y, Zhou F G, *et al.* Quadtree-based multi-focus image fusion using a weighted focus-measure [J]. Information Fusion, 2015, 22: 105-118.
- [9] Liu Y, Wang Z F. Multi-focus image fusion based on wavelet transform and adaptive block[J]. Journal of Image and Graphics, 2013, 18(11): 1435-1444.
刘羽, 汪增福. 结合小波变换和自适应分块的多聚焦图像快速融合[J]. 中国图象图形学报, 2013, 18(11): 1435-1444.
- [10] Nooshyar M, Abdipour M, Khajuee M. Multi-focus image fusion for visual sensor networks in wavelet domain [C]. International Symposium on Artificial Intelligence and Signal Processing, 2013: 23-31.
- [11] Cai M R, Yang J Y, Cai G H. Multi-focus image fusion algorithm using LP transformation and PCNN [C]. 6th IEEE International Conference on Software Engineering and Service Science, 2015: 237-241.
- [12] Zhou Z Q, Li S, Wang B. Multi-scale weighted gradient-based fusion for multi-focus images [J]. Information Fusion, 2014, 20: 60-72.
- [13] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [14] Lecun Y, Boser B, Denker J S, *et al.* Backpropagation applied to handwritten zip code recognition[J]. Neural Computation, 2014, 1(4): 541-551.
- [15] Simonyan K, Zisserman A. Verydeep convolutional networks for large-scale image recognition [J]. Computer Science, 2014, 246-259.
- [16] Szegedy C, Liu W, Jia Y Q, *et al.* Going deeper with convolutions[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1-9.
- [17] Nejati M, Samavi S, Shirani S. Multi-focus image fusion using dictionary-based sparse representation [J]. Information Fusion, 2015, 25: 72-84.