

基于鲁棒主成分分析的多聚焦图像融合

王书朋, 蒋 艺

(西安科技大学 通信与信息工程学院, 陕西 西安 710054)

摘 要:针对多聚焦图像融合提取聚焦区域不准确,融合图像的边缘存在伪影的问题,提出了一种基于鲁棒主成分分析的多聚焦图像融合方法。首先,用鲁棒主成分分析将源图像分解为低秩分量和稀疏分量。然后分析不同分量的特性,设计相应的融合规则。针对低秩分量,提出基于卷积神经网络的融合策略,用于提取源图像的纹理细节,构建权重图;对于稀疏分量,采用基于拉普拉斯能量和的方法生成决策图,为了使决策图的边缘与源图像保持一致,将多幅源图像的均值作为引导图像并用引导滤波器对决策图进行优化。最后,根据融合后的低秩分量和稀疏分量重构出最终的融合图像。实验结果表明,该方法提高了融合图像的对比度和边缘清晰度,在主观分析和客观评价指标上都优于其他七种方法。

关键词:多聚焦图像;图像融合;鲁棒主成分分析;卷积神经网络;融合规则

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2020)10-0053-06

doi:10.3969/j.issn.1673-629X.2020.10.010

Multi-focus Image Fusion Based on Robust Principal Component Analysis

WANG Shu-peng, JIANG Yi

(School of Communication and Information Engineering, Xi'an University of Science and Technology, Xi'an 710054, China)

Abstract: In order to solve the problems of inaccurate focusing region extraction and artifacts in edge of image fusion, we propose a new image fusion algorithm based on robust principal component analysis. Firstly, the source images are decomposed into low-rank components and sparse components by robust principal component analysis. Then the characteristics of different components is analyzed and the corresponding fusion rules are designed. For the low-rank components, a fusion rule based on convolutional neural network is proposed to extract the texture details of the source image and construct the weight map. For the sparse components, the Sum-Modified-Laplacian is used to generate the decision map. In order to keep the boundary of the decision map consistent with the source images, the mean value of multiple source images is used as the guided image in the guided filter to optimize the decision map. Finally, the final fused image is reconstructed from the fused low-rank component and the sparse component. The experiment shows that the proposed method improves the contrast and edge definition of the fused images, and is superior to other seven methods in both subjective and objective evaluation indicators.

Key words: multi-focus image; image fusion; robust principle component analysis; convolutional neural network; fusion rule

0 引 言

由于成像设备受光学镜头景深的限制,只能对聚焦区域内的物体产生清晰的图像,而聚焦区域以外的物体在图像中都是模糊的^[1]。为了克服物理设备的限制,多聚焦图像融合技术将来自同一场景的两幅或多幅聚焦区域不同的图像进行融合生成单幅全聚焦图像,使同一场景下的所有目标都清晰地呈现出来,从而便于人们后期对图像进行分析和处理。然而现有的多

聚焦图像融合算法通常难以准确地区分聚焦区域和非聚焦区域,导致融合图像中聚焦区域的边界处有伪影的现象。针对该问题,该文拟提出一种新的多聚焦图像融合方法。

现有的多聚焦图像融合方法可以分为三类:基于空间域的方法、基于变换域的方法和基于深度学习的方法。基于空间域的方法^[2-3]主要是依据某种清晰度指标,直接从源图像中选择清晰的部分组合成融合图

收稿日期:2019-12-09

修回日期:2020-04-10

基金项目:国家自然科学基金(61801373)

作者简介:王书朋(1975-),男,博士,副教授,研究方向为图像处理、模式识别、计算机视觉等;蒋 艺(1994-),女,硕士研究生,研究方向为图像融合、模式识别。

像,但这类方法容易受到噪声的影响,融合结果中通常存在块效应问题^[4]。基于变换域的方法是通过某种变换将图像分解成不同频带的系数,然后选取不同的融合规则得到融合系数,最后通过逆变换生成融合图像。如基于多尺度变换的方法^[5-6]、基于稀疏表示的方法^[7]等。这类方法比基于空间域的方法能更好地提取图像的边缘和轮廓等特征,但图像分解过程中会缺失部分高频分量的信息,融合结果中易产生振铃效应^[8]。近年来,机器学习及深度学习理论不断发展^[9-10]。Liu 等人^[11]将 CNN 引入到多聚焦图像融合中,通过训练 CNN 模型直接生成焦点图,克服了手动设计聚焦区域检测方法的难题。多聚焦图像融合的关键是准确区分源图像中的聚焦区域和非聚焦区域。鲁棒主成分分析(robust principal component analysis, RPCA)是一种新的聚焦点检测算法,该算法可以将图像分解成代表背景的低秩分量和代表图像显著性特征的稀疏分量。杨明伟等人^[12]用 RPCA 分解源图像得到稀疏分量,然后对稀疏分量进行三方向一致性和区域生长法处理。Zhang 等人^[13]用引导滤波器对稀疏分量进行处理得到增强图像,用增强图像与源图像之前的差值图像提取背景区域,从而确定聚焦区域的位置。尽管这类方法可以避免振铃效应,提高了融合决策图的准确性,但基于像素值取大和空间频率的融合规则没有得到最优的聚焦区域检测结果,融合图像中有严重的块效应,且聚焦区域的边界有晕影的现象。

为解决上述问题,该文提出了一种基于 RPCA 的多聚焦图像融合方法。该方法首先通过 RPCA 将源图像分解为低秩和稀疏分量。针对低秩分量,利用 CNN 构建权重图,可以较好地地区分聚焦区域和非聚焦区域。对于稀疏分量,采用基于拉普拉斯能量和的方法构建稀疏分量的融合决策图,然后用引导滤波器优化决策图,使决策图的边缘与源图像保持一致,避免伪轮廓。从主观和客观两个方面将所提算法与其他七种经典算法进行比较。实验结果表明,所提方法可以准确区分聚焦区域和非聚焦区域,融合图像中聚焦区域的边界清晰且不会引入伪影。

1 鲁棒主成分分析

为了解决主成分分析(PCA)鲁棒性不佳的问题,Wright 等人^[14]提出了鲁棒主成分分析(RPCA)理论,它的基本思想是数据矩阵在最优化准则下可以表示为一个低秩矩阵和一个系数矩阵的和。假设有一个输入矩阵 $I \in N^{H \times W}$,那么该矩阵可以分解为:

$$I = L + S \quad (1)$$

其中, L 是低秩矩阵, S 是稀疏矩阵,输入矩阵 I 的大小为 $H \times W$ 。

与其他稀疏表示方法类似,RPCA 采用核规范作为近似稀疏约束:

$$\min_{L, S} (\text{rank}(L) + \lambda \|S\|_0) \quad (2)$$

s. t. $L + S = I$

其中, $\text{rank}(\bullet)$ 是矩阵的秩, $\|\bullet\|_0$ 是 0 范数矩阵, λ 是加权参数且 $\lambda > 0$ 。

在一般情况下,这种分解是 NP 难问题。由于一个矩阵的秩与它的非零奇异值的个数相等,可以用矩阵的核范数近似代替矩阵的秩,用 0 范数等价于 1 范数,则稀疏矩阵可以转化为以下凸优化问题:

$$\min_{L, S} (\|L\|_* + \lambda \|S\|_1) \quad (3)$$

s. t. $L + S = I$

其中, $\|\bullet\|_*$ 是核范数,即矩阵的奇异值之和; $\|\bullet\|_1$ 是 1 范数矩阵,即矩阵的所有元素的绝对值之和。该文用增广的拉格朗日乘法实现图像的 RPCA 分解,如图 1 所示,参数 λ 的大小为 $1/\sqrt{M}$, M 为源图像的行数。



图 1 多聚焦图像的 RPCA 分解

2 所提算法

为了准确地检测源图像中的聚焦区域,提出了基于 RPCA 的多聚焦图像融合方法。该方法的框图如图 2 所示。

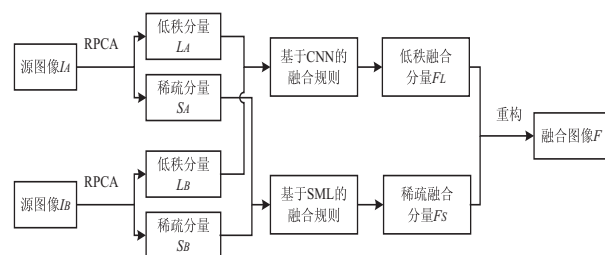


图 2 算法框图

首先,输入两幅聚焦区域不同的源图像 I_A 和 I_B ,用 RPCA 对这两幅源图像进行分解,得到低秩分量 L_A 、 L_B 和稀疏分量 S_A 、 S_B 。然后,针对低频分量包含源图像整体结构和细节的特性,采用基于 CNN 的融合规则构建决策图。针对稀疏分量包含聚焦区域的边缘和纹理特性,采取基于引导滤波^[15]改进的拉普拉斯能量和(SML)的融合规则。最后,将融合后的低秩分量 F_L 和稀疏分量 F_S 重构得到融合图像 F ,即 $F(i,j) = F_L(i,j) + F_S(i,j)$, (i,j) 为像素点的位置。

2.1 融合低秩分量

低秩图像包含源图像大部分的结构和能量,一般的融合规则很难准确地区分聚焦区域和非聚焦区域的边界。实际上,在多聚焦图像融合过程中,决策图的生成可以看成二分类问题,CNN 对解决这类问题是有效的^[11]。因此通过图 3 中的 CNN 模型对低秩分量进行

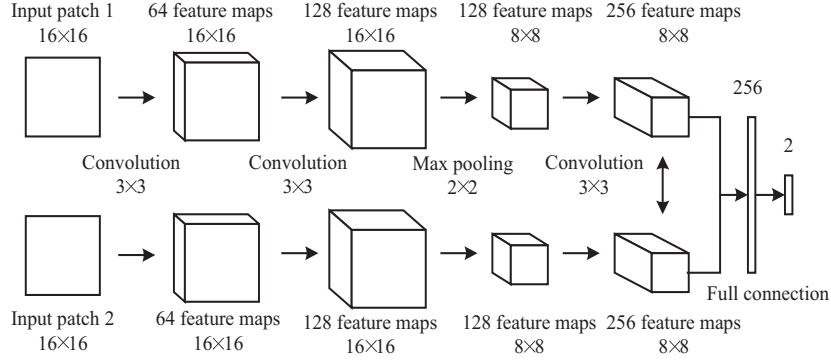


图 3 CNN 模型

为了降低网络训练的复杂度,文中的 CNN 模型为暹罗网络^[16],使两个分支获得相同的权重,每个分支包含了三个卷积层和一个最大池化层^[11],第三个卷积层的输出特征图可以表征不同源图像聚焦区域的特征。如果输入图像块过大,可能同时包含聚焦区域和非聚焦区域,会导致聚焦边界的误判。当图像块过小时,图像中包含的特征信息过少,可能会降低图像分类的准确性,所以该文将训练的图像块大小设为 16×16 。卷积核的大小及步长分别为 3×3 和 1,池化层的池化因子和跨度分别为 2×2 和 2,将每个分支得到的 256 个特征级联后与 256 维特征向量全连接,最后再与 2 维特征向量全连接。经 Softmax 层分类后,输出值的大小即为这一对输入图像块的聚焦属性。

基于 CNN 的低秩分量融合规则步骤如下:首先将两幅 RPCA 分解得到的低秩图像 L_A 和 L_B 输入到训练好的 CNN 模型中进行焦点检测得到得分图 S_{map} , S_{map} 中的每个系数表示来自两个低秩图像相同位置处的一对图像块的聚焦特性。当 $S_{\text{map}} \approx 1$ 时,说明 L_A 聚焦, L_B 散焦;当 $S_{\text{map}} \approx 0$ 时,说明 L_A 散焦, L_B 聚焦。

然后对得分图进行阈值分割得到初始的二值图像:

$$\hat{S}_{\text{map}}(i,j) = \begin{cases} 1 & S_{\text{map}}(i,j) \geq 0.5 \\ 0 & S_{\text{map}}(i,j) < 0.5 \end{cases} \quad (4)$$

其次用小区域移除策略矫正初始二值图像 \hat{S}_{map} 中错误分类的像素:

$$D_L = \text{bwareaopen}(\hat{S}_{\text{map}}, \text{area}) \quad (5)$$

其中, D_L 为低秩分量的融合决策图。 $\text{bwareaopen}(\hat{S}_{\text{map}}, \text{area})$ 表示删除二值图像 \hat{S}_{map} 中面积小于 area 的区域,

特征提取和分类,将低秩图像利用滑窗技术分成大小为 16×16 的图像块,卷积层和池化层用于特征提取,全连接层用于分类。然后判断两幅低秩图像相同位置处的图像块哪个是清晰的,哪个是模糊的。最终得到低秩分量的融合决策图。

$\text{area} = \lceil 0.01 \times M \times N \rceil$, $M \times N$ 为源图像的大小。

最后,利用所获得的决策图融合低秩图像,得到低秩融合图像 F_L 。

$$F_L(i,j) = D_L(i,j)L_A(i,j) + (1 - D_L(i,j))L_B(i,j) \quad (6)$$

流程如图 4 所示。

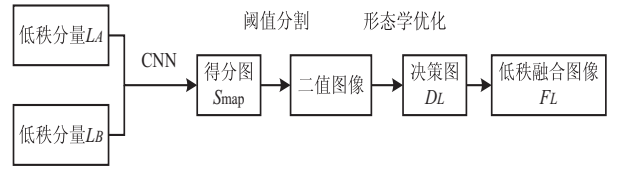


图 4 基于 CNN 的低秩图像融合

2.2 融合稀疏分量

传统的图像清晰度检测的方法有方差、空间频率和 SML 等,文献[17]从主观和客观评价两个方面证明,SML 比其他清晰度检测的方法具有更好的性能。所以该文对稀疏分量采用基于 SML 的融合规则。

首先分别计算稀疏分量 S_A 和 S_B 的 SML 值,得到 SML_A 和 SML_B 。像素点 (i,j) 处的 SML 值可通过如下公式计算:

$$\text{SML}(i,j) = \sum_{i=-m}^m \sum_{j=-n}^n |\text{ML}(i,j)|^2 \quad (7)$$

其中,局部窗口的大小为 $m \times n$,文中 $m = n = 3$ 。拉普拉斯算子 ML 定义为:

$$\text{ML}(i,j) = |2I(i,j) - I(i-1,j) - I(i+1,j)| + |2I(i,j) - I(i,j-1) - I(i,j+1)| \quad (8)$$

其中, $I(i,j)$ 是像素点 (i,j) 处的像素值。

然后根据如下公式得到稀疏分量的初始融合决策图 G :

$$G(i,j) = \begin{cases} 1 & \text{SML}_A(i,j) \geq \text{SML}_B(i,j) \\ 0 & \text{SML}_A(i,j) < \text{SML}_B(i,j) \end{cases} \quad (9)$$

然而,初始决策图中部分像素会出现不连贯的现象,因此需要对初始决策图进行优化处理。引导滤波^[7]是一种边缘保持滤波器,具有较好的保留边缘和去除噪声的能力。该文用引导滤波器对决策图 G 进行平滑滤波,提高决策图的空间一致性,公式如下:

$$D_s = \text{Guidedfilter}(I, G, R, \text{eps}) \quad (10)$$

其中, D_s 是通过引导滤波器处理后的稀疏分量决策图, I 是引导图像。为了使决策图的边缘与源图像保持一致,该文用两幅源图像的均值作为引导图像,即 $I = (I_A + I_B)/2$ 。 R 表示引导滤波器的半径, R 值越大,平滑效果越好,但边缘保持能力下降。 eps 表示正则化参数, eps 值越大,滤波效果越明显。该文将 R 设为 7, eps 设为 0.01。

最后,通过下式可得融合后的稀疏分量 F_s :

$$F_s(i,j) = D_s(i,j)S_A(i,j) + (1 - D_s(i,j))S_B(i,j) \quad (11)$$

3 实验结果及分析

为了验证所提算法的有效性,从主观视觉感知和客观评价指标两个方面将文中方法与七种方法进行比较,包括基于非下采样轮廓波变换的方法^[5](NSCT)、基于 NSCT 与 SR 相结合的方法^[7](NSCT-SR)、基于引导滤波的方法^[18](GF)、基于密度尺度不变特征变换的方法^[3](DSIFT)、基于稀疏分解和背景检测的方法^[13](RPCA)、基于脉冲耦合神经网络的方法^[10](PCNN)、基于卷积神经网络的方法^[11](CNN)。实验选取了 2 种不同类型的灰度图像^[18]进行对比分析。

CNN 模型的训练样本是由 ILSVRC 2012 中的自然图像生成的。将每个源图像使用标准偏差为 2,大

小为 7×7 的高斯滤波器处理后,获得五种具有不同模糊程度的模糊图像。对于每类模糊图像和源图像,随机采样 20 对大小为 16×16 的图像块,总共获得 100 万对清晰和模糊的图像块。用 Softmax 损失函数作为网络的目标函数,用随机梯度下降法最小化损失函数。在训练过程中,批处理的大小设置为 128。使用 Xavier 算法^[14]初始化每个卷积层的权重,学习率为 0.000 1。

3.1 主观分析

图 5 是“pepsi”图像及不同方法的融合结果。从图 5(c)和(d)中可以看出,NSCT 和 NSCT-SR 方法使融合图像丢失了部分边缘信息,可乐瓶左侧的轮廓模糊。GF 方法的融合图像中桌子的底部细节保留不完整。DSIFT 方法可以较好地提取源图像中大部分的细节。图 5(g)显示,基于 RPCA 的方法错误提取了源图像中条形码的聚焦区域,融合图像的视觉效果最差。PCNN 的融合图像有重影。图 6 是图 5 中方框区域的放大图。NSCT 和 NSCT-SR 算法在图 6(c)(d)中聚焦区域的伪轮廓是明显的。DSIFT 和 RPCA 算法使字母的边界有一些扩展,且 RPCA 算法的空间连续性较差,PCNN 的融合结果中字体严重模糊。CNN 和文中方法不会引入伪影,融合结果具有较高的视觉质量。

图 7 是源图像“office”及不同方法的融合结果。NSCT 和 NSCT-SR 方法在融合图像中的电脑区域引入了明显的伪影,DSIFT、RPCA 和 PCNN 方法使闹钟边界处有不同程度的模糊,部分细节丢失并引入较多的人造纹理。CNN 的融合结果中桌子边缘不清晰。图 8 是图 7 中方框区域的放大图。除所提算法以外,其余七种算法在人的头部都引入了伪边界,还有一些白色的伪影,人耳的轮廓模糊,图像视觉质量较差。文中方法的结果图对比度较高,人耳和头部细节保留完整,轮廓清晰,图像融合效果更好。

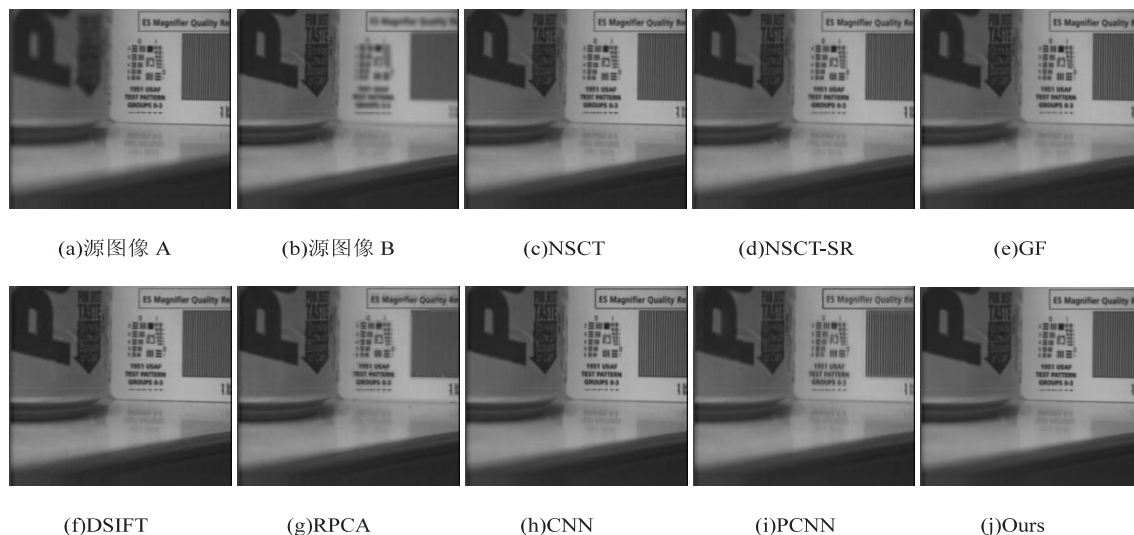


图 5 源图像“pepsi”及不同方法的融合结果



图 6 图 5 中方框区域的放大图

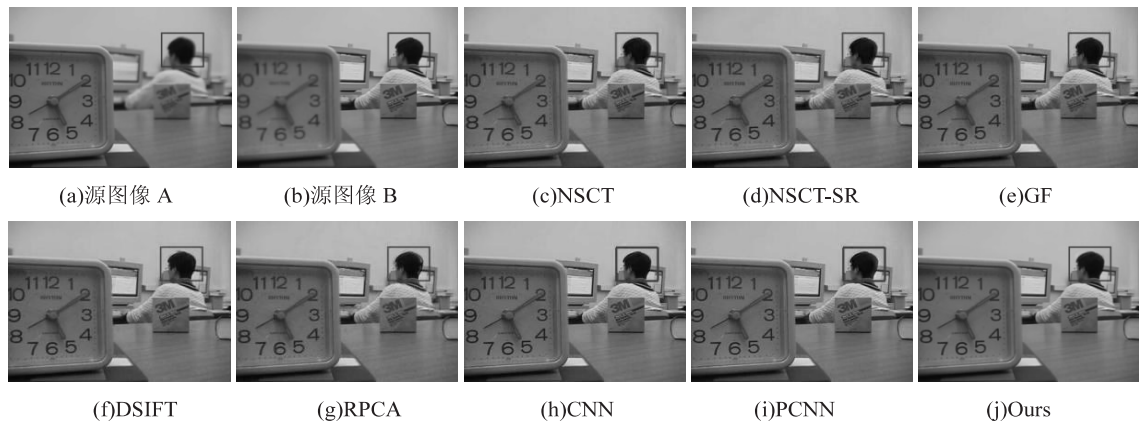


图 7 源图像“office”及不同方法的融合结果

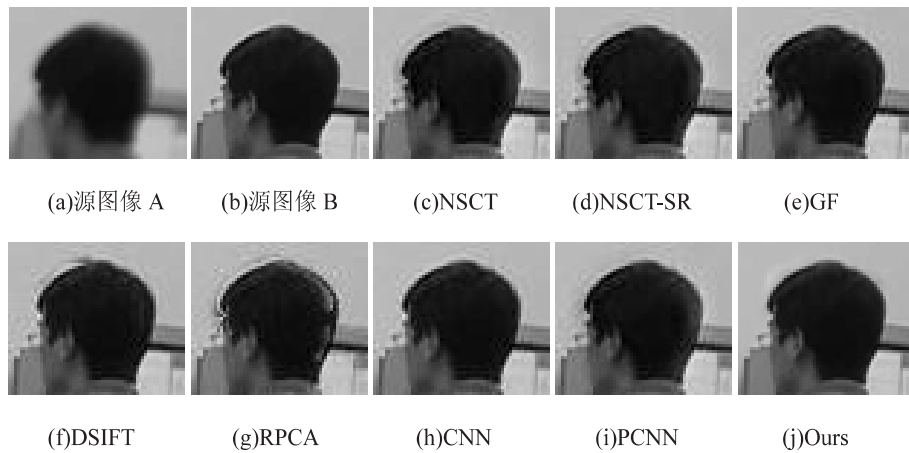


图 8 图 7 中方框区域的放大图

3.2 客观评价

为了定量评估不同融合方法的性能,选择三种客观评价指标对实验结果进行评估:结构相似性^[19](SSIM)、基于人类视觉感知的度量^[20](Q^{CB})、边缘梯度算子^[21]($Q^{AB/F}$)。其中,SSIM 根据图像的结构评估融合图像和源图像之间的相似性,SSIM 值越大,融合结果与源图像的结构相似度越高。 Q^{CB} 是描述图像视觉特性的度量, Q^{CB} 值越大,图像的对比度越高,视觉效果越好。 $Q^{AB/F}$ 通过度量融合结果包含源图像的边缘信息量来评估融合性能, $Q^{AB/F}$ 值越大,融合图像中包

含的边缘信息量越多。

表 1 列出了上述三种融合图像的客观评价结果。文中方法在 Q^{CB} 和 $Q^{AB/F}$ 两种指标上都取得了最大值,说明基于文中方法的融合结果中包含充分的纹理细节信息。其中有两幅图像在指标 SSIM 上并未取得最大值,但是与最大值之间的差距较小,说明融合图像与源图像在结构上保持良好的一致性。这是因为利用 RPCA 对图像进行分解,增强了算法的鲁棒性。同时设计了基于卷积神经网络和 SML 的融合规则,提高了决策图的准确性,使融合图像更符合人类视觉感知。

表 1 不同图像融合结果的客观指标

图像	客观指标	NSCT	NSCT-SR	GF	DSIFT	RPCA	CNN	PCNN	Ours
pepsi	SSIM	0.901 8	0.900 9	0.892 3	0.883 4	0.899 9	0.892 3	0.895 1	0.908 2
	$Q^{AB/F}$	0.644 2	0.647 3	0.641 9	0.474 6	0.647 5	0.644 6	0.597 5	0.783 6
	Q^{CB}	0.583 0	0.608 3	0.625 1	0.539 8	0.601 5	0.622 1	0.555 3	0.662 6
office	SSIM	0.893 9	0.893 7	0.892 9	0.890 1	0.886 7	0.895 5	0.903 5	0.903 1
	$Q^{AB/F}$	0.678 5	0.678 9	0.679 7	0.678 3	0.657 0	0.679 7	0.680 5	0.765 3
	Q^{CB}	0.610 1	0.610 6	0.605 4	0.607 6	0.592 3	0.604 3	0.600 5	0.638 2

4 结束语

提出了一种基于 RPCA 分解的多聚焦图像融合方法。首先利用 RPCA 将源图像分解为低秩和稀疏分量。然后利用基于 CNN 的融合规则得到低秩分量的融合决策图,可以更好地提取图像的细节信息。对于稀疏分量,采用基于 SML 值取大的方法构建决策图,再用引导滤波器对决策图进行优化,提高了决策图的空间一致性。最后通过重构得到最终的融合图像。将所提方法与七种经典方法进行比较,从主观和客观的分析结果可以表明,所提方法能准确地提取聚焦区域,充分保留了源图像的细节信息,融合结果更自然。

参考文献:

- [1] LI S, KANG X, FANG L, et al. Pixel-level image fusion: a survey of the state of the art [J]. *Information Fusion*, 2017, 33: 100–112.
- [2] XIA X, YAO Y, YIN L, et al. Multi-focus image fusion based on probability filtering and region correction [J]. *Signal Processing*, 2018, 153: 71–82.
- [3] QIU X, LI M, ZHANG L, et al. Guided filter-based multi-focus image fusion through focus region detection [J]. *Signal Processing: Image Communication*, 2019, 72: 35–46.
- [4] LIU Y, LIU S, WANG Z. Multi-focus image fusion with dense SIFT [J]. *Information Fusion*, 2015, 23: 139–155.
- [5] ZHANG Q, GUO B. Multi-focus image fusion using the non-subsampled contourlet transform [J]. *Signal Processing*, 2009, 89(7): 1334–1346.
- [6] ZHOU Z, WANG B, LI S, et al. Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters [J]. *Information Fusion*, 2016, 30: 15–26.
- [7] NEJATI M, SAMAVI S, SHIRANI S. Multi-focus image fusion using dictionary-based sparse representation [J]. *Information Fusion*, 2015, 25: 72–84.
- [8] 刘洋, 周国众, 王凡. 利用一致性低秩表示模型进行多聚焦图像融合 [J]. *指挥控制与仿真*, 2019, 41(3): 30–36.
- [9] GANASALA P, KUMAR V. Feature-motivated simplified adaptive PCNN-based medical image fusion algorithm in NSST domain [J]. *Journal of Digital Imaging*, 2016, 29(1): 73–85.
- [10] YIN M, LIU X, LIU Y, et al. Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampling shearlet transform domain [J]. *IEEE Transactions on Instrumentation and Measurement*, 2018, 68(1): 49–64.
- [11] LIU Y, CHEN X, PENG H, et al. Multi-focus image fusion with a deep convolutional neural network [J]. *Information Fusion*, 2017, 36: 191–207.
- [12] 杨明伟, 黄永东, 常霞. 基于 SML 特征检测的 RPCA 域多聚焦图像融合 [J]. *计算机工程与应用*, 2017, 53(22): 157–162.
- [13] ZHANG B, LU X, PEI H, et al. Multi-focus image fusion based on sparse decomposition and background detection [J]. *Digital Signal Processing*, 2016, 58: 50–63.
- [14] WRIGHT J, GANESH A, RAO S, et al. Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization [C] // *Proceedings of neural information processing systems*. Whistler: MIT Press, 2009: 2080–2088.
- [15] 杨艳春, 李娇, 党建武, 等. 基于冗余小波变换与引导滤波的多聚焦图像融合 [J]. *计算机科学*, 2018, 45(2): 301–305.
- [16] NAIR V, HINTON G E. Rectified linear units improve restricted Boltzmann machines [C] // *Proceedings of the 27th international conference on machine learning (ICML-10)*. Haifa, Israel: Omnipress, 2010: 807–814.
- [17] HUANG W, JING Z. Evaluation of focus measures in multi-focus image fusion [J]. *Pattern Recognition Letters*, 2007, 28(4): 493–500.
- [18] LI S, KANG X, HU J. Image fusion with guided filtering [J]. *IEEE Transactions on Image Processing*, 2013, 22(7): 2864–2875.
- [19] 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.
- [20] CHEN Y, BLUM R S. A new automated quality assessment algorithm for image fusion [J]. *Image and Vision Computing*, 2009, 27(10): 1421–1432.
- [21] HOSSNY M, NAHAVANDI S, CREIGHTON D. Comments on ‘Information measure for performance of image fusion’ [J]. *Electronics Letters*, 2008, 44(18): 1066–1067.