

DOI:10.11918/201909064

利用拉普拉斯能量和 CNN 的多聚焦图像融合方法

翟浩, 庄毅

(南京航空航天大学 计算机科学与技术学院, 南京 211106)

摘要: 多聚焦图像融合技术的目的是生成一幅全聚焦图像. 所谓全聚焦图像, 就是将不同源图像的清晰区域集成到一幅单一的图像中. 传统的图像融合方法通常存在块伪影、人造边、晕轮效果、振铃效果以及对比度下降等问题. 对此, 本文提出了一种利用拉普拉斯能量和 CNN 的多聚焦图像融合方法. 使用拉普拉斯能量算子可以有效的提取源图像的聚焦信息, 而训练后的卷积神经网络模型从聚焦信息图中提取的聚焦特征可以有效的进行聚焦子块和离焦子块的区分. 训练后的卷积神经网络模型不仅具有很好的提取活跃窗口相对聚焦度的能力, 而且可以获得精确的分割边界. 在经过多轮训练后, 卷积神经网络模型可以很好的在源图像和分值图之间建立一种有效的映射, 这对于生成一幅精准的聚焦图至关重要. 采用二值分割和小区域滤波技术来对聚焦图进行进一步的修正, 获得用于融合的最终决策图. 最后, 根据最终决策图提供的权值, 对多幅源图像进行融合形成最终的融合图像. 实验结果表明, 无论从视觉效果还是从定量评价方面, 提出的方法均优于目前已有的其它融合方法.

关键词: 多聚焦图像融合; 拉普拉斯能量; CNN; 深度学习; 聚焦度量

中图分类号: TP391.41

文献标志码: A

文章编号: 0367-6234(2020)05-0137-11

Multi-focus image fusion method using energy of Laplacian and convolutional neural network

ZHAI Hao, ZHUANG Yi

(College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: The aim of multi-focus image fusion technology is to produce an all-in-focus image in which the clear parts of different source images are integrated to a single image. Most of the existing fusion methods still suffer from the problems such as block artifacts, artificial edges, halo effects, ringing effects, and contrast reduction. To address these problems, a multi-focus image fusion method using the energy of Laplacian CNN is proposed in this paper. The focus information of source images can be extracted effectively by using Laplacian energy operator, and the focus feature extracted from focus information maps by the trained convolutional neural network model can effectively distinguish focused sub-blocks from defocused sub-blocks. The trained convolutional neural network model not only has a good ability to extract the relative focus degree of active windows, but also can obtain an accurate segmentation boundary. After multiple rounds of training, the convolutional neural network model can well establish an effective mapping between source images and a score map, which is essential to generate an accurate focus map. Then, the focus map is further modified using binary segmentation and small region filtering, and the final decision map for fusion is obtained. Finally, according to the weights provided by the final decision map, the final fusion image will be formed by fusing multiple source images. The experimental results show that the proposed method is superior to other existing fusion methods in terms of visual effects and quantitative evaluation.

Keywords: multi-focus image fusion; energy of Laplacian; convolutional neural network; deep learning; focus measurement

在数码摄影中, 由于光学透镜的有限景深, 很难获得一幅所有相关目标都聚焦的全聚焦图像. 一般情况下, 只有位于焦平面上的目标是聚焦的或者清晰的, 而位于焦平面其它位置例如焦平面之前或者

之后的目标则都是离焦的或者模糊的. 多聚焦图像融合技术是一种有效的可以合成相同场景下不同聚焦设置采集到的多幅源图像, 获得一幅单一的具有扩展景深的融合图像. 目前, 多聚焦图像融合技术已经应用到了大量的应用领域, 例如显微成像、数码摄影、视频监控以及计算机视觉等领域^[1]. 近些年, 出现了大量的多聚焦图像融合方法. 根据空间域的不同, 这些融合方法大致可以分为两大类: 基于变换域的方法和基于空域的方法^[2].

收稿日期: 2019-09-05

基金项目: 国家自然科学基金(61572253); 航空科学基金(2016ZC52030); 江苏省科技创新基金(KYLX16_0381)

作者简介: 翟浩(1987—), 男, 博士研究生;
庄毅(1965—), 女, 教授, 博士生导师

通信作者: 庄毅, zyl16@nuaa.edu.cn

基于变换域的融合方法一般主要包括三个步骤:分解、融合和重构. 基于稀疏特征矩阵分解和形态学滤波的多聚焦图像融合方法被提出^[3]. 该方法通过将源图像进行稀疏特征分解形成稀疏特征矩阵,然后对该稀疏特征矩阵进行加权形成一个临时矩阵,最后对该临时矩阵进行形态学滤波,再结合基图像,形成最终的融合图像. 接着,基于图像分解的多聚焦图像融合方法被提出^[4],该方法通过使用改进的迭代重加权分解算法将源图像分解成卡通分量和纹理分量,然后对卡通分量和纹理分量分别进行融合,形成最终的融合图像. 此外,基于多级形态分量分析和支持向量机的多聚焦图像融合方法被提出^[5],该方法通过使用多级形态分量分析分解源图像,并从自然层提取特征向量. 然后使用训练过的二类支持向量机对这些特征向量进行分类,最后融合系数形成最终的融合图像. 这类融合方法极易导致最终的融合图像产生对比度或者清晰度下降的问题.

同基于变换域的方法相比,在基于空域的融合方法中,融合规则直接应用到像素、子块和区域的. 根据聚焦度量方法,选择具有较高清晰度的像素、子块或者区域进行重构来形成最终的融合图像. 一种基于四叉树结构利用权重聚焦度量的多聚焦图像融合方法被提出^[6],根据有效的四叉树分解策略,把源图像分解成四叉树结构下最优尺寸的子块,采用权重修改拉普拉斯和来检测清晰区域,并将这些清晰区域组合成全聚焦图像. 接着,基于多尺度形态学聚焦度量寻找边界的多聚焦图像融合方法被提出^[7]. 该方法通过采用一种形态学聚焦度量方法来获取边界区域和非边界区域,然后采用融合方法对边界区域和非边界区域分别进行融合. 其中,最具代表性的是基于梯度信息的方法^[7-8]和基于深度学习的方法^[9-11]. 在基于空域的融合方法中,由于分割区域的形状或者聚焦度量准则难以确定,所以基于空域的融合方法通常很难精确的获得精确的分割边界,并且最终的融合图像在分割边界附近将会产生严重的伪影效果.

针对以上问题,利用拉普拉斯能量和 CNN 的多聚焦图像融合方法被提出来,该方法的优势主要体现在以下三方面. 第一,直接采用源图像中的像素进行融合,克服了融合图像中清晰度或者对比度下降的问题. 第二,采用拉普拉斯能量作为聚焦度量法则,可以有效的提取源图像中的细节信息,这对于选择更高清晰度的子块至关重要. 第三,由于训练后的卷积神经网络模型具有很强的学习子块相对聚焦度的能力,因此该方法可以获得非常精确的分割边界.

1 提出的方法

目前,卷积神经网络已经广泛应用于计算机视觉的大量应用领域,例如脸部检测^[12]、脸部识别^[13]、超分辨率^[14]以及块相似性对比^[15]等等. 受卷积神经网络模型在特征提取和分类中具有强大能力的启发,利用拉普拉斯能量和 CNN 的多聚焦图像融合方法被提出来. 在本研究中主要探讨两幅源图像的融合问题,多幅源图像的融合问题可以按照两幅源图像的融合顺序依次进行融合. 提出方法的框架如图 1 所示. 从图 1 中可以看出,提出的方法主要包括 5 个阶段:聚焦度量、CNN 聚焦、二值分割、小区域滤波和融合阶段.

接下来,详细的展示多聚焦图像的融合过程,将采用 Lytro 数据集的“child”图像作为实验对象,提出方法的融合过程如图 2 所示. 首先对源图像进行聚焦度量,分别获得两幅源图像对应的聚焦信息图. 根据栅栏扫描的顺序,采用滑动窗口依次获取聚焦信息图的对应子块. 同理,通过滑动窗口的不断采样,用于获得聚焦信息图中所有的样本数据. 把这些样本数据反馈至训练后的 CNN 模型,获得这些样本数据的预测标签值. 对这些标签值重新排列成源图像的尺寸,获得聚焦图. 最后,对聚焦图进行二值分割和小区域滤波的处理,获得一个正则化的最终决策图. 根据最终决策图对两幅源图像进行融合,形成最终的融合图像.

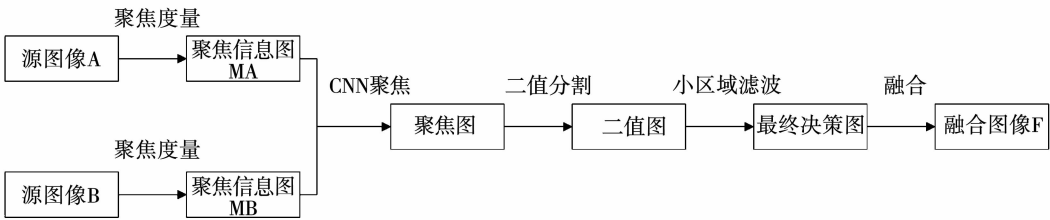


图 1 提出方法的框架

Fig. 1 Framework of the proposed method

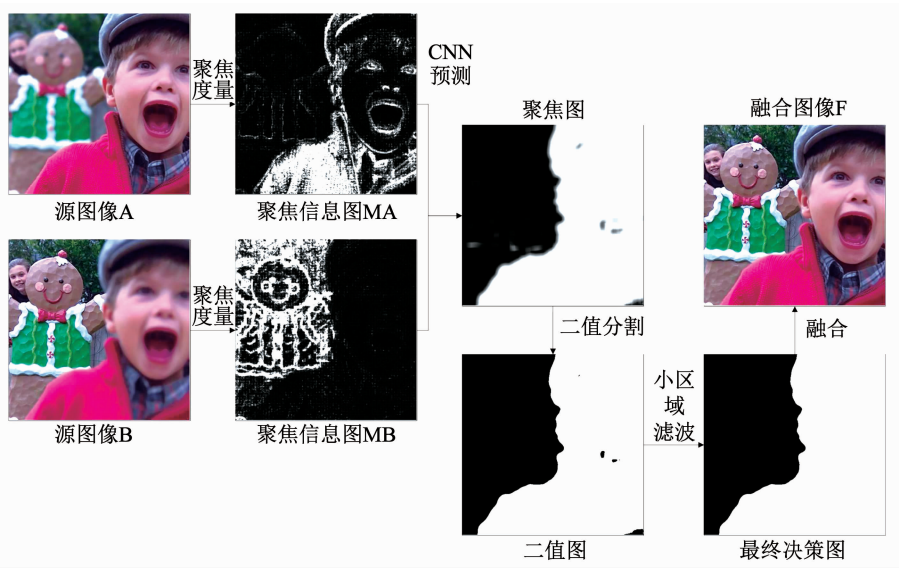


图 2 提出方法的融合过程

Fig. 2 Fusion process of the proposed method

1.1 训练样本集

根据文献[16]中的客观度量,EOL 是空域融合方法中首选的聚焦度量方法. 对于每个像素,计算获得的 EOL 作为该子块拉普拉斯的局部平均值. 因此,获得聚焦信息图的公式如式(1)所示. 其中 H 是平均滤波器, $\otimes$ 和 ∇^2 分别表示卷积运算和拉普拉斯运算. 对于滤波器 H ,高斯滤波器的尺寸为 5×5 ,并且靠近局部窗口中心的拉普拉斯值大于远离局部窗口中心的拉普拉斯值. 具体的计算公式如式(2)所示. 对于由两个变量 i 和 j 构成的一幅数字图像,拉普拉斯是最简单的各向异性的导数运算.

$$M = H \otimes (\nabla^2 \times I), \tag{1}$$

$$\nabla^2 I = \frac{\partial^2 I}{\partial i^2} + \frac{\partial^2 I}{\partial j^2}. \tag{2}$$

聚焦信息图展示的是源图像的高频信息,在清晰区域具有较大的值而在模糊区域具有较小的值. 两幅源图像的聚焦信息图如图 2 所示,清晰区域包含更多的细节信息,而模糊区域则是相反的. 在对聚焦信息图计算完后,创建训练样本集来对 CNN 模型进行优化. 对于每一对源图像,采用滑动窗口的方式依次获取对应的聚焦信息图子块.

多聚焦图像融合问题可看作是一个二类分类的问题. 为此,使用大量的训练样本数据作为训练对象,每个相同场景的对应子块作为一组训练数据. 对于一个样本 $\{p_A, p_B\}$ 当 p_A 比 p_B 更加清晰,则该样本被认定为正向样本,并且该样本的标签值设定为 1. 反之,当 p_B 比 p_A 更加清晰,则该样本被认定为负向样本,并且该样本的标签值设定为 0. 使用常用的灰度多聚焦图像集作为训练样本数据. 对于彩色图像,首先要转化成灰度图像. 对于每对多聚焦图像对,随

机从多聚焦图像集中选取尺寸为 16×16 的块作为采样数据. 在本研究中,最终的训练样本数据共包含 500 000 个正向样本和 500 000 个负向样本.

1.2 网络结构

首先,面临的重要问题就是窗口尺寸的问题. 当窗口尺寸过大或者过小,生成的聚焦图都不会很理想,只有合适的窗口尺寸才能获得最好的效果. 当窗口尺寸过大,活跃窗口更容易同时包含聚焦区域和离焦区域,导致融合图像的分割边界不精确. 当窗口尺寸过小,CNN 模型的分类准确率会降低,这样会导致大量的标签误选. 基于以上考虑以及窗口尺寸的实验,当设置窗口的尺寸为 16×16 时,获得的聚焦图效果最好.

其次,当使用 CNN 时,需要先确定 CNN 的网络结构. 提出方法的网络结构如图 3 所示. 该 CNN 的网络结构共包含 5 层,其中包含 1 个输入层,3 个卷积层和 1 个最大池层. 第 1 层是输入层. 输入是聚焦信息图对应的子块,其尺寸为 16×16 . 第 2 层是卷积层. 使用尺寸为 3×3 和幅度为 1 的卷积核对输入子块进行卷积. 卷积后获得 16 个特征图,其中特征图的尺寸为 14×14 . 第 3 层是卷积层. 同样使用尺寸为 3×3 和幅度为 1 的卷积核对第 2 层输出的特征图进行卷积运算. 卷积后获得 32 个特征图,其中特征图的尺寸为 12×12 . 第 4 层是最大池层. 使用尺寸为 2×2 和幅度为 2 的下采样核来对第 3 层输出的特征图进行最大池运算,池运算后可以获得 32 个特征图,其中特征图的尺寸为 6×6 . 第 5 层是卷积层,使用尺寸为 3×3 和幅度为 1 的卷积核对第 4 层的特征图进行卷积运算. 卷积后获得 64 个特征图,其中特征图的尺寸为 4×4 . 接下来,对每个分支

的 64 个特征图进行连接形成一个尺寸为[2 048,1]的特征向量 $F1$. 然后对该特征向量进行全连接运算,形成一个尺寸为[2,1]的特征向量 $F2$. 最后使用 Softmax 函数对该特征向量 $F2$ 进行计算,形成尺寸为[2,1]的特征向量 $F3$. 该特征向量 $F3$ 的每个值分别表示属于各个类的概率.

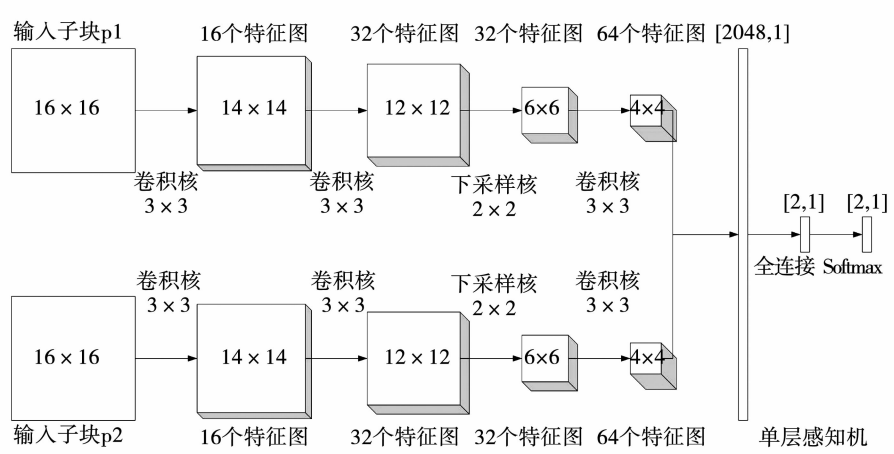


图 3 提出方法的 CNN 网络结构

Fig. 3 CNN structure of the proposed method

1.3 训练模型

在训练过程中,对于一个样本数据,首先把该样本数据先转化成活跃窗口 WA 和 WB,然后采用前向传播和反向传播两个阶段来进行卷积神经网络模型的参数优化. 卷积层的前向传播见式(3). 其中, X_i 为第 i 个输入特征图, K_{ij} 为第 i 个输入特征图和第 j 个输出特征图对应的卷积核,而 b_j 为第 j 个输出特征图的偏置值. 此外, n 为输入特征图的个数, f 函数为 Relu 激活函数,并且符号 $\times$ 表示的是卷积运算.

$$X_j^l = f(\sum_{i=1}^n X_i^{l-1} \times K_{ij}^l + b_j^l). \quad (3)$$

前向传播过程中的池化层可以看作是一个下采样层. 对于下采样层来说,输入特征图和输出特征图的个数是相同的,只是输出特征图的尺寸变小了. 池化层的运算一般有两种:最大池和平均池. 池化层的运算一般是对输入图像的不同子块获取其最大值或者平均值,来作为该子块的输出值. 每个输出图都有一个乘性偏置 β 和一个加性偏置 b . 池化层的前向传播见式(4). 其中, $\text{down}()$ 表示一个下采样函数.

$$X_j^l = f(\beta_j^l \text{down}(X_j^{l-1}) + b_j^l). \quad (4)$$

在训练过程中,常使用的是一个平方误差的损失函数. 该函数的目的主要是获得网络预测数值与真实数值之间的差异,而训练阶段的目的则是进一步缩小该差异. 假设有 N 个带标签的样本 $\{z^1, y^1\}, \{z^2, y^2\}, \dots, \{z^n, y^n\}$, 标签形式为“one - of - c”的格式. 对于单个样本 z^n , 每个样本共有 c 类,那么单个样本 z^n 的平方误差代价函数如式(5)所示,其中 t_k^n 表示第 n 个样本属于第 k 类的预测概率值,而 y_k^n 表示第 n 个样本属于第 k 类的真实概率值.

$$E^n = \frac{1}{2} \sum_{k=1}^c (t_k^n - y_k^n)^2. \quad (5)$$

全部训练集上的误差其实就是每个样本误差的总和,因此 N 个样本的总误差如式(6)所示.

$$E^N = \sum_{n=1}^N E^n = \frac{1}{2} \sum_{n=1}^N \sum_{k=1}^c (t_k^n - y_k^n)^2. \quad (6)$$

假设 W 表示所有待优化参数组成的向量. 若能从中找出一组参数 W^* 来最小化 E^N , 那么该组参数 W^* 则为所要寻找的最优参数,如式(7)所示.

$$W^* = \arg \min_W E^N. \quad (7)$$

由于 E^N 中的参数过多,所以一般的方法很难求得其最小值. 对此,一般都是使用梯度下降的方法来进行求解,具体求解过程如式(8)所示. 其中, α 为学习效率. E^N 直接对参数求导是十分复杂的,所以先用 E^N 对输入加权和求导,在用输入加权和对参数求导. 至此,卷积神经网络的所有参数优化完成.

$$W^{(k)} = W^{(k-1)} - \alpha \frac{\partial E^N}{\partial W} |_{W=W^{(k)}}. \quad (8)$$

1.4 决策图正则化

将多聚焦源图像对反馈至训练后的 CNN 模型, 获得一个聚焦图. 此时的聚焦图中还存在一些误判的标签. 采用“选择最大”的策略来对聚焦图进行进一步的修正,有助于去除聚焦图中一些误判的标签,进一步完善聚焦图. 使用设定的阈值来对聚焦图进行分割形成二值图,详细的计算过程如式(9)所示. 其中, M 表示聚焦图, B 表示二值图,并且 T 表示设定的阈值. 本文中设定的阈值为 0.5.

$$B(x, y) = \begin{cases} 1, & M(x, y) \geq T; \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

此时,二值图中仍然存在一些类似“孔洞”的小区域,通过小区域滤波可以有效的处理这些“孔洞”问题.具体来说,在二值图中小于阈值面积的区域需要进行反转.通常情况下,当图像中区域的面积 $< 0.01 \times H \times W$ 时,该区域相对于整幅图像来说属于较小的区域.其中, H 和 W 分别表示每幅源图像的高度和宽度.因此,把区域面积尺寸 $< 0.01 \times H \times W$ 的区域认定为小区域,通过对这些小区域的去除形成最终决策图.

最后,根据最终决策图提供的权重分配图,对多聚焦源图像进行融合.具体的融合规则如式(10)所示.其中,多聚焦源图像分别为 A 、 B ,最终决策图为 D ,而最终的融合图像为 F .如果多聚焦源图像是彩色图像,使用最终决策图的分配权重分别对彩色图像的各个通道进行融合.最后,重新组合各个融合后的通道形成最终的融合图像.

$$F(x, y) = A(x, y) \times D(x, y) + B(x, y) \times (1 - D(x, y)). \quad (10)$$

2 实验

2.1 实验设置

在本实验中,提出方法的有效性主要是通过两组不同的公开图像数据集进行验证.第一组采用的是常用的灰度图像数据集,包含 20 对不同尺寸的常用多聚焦灰度图像对.第二组采用的是 Lytro 彩色多聚焦图像数据集,共包含 20 对尺寸为 520×520 的

彩色图像.图 4 展示了本文所采用的部分灰度多聚焦图像数据集和 Lytro 彩色多聚焦图像数据集.与最近提出的一些图像融合方法 PCNN^[17]、PC^[18]、FFIF^[19]、BFM^[7] 和 CNN^[9] 进行对比,来展示提出方法的优越性.

近年来,对融合图像的质量进行评价,提出了许多客观度量指标.本文所采用的客观度量指标包括以下 4 个.第 1 个指标采用的是归一化互信息指标^[20] Q_{NMI} .该度量指标主要是测量源图像和融合图像的相关程度,即源图像传递到融合图像的原始信息量. Q_{NMI} 的公式如式(11)所示,其中 $H(I_A)$ 、 $H(I_B)$ 和 $H(I_F)$ 分别表示源图像 A 、 B 和融合图像 F 的信息熵,而 $MI(I_A, I_F)$ 和 $MI(I_B, I_F)$ 表示的是源图像 A 、 B 与融合图像 F 的互信息.

$$Q_{NMI} = 2 \left[\frac{MI(I_A, I_F)}{H(I_A) + H(I_F)} + \frac{MI(I_B, I_F)}{H(I_B) + H(I_F)} \right]. \quad (11)$$

第 2 个指标采用的是基于梯度的度量指标^[21] Q_G .该指标主要测量的是从源图像传递到融合图像的边缘信息量,即源图像和融合图像的边缘相似度. Q_G 的公式如式(12)所示,其中 $Q^{AF}(i, j) = Q_g^{AF}(i, j) Q_\alpha^{AF}(i, j)$,而 $Q_g^{AF}(i, j)$ 和 $Q_\alpha^{AF}(i, j)$ 分别表示源图像 A 和融合图像 F 在位置 (i, j) 的相对边缘强度和边缘方向. $Q^{BF}(i, j)$ 的定义与 $Q^{AF}(i, j)$ 是相同的. $W^A(i, j)$ 分别是 $Q^{AF}(i, j)$ 和 $Q^{BF}(i, j)$ 的权重因子,并且 W 和 H 分别是源图像的宽度和高度.



图 4 部分多聚焦图像数据集

Fig. 4 Partial multi-focus image set

$$Q_G = \frac{\sum_{i=1}^H \sum_{j=1}^W (Q^{AF}(i,j)w^A(i,j) + Q^{BF}(i,j)w^B(i,j))}{\sum_{i=1}^H \sum_{j=1}^W (w^A(i,j) + w^B(i,j))}. \quad (12)$$

$$Q_s = \begin{cases} \lambda(w)S(A,F|w) + (1 - \lambda(w))S(B,F|w), S(A,B|w) \geq 0.75; \\ \max\{S(A,F|w), S(B,F|w)\}, S(A,B|w) < 0.75. \end{cases} \quad (13)$$

$$\lambda(w) = \frac{s(A|w)}{s(A|w) + s(B|w)}. \quad (14)$$

第 4 个指标采用的是基于人类视觉度量的指标^[23] Q_{HP} , 该指标主要是用来度量源图像和融合图像的人类视觉系统中主要特征的相似性. 全局质量图 Q_{GDM} 的公式如(15)所示, 其中, $Q_{AF}(i, j)$ 表示从源图像 A 传递到融合图像 F 的对比度, $Q_{BF}(i, j)$ 与 $Q_{AF}(i, j)$ 的定义是相同的. λ_A 和 λ_B 分别是 $Q_{AF}(i, j)$ 和 $Q_{BF}(i, j)$ 的显著图. 全局质量图 Q_{GDM} 的平均值作为 Q_{HP} 的度量值, 如式(16)所示.

$$Q_{GQM}(i,j) = \lambda_A(i,j)Q_{AF}(i,j) + \lambda_B(i,j)Q_{BF}(i,j), \quad (15)$$

$$Q_{CB} = \overline{Q_{GQM}}. \quad (16)$$

2.2 实验结果与讨论

2.2.1 视觉效果对比

展示提出的方法可以生成具有良好视觉效果

第 3 个指标采用的是基于结构相似性的度量指标^[22] Q_{SSIM} . 该指标主要是用来度量源图像和融合图像的结构相似性. Q_{SSIM} 的式如(13)所示, 并且 $\lambda(w)$ 的式如(14)所示. 其中, $s(A|w)$ 和 $s(B|w)$ 是 A 和 B 在局部窗口 w 的活跃程度.

融合图像, 分别选择具有代表性的灰度图像“disk”和 Lytro 彩色图像“child”进行实验. 以“disk”为实验对象, 本文方法与其它图像融合方法的实验结果如图 5 所示. 从图 5 中可以看出, PCNN 方法的融合图像在钟表的边框产生了较为严重的失真, 并且也伴有明显的对比度下降. PC 方法中的融合图像在钟表的边框上方产生了大面积的模糊效果. QT 方法的融合图像则是在钟表的上边框产生了小面积的模糊效果, 而且在钟表的左表框产生了两个严重的缺口. BFM 方法的融合图像在钟表边框上方产生了较为严重的伪影. FFIF 方法的融合图像在钟表的左边框产生了两个较为严重的缺口. CNN 方法的融合图像在钟表的左边框产生了明显的伪影效果. 同其它的图像融合方法相比, 提出的方法则是展现了较好的融合效果.



图 5 不同图像融合方法的“disk”融合效果

Fig. 5 Fusion results of different methods for the “disk” image

对本文方法的融合性能进行测试,有必要对彩色图像数据集 Lytro 进行实验验证. 以“child”图像作为实验对象,不同图像融合方法的实验结果如图 6 所示. 从图 6 中可以看出,PCNN 方法的融合图像具有明显的失真. PC 方法的融合图像在耳朵上方和分界区域出现了大面积的模糊效果. QT 方法的融合图像在一些区域产生了明显的伪影效果. BFM 方法的融合图像不仅在耳朵上方产生了严重的模糊效

果,而且在分界区域也产生了明显的伪影效果. FFIF 方法的融合图像在很多区域都产生了严重的伪影效果. CNN 和提出方法的融合效果较好,具有良好的视觉效果. 不同图像融合方法的性能通过对不同融合方法的最终决策图进行展示,如图 7 所示. 从图 7 中可以看出,与其它图像融合方法相比,CNN 和提出的方法获得了更精确的最终决策图.

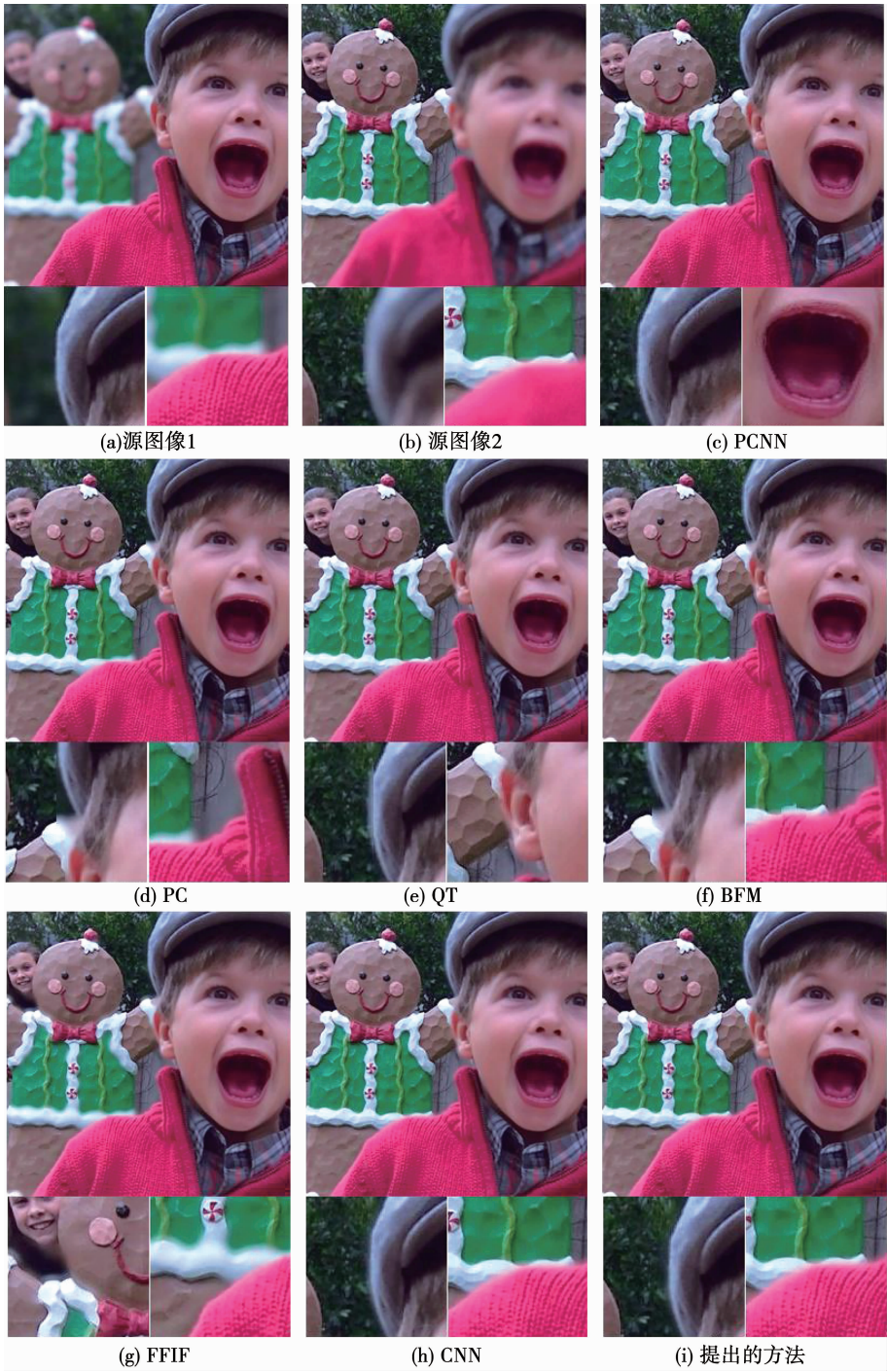


图 6 不同图像融合方法的“child”融合效果.

Fig. 6 Fusion results of different methods for the “child” image

2.2.2 定量比较

更加客观的对不同图像融合方法的性能进行对比,使用 Q_{NMI} 、 Q_G 、 Q_{SSIM} 和 Q_{HP} 等 4 个评价指标来评价不同图像融合方法的性能优劣. 对于灰度多聚焦

图像数据集,不同图像融合方法的客观度量指标的平均值如表 1 所示,每一项评价指标中最好的结果用粗体表示. 从表 1 中可以看出,本文提出方法的实验结果优于其它图像融合方法.

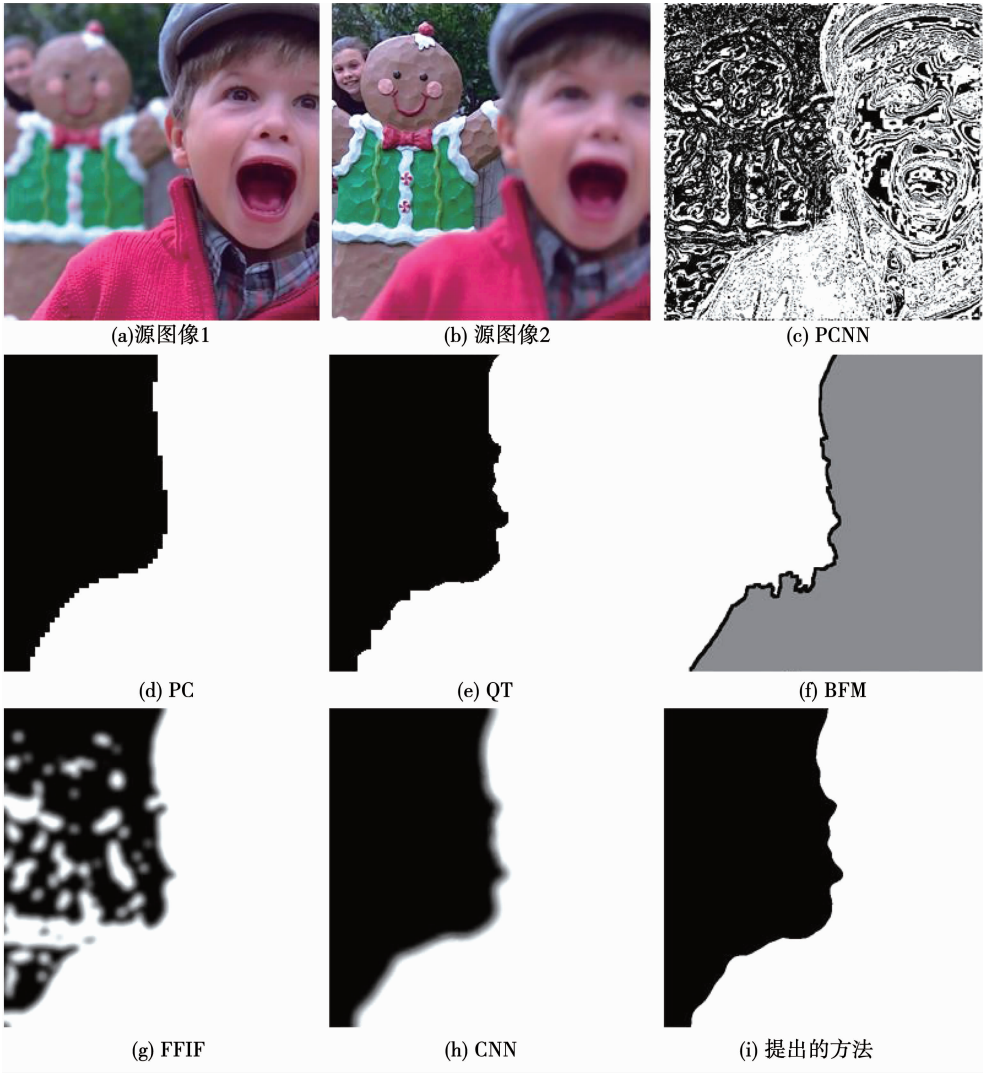


图 7 不同图像融合方法的最终决策

Fig. 7 Final decision maps of different image fusion methods for the “child” image

表 1 不同图像融合方法在灰度多聚焦图像集的客观度量结果

Tab. 1 Evaluation results of different image fusion methods on the grayscale dataset

评价指标	PCNN	PC	QT	BFM	FFIF	CNN	Ours
Q_{NMI}	1.168 8	1.206 5	1.210 1	1.208 2	1.176 0	1.184 4	1.210 9
Q_G	0.672 4	0.735 2	0.749 5	0.747 9	0.717 9	0.749 8	0.751 0
Q_{SSIM}	0.874 5	0.981 1	0.965 6	0.980 5	0.963 8	0.980 0	0.981 7
Q_{HP}	0.684 2	0.778 8	0.765 7	0.781 0	0.742 4	0.781 8	0.782 1

对于 Lytro 彩色图像数据集,不同图像融合方法的客观度量指标的平均值如表 2 所示. 从表 2 可以看出,提出的融合方法在 Lytro 数据集上也展现出了较好的实验结果. 更清晰的展示提出方法的优越性,对两个数据集进行整体分析. 其中,不同图像融合方法在整个数据集上的客观度量指标的平均值如表 3

所示.

2.2.3 中间融合结果以及多幅源图像的融合

提出的方法在多聚焦图像融合的有效性,对 6 对多聚焦源图像的中间融合结果进行展示,如图 8 所示. 从二值图可以看出,聚焦图中的大部分像素都进行了准确的分类,只有平滑区域的一些标签做出

了误判. 采用小区域滤波技术可以完美的去除二值图中的误判标签. 从最终的决策图可以看出, 不仅聚焦图中的所有像素都进行了准确的分类, 而且在聚焦区域和离焦区域之间获得了精准的分割边界.

表 2 不同图像融合方法在 Lytro 数据集的客观度量结果

评价指标	PCNN	PC	QT	BFM	FFIF	CNN	Ours
Q_{NMI}	1.145 4	1.178 8	1.187 3	1.181 2	1.141 0	1.151 2	1.188 1
Q_G	0.700 1	0.745 7	0.762 5	0.754 5	0.756 7	0.762 5	0.762 8
Q_{SSIM}	0.921 4	0.988 7	0.989 0	0.988 6	0.985 5	0.988 6	0.989 3
Q_{HP}	0.720 4	0.801 1	0.808 5	0.805 7	0.795 3	0.808 4	0.808 8

表 3 不同图像融合方法对整个数据集的客观度量结果

评价指标	PCNN	PC	QT	BFM	FFIF	CNN	Ours
Q_{NMI}	1.157 1	1.192 7	1.198 7	1.194 7	1.158 5	1.167 8	1.199 5
Q_G	0.686 3	0.740 5	0.756 0	0.751 2	0.737 3	0.756 2	0.756 9
Q_{SSIM}	0.898 0	0.984 9	0.977 3	0.984 6	0.974 7	0.984 3	0.985 5
Q_{HP}	0.702 3	0.790 0	0.787 1	0.793 4	0.768 9	0.795 1	0.795 5

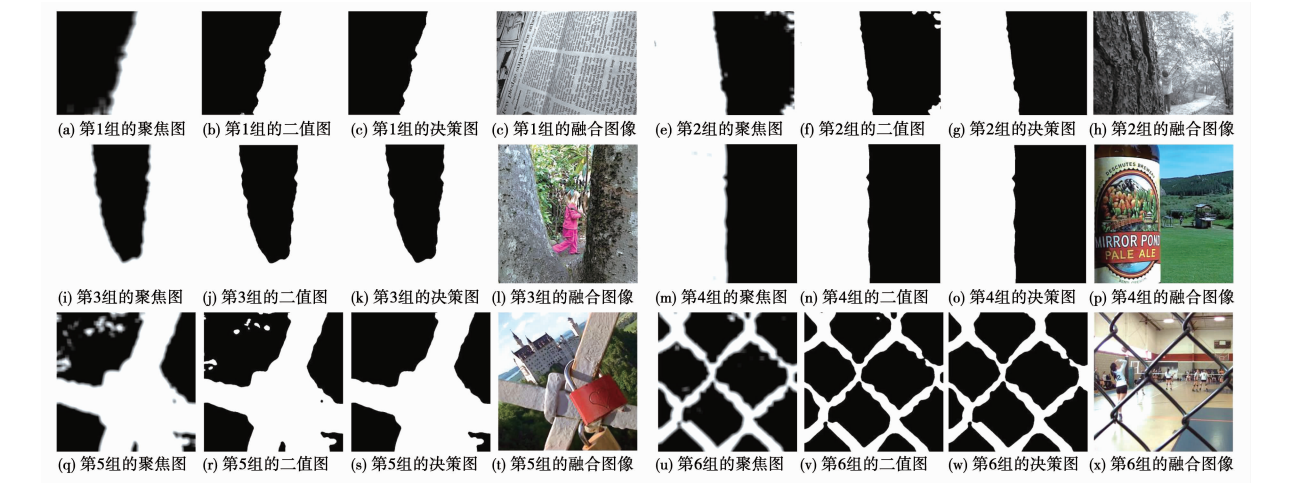


图 8 部分多聚焦图像的中间融合结果

Fig. 8 Intermediate fusion results of partial multi-focus images

对于两幅以上的多聚焦源图像进行融合, 提出的方法同样是适用的. 展示提出的方法对于两幅以上的多聚焦源图像的融合问题, 采用 Lytro 数据集中的三幅源图像进行实验. 三幅多聚焦源图像和最终

的融合图像如图 9 所示. 从图 9 中可以看出, 在没有引入伪影的前提下, 所有源图像中的聚焦区域都集成到了最终的融合图像当中.



图 9 三幅多聚焦源图像的融合

Fig. 9 Fusion of three multi-focus source images

2.2.4 参数讨论

对于窗口尺寸的问题,选择常用的滑动窗口尺寸 $4 * 4$ 、 $8 * 8$ 、 $16 * 16$ 和 $32 * 32$ 进行实验. 当窗口尺寸为 $4 * 4$ 和 $8 * 8$ 时,由于活跃窗口包含的信息较少,所以在聚焦图中存在大量的标签误选. 反之,当窗口尺寸为 $32 * 32$ 时,虽然活跃窗口包含的信息更加丰富,有助于提高聚焦图中预测标签的准确率,但是较大的窗口容易导致聚焦区域之间的分割边界不精确. 只有当窗口尺寸为 $16 * 16$ 时,聚焦区域之间的分割边界才很精确. 此时,虽然在聚焦区域内仍然存在少量的误选标签,但是这些误判的标签可以通过小区域滤波进行去除. 对于整个数据集,不同窗口尺寸的实验结果如表 4 所示. 从表 4 可以看出,当窗口尺寸为 $16 * 16$ 时,提出的方法获得的融合效果是最好的.

表 4 不同窗口尺寸的实验结果

Tab. 4 Experimental results of different window sizes

评价指标	4 * 4/像素	8 * 8/像素	32 * 32/像素	16 * 16/像素
Q_{NMI}	1.025 2	1.189 2	1.196 2	1.199 5
Q_G	0.723 1	0.738 5	0.735 1	0.756 9
Q_{SSIM}	0.940 3	0.971 4	0.979 4	0.985 5
Q_{HP}	0.737 1	0.791 8	0.793 4	0.795 5

对于二值分割的阈值问题,选取主要的分割阈值 0 、 0.25 、 0.5 、 0.75 和 1 进行实验. 当分割阈值为 0.25 或者 0.75 时,与分割阈值 0.5 相比,不仅聚焦区域之间的分割边界不精确,而且聚焦区域内存在更多的误判标签. 当分割阈值为 0 或者 1 时,与分割阈值 0.25 或 0.75 相比,获得聚焦图的实验结果更差. 对于整个数据集,不同分割阈值的实验结果如表 5 所示. 从表 5 可以看出,当分割阈值为 0.5 时,提出的方法可以获得最好的融合效果.

表 5 不同分割阈值的实验结果

Tab. 5 Experimental results of different segmentation thresholds

评价指标	0	0.25	0.75	1	0.5
Q_{NMI}	1.190 8	1.198 3	1.196 9	1.191 8	1.199 5
Q_G	0.747 3	0.754 4	0.755 3	0.749 8	0.756 9
Q_{SSIM}	0.976 8	0.981 9	0.982 1	0.977 5	0.985 5
Q_{HP}	0.770 9	0.791 7	0.792 5	0.771 5	0.795 5

3 结 论

在本文中,利用拉普拉斯能量和卷积神经网络

的多聚焦图像融合方法被提出来. 该方法的新颖之处在于拉普拉斯能量算子可以有效的提取源图像中的聚焦信息,并且训练后的 CNN 模型根据提取的聚焦信息可以准确的区分清晰子块和模糊子块. 清晰子块和模糊子块的聚焦特征完全不同,并且训练后的 CNN 模型具有很好的分类能力. 在经过大量的样本训练后,对于给定的多聚焦源图像对,提出的方法可以产生精确度很高的聚焦图. 在经过二值分割和小区域滤波后获得精准的决策图. 根据最终的融合图像可以看出,提出的方法不仅精确的提取了源图像中的聚焦区域,而且获得了精确的分割边界. 实验结果表明无论从融合图像的视觉效果还是从客观评价指标来看,提出的方法均优于已有的其它融合方法. 在未来,深度学习的方法会进一步促进多聚焦图像融合技术的发展.

参考文献

[1] LI Shutao, YANG Bin, HU Jianwen. Performance comparison of different multi-resolution transforms for image fusion[J]. Information Fusion, 2011, 12(2): 78. DOI: 10.1016/j.inffus.2010.03.002

[2] 李娇, 杨艳春, 党建武, 等. NSST 与引导滤波相结合的多聚焦图像融合方法[J]. 哈尔滨工业大学学报, 2018, 50(11): 145

LI Jiao, YANG Yanchun, DANG Jianwu, et al. NSST and guided filtering for multi-focus image fusion algorithm[J]. Journal of Harbin Institute of Technology, 2018, 50(11): 145. DOI: 10.11918/j.issn.0367-6234.201805006

[3] LI Hui, LI Li, ZHANG Jixiang. Multi-focus image fusion based on sparse feature matrix decomposition and morphological filtering[J]. Optics Communications, 2015, 342: 1. DOI: 10.1016/j.optcom.2014.12.048

[4] LIU Zhaodong, CHAI Yi, YIN Hongpeng, et al. A novel multi-focus image fusion approach based on image decomposition[J]. Information Fusion, 2017, 35: 102. DOI: 10.1016/j.inffus.2016.09.007

[5] LI Xiongfei, WANG Lingling, WANG Jing, et al. Multi-focus image fusion algorithm based on multilevel morphological component analysis and support vector machine[J]. IET Image Processing, 2017, 11(10): 919. DOI: 10.1049/iet-ipr.2016.0661

[6] BAI Xiangzhi, ZHANG Yu, ZHOU Fugen, et al. Quadtree-based multi-focus image fusion using a weighted focus-measure[J]. Information Fusion, 2015, 22: 105. DOI: 10.1016/j.inffus.2014.05.003

[7] ZHANG Yu, BAI Xiangzhi, WANG Tao. Boundary finding based multi-focus image fusion through multi-scale morphological focus-measure[J]. Information Fusion, 2017, 35: 81. DOI: 10.1016/j.inffus.2016.09.006

[8] ZHOU Zhiqiang, LI Sun, WANG Bo. Multi-scale weighted gradient-based fusion for multi-focus images[J]. Information Fusion, 2014, 20(1): 60. DOI: 10.1016/j.inffus.2013.11.005

[9] LIU Yu, CHEN Xun, PENG Hu, et al. Multi-focus image fusion

with a deep convolutional neural network[J]. Information Fusion, 2017, 36: 191. DOI: 10.1016/j.inffus.2016.12.001

[10] YAN Xiang, GILANI S Z, QIN Hanlin, et al. Unsupervised deep multi-focus image fusion[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2018: 1

[11] TANG Han, XIAO Bin, LI Weisheng, et al. Pixel convolutional neural network for multi-focus image fusion [J]. Information Sciences, 2018, 433: 125. DOI: 10.1016/j.ins.2017.12.043

[12] FARFADE S S, SABERIAN M, LI Lijia. Multi-view face detection using deep convolutional neural networks[C]//Proceedings of the 2015 ACM on International Conference on Multimedia Retrieval. Shanghai: ACM, 2015: 643. DOI: 10.1145/2671188.2749408

[13] SUN Yi, WANG Xiaogang, TANG Xiaou. Deep learning face representation from predicting 10,000 classes[C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Columbus: IEEE, 2014: 1891. DOI: 10.1109/CVPR.2014.244

[14] DONG Chao, LOY C C, HE Kaiming, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2): 295. DOI: 10.1109/TPAMI.2015.2439281

[15] ZAGORUYKO S, KOMODAKIS N. Learning to compare image patches via convolutional neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 4353

[16] HUANG Wei, JING Zhongliang. Evaluation of focus measures in multi-focus image fusion[J]. Pattern Recognition Letters, 2007, 28(4): 493. DOI: 10.1016/j.patrec.2006.09.005

[17] CAI Mingrui, YANG Junyi, CAI Guanghui. Multi-focus image fusion algorithm using LP transformation and PCNN [C]//Proceedings of 2015 6th IEEE international conference on software engineering and service science. Beijing: IEEE, 2015: 237

[18] ZHAN Kun, LI Qiaoqiao, TENG Jicai, et al. Multifocus image fusion using phase congruency[J]. Journal of electronic imaging, 2015, 24(3): 033014. DOI: 10.1117/1.JEI.24.3.033014

[19] ZHAN Kun, XIE Yuange, WANG Haibo, et al. Fast filtering image fusion[J]. Journal of Electronic Imaging, 2017, 26(6): 063004. DOI: 10.1117/1.JEI.26.6.063004

[20] HOSSNY M, NAHAVANDI S, CREIGHTON D. Comments on information measure for performance of image fusion[J]. Electronic Letters, 2008, 46(18): 1066. DOI: 10.1049/el.20081754

[21] XYDEAS C S, PETROVIC V. Objective image fusion performance measure[J]. Electronics Letters, 2000, 36(4): 308. DOI: 10.1049/el.20000267

[22] YANG Cui, ZHANG Jianqi, WANG Xiaorui, et al. A novel similarity based quality metric for image fusion [J]. Information Fusion, 2008, 9(2): 156. DOI: 10.1016/j.inffus.2006.09.001

[23] CHEN Yin, BLUM R S. A new automated quality assessment algorithm for image fusion[J]. Image and Vision Computing, 2009, 27: 1421. DOI: 10.1016/j.imavis.2007.12.002

(编辑 苗秀芝)