

文章编号:1007-2780(2021)09-1283-11

基于多尺度细节的孪生卷积神经网络 图像融合算法

刘 博^{1,2}, 韩广良^{1*}, 罗惠元^{1,2}

(1. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033;

2. 中国科学院大学, 北京 100049)

摘要:图像融合将来自不同捕获条件或不同传感器的互补图像进行融合以提高图像的视觉质量。针对这一任务,本文提出一种改进的滚动引导滤波与神经网络相结合的多尺度融合算法。首先,使用孪生卷积神经网络学习图像特征,并以此获得包含源图像显著特征的权值映射图。随后,使用改进的滚动引导滤波对图像进行多尺度分解,结合信息熵使滚动引导滤波权重参数自适应化来实现多尺度自适应分解,并结合非线性映射增强图像细节信息。最后,采用局部能量与权值图相结合的自适应调整融合模式对多尺度图像进行融合。经实验对比,所提方法能够避免出现图像边缘圆晕效应,且能够更好地突出图像边缘、细节纹理特征。另外,与其他算法相比,本文所提出的算法在平均梯度、信息熵、视觉信息保真度以及空间频率等客观评价指标项上均取得了更优的性能表现。

关键词:图像处理;孪生卷积神经网络;图像融合;滚动引导滤波;多尺度图像

中图分类号:TP391.4 文献标识码:A doi:10.37188/CJLCD.2020-0339

Image fusion algorithm based on multi-scale detail siamese convolutional neural network

LIU Bo^{1,2}, HAN Guang-liang^{1*}, LUO Hui-yuan^{1,2}(1. *Changchun Institute of Optics, Fine Mechanics and Physics,
Chinese Academy of Science, Changchun 130033, China;*2. *University of Chinese Academy of Sciences, Beijing 100049, China*)

Abstract: Image fusion is to fuse the complementary images under different acquisition conditions or sensors for improving the visual quality of images. For this task, this paper proposes an improved multi-scale fusion algorithm based on the combination of rolling guided filtering and neural network. Firstly, the siamese convolutional neural network is used to learn the image features and further obtain the weight map containing the significant features of the source image. Then, the image is decomposed by the improved rolling guide filter, and the weight parameters of the rolling guide filter are adaptive to achieve multi-scale adaptive decomposition by combining the information entropy, and the

收稿日期:2020-12-09;修订日期:2021-01-26.

基金项目:国家自然科学基金青年基金(No.61401425, No.61602432)

Supported by National Natural Science Foundation of China(No. 61401425, No.61602432)

* 通信联系人, E-mail: hangl@ciomp.ac.cn

details of the image are enhanced by combining the nonlinear mapping. Finally, based on the local energy and weight map, an adaptive fusing strategy is designed to combine the multi-scale images adaptively. The experimental comparison shows that the proposed method can avoid the halo effect of image edge and better highlight the edge and detail texture features of image. In addition, compared with other algorithms, our algorithm can achieve a better performance according to the objective evaluation index items, such as average gradient, information entropy, visual information fidelity and spatial frequency.

Key words: image processing; siamese convolution neural network; image fusion; rolling guided filtering; multiscale image

1 引言

图像融合将同一场景下两幅或多幅源图像的相关有效信息整合成一幅综合图像,实现了不同模态图像的信息互补。经融合处理得到的图像能够弥补单一图像信息量不足问题,更符合人眼视觉感知和工业视觉领域需求,因此图像融合在许多领域都有广泛的应用。

近些年,图像融合技术发展迅速,相关学者针对像素级图像融合进行了深入研究。其中基于多尺度的融合方法因有着鲁棒性强、具有局部时频特性等特点,一直是一个热门的研究课题。经典融合方法包括基于塔形分解的变化融合方法、基于小波变换的融合方法、基于轮廓波变换方法^[1]等。基于塔形分解的变换方法包括:基于拉普拉斯金字塔变换(LP)^[2]、基于对比度金字塔变换^[3]以及基于梯度金字塔变换^[4]等。基于小波变换方法包括:基于方向小波变换^[5]以及双树复小波变换(DTCWT)^[6]等。然而以上方法在方向性上存在较大限制,且不具备平移不变性,融合结果中易出现边缘特征捕获不足和边缘圆晕效应问题。而非下采样轮廓波(NSCT)^[7]和非下采样剪切波(NSST)^[8]具有较好的方向性和平移性,能够在一定程度上解决上述问题,但融合过程中没有充分考虑空间一致性,对全局结构的边缘轮廓特征提取能力不强。边缘保留滤波具有平移不变性和边缘保留性,能够在平滑图像的同时保持边缘清晰。因此,很多方法将边缘保留滤波引入到图像融合中,以克服上述缺陷。文献[9]构造了一种多尺度方向双边滤波器,结合双边滤波器和非采样方向滤波器组进行图像融合。文献[10]采用基于L0光滑滤波的多尺度边缘保持分解,提出了一种

增强融合结果细节的融合策略。文献[11]提出了基于滚动引导滤波和混合多尺度分解的图像融合方法,其利用滚动引导滤波将图像分解到基本层、小尺度层和大尺度层,通过对不同层级进行融合使融合图像具有更好的纹理边缘特征。滚动引导滤波作为保边滤波器的一种,不仅能够平滑细节,而且能够保留图像的边缘特征,避免边缘位置出现圆晕现象,这对保证融合图像质量具有重要意义。

此外,随着机器学习技术的发展,一些基于学习的融合方法也被相继提出。文献[12]设计一种基于编码和解码体系结构的网络框架,通过无监督训练实现端到端图像融合。文献[13]结合多尺度模块设计融合网络,通过对不同尺度的特征进行提取提高融合效果。文献[14]结合图像块融合思想,通过设计神经网络来指导源图像块进行融合。神经网络具有强大的特征提取能力,其可以捕捉图像中最显著物体的内部特征。但随着网络层数增多,不断进行池化的操作,可能会使特征图损失很多细节信息,导致在边缘部分的信息丢失。

基于边缘滤波的多尺度分解具有较好的细节信息与边缘轮廓信息提取能力,而神经网络则对物体的内部区域具有较强的特征提取能力。本文中这两个模块结合起来使用,并提出了一种基于多尺度细节的卷积神经网络图像融合方法。通过将神经网络与改进的滚动引导滤波相结合,以多尺度的方式结合神经网络得到的权值图共同处理信息,使融合图像更好地满足人类视觉效果。滚动引导滤波的联合滤波采用联合双边滤波器来实现,并采用前次迭代图像作为滤波器引导图像。通过结合信息熵原理对空间平滑参数与灰度权重参数进行改进,使其实现针对不同源图像的参数自适应化。这

种方法能够保留边缘信息的同时使多尺度图像充分提取细节信息。另外,通过非线性映射变换对构建的多尺度图像细节进行非线性映射,从而丰富突出细节层内细节纹理信息。

2 本文算法原理

本文融合框架主要分为3部分,如图1所示。

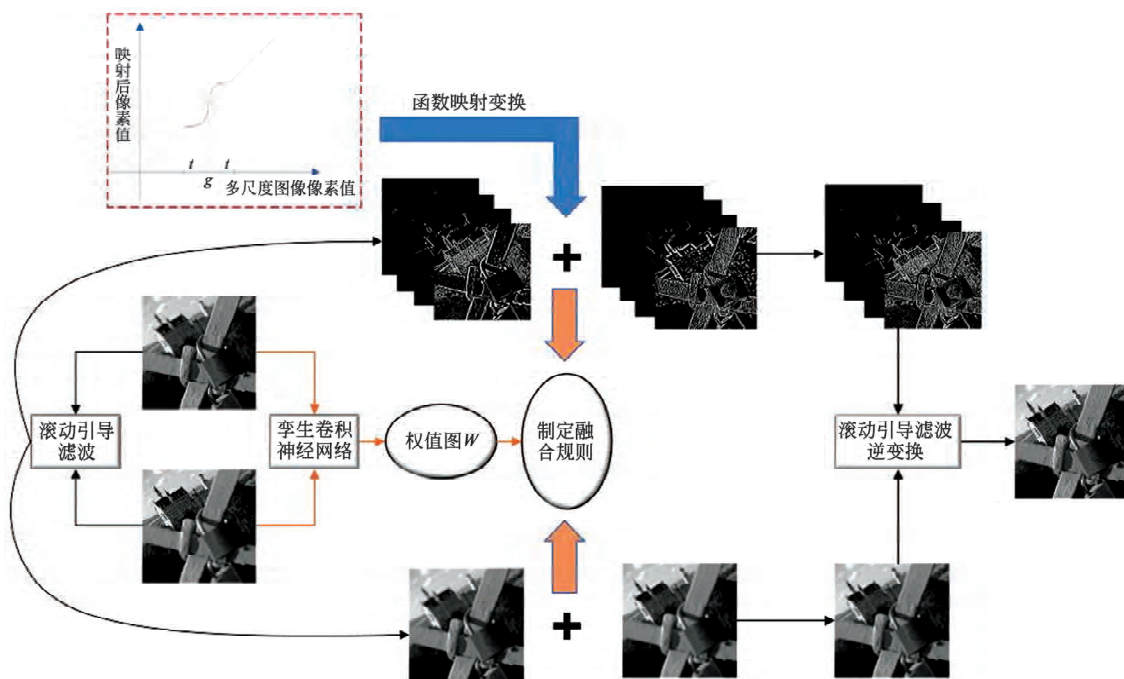


图1 多聚焦图像融合框架

Fig.1 Multi-focus image fusion framework

2.1 孪生卷积神经网络

孪生网络是由两个完全相同结构的卷积神经网络组成,两个CNN能够将输入图像映射到新的空间中进行表示,然后通过损失函数计算评价两者之间的相似度。Sumit等^[15]提出了孪生网络,通过从数据中学习相似性度量来对不同类别样本加以区分,从而实现多分类任务。本文以孪生网络作为神经网络框架提取融合图像的权值图 W ,如图2所示。源图像通过滑窗的方式得到 16×16 的patch作为网络的输入图像输入到训练好的网络中。网络对图像进行特征提取,然后将特征图像连接起来组成特征向量。特征向量经过两个全连接层,得到一个二分类的特征向量,该向量代表清晰度分类的得分情况。将该二维向量经

首先,两张源图像被送入孪生卷积神经网络中,经过网络特征提取后,得到与源图像特征相关的权值映射图。其次,对源图像进行滚动引导滤波,并对相邻迭代的细节层图像进行差值,获得多级差分细节图像。然后对多级差分图像进行函数映射变换,使图像细节得到增强。最后,通过局部能量与权值图相结合,自适应调整融合模式对多尺度图像进行融合,从而得到融合图像。

softmax层进行归一化后可得到所属类别的概率值,它可表示为图像的清晰度度量值,依据此值来组成权值图。该CNN结构主要由3个卷积层、一个池化层、两个全连接层以及softmax层组成。具体参数如表1所示,3个卷积层的卷积核大小为 3×3 ,步长设定为1,最大池化层采用 2×2 大小的内核,步长设定为2。

输入网络的patch的大小影响着算法的速度与精度,为保证两者之间的平衡,本文中patch采用 16×16 的尺寸。源图像通过滑窗得到 16×16 的patch,由于网络结构中存在一个pooling层,因此滑窗的步长为2。将取得的patch对送入网络后会得到与清晰度特征信息相关的二分类概率值,将此值作为融合系数用来指导源图像融合。

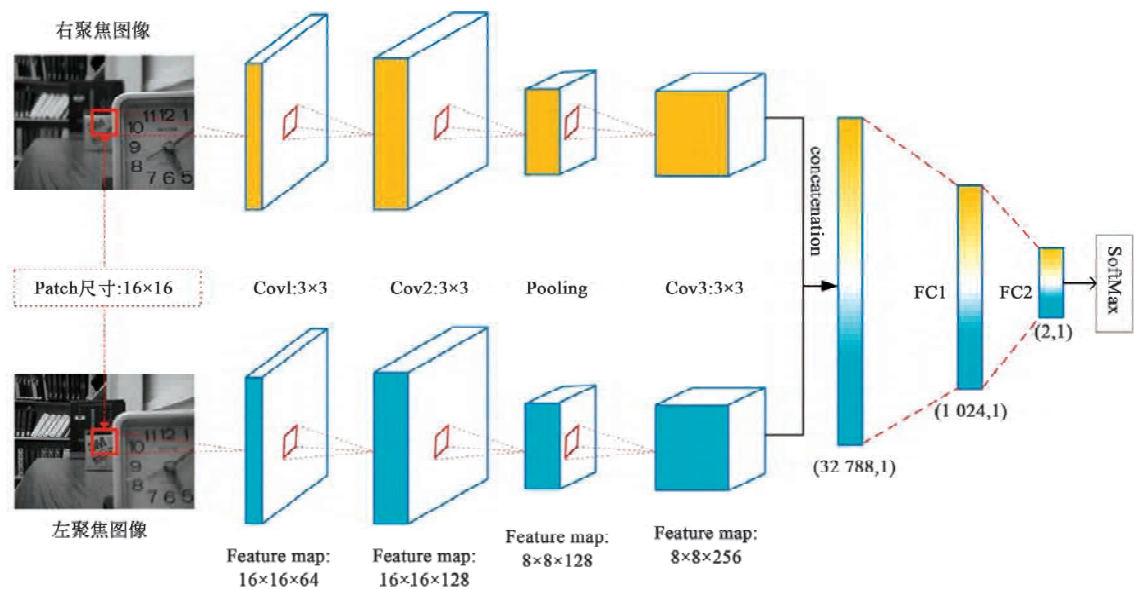


图 2 孪生网络框架

Fig.2 Architecture of siamese network

表 1 CNN 网络参数

Tab.1 CNN network parameters

模块	patch/向量尺寸	卷积核尺寸	步长	卷积核个数
Cov1	16×16	3×3	1	64
Cov2	16×16	3×3	1	128
Pooling	8×8	3×3	2	128
Cov3	8×8	3×3	1	256
Concatenation	(32 788,1)	/	/	/
FC1	(1 024,1)	/	/	/
FC2	(2,1)	/	/	/

为得到与源图像大小尺寸相对应的权值图,需要对其进行填充。如图 3 所示,将 1×1 的权值进行填充使其与 patch 的尺寸相同,对于滑窗相邻的 patch,其权值图在其相对应的重叠部分采用均值法来得到对应权值。通过这种方式可以重组出与源图像大小相对应的权值图。

2.2 多尺度图像构建

传统金字塔分解多尺度的过程中对边缘细节信息并不敏感,对细节的捕捉能力差,因此,利用保边滤波器进行多尺度图像构建能够更好地保护图像的边缘信息。其中滚动引导滤波^[16]能够在平滑细节纹理等信息的同时,通过不断迭代的方式实现边缘信息的恢复。

滚动引导滤波过程主要可分为小结构消除和

边缘恢复两部分。小结构消除过程主要通过高斯滤波来实现,如式(1)所示。

$$G_m = \frac{1}{K_m} \sum_{n \in \mathbb{S}_m} \exp\left(-\frac{\|m-n\|^2}{2\sigma_s}\right) \cdot I_n, \tag{1}$$

式中, $K_m = \sum_{n \in \mathbb{S}_m} \exp\left(-\frac{\|m-n\|^2}{2\sigma_s}\right)$ 为归一化系数, σ_s 作为空间标准差决定着图像的平滑度, m, n 是像素的空间索引坐标, $\mathbb{S}_m$ 是 m 点坐标的邻域集合。源图像 I_n 经过高斯滤波后得到滤波图像 G_m , 图像中的噪声、细节纹理等信息将会被消除,其边缘也会出现一定程度的模糊现象。边缘恢复过程主要针对边缘模糊现象,其过程通过联合滤波来实现。即

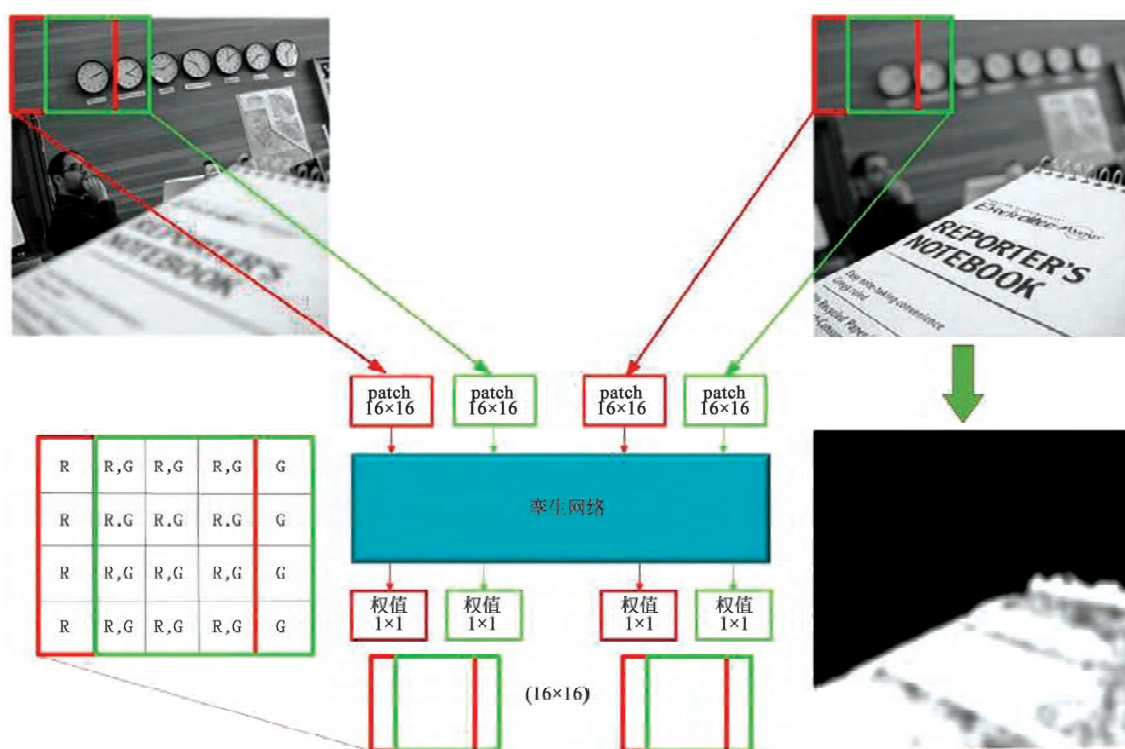


图3 权值图生成机制

Fig.3 Mechanism for generating the weight map

$$J_m^{t+1} = \frac{1}{K_m} \sum_{n \in \mathbb{S}_m} \exp\left(-\frac{\|m-n\|^2}{2\sigma_s^2} - \frac{\|J^t(m) - J^t(n)\|}{2\sigma_r^2}\right) \cdot I_n, \quad (2)$$

式中:

$$K_m = \sum_{n \in \mathbb{S}_m} \exp\left(-\frac{\|m-n\|^2}{2\sigma_s^2} - \frac{\|J^t(m) - J^t(n)\|}{2\sigma_r^2}\right)$$

同样作为归一化系数; $J^t(m)$ 、 $J^t(n)$ 是第 t 次迭代得到的引导图像在空间索引 m 、 n 处的像素值, σ_r 是灰度域的权重, 它决定图像边缘恢复的强度。该过程采用联合双边滤波的形式将第 t 次迭代的输出图像 J^t 作为引导图像, 在滤波过程中它能够在梯度变化处起主导作用, 提高滤波器处理边缘特征信息的强度, 经滤波后得到输出图像 J^{t+1} 。通过迭代使输出图像不断更新, 图像的边缘特征也被不断恢复。

参数 σ_s 和 σ_r 是表征滚动引导滤波性能的重要参数, 它们决定着图像滤波的最终效果。采用固定值作为 σ_s 和 σ_r 的输入往往需要大量实验和经验总结, 并且随着滤波迭代更替图像对参数的要求也会发生变化, 不具备自适应性。图像的信息熵能够反映图像内信息的丰富程度, 熵的本质

体现了一个系统的混乱程度, 而图像的滤波正是对图像混乱程度进行改变的过程。因此, 图像的信息熵与参数 σ_s 和 σ_r 存在联系, 通过图像的信息熵来决定参数 σ_s 和 σ_r 可以使滚动引导滤波具备较强的自适应性。

参数 σ_s 主要用于控制滤波的尺度, 通过式(3)、式(4)和式(5)对各层滤波参数进行设置。

$$\sigma_s^t = \frac{3}{2} \times \sigma_s^{t-1}, (t=1, 2, \dots, J), \quad (3)$$

$$\sigma_s^0 = \max(E_n - 1, \frac{3}{2}), \quad (4)$$

$$E_n = -\frac{1}{2} \left(\sum_{m=0}^{L-1} p_m \ln p_m \right), \quad (5)$$

其中: E_n 代表图像的信息熵之和, L 是图像的总灰度级数, p_m 是灰度值为 m 的像素点在图像中出现的概率, σ_s^0 是初始图像的滤波权重, σ_s^t 是第 t 次迭代时的滤波权重, J 为最终分解尺度。

图像的每次迭代会在联合滤波的作用下进行

不同程度的保边滤波,对相邻迭代图像进行差值处理可以得到差分图像。图 4 展示了相邻两次迭代图像间被模糊掉的细节信息。 σ_s 的每次迭代更新都采取公式(3)进行更新,这使得图像的纹理平

滑程度将随之变化,不同尺度的细节层包含的纹理信息也将由细到粗变化,从而实现细节轮廓特征的多尺度自适应提取。

参数 σ_r 主要用于滤波的边缘恢复与保留,其

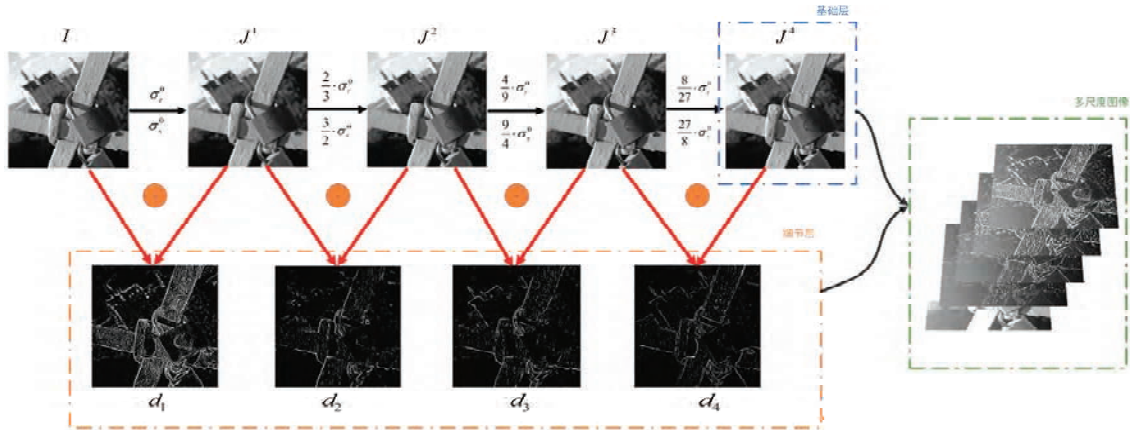


图 4 滚动引导滤波多尺度构建

Fig.4 Multi-scale construction of rolling guided filtering

设置如式(6)、式(7)和式(8)所示:

$$\sigma_r^t = \frac{2}{3} \times \sigma_r^{t-1}, (t=1, 2, \dots, J), \quad (6)$$

$$\sigma_r^0 = \left(\frac{3}{2}\right)^J \lambda_{En}, \quad (7)$$

$$\lambda_{En} = -\frac{2}{\sqrt{M \times N}} \left(\sum_{m=0}^{L-1} p_m \ln p_m \right), \quad (8)$$

其中: λ_{En} 是归一化的信息熵之和, M 与 N 是图片的尺寸大小, σ_r^0 表示初始的空间范围权重。受最终分解尺度 J 与迭代次数 t 的影响, σ_r^t 随 σ_s 逐渐增大的同时按公式(6)进行衰减,以保证滤波后的图像保留边缘细节。如图 5 所示,细节层 d_4 相较于 d_3 所提特征有所减少,表明 $J=4$ 时图像已趋于收敛,具有稳定的输出。因此,图像的空间多尺度特征由差分图像 d 构成的细节层图像与滚动引导滤波分解得到基础层图像共同构成。

局部拉普拉斯滤波^[17]是一种边缘保护滤波。它通过对图像进行逐个像素点的像素映射变换使图像的细节信息得到增强。本文参考其原理对滚动引导滤波分解的细节层图像进行非线性映射变换,实现图像细节增强,图 5 大致展示了其变换过程。本文首先对多尺度细节层图像进行函数映射变换,以 t 作为阈值用于界定细节特征像素与边

缘特征像素。对于 t 范围内的细节像素,通过映射变换对其进行非线性增强,而对边缘像素保持线性不变。

图 5 显示了细节增强的效果图,图 5(a)和图 5(b)是一组多聚焦源图像,图 5(c)和图 5(e)经滚动引导滤波分解的多尺度细节层图像,图 5(d)和图 5(f)是与之对应的映射增强后细节层图像。具体映射变换如式(9)、式(10)所示。

$$\begin{aligned} \text{Iremap} &= \text{fact} \times (I - \text{ref}) \times \\ &\exp\left(-\frac{(I - \text{ref})^2}{2\sigma^2}\right), \end{aligned} \quad (9)$$

$$\begin{aligned} I_{\text{out}} &= I + (|I - \text{ref}| < \text{threshold}) \times \\ &\text{Iremap} \times \frac{(1 - |\text{ref}|)}{\text{threshold}}, \end{aligned} \quad (10)$$

Iremap 是图像映射函数。fact 是图像增强系数,当其大于 0 时起到细节增强作用,相反小于 0 时起平滑作用。较大的 fact 会使图像中边缘轮廓产生亮斑,而较小的 fact 会导致增强效果不足。通过对增强系数进行对比实验发现, fact 取 0.7 时对细节的增强更佳。参数 σ 用于调整图像小结构细节的增强幅度,其值越高对细节的增强幅度越大,但过大的 σ 会导致图像局部图案亮度偏高,出现失真现象。通过对多种模态图像进行试验,参数 σ 在取 0.15 时具有优异效果。ref 是像素灰

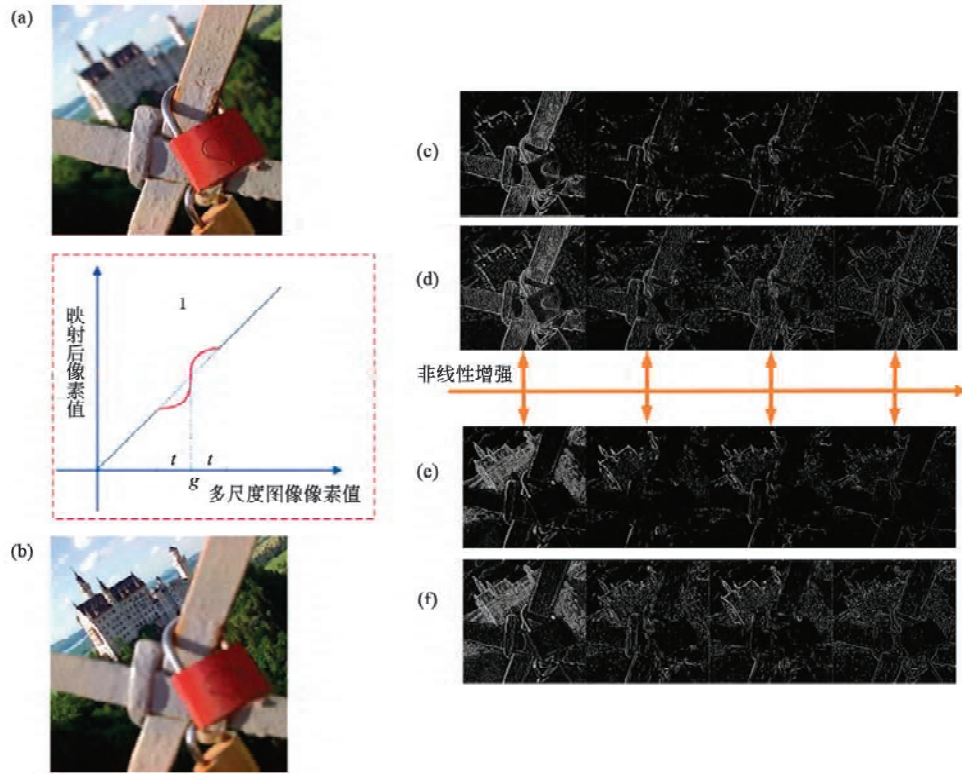


图5 非线性映射增强。(a),(b)一组多聚焦源图像;(c),(e)多尺度分解后的细节层图像;(d),(f)映射增强后细节层图像。

Fig.5 Nonlinear mapping enhancement. (a), (b) a group of multi-focus source images; (c), (e) detail layer images after multi-scale decomposition, (d), (f) detail layer images after mapping enhancement.

度进行 N 采样后每个灰度级的临界值。通过将 N 个灰度级进行迭代使图像 I 在不同灰度级进行计算,通常 N 设置为 12。threshold 是界定细节像素的阈值,在此范围内的认定为细节像素。根据 Iremap,这些细节像素会进行相关的映射增强。 $\frac{(1-|I-\text{ref}|)}{\text{threshold}}$ 影响映射强度差异,增强后的图像为 I_{out} 。通过这种映射变化的方式能够在保持边缘的同时对图像的细节信息进行增强,丰富图像细节特征信息,提高融合图像视觉效果。

2.3 融合规则制定

源图像经多尺度分解以及细节层增强后与基础层共同组成多尺度图像。图像的融合主要针对各层之间进行融合然后逐层叠加。左、右聚焦多尺度图像分别记作 $D^l\{L\}$ 、 $D^l\{R\}$,其中 l 代表第 l 层多尺度图像,并将基础层记作 $D^0\{L\}$ 、 $D^0\{R\}$ 。各层之间采用局部能量与权值图 W 相结合的方式融合。局部能量特征的定义是

$$E(x, y) = \sum_m \sum_n S(m, n) \cdot$$

$$F(x + m, y + n)^2, \quad (11)$$

式中, $S(m, n)$ 是局部区域的加权系数,其核大小为 3×3 。 $E(x, y)$ 代表点 (x, y) 在 m, n 区域像素值的能量值。反映到图像特征上, $E(x, y)$ 值高代表该区域的特征更加明显,即此区域对应的不同尺度图像细节信息更丰富或是低频亮度特征更充分。所以,对 $D^l\{L\}$ 、 $D^l\{R\}$ 求区域能量,即:

$$E_L^l(x, y) = \sum_{m=-1}^1 \sum_{n=-1}^1 S(m, n) \cdot D^l\{L\}(x + m, y + n)^2, \quad (12)$$

$$E_R^l(x, y) = \sum_{m=-1}^1 \sum_{n=-1}^1 S(m, n) \cdot D^l\{R\}(x + m, y + n)^2, \quad (13)$$

$$M^l(x, y) = \left[\sum_{m=-1}^1 \sum_{n=-1}^1 S(m, n) \cdot D^l\{L\}(x + m, y + n) \times D^l\{R\}(x + m, y + n) \right] / (E_L^l(x, y) + E_R^l(x, y)), \quad (14)$$

$M^l(x, y)$ 是估算 $E_L^l(x, y)$ 、 $E_R^l(x, y)$ 之间的相

似度函数, 计算如下: $M^l(x, y)$ 的取值范围在 $[-1, 1]$ 之间, 取值越大说明两者之间能量值越接近, 相似度越高。设定一个阈值 th , 当 $M^l < th$ 时, 说明 $D^l\{L\}$, $D^l\{R\}$ 之间相似度较低, 此时采用局部能量最大的图像部分作为融合图像, 如式(15)所示:

$$D^l\{F\} = \begin{cases} D^l\{L\}(x, y), E_L^l(x, y) \geq E_R^l(x, y) \\ D^l\{R\}(x, y), E_L^l(x, y) < E_R^l(x, y) \end{cases}, \quad (15)$$

当 $M^l \geq th$ 时, 此时认定两图像的相似度较高, 都存在重要特征信息, 故采用基于权值图 W 进行融合。计算如下:

$$D^l\{F\} = W(x, y) \cdot D^l\{L\}(x, y) + (1 - W(x, y)) \cdot D^l\{R\}(x, y). \quad (16)$$

通过这种融合规则对各层多尺度图像进行融合, 最后对各层的融合图像进行叠加, 得到源图像的融合图像。

3 实验与结果

为验证算法的有效性, 本部分将通过主观评价和客观评价相结合的方式对算法进行比较分析。所对比方法分别为拉普拉斯变换融合方法(LP)^[2], 基于稀疏表示与轮廓波的融合方法



图 6 采用 lytro 数据集中多聚焦图像“小孩”进行融合对比实验。(a)~(f)所采用方法依次是: CVT_SR、DTCWT、GFF、LP、NSCT_SR 以及本文方法。

Fig.6 Multi-focus image “child” in inlytro data set were used for the fusion comparison experiment. The methods used in (a) ~ (f) are CVT_SR, DTCWT, GFF, LP, NSCT_SR and Ours.

(CVT_SR)^[18], 双树复小波变换融合方法(DTCWT)^[6], 基于稀疏表示与非下采样轮廓波的融合方法(NSCT_SR)^[19]以及滚动引导滤波(GFF)^[11]的融合方法。主观评价由人眼的视觉效果作为评价指标, 具有直观、简单等优点。图 6 是采用 lytro 数据集中的一组多聚焦图像, 通过不同融合算法进行的实验效果对比图。左上部分为图中黄色框内的局部放大图像用于对比各种算法的细节纹理效果。图 6(f)为本文算法融合结果, 可以看到相比其他算法, 黄色框内的衣物对比度更高, 纹理特征更加清晰。为便于分辨与观测, 本文将图 6(d), 图 6(e)和图 6(f)分别与源图像进行差值处理, 通过归一化后取图中黄色框区域的差分图像进行对比, 如图 7 所示。

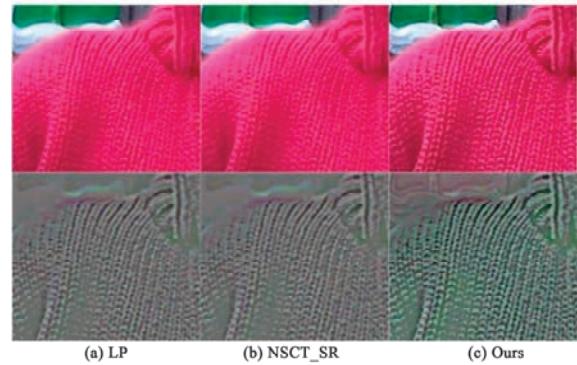


图 7 “小孩”的细节放大效果对比图

Fig.7 “Child” local details effect comprison diagram

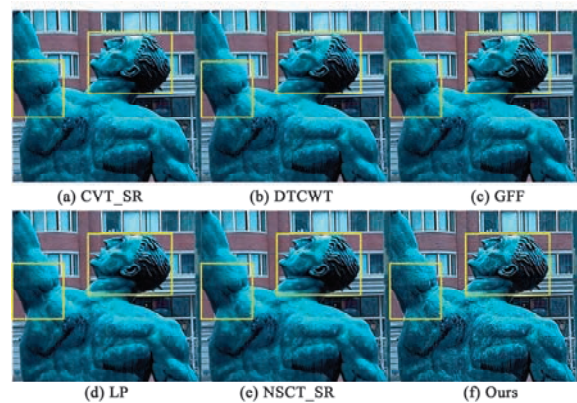


图 8 采用 lytro 数据集中多聚焦图像“雕塑”进行融合对比实验。(a)~(f)所采用方法依次是: CVT_SR、DTCWT、GFF、LP、NSCT_SR 以及本文方法。

Fig.8 Multi-focus image “sculpture” in inlytro data set were used for the fusion comparison experiment. The methods used in (a) ~ (f) are CVT_SR, DTCWT, GFF, LP, NSCT_SR and Ours.

可以看到,图 7(c)的纹理更加清晰,边缘部分没有模糊现象。图 8 也是采用 lytro 数据集中的一组多聚焦图像进行对比实验,图 9 是图 8(d),图 8(e)和图 8(f)中黄色框内的局部放大效果。经对比可以看到,图 9(c)中雕像头部与手臂的轮廓边缘清晰完整,细节特征更加丰富。

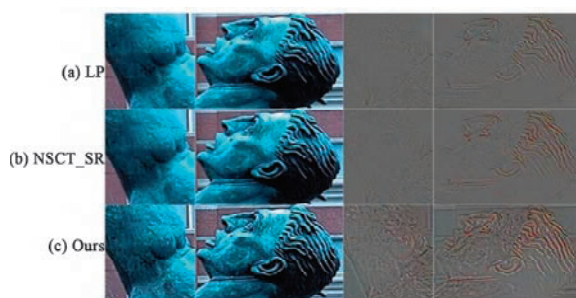


图 9 “雕塑”的细节放大效果对比图

Fig.9 “Sculpture” local details effect comparison diagram

图 10 是 TNO 数据集中一组红外可见图像的各算法实验结果。

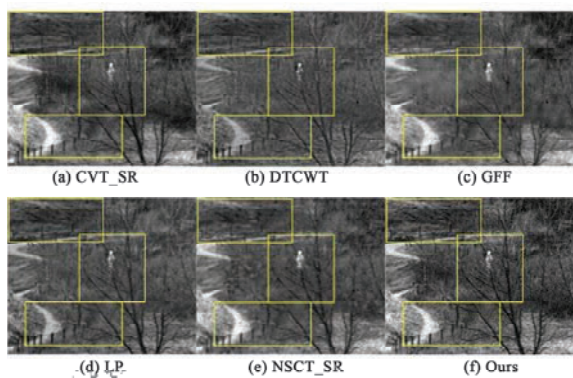


图 10 各算法全局效果对比图

Fig.10 Comparison of the global effect of each algorithm

类似地,本文在图 11 中展示了融合图像的局部放大图。可以看到,相较于其他算法效果图,图 11(c)人物目标显著,坡地草坪、树枝纹理以及木桩轮廓可分辨度更高,具有更好的视觉效果。

由于人的视觉感知存在差异,对评价结果存在一定的影响,因此需要通过构建与人眼视觉效果相近的客观评价指标模型对融合图像进行进一步的客观评价。

本文采用 6 种重要的评价指标对算法进行测试。

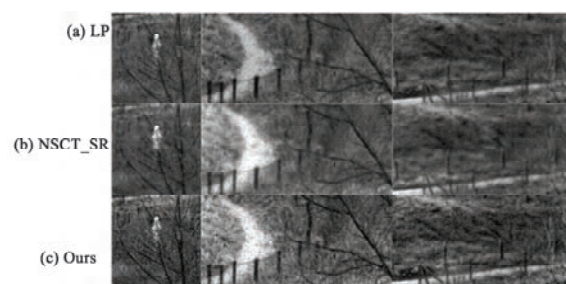


图 11 融合图像的细节放大效果对比图

Fig. 11 Comparison of details enlarged of the fusion images

这 6 种评价指标分别是:平均梯度(AG)、 $Q^{AB/F}$ 度量($Q^{AB/F}$)、空间频率(SF)、信息熵(EN)、结构相似性(SSIM)以及视觉信息保真度(VIFF)。其中 AG 也称作图像锐度值,能够衡量图像细节纹理特征清晰度,其值越高代表图像细节纹理越清晰。 $Q^{AB/F}$ 利用局部度量来估计输入图像的显著信息在融合图像中的表现程度,能够反映融合图像从输入图像中获得的视觉信息的质量。SF 能够度量图像空间域的总活跃度,可反映图像对微小细节反差表达的能力。EN 是用于测量融合图像中所包含信息量的重要指标。SSIM 通过计算融合图像与参考图像之间的结构相关信息来判断结构的相似度。VIFF 通过模拟人类视觉系统原理从通信和信息共享的角度评判图像质量,具有图像视觉效果的判定能力,其值越大,图像的保真度越高。

表 2、表 3 分别是多聚焦图像集和红外图像集在各种算法下所得客观指标的均值表。表 2 采用 lytro 数据集中的 17 对多聚焦图像作为源图像数据进行测试,而表 3 采用 TNO 数据集和红外图像数据集中的 29 对红外可见光图像作为源图像数据进行测试。由表 2 和表 3 可见,本文算法在 MG、SF、IE 和 VIFF 上具有明显优势,代表着本文算法在细节纹理分辨度和视觉效果保真度方面效果更佳。但在 SSIM 上略低于其他算法,究其原因在于进行多尺度非线性增强使图像的照度、对比度以及图像结构部分得到一定增强。用源图像作为参考图像时,由于结构发生一定变化,使得此项指标低于其他算法结果。综合主观评价与客观评价,本文算法在细节纹理、轮廓清晰以及视觉保真效果上具有明显优势。

表 2 多聚焦融合图像客观指标比较

Tab.2 Objective index comparison of multi-focus fusion images

指标	算法					
	CVT_SR	DTCWT	GFF	LP	NSCT_SR	Ours
AG	7.892 3	7.913 9	7.922 6	7.957 3	7.932 6	11.641 5
$Q^{AB/F}$	0.710 9	0.721 9	0.724 5	0.718 2	0.691 7	0.721 3
SF	19.047 2	19.084 4	19.138 3	19.151 2	19.109 2	25.960 8
EN	7.503 0	7.504 9	7.504 0	7.515 4	7.503 9	7.547 1
SSIM	0.825 1	0.827 4	0.848 6	0.832 1	0.833 8	0.815 9
VIFF	0.928 6	0.919 6	0.939 2	0.956 8	0.935 9	0.974 6

表 3 红外可见光融合图像客观指标比较

Tab.3 Comparison of objective indexes of infrared visible light fusion images

指标	算法					
	CVT_SR	DTCWT	GFF	LP	NSCT_SR	Ours
AG	6.132 5	5.952 3	5.931 8	6.178 3	6.899 6	9.661 5
$Q^{AB/F}$	0.485 0	0.528 3	0.557 6	0.579 7	0.516 0	0.571 4
SF	12.347 4	12.133 7	12.191 4	12.484 8	13.537 1	18.513 0
EN	7.125 2	6.865 5	7.063 8	6.798 6	7.130 0	7.157 1
SSIM	0.481 2	0.550 7	0.547 8	0.552 3	0.438 3	0.481 0
VIFF	0.461 6	0.381 0	0.344 9	0.483 9	0.385 6	0.507 4

4 结 论

本文提出一种基于多尺度细节的卷积神经网络图像融合算法。通过孪生网络提取图像特征构建权值图,并通过改进的滚动引导滤波对源图像进行自适应多尺度分解,并进一步使用

映射函数增强了图像细节,最后基于局部能量原理与权值图,设计自适应调整融合模式,将多尺度特征图进行融合。该方法能够更好地突出细节纹理特征,提高视觉感知特性。实验结果表明,所提方法融合结果具有更加丰富的细节特征,避免了图像出现圆晕效应,更符合人类的视觉感知系统。

参 考 文 献:

[1] 武治国,王延杰.一种基于边缘的曲波变换图像融合方法[J].光学技术,2009,35(5):682-685,690.
WU Z G, WANG Y J. Image fusion algorithm using Curvelet transform based on the edge detection [J]. *Optical Technique*, 2009, 35(5): 682-685, 690. (in Chinese)

[2] TOET A. Image fusion by a ratio of low-pass pyramid [J]. *Pattern Recognition Letters*, 1989, 9(4): 245-253.

[3] LI M J, ZHOU J F, LI J. Research on image fusion rules based on contrast pyramid [C]//2018 Chinese Automation Congress. Xian, China: IEEE, 2018: 1361-1364.

[4] 李建林,俞建成,孙胜利.基于梯度金字塔图像融合的研究[J].科学技术与工程,2007,7(22):5818-5822.
LI J L, YU J C, SUN S L. Study of image fusion based on grad pyramid algorithm [J]. *Science Technology and Engineering*, 2007, 7(22): 5818-5822. (in Chinese)

[5] 万智萍.结合方向小波的多光谱与全色遥感图像融合算法[J].遥感技术与应用,2014,29(4):660-668.

- WAN Z P. Multi-spectral and panchromatic remote sensing image fusion algorithm based on directional wavelet [J]. *Remote Sensing Technology and Application*, 2014, 29(4): 660-668. (in Chinese)
- [6] 陈清江,魏冰蔗,柴昱洲,等.非抽样双树复小波域的 BPP-NMF 图像融合[J].液晶与显示,2016,31(8):784-792.
CHEN Q H, WEI B Z, CHAI Y Z, *et al.* Image fusion based on block principal pivoting nonnegative matrix in undecimated dual-tree complex wavelet domain [J]. *Chinese Journal of Liquid Crystals and Displays*, 2016, 31(8): 784-792. (in Chinese)
- [7] 殷明,段普宏,褚标,等.基于非下采样双树复轮廓波变换和稀疏表示的红外和可见光图像融合[J].光学 精密工程, 2016,24(7):1763-1771.
YIN M, DUAN P H, CHU B, *et al.* Fusion of infrared and visible images combined with NSDCT and sparse representation [J]. *Optics and Precision Engineering*, 2016, 24(7): 1763-1771. (in Chinese)
- [8] 陈贞,邢笑雪.基于非下采样 Shearlet 变换与压缩感知的图像融合[J].液晶与显示,2015,30(6):1024-1031.
CHEN Z, XING X X. Image fusion algorithm based on non-subsampled Shearlet transform and compressed sensing [J]. *Chinese Journal of Liquid Crystals and Displays*, 2015, 30(6): 1024-1031. (in Chinese)
- [9] DA CUNHA A L, ZHOU J, DO M N. The nonsubsampling contourlet transform: theory, design, and applications [J]. *IEEE Transactions on Image Processing*, 2006, 15(10): 3089-3101.
- [10] ZHAO J F, FENG H J, XU Z H, *et al.* Detail enhanced multi-source fusion using visual weight map extraction based on multi scale edge preserving decomposition [J]. *Optics Communications*, 2013, 287: 45-52.
- [11] 赵程,黄永东.基于滚动导向滤波和混合多尺度分解的红外与可见光图像融合方法[J].激光与光电子学进展, 2019,56(14):141007.
ZHAO C, HAUNG Y D. Infrared and visible image fusion via rolling guidance filtering and hybrid multi-scale decomposition [J]. *Laser & Optoelectronics Progress*, 2019, 56(14): 141007. (in Chinese)
- [12] LI H, WU X J. DenseFuse: a fusion approach to infrared and visible images [J]. *IEEE Transactions on Image Processing*, 2019, 28(5): 2614-2623.
- [13] LIANG X C, HU P Y, ZHANG L G, *et al.* MCFNet: multi-layer concatenation fusion network for medical images fusion [J]. *IEEE Sensors Journal*, 2019, 19(16): 7107-7119.
- [14] 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.
- [15] CHOPRA S, HADSELL R, LECUN Y. Learning a similarity metric discriminatively, with application to face verification [C]//*IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Diego, CA: IEEE, 2005: 539-546.
- [16] ZHANG Q, SHEN X Y, XU L, *et al.* Rolling guidance filter [C]//*13th European Conference on Computer Vision (ECCV)*. Zurich, Switzerland: Springer, 2014: 815-830.
- [17] PARIS S, HASINOFF S W, KAUTZ J. Local Laplacian filters: edge-aware image processing with a Laplacian pyramid [J]. *Communications of the ACM*, 2015, 58(3): 81-91.
- [18] HE G Q, XING S Y, HE X J, *et al.* Image fusion method based on simultaneous sparse representation with non-subsampled contourlet transform [J]. *IET Computer Vision*, 2019, 13(2): 240-248.
- [19] 赵春晖,郭蕴霆.一种快速的基于稀疏表示和非下采样轮廓波变换的图像融合算法[J].电子与信息学报,2016,38(7):1773-1780.
ZHAO C H, GUO Y T. Fast image fusion algorithm based on sparse representation and non-subsampled contourlet transform [J]. *Journal of Electronics & Information Technology*, 2016, 38(7): 1773-1780. (in Chinese)

作者简介:



刘 博(1993—),男,吉林长春人,硕士研究生,2016 年于长春大学获得学士学位,主要从事计算机视觉、图像处理与信息处理方面的研究。E-mail:lbciomp@163.com



韩广良(1968—),男,山东嘉祥人,博士,研究员,2003 年于中国科学院长春光学精密机构与物理研究所获得博士学位,研究方向为图像和视频信息处理、计算机视觉、机器视觉和人工智能等。E-mail: hangl@ciomp.ac.cn