

控制与决策

Control and Decision

基于混合模型驱动的红外与可见光图像融合

沈瑜, 陈小朋, 刘成, 张泓国, 王霖

引用本文:

沈瑜, 陈小朋, 刘成, 等. 基于混合模型驱动的红外与可见光图像融合[J]. *控制与决策*, 2021, 36(9): 2143–2151.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.1749>

您可能感兴趣的其他文章

Articles you may be interested in

小样本下多稀疏表示分类器的决策融合方法

Decision fusion of multiple sparse representation-based classifiers in case of small samples

控制与决策. 2021, 36(8): 1984–1990 <https://doi.org/10.13195/j.kzyjc.2019.1839>

基于低秩矩阵恢复的视觉显著性目标检测与细化

Saliency object detection and refinement based on low rank matrix recovery

控制与决策. 2021, 36(7): 1707–1713 <https://doi.org/10.13195/j.kzyjc.2019.1795>

基于双边分解与 L_1 暗通道的战场图像增强算法

A battlefield image enhancement algorithm based on bilateral decomposition and L_1 dark channel prior

控制与决策. 2021, 36(5): 1165–1172 <https://doi.org/10.13195/j.kzyjc.2019.1213>

融合稀疏编码与深度学习的草图特征表示

A feature representation of sketch based on fusion of sparse coding and deep learning

控制与决策. 2021, 36(3): 699–704 <https://doi.org/10.13195/j.kzyjc.2019.0941>

一种基于稀疏系数匹配学习的图像去雾算法

An image dehazing method based on learning framework with sparse coefficient matching

控制与决策. 2020, 35(11): 2797–2802 <https://doi.org/10.13195/j.kzyjc.2018.1764>

基于混合模型驱动的红外与可见光图像融合

沈瑜, 陈小朋[†], 刘成, 张泓国, 王霖

(兰州交通大学 电子与信息工程学院, 兰州 730070)

摘要: 为了解决红外与可见光图像融合中显著特征不突出、图像对比度低的问题, 提出一种混合模型驱动的融合算法. 首先, 采用潜在低秩表示模型分别提取红外与可见光图像的基础子带、显著子带及稀疏噪声子带; 其次, 采用非下采样剪切波变换模型将基础子带分解为低频系数和高频系数, 对低频系数采用字典学习和稀疏表示进行精确拟合, 对高频系数采用局部窗口结合逻辑加权进行选择; 再次, 显著子带采用区域能量比阈值自适应加权法进行融合; 最后, 对融合后的低频系数和高频系数进行一级重建, 得到融合基础子带, 舍弃稀疏噪声子带, 再结合融合显著子带进行二级重建, 得到融合图像. 实验结果表明: 所提出算法能够得到蕴含丰富信息且较为清晰的融合图像, 具有可行性; 融合结果的对比较高, 目标轮廓显著, 能够提升场景的辨识度, 具有有效性.

关键词: 图像融合; 潜在低秩表示; 非下采样剪切波变换; 区域能量比; 稀疏表示; 逻辑加权

中图分类号: TP391

文献标志码: A

DOI: 10.13195/j.kzyjc.2019.1749

开放科学(资源服务)标识码(OSID):



引用格式: 沈瑜, 陈小朋, 刘成, 等. 基于混合模型驱动的红外与可见光图像融合[J]. 控制与决策, 2021, 36(9): 2143-2151.

Infrared and visible image fusion based on hybrid model driving

SHEN Yu, CHEN Xiao-peng[†], LIU Cheng, ZHANG Hong-guo, WANG Lin

(School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: In order to solve the problems of less-prominent features and low image contrast in infrared and visible image fusion, the paper proposes a hybrid model-driven fusion algorithm. Firstly, the latent low-rank representation model is used to extract the base sub-band, significant sub-band and sparse noise sub-band of the infrared and visible image respectively. Then, the base sub-band is decomposed into low frequency coefficients and high frequency coefficients by the non-subsampled Shearlet transform model. The low frequency coefficients are accurately fitted by dictionary learning and sparse representation, while the high frequency coefficients are selected by local window combined with logic weighting. The significant sub-band is fused by adaptive weighting of regional energy ratio threshold. Finally, the fused low frequency coefficients and high frequency coefficients are 1st-order reconstructed to obtain the fused base sub-band and discard the sparse noise sub-band. Then the 2nd-order reconstruction is conducted with the fused significant sub-band to get the fused image. The experiment results show that the proposed algorithm can obtain clear fused image with rich information. The fused results have high contrast and significant target contour which can improve the recognition of the scene.

Keywords: image fusion; latent low-rank representation; non-subsampled shearlet transform; regional energy ratio; sparse representation; logic weighting

0 引言

图像融合是通过特定的方法对不同传感器从同一场景采集的两幅或多幅图像蕴含的有效信息进行提取并有机整合, 得到信息丰富且视觉效果更好的复

合图像, 以满足后续处理与研究需要. 可见光传感器采集到的图像具有较高的空间分辨率且蕴含丰富的背景信息, 但成像质量易受外界环境影响; 而红外传感器采集到的图像虽然质量较差, 但成像却较为稳

收稿日期: 2019-12-04; 修回日期: 2020-05-25.

基金项目: 国家自然科学基金项目(61861025, 61562057, 61761027, 51669010); 教育部长江学者和创新团队发展计划项目(IRT_16R36); 甘肃省教育厅高等学校科研项目(216130); 光电技术与智能控制教育部重点实验室(兰州交通大学)开放课题(KFKT2018-9); 兰州市人才创新创业项目(2018-RC-117); 兰州交通大学青年基金项目(2015005).

责任编辑: 程龙.

[†]通讯作者. E-mail: 3064683191@qq.com.

定^[1-2]. 红外与可见光图像的融合正是利用源图像在时空上的相关性及场景描述时的信息互补性,使融合图像对场景的描述更加细致、全面,从而更有利于人眼的识别和机器的自动检测,现已在视频监控^[3]、目标检测^[4]、脸部识别等多个方面得到广泛应用.

红外与可见光图像的融合一般为像素级融合,分为基于数据驱动与基于模型驱动两类,前者通过操作源图像的像素值直接对图像进行融合,后者通过模型变换间接实现图像的融合^[5]. 基于模型驱动的多尺度几何分析法因其独特的多尺度分析特性已得到广泛的研究与应用. 冯鑫^[6]对Tetrolet分解后的低频子带采用活动水平指导稀疏系数进行精确拟合,高频子带根据脉冲耦合神经网络中神经元点火次数进行选择,能够有效保留源图像的边缘信息和细节特征. Zhang等^[7]对非下采样轮廓波变换分解后的低频子带采用区域能量自适应准则进行融合,对高频子带中的最高频系数采用绝对值取大法进行融合,其余子带采用自适应高斯区域标准差进行融合,融合图像的目标更为突出且对比度有所提升. 王峰等^[8]对剪切波变换分解后的低频子带采用标准差与物理特征相结合的加权法进行融合,对高频子带采用局部标准差加权与绝对值取大法进行融合,融合图像清晰度较高且纹理信息丰富. Liu等^[9]对非下采样剪切波变换分解后的低频子带采用导频滤波算法计算其显著映射,对高频子带采用相位一致性与场景一致性构造权重系数进行融合,能够较好地保留源图像的信息且融合图像更加平滑.

随着稀疏表示理论的不完善,其通过学习获取到的稀疏表示能力对图像融合效果的改善也越来越显著. Liu等^[10]通过Two-scale法与作差法得到图像基础子带和细节子带,对不同的基础子带采用取大法及均值法进行融合,对细节子带采用卷积稀疏表示进行编码融合,克服了稀疏表示融合中细节保存能力有限的缺点. 常莉红等^[11]对四元数小波变换分解后的低频子带采用联合稀疏表示进行融合,对高频子带采用绝对值取大法进行融合,克服了融合中边界处理过于平滑及耗时过长的问題. 王昕等^[12]通过显著性检测指导区域生长法对图像进行目标分割,对目标区域的高、低频子带分别采用替代法、模值取大法融合,对背景区域的高、低频子带分别采用多奇异值分解与离散傅里叶变换进行融合,解决了图像目标信息弱化且易受噪声干扰的问题.

目前,红外与可见光图像融合仍存在以下问题:

1) 红外与可见光传感器的成像过程中普遍存在噪声,

影响融合结果;2) 常见的多尺度分析方法的变换过程过于复杂且分解方向数目有限,无法进一步实现图像的稀疏表示;3) 现有的融合方法存在图像显著性特征不突出的问题,融合图像不符合人眼视觉的观察习惯,融合效果有待提升.

针对上述问题,本文提出一种潜在低秩表示模型与非下采样剪切波变换模型相结合的红外与可见光图像的融合算法. 现有的图像融合算法大多基于两种模型中的一种对图像进行分解,进而设计融合规则以实现图像的融合. 本文根据两种模型各自的特点(即非下采样剪切波变换能够对源图像存在的空间校正误差进行一定程度的修正,且对于剪切后的方向数没有限制,逆变换不需要合成方向滤波器,能够减小融合误差,降低运算的复杂度;潜在低秩表示能够通过隐藏的数据信息获取图像的整体结构与部分结构,提取图像的特征信息),将两种框架进行嵌套,实现图像的多层次分解,不仅能够得到源图像的稀疏表示,而且可以获得源图像的低秩特征;通过增加掩模卷积设计融合规则,得到高质量的融合图像. 为了保留源图像基础子带低频所包含的背景信息,基于稀疏表示向量的元素和来设计融合规则;为了最大限度保留基础子带高频中的边缘轮廓,提出基于局部窗口最大滤波的逻辑加权法对其进行融合;为了突出显著子带中的显著特征与区域,提出基于区域能量比的阈值自适应加权融合规则;为了降低噪声对融合图像的干扰、实现更好的融合效果,将多层次分解后的稀疏噪声子带舍弃,达到简单滤波的目的.

1 图像分解模型

1.1 潜在低秩分解

低秩表示(low-rank representation, LRR)理论是在确定学习字典的情况下,将原始数据矩阵 Q 表示为字典矩阵 A 下的线性组合且表示系数矩阵低秩,以分割干净数据与噪声. LRR的数学模型为

$$\min_Z \|Z\|_*; \text{ s.t. } Q = AZ. \quad (1)$$

其中: $\|\cdot\|_*$ 为核范数, Z 为原始数据最优的低秩表示矩阵. 一般选择原始数据本身作为字典,即 $A = Q$.

虽然LRR可以表示数据的整体结构,但是却无法保留图像的局部结构信息. 2011年,Liu等^[13]在LRR的基础上提出了潜在低秩表示(latent low-rank representation, LatLRR). LatLRR不仅继承了低秩表示的优良特性,而且考虑隐藏的数据信息对学习字典的影响,可以提取图像的全局结构和局部结构,从而具有更强的特征信息提取能力与去噪能力. LatLRR

的数学模型表示为

$$\begin{aligned} \min_{Z,L,E} & \|Z\|_* + \|L\|_* + \lambda\|E\|_1; \\ \text{s.t.} & Q = QZ + LQ + E. \end{aligned} \quad (2)$$

其中: $\lambda > 0$ 为平衡系数, $\|\cdot\|_1$ 为 L_1 范数, Q 与 Z 的意义与式(1)一致, L 为显著性系数矩阵, E 为稀疏噪声. 式(2)可以看作带有核范数的凸优化问题, 可由增广拉格朗日乘子(ALM)求解, 得到原始数据矩阵的低秩部分 QZ 、显著部分 LQ 及稀疏噪声 E . 当 LatLRR 用于图像分解时, QZ 、 LQ 与 E 分别对应图像的基础子带、显著子带及稀疏噪声子带. 以“Camp”红外与可见光图像对为例, LatLRR 的分解效果如图1所示.

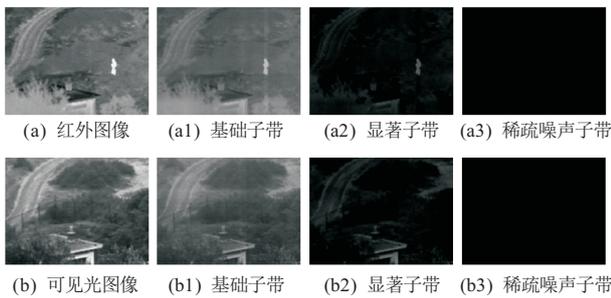


图1 LatLRR分解图示

1.2 非下采样剪切波变换

为了稀疏表示图像, 出现了具有多尺度性、局部性和方向性等多种优良特性的剪切波变换, 通过仿射理论的应用将多尺度分析与几何分析巧妙地结合在一起, 但是该变换过程中涉及抽取操作, 因而具有移变性, 进行图像处理时容易造成吉布斯失真.

非下采样剪切波变换 (non-subsampled shearlet transform, NSST)^[14] 的出现有效弥补了该不足, 不仅继承了剪切波变换的优良特性, 而且具备较好的平移不变性, 能够更加有效地提取图像的边缘细节信息. 此外, 在方向局部化的过程中, 可通过选择长宽比可变的 Meyer 窗来自定义分解方向数量, 打破了分解方向数目的局限性. NSST 对图像的分解过程分为多尺度分解和方向局部化两部分: 多尺度分解的对象为源图像, 通过非下采样塔式滤波器组实现; 方向局部化的对象为源图像经多尺度分解后的低频子带, 通过剪切波滤波器实现.

2 图像融合

设待融合的红外图像 I 与可见光源图像 V , 经 LatLRR 分解得到显著子带 $\{S_I, S_V\}$ 、基础子带 $\{B_I, B_V\}$ 及稀疏噪声子带. 显著子带融合后为 S_F ; 基础子带经 j 级 NSST 分解得到其低频系数 $\{L_I^j, L_V^j\}$ 与高频系数 $\{H_I^j, H_V^j\}$, 分别融合得到 L_F^j 与 H_F^j , 逆 NSST 变

换得到融合后的基础子带 B_F ; 舍弃稀疏噪声子带后 LatLRR 重建得到融合图像 F . 算法流程如图2所示.

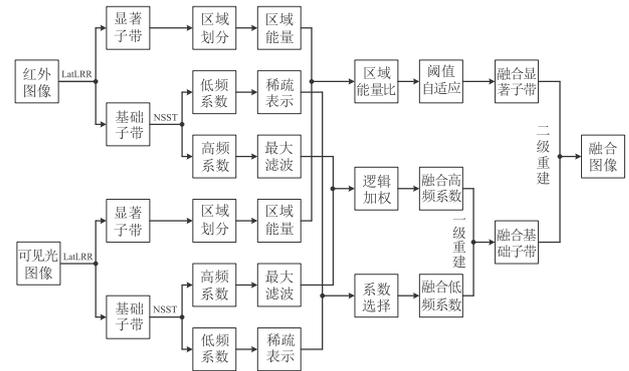


图2 算法流程

2.1 显著子带融合

传统的加权融合虽然能够得到低噪声且较为稳定的融合子带, 但是待融合子带的本身特征 (如统计特征和幅度特征) 均为随机分布且对加权系数有直接的影响, 简单的加权融合并不能完美综合二者的本身特征. 基于区域的融合方法虽然在一定程度上能够改善此种状况, 但是在基于区域的梯度、方差及能量3种方法中, 仅有基于区域能量的融合方法考虑到局部特征由该区域内多个像素共同表征且同一区域内的点具有较强的相关性, 能够进一步体现图像的特征^[15]. 因此, 本文提出一种新的基于区域能量比的阈值自适应加权融合规则, 根据区域中心像素及对应区域能量的不断变化, 结合阈值调整加权系数自适应的变化, 达到融合图像的本身特征并充分保留细节的目的, 详细过程如下.

首先, 对 LatLRR 分解后的显著子带系数 S_I 和 S_V , 分别求得以像素 (m, n) 为中心的局部区域能量 $E_I(m, n)$ 和 $E_V(m, n)$, 计算公式如下:

$$\begin{aligned} E_I(m, n) &= \sum_{m' \in X, n' \in Y} \omega \times [S_I(m + m', n + n')]^2, \\ E_V(m, n) &= \sum_{m' \in X, n' \in Y} \omega \times [S_V(m + m', n + n')]^2. \end{aligned} \quad (3)$$

其中: X, Y 表示局部区域窗口的最大行、列坐标, 局部区域窗口大小一般取 3×3 ; ω 为窗口系数, 取

$$\omega = \frac{1}{16} \times \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}.$$

然后, 根据区域能量计算区域能量比 $E_{\text{ratio}}(m, n)$, 即

$$E_{\text{ratio}}(m, n) = \frac{E_I(m, n)}{E_V(m, n)}. \quad (4)$$

最后,通过加权法计算 (m, n) 处的融合系数 $S_F(m, n)$, 即

$$S_F(m, n) = w_1 \times S_I(m, n) + w_2 \times S_V(m, n). \quad (5)$$

其中: w_1, w_2 分别为红外与可见光显著子带在中心像素 (m, n) 处的加权系数, 计算公式如下:

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{cases} [1 \ 0]^T, & E_{\text{ratio}} < \text{th}_1; \\ \begin{bmatrix} E_I/(E_I + E_V) \\ E_V/(E_I + E_V) \end{bmatrix}, & \text{th}_1 < E_{\text{ratio}} < \text{th}_2; \\ [0 \ 1]^T, & E_{\text{ratio}} > \text{th}_2. \end{cases} \quad (6)$$

这里: th_1, th_2 为阈值系数, 根据图像的整体能量分布确定. 经过遍历计算得到融合后的显著子带 S_F . 由式 (5) 和 (6) 可知: 若区域能量比过小或过大, 则意味着该区域对应的两个能量值差异较大, 此时将该区域能量较大者对应的权值设为 1, 较小的设为 0; 若区域能量比在阈值区间内, 则说明该区域的两个能量值较为接近, 此时根据能量占比自适应地计算权重系数. 即区域能量越大, 与之相对应的加权系数越大, 在融合结果中占的比重越高; 反之能量越小, 对融合结果的贡献也越小.

2.2 基础子带低频系数融合

字典学习是稀疏表示的基础, 通过字典中的原子可以对原始样本数据进行稀疏线性表示, 因此, 字典决定了稀疏编码对原始数据的表示能力. 过完备字典的获取方法包括基于分析模型和基于机器学习两种, 前者得到的字典虽然结构良好且计算速度较快, 但是比较依赖于图像的几何特征, 具有一定的局限性; 而后者通过将图像进行分割、训练, 对图像块进行学习得到过完备字典^[16], 不仅能够捕捉图像的几何特征, 而且具有较好的适应性, 因此本文通过训练法获取过完备字典.

以过完备字典 D 为基础, 对基础子带的低频系数 $\{L_I^j, L_V^j\}$ 进行稀疏表示融合, 具体步骤如下.

首先, 采用滑动窗口将低频系数按顺序划分成大小为 $\sqrt{d} \times \sqrt{d}$ 的图像块, 滑动步长为 s , 分别得到 M 个图像块 $\{P_I^i, P_V^i\} (i = 1, 2, \dots, M)$.

其次, 将图像块的元素重新排列成列向量 $\{C_I^i, C_V^i\}$, 将其均值归一化得以更好地提取图像块的结构信息, 计算公式如下:

$$\begin{aligned} \hat{C}_I^i &= C_I^i - \bar{C}_I^i \cdot W, \\ \hat{C}_V^i &= C_V^i - \bar{C}_V^i \cdot W. \end{aligned} \quad (7)$$

其中: W 为元素值均为 1 的 $d \times 1$ 维向量, \bar{C}_I^i 和 \bar{C}_V^i 为

列向量 C_I^i, C_V^i 的平均值.

再次, 通过 OMP (orthogonal matching pursuit) 算法和过完备字典计算 $\{\hat{C}_I^i, \hat{C}_V^i\}$ 的稀疏表示向量, 计算公式如下:

$$\begin{aligned} \zeta_I^i &= \arg \min_{\zeta} \|\zeta\|_0, \\ \text{s.t. } & \|\hat{C}_I^i - D\zeta\|_2 < \varepsilon; \\ \zeta_V^i &= \arg \min_{\zeta} \|\zeta\|_0, \\ \text{s.t. } & \|\hat{C}_V^i - D\zeta\|_2 < \varepsilon. \end{aligned} \quad (8)$$

接着, 对稀疏表示向量 $\{\zeta_I^i, \zeta_V^i\}$ 的元素求累和, 依据和的大小获得融合稀疏表示向量 ζ_F^i , 计算公式如下:

$$\zeta_F^i = \begin{cases} \zeta_I^i, & \sum_{l=1}^n |\zeta_I^i[l]| > \sum_{l=1}^n |\zeta_V^i[l]|; \\ \zeta_V^i, & \sum_{l=1}^n |\zeta_I^i[l]| < \sum_{l=1}^n |\zeta_V^i[l]|. \end{cases} \quad (9)$$

最后, 根据融合后的稀疏表示向量 ζ_F^i 计算原始列向量的融合结果 C_F^i , 计算公式如下:

$$C_F^i = \begin{cases} D\zeta_F^i + \bar{C}_I^i \cdot W, & \zeta_F^i = \zeta_I^i; \\ D\zeta_F^i + \bar{C}_V^i \cdot W, & \zeta_F^i = \zeta_V^i. \end{cases} \quad (10)$$

对所有的图像块重复上述步骤, 得到融合后的向量 C_F^i , 然后将所有的列向量恢复为图像块 P_F^i 并按原位置填充得到融合后的基础子带低频系数 L_F^j . 图像块填充时, 重叠部分的像素取累计平均.

2.3 基础子带高频系数融合

图像的低频系数反映了图像块的亮度, 高频系数代表了图像的细节、边缘, 反映了图像块的纹理特性. 对基础子带的高频系数而言, 两者的融合主要是为了得到清晰的边缘区域及丰富的纹理信息, 以提高融合图像的清晰度. 由于红外与可见光图像的高频部分差异较大, 高频系数的融合策略应以增强融合图像的细节信息为目的. 本文提出一种基于最大滤波的逻辑加权法对基础子带的高频系数进行融合, 具体步骤如下.

首先, 对高频系数 $\{H_I^j, H_V^j\}$ 取绝对值, 采用尺寸为 3×3 的局部窗口 w 进行最大值滤波, 数学表达为

$$\begin{aligned} Z_I^j &= F_{\max}(|H_I^j|), \\ Z_V^j &= F_{\max}(|H_V^j|), \end{aligned} \quad (11)$$

其中 $F_{\max}(\cdot)$ 为最大滤波函数.

其次, 通过比较位置 (\tilde{m}, \tilde{n}) 处 $Z_I^j(\tilde{m}, \tilde{n})$ 、 $Z_V^j(\tilde{m}, \tilde{n})$ 的差值 $U(\tilde{m}, \tilde{n})$ 与阈值 T 得到逻辑图, 计算公式如下:

$$\text{map}(\tilde{m}, \tilde{n}) = \begin{cases} 1, & U(\tilde{m}, \tilde{n}) > T; \\ 0, & U(\tilde{m}, \tilde{n}) \leq T. \end{cases} \quad (12)$$

其中 $U(\tilde{m}, \tilde{n}) = |Z_I^j(\tilde{m}, \tilde{n}) - Z_V^j(\tilde{m}, \tilde{n})|$.

最后, 根据逻辑图进行融合得到对应的高频系数, 计算公式如下:

$$H_F^j(\tilde{m}, \tilde{n}) = \text{map}(\tilde{m}, \tilde{n}) \times H_I^j(\tilde{m}, \tilde{n}) + [\sim \text{map}(\tilde{m}, \tilde{n})] \times H_V^j(\tilde{m}, \tilde{n}). \quad (13)$$

其中: $\text{map}(\tilde{m}, \tilde{n})$ 的元素为逻辑值 0 或 1, \sim 为逻辑值取反运算符. 从式(12)和(13)可知, 高频系数经过最大滤波, 如果对应位置的系数差大于预设的阈值, 则逻辑图中该点所对应的逻辑值为 1, 反之为 0. 进行加权融合时, 特征相对明显的点对应逻辑权重 1, 与之对应点的权重随之取反, 达到保留细节的目的.

3 实验评估

3.1 参数设置

为验证本文算法的可行性与有效性, 从 TNO image fusion dataset 中选取红外与可见光图像进行融合实验. 实验仿真平台采用配置为 Intel Core i7-9700K CPU, 主频 3.6 GHz, 运行内存 64 GB, 搭载 64 位 Windows 10 系统的台式机, 编程环境为 Matlab 2018b. 本文算法的参数设置如下: LatLRR 的分解级

数为 1; NSST 分解级数为 4, 滤波器参数为 “maxflat”, 分解方向数为 [8 8 16 16]; 区域能量比阈值系数 $\text{th}_1 = 2, \text{th}_2 = 4$; 稀疏表示的图像块大小为 8×8 , 滑动步长 $s = 1$, 误差容限 $\varepsilon = 0.1$, 阈值 $T = 4$.

3.2 可行性评估

可行性评估选取 Camp 序列中的 20 组红外与可见光图像进行实验, 并将本文算法的融合结果与文献 [10]、文献 [17] 及文献 [18] 的融合结果进行对比. 通过观察融合后的图像效果, 并计算源图像与融合图像的熵 (EN)、平均梯度 (AvG)、空间频率 (SF) 及标准差 (Std) 四种无参考图像质量评价指标进行综合评估. EN 用于衡量图像中所包含的信息量, AvG 用于衡量图像的层次性, SF 用于衡量图像的灰度变化率, Std 用于衡量图像像素灰度值的离散程度. 以上 4 种均为正向指标, 其值越大越好.

Camp 序列 20 组红外与可见光图像的融合结果如图 3 所示. 其中: 图 (a) 和图 (g) 为红外图像, 用于捕捉场景中的热源物体; 图 (b) 和图 (h) 为可见光图像, 用于捕捉场景中的房屋、围栏及树木等细节信息; 图 (c) 和图 (i)、图 (d) 和图 (j)、图 (e) 和图 (k) 分别为文献 [10]、文献 [17] 和文献 [18] 的融合图像; 图 (f) 和图 (l) 为本文所提出算法的融合图像. 从融合结果来看, 相

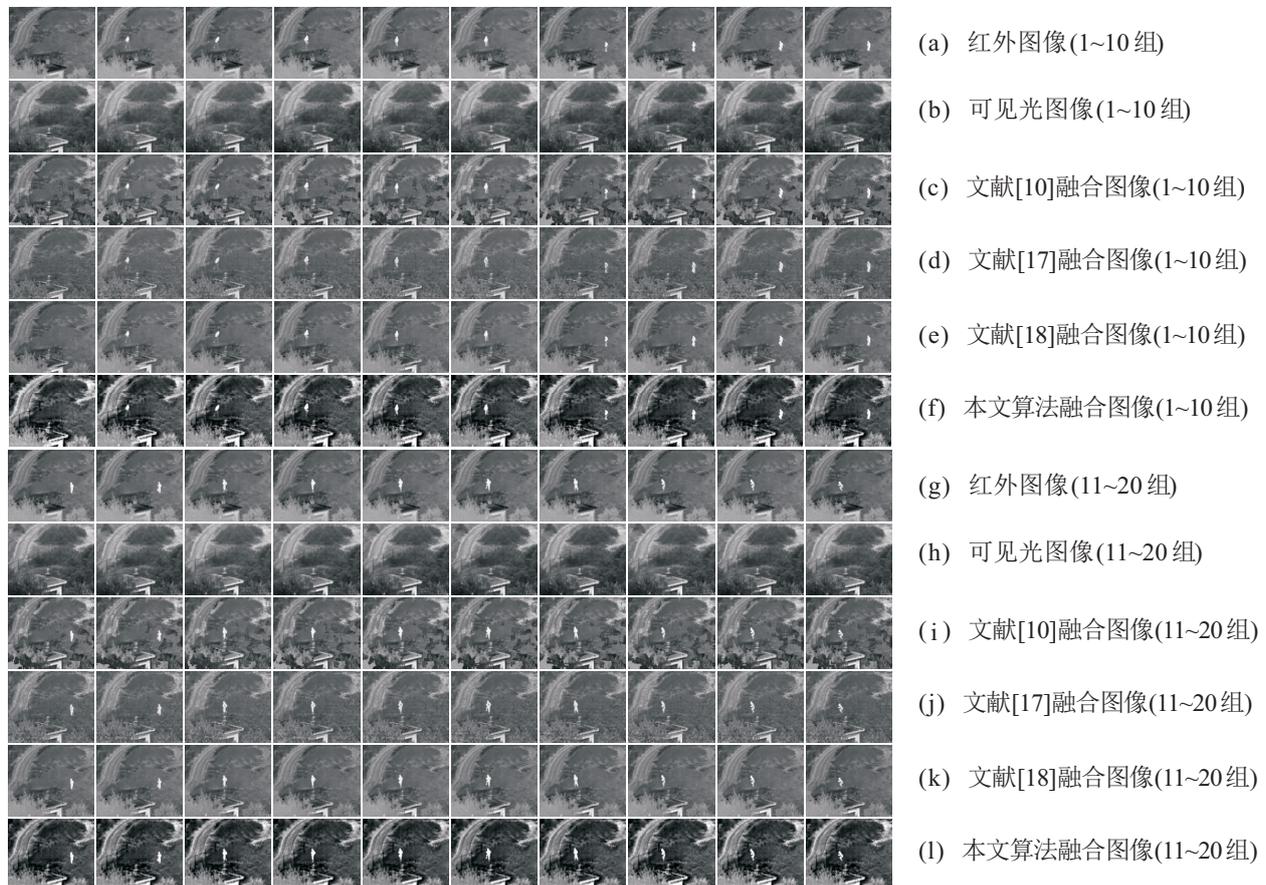


图 3 图像序列融合

较于对比算法融合图像中出现的融合不完全、受噪声干扰严重等明显不足,本文的融合算法能够有效地将红外图像中热源物体的轮廓特征与可见光图像的背景细节信息进行整合,使得融合后的图像清晰度更高,对比度也有所增强,更符合人眼视觉的观察效果。

Camp序列的源图像与融合图像的4种指标如图4所示.从曲线图上可以看出:由于整个序列图像的场景变化不大,单一图像指标的曲线并无较大波动;4种算法的融合图像与源图像相比,在4种指标上均有

所领先,表明这几种融合算法确实能够起到整合源图像信息的作用,使得场景信息更加丰富;而通过对不同算法的曲线可以看出,本文算法在4种指标上均具有较大的优势,说明本文算法的融合图像对比度较高,所含信息量也更加丰富,整体视觉效果更好,更有利于场景的认知,与人眼主观的视觉评价结论一致.综上所述,本文算法能够综合红外与可见光图像中的有效信息进而实现图像的融合,且融合图像视觉效果较好,能够进一步提高场景的辨识度,切实可行。

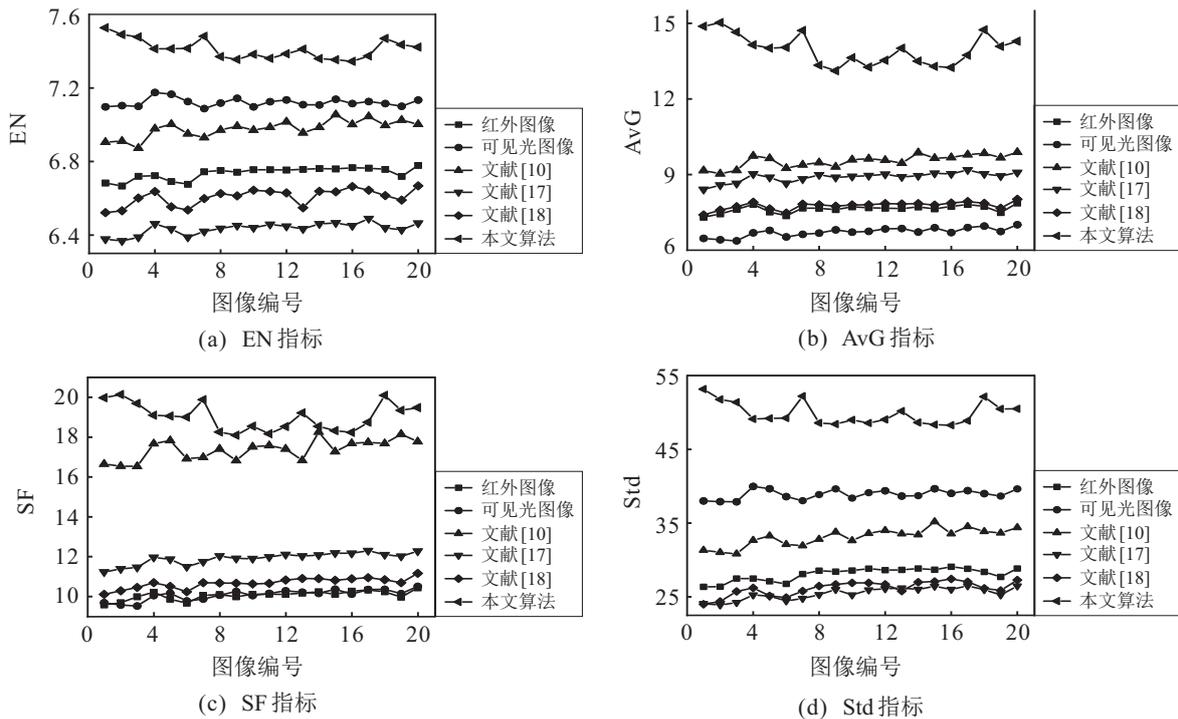


图4 Camp图像序列的客观指标

3.3 有效性分析

选取3组不同的红外与可见光图像,将本文算法的融合图像与稀疏表示(SR)、卷积稀疏表示(CSR)^[10]、交叉双边滤波(CBF)^[17]、引导滤波(GF)^[18]、联合稀疏表示(JSR)^[19]、显著性检测与联合稀疏表示(JSRSD)^[20]、轮廓波变换(Contourlet)、非下采样轮廓波变换(NSCT)、非下采样剪切波变换(NSST)及双树复小波变换(DTCWT)^[21]、Curvelet变换^[21]等多种方法的融合结果进行对比,从主观和客观两个方面进行全方位评价.对比SR的参数设置与本文算法一致,SSR、JSR、JSRSD、GF、CBF、DTCWT及Curvelet的参数设置均与原文献保持一致. Counterlet、NSCT与NSST三种变换域方法的分解层数均为4,其分解方向数分别为[4, 8, 8, 16]、[4, 8, 8, 16]和[8, 8, 16, 16],其滤波器参数分别为“pkva”“pyrexc”“vk”和“maxflat”。

主观评价根据人眼视觉对融合结果的优劣进行直观判断.3组红外与可见光图像的融合结果分别如图5所示.其中:(a)为红外图像,(b)为可见光图像,(c)~(m)为各种对比方法的融合结果,(n)为本文算法的融合结果.通过对比可以看出:SR、CSR、JSR和JSRSD四种单一的基于稀疏表示的融合方法与基于滤波的CBF方法所得到的图像或多或少地会出现融合不完全现象,使得场景信息出现不同程度的丢失,视觉效果并不理想;基于变换域的Contourlet方法得到的融合图像整体偏暗,场景辨识度较低;NSCT、NSST和Curvelet方法的融合结果对热源物体的表现不够突出,观察效果不佳;相对而言,GF和DTCWT方法的融合图像视觉效果相对较好,能够较为全面地描绘场景信息,但是与本文算法相比,融合图像在显著特征与细节轮廓的表达上稍逊一筹.综合3组图像的主观评价,本文算法的融合图像在视觉效果上更好,

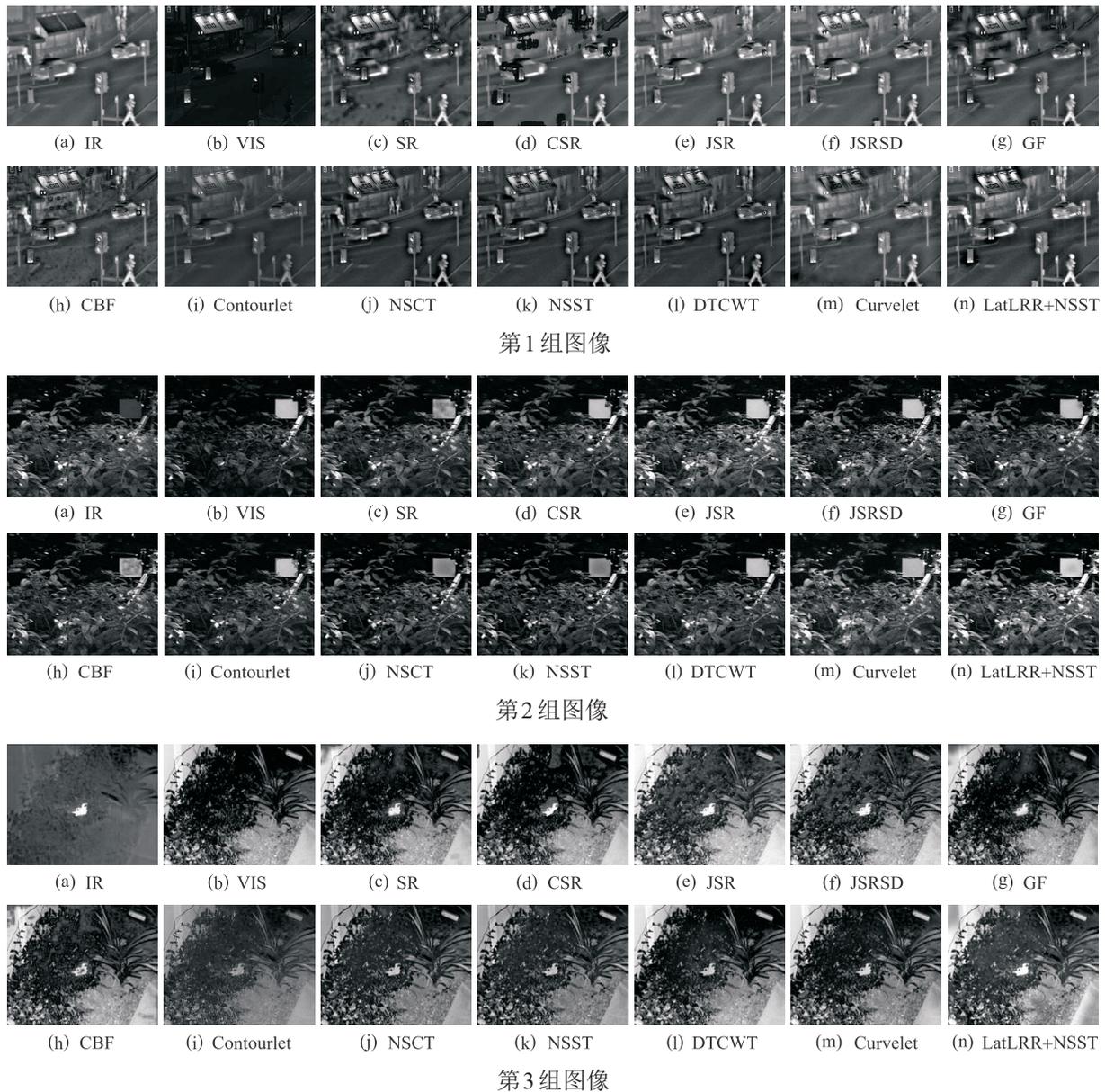


图5 3组图像的融合对比

更加符合人眼视觉系统的观察习惯。

不同的人视觉敏感度存在差异,故主观评判结果具有片面性,应综合客观评价指标进行综合评判.在原有4种指标的基础上,另计算峰值信噪比(PSNR)、相关差异和(SCD)^[22]及结构相似度(SSIM)三种正向评价指标.其中:PSNR用于衡量融合图像在融合过程中的失真程度,SCD用于衡量融合图像与源图像总的相关性,SSIM用于衡量融合图像与源图像的结构相似度.如表1所示.

从表1可以看出:本文算法在SF、AvG、Std和EN四种无参考指标方面表现良好,相比于大部分对比方法均有不同程度的领先,表明融合图像信息量丰富,整体层次鲜明且清晰度较高;SCD指标较之对比方法有较大的领先,表明从源图像转移到融合图像的

信息量较大;SSIM指标的领先幅度不大,说明所有的融合图像与源图像的整体结构大体相似;而PSNR指标领先幅度较小,甚至出现低于对比方法的现象,是因为该指标由两幅图像所对应像素点间的误差决定,并未考虑到人眼的视觉特性,而且人眼对误差的敏感度并不绝对,因此会出现客观评价结果与主观感觉不一致的情况.综上所述,客观评价的结果与主观评价大体一致,本文算法较之对比算法具有一定优势,切实有效.

4 结论

本文提出了一种基于LatLRR与NSST混合驱动的红外与可见光图像融合算法,采用LatLRR能够提取图像特征并进行简单滤波,NSST的多尺度、多方

表1 图像融合指标

组别	方法与指标	SF	AvG	Std	EN	PSNR	SCD	SSIM
第1组	SR	19.861 2	11.045 6	36.613 3	6.814 8	15.056 6	1.195 4	0.526 6
	CSR	22.836 7	10.907 3	41.067 2	6.996 2	14.492 7	1.119 3	0.542 4
	JSR	18.075 9	9.931 7	36.477 6	6.804 3	16.017 0	1.666 5	0.583 1
	JSRSD	18.075 9	9.931 7	36.477 6	6.804 3	16.017 0	1.666 5	0.583 1
	GF	18.927 6	10.034 9	37.265 6	6.833 3	14.914 0	1.150 5	0.565 2
	CBF	20.213 0	12.016 9	36.243 5	6.709 7	15.636 2	1.346 4	0.508 9
	Contourlet	13.805 8	6.580 1	21.736 1	5.952 4	16.797 8	1.462 8	0.616 1
	NSCT	18.643 2	10.003 2	31.653 0	6.485 3	16.328 5	1.722 9	0.601 8
	NSST	18.414 1	9.947 5	31.129 7	6.457 3	16.350 0	1.704 2	0.592 9
	DTCWT	18.358 2	9.795 4	30.410 3	6.436 6	16.331 9	1.642 2	0.593 0
	Curvelet	17.874 5	9.944 7	36.438 5	6.867 3	16.812 6	1.398 8	0.577 8
proposed	21.605 5	11.830 4	43.191 5	6.954 3	16.846 3	1.734 1	0.593 4	
第2组	SR	22.728 9	14.119 3	40.154 2	6.662 4	21.030 7	1.319 3	0.810 9
	CSR	23.629 0	14.470 7	43.095 1	6.741 3	19.849 5	1.352 2	0.780 4
	JSR	29.415 3	19.737 3	50.598 8	7.096 7	17.655 1	1.711 1	0.688 9
	JSRSD	29.093 9	19.718 1	46.932 1	6.878 0	18.242 7	1.201 6	0.609 4
	GF	23.088 8	14.531 5	42.487 4	6.758 9	20.224 7	1.410 0	0.802 7
	CBF	22.493 5	14.338 5	38.291 2	6.655 5	21.332 3	1.208 4	0.805 8
	Contourlet	22.250 6	13.312 1	38.890 1	6.506 6	21.250 6	1.232 2	0.835 8
	NSCT	23.956 2	15.005 8	38.598 2	6.536 3	21.637 3	1.309 4	0.790 6
	NSST	24.363 8	15.432 1	38.556 1	6.542 3	21.751 1	1.340 9	0.800 5
	DTCWT	23.853 8	15.072 7	42.871 5	6.776 7	20.159 5	1.411 7	0.796 6
	Curvelet	27.970 7	18.141 1	48.998 5	7.049 1	17.572 1	1.609 8	0.708 4
proposed	29.897 5	19.288 5	54.262 2	7.141 6	21.121 2	1.893 7	0.833 8	
第3组	SR	31.986 4	22.820 2	65.607 5	7.482 7	17.092 2	1.246 1	0.523 8
	CSR	31.932 4	21.871 2	68.549 4	7.396 2	17.511 5	1.296 4	0.535 3
	JSR	32.072 7	24.129 8	69.008 9	7.741 2	14.330 7	1.525 5	0.456 3
	JSRSD	33.787 0	23.347 1	64.546 7	7.579 2	15.515 0	1.249 8	0.430 0
	GF	31.739 7	22.429 1	65.449 1	7.475 5	17.241 7	1.238 3	0.532 6
	CBF	31.957 9	23.404 9	60.841 7	7.551 7	16.777 6	1.081 1	0.492 2
	Contourlet	22.598 6	15.683 6	40.559 5	7.042 9	17.637 8	0.931 0	0.539 5
	NSCT	31.473 8	22.367 1	48.920 4	7.275 4	17.110 6	0.931 7	0.520 4
	NSST	31.487 5	22.481 9	48.615 1	7.277 0	17.131 5	0.946 4	0.528 8
	DTCWT	31.841 6	22.360 5	66.464 2	7.507 1	18.529 3	1.279 0	0.526 8
	Curvelet	32.048 6	22.825 6	61.955 5	7.490 4	17.720 6	1.133 4	0.525 2
proposed	33.914 5	23.365 0	67.883 5	7.576 9	17.029 9	1.284 8	0.537 5	

向特性能够保留源图像的信息,融合后的图像不仅清晰度较高,而且显著目标也更加突出. 对比源图像与融合图像的视觉效果以及四种评价指标值,所得结果表明本文算法切实可行. 对比本文算法与其他7种方法的融合图像表明,本文算法能够充分整合源图像的信息并突出显著特征,融合图像具有较高的对比度,符合人眼视觉的观察效果,切实有效.

参考文献(References)

- [1] Jin X, Jiang Q, Yao S W, et al. A survey of infrared and visual image fusion methods[J]. *Infrared Physics & Technology*, 2017, 85: 478-501.
- [2] Ma J Y, Ma Y, Li C. Infrared and visible image fusion methods and applications: A survey[J]. *Information Fusion*, 2019, 45: 153-178.
- [3] Zhang Q, Wang Y B, Martin D L, et al. Multisensor video fusion based on higher order singular value decomposition[J]. *Information Fusion*, 2015, 24: 54-71.
- [4] Gao S B, Cheng Y M, Zhao Y Q. Method of visual and infrared fusion for moving object detection[J]. *Optics Letters*, 2013, 38(11): 1981-1983.
- [5] Li S T, Kang X D, Fang L Y, et al. Pixel-level image fusion: A survey of the state of the art[J]. *Information Fusion*, 2017, 33: 100-112.

- [6] 冯鑫. Tetrolet框架下红外与可见光图像融合[J]. 光子学报, 2019, 48(2): 76-84.
(Feng X. Fusion of infrared and visible images based on tetrolet framework[J]. Acta Photonica Sinica, 2019, 48(2): 76-84.)
- [7] Zhang Q, Maldague X. An adaptive fusion approach for infrared and visible images based on NSCT and compressed sensing[J]. Infrared Physics & Technology, 2016, 74: 11-20.
- [8] 王峰, 程咏梅. 基于Shearlet变换域改进的IR与灰度VIS图像融合算法[J]. 控制与决策, 2017, 32(4): 703-708.
(Wang F, Cheng Y M. Improved infrared and gray visible light image fusion algorithm based on Shearlet transform[J]. Control and Decision, 2017, 32(4): 703-708.)
- [9] Liu Z W, Feng Y, Chen H, et al. A fusion algorithm for infrared and visible based on guided filtering and phase congruency in NSST domain[J]. Optics and Lasers in Engineering, 2017, 97: 71-77.
- [10] Liu Y, Chen X, Ward R K, et al. Image fusion with convolutional sparse representation[J]. IEEE Signal Processing Letters, 2016, 23(12): 1882-1886.
- [11] 常莉红, 冯象初, 张瑞. 四元数小波变换联合稀疏表示的图像融合[J]. 系统工程与电子技术, 2017, 39(7): 1633-1639.
(Chang L H, Feng X C, Zhang R. Image fusion scheme based on quaternion wavelet transform and sparse representation[J]. Systems Engineering and Electronics, 2017, 39(7): 1633-1639.)
- [12] 王昕, 吉桐伯, 刘富. 结合目标提取和压缩感知的红外与可见光图像融合[J]. 光学精密工程, 2016, 24(7): 1743-1753.
(Wang X, Ji T B, Liu F. Fusion of infrared and visible images based on target segmentation and compressed sensing[J]. Optics and Precision Engineering, 2016, 24(7): 1743-1753.)
- [13] Liu G C, Yan S C. Latent low-rank representation for subspace segmentation and feature extraction[C]. IEEE International Conference on Computer Vision. Barcelona: IEEE, 2011: 1615-1622.
- [14] Easley G, Labate D, Lim W Q. Sparse directional image representations using the discrete shearlet transform[J]. Applied and Computational Harmonic Analysis, 2008, 25(1): 25-46.
- [15] Srivastava R, Prakash O, Khare A. Local energy-based multimodal medical image fusion in curvelet domain[J]. Iet Computer Vision, 2016, 10(6): 513-527.
- [16] Zhang Q, Shi T, Wang F, et al. Robust sparse representation based multi-focus image fusion with dictionary construction and local spatial consistency[J]. Pattern Recognition, 2018, 83: 299-313.
- [17] Shreyamsha Kumar B K. Image fusion based on pixel significance using cross bilateral filter[J]. Signal, Image and Video Processing, 2015, 9(5): 1193-1204.
- [18] Ma J Y, Chen C, Li C, et al. Infrared and visible image fusion via gradient transfer and total variation minimization[J]. Information Fusion, 2016, 31: 100-109.
- [19] Zhang Q H, Fu Y L, Li H F, et al. Dictionary learning method for joint sparse representation-based on image fusion[J]. Optical Engineering, 2013, 52(5): 057006.
- [20] Liu C H, Qi Y, Ding W R. Infrared and visible image fusion method based on saliency detection in sparse domain[J]. Infrared Physics & Technology, 2017, 83: 94-102.
- [21] Liu Y, Liu S P, Wang Z F. A general framework for image fusion based on multi-scale transform and sparse representation[J]. Information Fusion, 2015, 24: 147-164.
- [22] Aslantas V, Bendes E. A new image quality metric for image fusion: The sum of the correlations of differences[J]. International Journal of Electronics and Communications, 2015, 69(12): 1890-1896.

作者简介

沈瑜(1982—), 女, 教授, 博士, 从事机器学习、智能信息处理等研究, E-mail: 18609311366@163.com;

陈小朋(1994—), 男, 硕士生, 从事深度学习、图像融合的研究, E-mail: 3064683191@qq.com;

刘成(1993—), 男, 硕士生, 从事深度学习、图像去模糊的研究, E-mail: 1290643579@qq.com;

张泓国(1997—), 男, 硕士生, 从事图像融合的研究, E-mail: 2466896877@qq.com;

王霖(1996—), 男, 硕士生, 从事遥感图像分割的研究, E-mail: 1985405357@qq.com.

(责任编辑: 李君玲)