

## 红外与可见光图像融合技术的研究进展

沈 英, 黄春红, 黄 峰, 李 杰, 朱梦娇, 王 舒\*

(福州大学 机械工程及自动化学院, 福建福州 350116)

**摘要:** 红外与可见光融合图像既具有红外的辐射信息又具有可见光的细节信息, 在生产生活、军事监视等场景得到广泛应用, 已然成为图像融合领域的重点研究方向。根据图像融合方法的核心思想、融合框架、研究进展对基于多尺度变换、稀疏表示、神经网络等融合方法进行详细阐述对比, 并综述了红外与可见光图像融合在各领域内的应用现状, 以及常用的评价指标。并选择具有代表性的多种融合方法与评价指标, 应用于六个不同场景, 验证各方法的优势与不足。最后, 实验分析并总结现有红外与可见光图像融合方法存在的问题, 对红外与可见光图像融合技术的发展趋势进行展望。

**关键词:** 红外图像; 可见光图像; 图像融合; 多尺度变换; 稀疏表示; 神经网络  
**中图分类号:** TP391.4 **文献标志码:** A **DOI:** 10.3788/IRLA20200467

## Research progress of infrared and visible image fusion technology

Shen Ying, Huang Chunhong, Huang Feng, Li Jie, Zhu Mengjiao, Wang Shu\*

(College of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350116, China)

**Abstract:** Infrared and visible image fusion combines the infrared thermal radiation information and visible detail information. The image fusion technique has facilitated development in numerous fields, including production, life sciences, military surveillance and others, and has become a key research direction in the field of image technology. According to the core idea, fusion framework and research progress of image fusion methods, the fusion methods based on multi-scale transformation, sparse representation, neural network, etc. are elaborated and compared, and the application status of infrared and visible light image fusion in various fields and the commonly used the evaluation index. The most representative methods and evaluation indicators are selected and applied to six different scenes in order to verify the advantages and disadvantages of each one. Finally, the existing problems of infrared and visible image fusion methods are experimentally analyzed and summarized, the development prospects of infrared and visible image fusion technology are presented.

**Key words:** infrared image; visible image; image fusion; multi-scale transform; sparse representation; neural network

收稿日期:2021-04-10; 修订日期:2021-05-15

基金项目:国家自然科学基金(62005049);福建省自然科学基金(2020J01451);福建省教育厅中青年教师教育科研项目(JAT190003)

作者简介:沈英,女,教授,博士,主要从事光机电一体化方面的研究。

通讯作者:王舒,男,助理研究员,博士,主要从事光学成像方面的研究。

## 0 引言

在图像处理领域中,单一传感器获取的图像通常只具备某一方面的信息,已无法满足市场需求,因此图像融合技术应运而生。红外图像主要依靠物体自身的热辐射进行成像,突出背景中隐藏的热目标,其不受光照条件、天气的影响,但对比度较低,纹理细节不丰富<sup>[1]</sup>。可见光图像通过反射可见光进行成像,纹理细节和对比度更适合人类的视觉感知,但可见光图像在烟雾、夜间等条件下的成像效果差<sup>[2]</sup>。基于此,两者融合后能够获得一幅既有可见光图像边缘、细节信息又有红外热辐射目标信息的互补融合图像。随着目标检测与识别、军事监视等应用需求的不断提高,红外与可见光图像的融合技术成为该领域研究的热点方向。安防领域,红外与可见光融合图像可准确识别黑暗环境、化妆打扮<sup>[3]</sup>、佩戴眼镜<sup>[4]</sup>等条件下的人脸,为商业应用、公安执法等提供便利需求;军事领域,红外与可见光融合图像可实现恶劣环境下隐藏目标的识别与跟踪<sup>[5]</sup>;智能交通领域,红外与可见光融合图像应用于行人检测<sup>[6]</sup>、车辆识别与车距检测<sup>[7]</sup>、道路障碍物分类<sup>[8]</sup>;农业生产领域,红外与可见光融合图像可应用于水果的成熟度检测<sup>[9]</sup>、植物的病态检测<sup>[10]</sup>等。

近几十年来,大量红外与可见光图像的融合方法相继被提出,并在实际应用中得到推广。目前的图像融合综述中,大部分文献都选择对整个图像融合领域进行综述,较少文献针对红外与可见光图像融合方法进行详细阐述;在部分阐述红外与可见光图像融合的综述中,只对现阶段的图像融合方法进行简要分析,没有与实际应用相结合,缺少应用实例。文中首先综述了红外与可见光图像常用融合方法的研究现状,其次概述了红外与可见光图像融合的主要应用,以及用于评价融合质量的性能指标,并针对选定的六个应用场景,选择九种该领域典型的融合方法和六个图像质量评价指标进行实验分析,最后对红外与可见光图像融合技术的发展与应用进行总结与展望。

## 1 红外与可见光图像融合方法

### 1.1 基于多尺度变换的融合方法

通过分辨率分解方法,基于变换域的图像融合方法可获取一系列包含不同层次的子图像,以保留更多的图像细节信息。多尺度变换是应用最为广泛的基于变换域的融合方法,其融合方法主要分为三个步骤:多尺度的正逆变换、融合规则的设计,具体融合框架流程如图 1 所示。

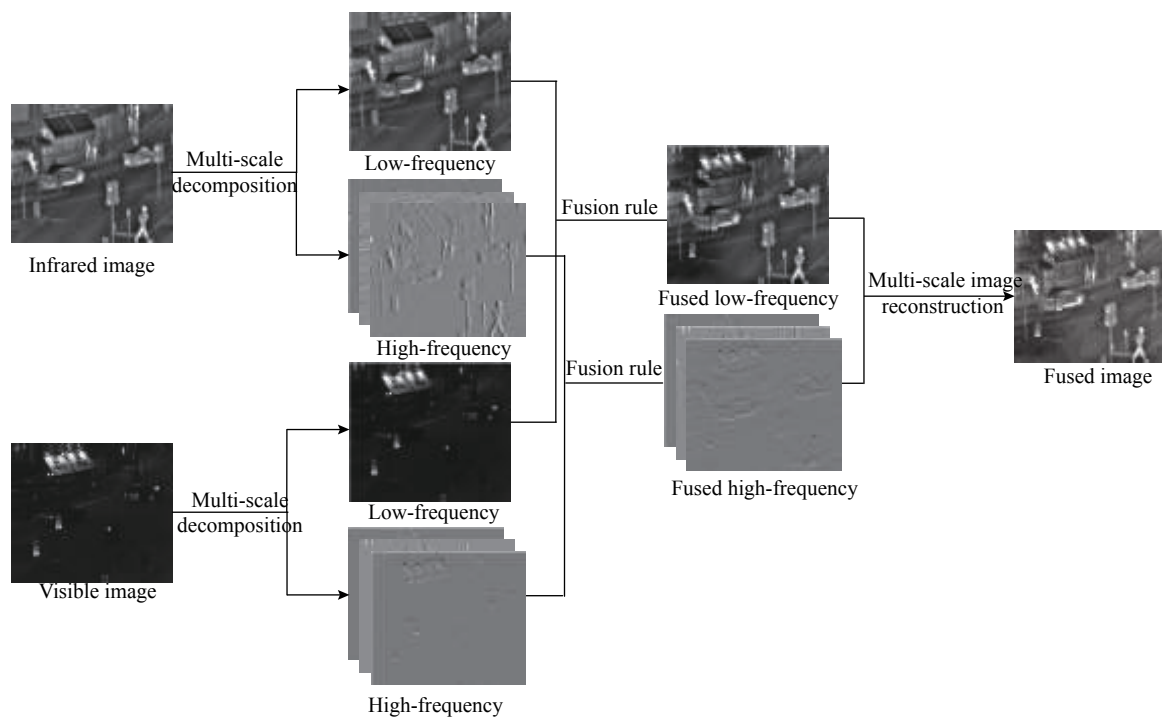


图 1 基于多尺度变换的红外与可见光图像融合框架

Fig.1 Multi-scale transform-based infrared and visible image fusion frame

### 1.1.1 金字塔变换

金字塔变换将图像分解成不同尺度且呈金字塔状的子带图像进行融合。拉普拉斯金字塔 (Laplacian Pyramid, LP) 变换<sup>[11]</sup>是最早提出的基于金字塔变换的融合方法, 基于 LP 变换的成功应用, 比率低通金字塔<sup>[12]</sup>、对比度金字塔 (Contrast Pyramid, CP)<sup>[13]</sup>、形态学金字塔<sup>[14]</sup>和可控金字塔<sup>[15]</sup>等融合方法相继被提出。LP 变换基于高斯金字塔获取的一系列差值图像虽然凸显了高频子带图像的细节特征信息, 但存在图像对比度低, 信息冗余等问题。基于彩色参考图像的假彩色融合<sup>[16]</sup>以及具备优越边缘表达能力的模糊逻辑<sup>[9]</sup>可改善 LP 变换的不足, 增强图像融合效果。此外, 具备人眼视觉系统图像感知特性的 CP 变换, 可弥补 LP 变换图像对比度低的缺陷。但 CP 变换不具备方向不变性, 可结合方向滤波器予以解决<sup>[17]</sup>。CP 变换的改进还可从图像融合规则入手, 例如基于改进区域能量的对比度金字塔算法<sup>[18]</sup>。

相比基于空间域的图像融合方法, 金字塔变换作为最先发展起来的多尺度变换方法, 在图像细节信息保留方面有较大提升。但金字塔变换属于冗余变换, 各层数据之间相关性大, 源图像差异较大的区域融合后易产生块效应, 降低了算法的鲁棒性。同时, 金字塔变换还存在源图像结构信息丢失严重和图像信噪比低的问题。

### 1.1.2 小波变换

小波变换的概念最早由 Grossman 和 Morlet<sup>[19]</sup>于 1984 年提出, 之后 Mallet<sup>[20]</sup>根据信号分解和重建的塔式算法建立了基于小波变换的多分辨率分解理论。小波变换具体包括离散小波变换 (Discrete wavelet Transform, DWT)<sup>[21]</sup>、双树离散小波变换 (Dual Tree Discrete Wavelet Transform, DT-DWT)<sup>[22]</sup>、提升小波变换<sup>[23]</sup>、四元数小波变换<sup>[24]</sup>和谱图小波变换<sup>[25]</sup>等。DWT 通过滤波器组合实现源图像的多尺度分解, 各尺度间独立性高, 纹理边缘信息保留度高<sup>[26]</sup>。但 DWT 存在一些缺陷, 具体包括振荡、移位误差、混叠以及缺乏方向选择性等<sup>[27]</sup>。DT-DWT 利用可分离的滤波器组合对图像进行分解, 解决了 DWT 缺乏方向性的问题, 且具有冗余信息少, 计算效率高的优势。但作为第一代小波变换, DT-DWT 不适用于非欧式空间。提升小波变换是构造第二代小波变换的理想方法, 可

完全视为空间域的变换, 其具有自适应设计强、可不规则采样等优点, 融合视觉效果较好<sup>[23]</sup>。

与金字塔变换相比, 小波变换不会产生图像块效应, 具有高信噪比。此外, 小波变换还具有完备的图像重构能力, 降低了图像分解过程中的信息冗余。然而, 其表达的是源图像中部分方向信息, 仍会造成图像细节信息的丢失。

### 1.1.3 非下采样多尺度多方向几何变换

为获取图像方向信息, 消除吉布斯现象, 解决平移不变性等问题, Da 等人<sup>[28]</sup>提出了非下采样轮廓波变换 (Non-sampled Contourlet Transform, NSCT), 由于不存在采样过程, 其解决了轮廓波变换的频谱混叠问题。此外, NSCT 与模糊逻辑相结合能有效增强红外目标并保留可见光图像的细节<sup>[29]</sup>。NSCT 与感兴趣区域的提取相结合, 可成功凸显红外目标<sup>[30]</sup>。但是, NSCT 的计算效率较低, 无法满足高实时性的应用需求。

为满足高实时性需求, Guo 等人<sup>[31]</sup>提出多尺度多方向性的剪切波变换, 但其仍不具备平移不变特性。而非下采样剪切波变换 (Non-sampled Shearlet Transform, NSST)<sup>[32]</sup>不仅可以满足上述需求, 且相比于 NSCT, 拥有更高的计算效率。优越的信息捕获和表示能力使 NSST 成为红外与可见光图像融合方法中的一种流行方法。Kong 等人<sup>[33]</sup>在 NSST 融合方法的基础上, 引入区域平均能量和局部对比度的融合规则, 将空间域分析和多尺度几何分析的优点最大程度地融合在一起。NSST 实质上属于冗余变换, 为克服该不足, Kong 等人<sup>[34]</sup>进一步在 NSCT 融合方法中提出快速非负矩阵分解, 最终融合图像在保留源图像全部特征的同时降低了图像的冗余信息。

非下采样多尺度多方向几何变换所分解出的子带图像与源图像尺寸相同, 这有利于源图像细节、纹理特征的提取, 同时简化了后续融合规则的设计。但 NSCT 分解过程复杂, 难以应用于实时性要求高的场景。NSST 分解时由于引入了非下采样金字塔变换的方法, 容易造成高频子带图像细节的丢失, 还会降低图像亮度。

### 1.1.4 其他方法

常用的多尺度变换的融合方法还有保边滤波器、Tetrolet 变换、顶帽变换、哈尔小波变换等, 其中保边

滤波器的应用最为广泛,该方法将源图像分解为一个基本层以及一个或多个的细节层,在保持空间结构的同时还减少了融合图像的伪影现象。其中,导向滤波器在保留源图像边缘细节信息的同时还消除其块效应<sup>[35]</sup>,局部边缘保持滤波器在保证图像全局特征的前提下,有效突显图像的细节信息<sup>[36]</sup>。

### 1.2 基于稀疏表示的融合方法

2010年, Yang 和 Li<sup>[37]</sup> 提出了基于稀疏表示 (Sparse Representation, SR) 图像融合方法 (融合框架如图 2 所示), 其重点在于过完备字典的构造和稀疏系数分解算法<sup>[38]</sup>。

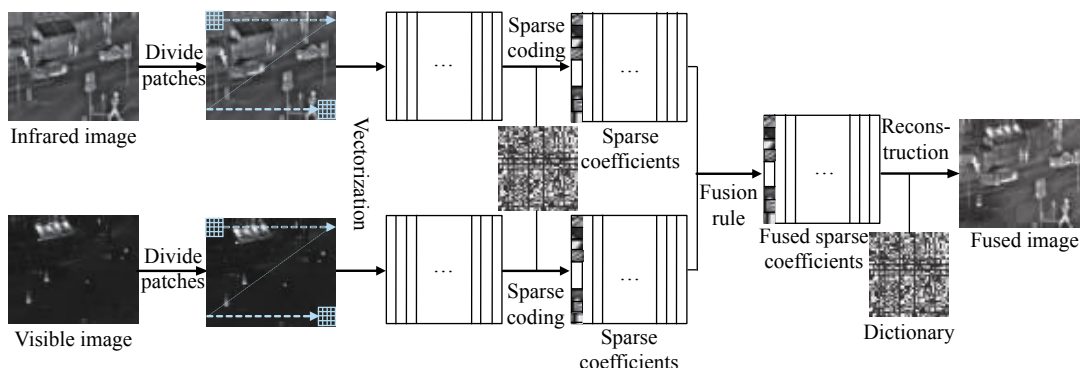


图 2 基于稀疏表示的红外与可见光图像融合框架

Fig.2 Sparse representation-based infrared and visible image fusion frame

过完备字典的构造方式主要有两种: 基于数据模型和基于学习算法的过完备字典。基于数据模型的过完备字典是利用特定的数学模型进行构建的, 该方法虽然高效但难以应对复杂数据, 可利用基于联合学习策略的平稳小波多尺度字典<sup>[39]</sup> 和由过完备离散余弦字典与基函数相结合的混合字典<sup>[40]</sup> 来解决此问题。而基于学习算法<sup>[38, 40-47]</sup> 的过完备字典是通过训练样本集的方式构造的, 常用的是基于最优方向法 (Method of Optimal Directions, MOD) 字典<sup>[48]</sup> 和奇异值分解的稀疏字典训练 (Kernel Singular Value Decomposition, K-SVD) 字典<sup>[49]</sup>。该类字典具有较强的自适应性但也带来较高的计算量。因此, 目前过完备字典的研究趋势倾向于综合两种构造方式优点的算法。起初, Rubinstein 等人<sup>[50]</sup> 设计了一种综合固定字典和学习字典的用于学习稀疏字典的稀疏 K-SVD 方法。后来, 聚类补丁和 PCA<sup>[51-53]</sup>、最优方向<sup>[46]</sup>、自适应稀疏表示<sup>[41]</sup>、在线学习方法<sup>[44, 54]</sup>、多尺度几何分析领域中的 K-SVD 字典<sup>[38, 55]</sup> 等方法也成功应用于图像融合领域。

稀疏系数分解算法中, 匹配跟踪 (Matching Pursuit, MP) 算法利用原子向量的线性运算从训练好的过完备字典中选择最佳线性组合的原子以表征图

像信息, 但其迭代结果只是次最优。正交匹配追踪 (Orthogonal Matching Pursuit, OMP) 算法<sup>[39, 56-57]</sup> 在此基础上对 MP 算法进行改进, 经 OMP 算法处理后的原子组合都已处于正交化状态。在精度要求相同的情况下, OMP 算法的计算效率高于 MP 算法。此外, 为解决 OMP 算法和 MP 算法融合规则设计难的问题, 同步正交匹配追踪 (Synchronous Orthogonal Matching Pursuit, SOMP) 算法<sup>[39, 51, 58]</sup> 基于 OMP 算法进行了原子集的改进, 其可从不同的源图像中分解出相同子集的字典, 从而简化了图像融合规则设计, 在图像融合领域得到广泛应用。

稀疏表示与传统多尺度变换的图像融合方法相比有两大区别<sup>[59]</sup>: 一是多尺度融合方法一般都是基于预先设定的基函数进行图像融合, 这样容易忽略源图像某些重要特征; 而基于稀疏表示的融合方法是通过学习超完备字典来进行图像融合, 该字典蕴涵丰富的基原子, 有利于图像更好的表达和提取。二是基于多尺度变换的融合方法是利用多尺度的方式将图像分解为多层图像, 再进行图像间的融合, 因此, 分解层数的选择就尤为关键。一般情况下, 为从源图像获取丰富的空间信息, 设计者都会设置一个相对较大的分解层数, 但随着分解层数的增加, 图像融合对噪声和配

准的要求也越来越严格;而稀疏表示则是利用滑窗技术将图像分割成多个重叠小块并将其向量化,可减少图像伪影现象,提高抗误配准的鲁棒性。

基于稀疏表示的图像融合方法虽然能够改善多尺度变换中特征信息不足和配准要求高的问题,但其自身仍存在一些不足。主要体现在:(1)过完备字典的信号表示能力有限,容易造成图像纹理细节信息的丢失;(2)“max-L1”融合规则对随机噪声敏感,这会降低融合图像信噪比;(3)滑窗技术分割出的重叠小块,降低了算法的运行效率。

### 1.3 基于神经网络的融合方法

神经网络在图像融合的应用始于脉冲耦合神经网络(Pulse Coupled Neural Network, PCNN)模型。与其他神经网络模型相比较,PCNN模型无需训练与学习过程,就可有效提取图像信息<sup>[60]</sup>。基于PCNN的红外与可见光图像融合方法通常会与多尺度变换方法结合,例如NSCT<sup>[61-66]</sup>、NSST<sup>[60]</sup>、曲波变换<sup>[67]</sup>、轮廓波变换<sup>[68]</sup>等。PCNN主要应用于图像融合策略中,最常见的两种方式是单独作用于图像的高频子带和同时作用于图像的高低频子带。例如在NSCT融合框架中,可采用自适应双通PCNN融合图像的高低频子带系数,也可将PCNN模型作为高频子带图像的融合规则,加权平均<sup>[62]</sup>、区域方差积分<sup>[63]</sup>等作为低频子带图像的融合策略。

基于深度学习的图像融合方法渐渐成为图像融合领域的主要研究方向,但其在异源图像融合的研究还处于一个初步发展的水平。深度学习方法是將源图像的深度学习特征作为一种显著特征用于融合图像的重建,卷积神经网络是目前最常用于图像融合的深度学习方法。Li<sup>[69]</sup>、Ren<sup>[70]</sup>都提出一种基于预训练的VGG-19(Visual Geometry Group-19)网络,以提取源图像的深层特征,获得较好的融合效果。2019年, Ma等人<sup>[71]</sup>首次将端到端的生成对抗网络模型用于解决图像融合问题,避免了传统算法手动设计复杂的分解级别和融合规则,并有效保留源图像信息。

PCNN模型中神经元与图像像素一一对应,解决了传统方法中图像细节易丢失的问题。但PCNN模型网络结构复杂,参数设置繁琐。同时,其与多尺度变换组合的方法,也只能实现模型的局部自适应,且计算速率和泛化能力仍有待提高。基于深度学习的

图像融合可从图像数据中提取深层特征,实现模型的自适应融合,具有较强的容错性和鲁棒性。但其仍未得到广泛应用,原因主要有:(1)卷积神经网络的标签数据集制作难度大。在红外与可见光图像融合领域中,一般不存在标准参考图像,模型的ground truth无法得到准确的定义;(2)端到端模型损失函数的针对性不强。端到端模型虽然解决了卷积神经网络模型图像清晰度不高和需要标准参考图像的问题,但仍缺乏针对性损失函数的设计。

### 1.4 其他方法

红外与可见光图像还可使用基于子空间的融合方法,主成分分析法<sup>[72]</sup>、鲁棒主成分分析法<sup>[73]</sup>、独立成分分析法<sup>[74]</sup>、NMF<sup>[75]</sup>等,一般而言,大部分源图像都存在着冗余信息,基于子空间的融合方法通过将高维的源图像数据投影至低维空间或子空间中,以较少的计算成本获得图像的内部结构。

混合模型通过结合各方法的优点以提高图像融合的性能,常见的混合模型有多尺度变换与显著性检测、多尺度变换与SR、多尺度变换与PCNN结合等。多尺度变换与显著性检测相结合的图像融合方法一般是將显著性检测融入多尺度变换的融合框架中,以增强感兴趣区域的图像信息。显著性检测的应用方式主要两种,权重计算<sup>[76-80]</sup>和显著目标提取<sup>[81-82]</sup>。权重计算是在高低频子带图像中获得显著性图,并计算出对应权重图,最终将其应用于图像重构部分。目标检测与识别等监视应用中常会采用显著目标提取,较有代表性的是Zhang等人<sup>[81]</sup>在NSST融合框架的基础上,利用显著性分析提取红外图像的目标信息。

多尺度变换存在图像对比度低,多尺度分解级别不易确定的问题;稀疏表示则表现出源图像的纹理和边缘信息趋于平滑,计算效率低等不足。基于此,将多尺度变换和稀疏表示结合的混合模型,通常可以取得这两者的最佳平衡,稀疏表示模型通常应用于多尺度分解后的低频子带图像<sup>[83]</sup>。此外,根据PCNN可充分提取图像细节信息的优点,多尺度变换还常与SR、PCNN结合在一起,低频选择基于SR的融合规则,高频选择基于PCNN的融合规则<sup>[84]</sup>。该混合模型有效提高了融合图像的清晰度和纹理特征,但在设计融合模型时,需统筹SR和PCNN的优缺点,以免模型过于复杂,增加计算成本。

### 1.5 图像融合规则

图像融合规则的优劣直接影响图像融合效果。传统方法所采用的融合规则均为“高频绝对值取大, 低频加权平均”。这两类融合规则均属于像素级的融合规则, 绝对值取大法运算成本低, 但容易造成图像信息的缺失, 边缘尖锐等问题。加权平均法也会使得图像某些重要特征丢失, 同时, 如果算法的权重设置过于简单会导致融合图像的灰度值与源图像差异过大。许多学者会基于像素级的图像融合规则进行一定的改进, 例如, 基于对比度和能量的绝对值取大法<sup>[24]</sup>, 根据图像空间频率确定权重值的加权平均法<sup>[82]</sup>, 基于双边滤波器的加权平均法<sup>[25]</sup>等。

此外, 常见的还有基于区域、基于矩阵分析的融合规则。由于人眼在观察图像时, 关注的往往是区域

级别的, 区域分类法会更符合人眼视觉感知系统, 同时该方法充分考虑各像素间的关系, 图像的局部特征可得到进一步的体现。常见的有区域平均能量<sup>[85]</sup>、区域交叉梯度<sup>[86]</sup>等。矩阵分析法主要分为模糊逻辑<sup>[56, 87]</sup>和非负矩阵分解法<sup>[34]</sup>, 模糊逻辑常被应用到加权平均融合规则中以融合图像的高频子带, 非负矩阵分解法利用分解后图像的像素值为非负值这一特点将图像分解为非负分量, 以达到低运算复杂度和高清晰度的效果, 其通常应用于图像的低频子带融合规则。

表 1 列举了红外与可见光图像常用的融合方法、融合策略、优缺点以及所适用的场景。表中根据各方法提出时间的顺序介绍金字塔变换、小波变换、非下采样多尺度多方向几何变换、稀疏表示、神经网络、混合方法六大类融合方法。

表 1 红外与可见光图像融合方法的对比

Tab.1 Comparison of infrared and visible image fusion methods

Fusion methods	Specific methods	Fusion strategies	Advantages	Limitations	Applicable scenes
Pyramid transforms	Laplacian pyramid	Fuzzy logic <sup>[9]</sup>	Smoothing image edge; Less time consumption; Less artifacts	Losing image details; Block phenomenon; Redundancy of data	
	Contrast pyramid	Clonal selection algorithm <sup>[17]</sup> ; Teaching learning based optimization <sup>[88]</sup> ; Multi-objective evolutionary algorithm <sup>[89]</sup>	High image contrast; Abundant characteristic information	Low computing efficiency; Losing image details	Short-distance scenes with sufficient light, such as equipment detection
	Steerable pyramid	The absolute value maximum selection(AVMS) <sup>[90]</sup> ; The expectation maximization(EM) algorithm <sup>[91]</sup> ; PCNN and weighting <sup>[92]</sup>	Abundant edge detail; Inhibiting the Gibbs effect effectively; Fusing the geometrical and thematic feature available	Increasing the complexity of algorithm; Losing the image details	
Wavelet transform	Discrete wavelet transform	Regional energy <sup>[93]</sup> ; Target region segmentation <sup>[21]</sup>	Significant texture information; Highly independent scale information; Less blocking artifacts; Higher signal-to-noise ratios	Image aliasing; Ringing artifacts; Strict registration requirements	Short-distance scenes, such as face recognition
	Dual tree discrete wavelet transform	Particle swarm optimization <sup>[22]</sup> ; Fuzzy logic and population-based optimization <sup>[94]</sup>	Less redundant information; Less time consumption	Limited directional information	
Nonsampled multi-scale and multi-direction geometrical transform	Lifting wavelet transform	Local regional energy <sup>[23]</sup> ; PCNN <sup>[85]</sup>	High computing speed; Low space complexity;	Losing image details; Distorting image	
	NSCT	Fuzzy logic <sup>[29]</sup> ; Region of interest <sup>[30]</sup>	Distinct edge features; Eliminating the Gibbs effect; Better visual perception	Losing image details; Low computing efficiency; Poor real-time	Scenes with a complex background, such as rescue scenes
Sparse representation	NSST	Region average energy and local directional contrast <sup>[33]</sup> ; FNMF <sup>[34]</sup>	Superior sparse ability; High real-time performance	Losing luminance information; Strict registration requirement; Losing image details of high frequency	Cases need real-time treatment, such as intelligent traffic monitoring
		Saliency detection <sup>[44, 86-87]</sup> ; PCNN <sup>[56, 95]</sup>	Better robustness; Less artifacts; Reducing misregistration; Abundant brightness information	Smoothing edge texture information; Complex calculation; Losing edge features of high frequency images	Scenes with little feature points, such as the surface of the sea

续表 1

Tab.1 Continued

	PCNN	Multi-scale transform and sparse representation; Multi-scale transform	Superior adaptability; Higher signal-to-noise ratios; High fault tolerance	Model parameters are not easy to set; Complex and time-consuming algorithms	
Neural network	Deep learning	VGG-19 and multi-layer fusion <sup>[69]</sup> ; VGG-19 and saliency detection <sup>[70]</sup>	Less artificial noise; Abundant characteristic information Less artifacts	Requiring the ground truth in advance	Automatic target detection and localization
		GAN <sup>[71]</sup>	Avoiding manually designing complicated activity level measurements and fusion rules	The visual information fidelity and correlation coefficient is not optimal	
Hybrid methods	Multi-scale transform and saliency	Weight calculation <sup>[76-80]</sup> ; Salient object extraction <sup>[81, 82]</sup>	Maintaining the integrity of the salient object region; Improving the visual quality of the fused image; Reducing the noise	Highlighting saliency area inconsistently; Losing the background information	The surveillance application, such as object detection and tracking
	Multi-scale transform and SR	The absolute values of coefficient and SR <sup>[38]</sup> ; The fourth-order correlation coefficients match and SR <sup>[83]</sup>	Retaining luminance information; Excellent stability and robustness	Poor real-time Losing the image details	

## 2 应用

### 2.1 目标识别

目标识别一直是计算机视觉重要且热门的研究领域, 图像融合技术能够对图像信息进行多层次、多方位的处理, 可获得更为准确的图像信息和最佳视觉效果, 为后期图像处理提供强有力支撑。随着商业应用、公安执法等对安全需求的不断提高, 人脸识别技术已成为一个重要的研究领域。应用于面部的材料(如化妆品)会改变面部特征, 从而导致在可见光下工作的生物特征人脸识别系统不能正常运作, 但是红外热辐射是从皮肤组织自然散发出来的, 不受化妆品的影响<sup>[3, 96]</sup>。如图 3 所示, 化妆品严重影响可见光图像

的成像, 但是在红外热成像与偏振热成像中几乎没有影响, 通过图像融合可以获得较好的人脸识别结果。实际应用中, 安防领域还需要对中远距离的人脸进行识别, Omri 等人提出基于奇异值分解和主成分分析的小波变换, 该方法已被成功应用于远程自动人脸识别的产品中。类似的, 遗传算法<sup>[97]</sup>也成功应用于该领域。此外, 针对佩戴眼镜的特殊群体, 由于红外图像对玻璃不具有穿透能力, 所以常结合椭圆拟合方法<sup>[98]</sup>进行人脸识别。除基于脸部进行人脸识别外, 红外与可见光图像融合后的图像还可根据其他特征实现人脸识别。例如, 对人体进行全天候监测的基于人耳的全自动生物识别系统<sup>[99]</sup>; 人脸面部被遮挡情况下基于眼周区域的面部识别技术<sup>[100]</sup>。

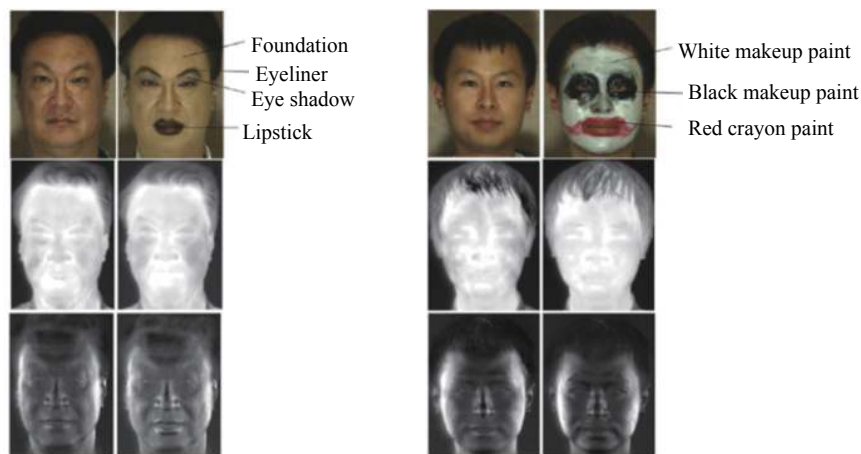


图 3 可见光、红外与偏振进行人脸识别的差异<sup>[3]</sup>

Fig.3 Difference between visible light, infrared and polarization for face recognition<sup>[3]</sup>

## 2.2 目标检测与跟踪

目标检测与目标识别的最大区别在于目标检测需要确定目标的准确位置。图像融合方法在目标检测的应用主要有两种：融合后检测、检测后融合。目前较为流行的方法为前者，可应用于提高农作物自动

化采摘精度，克服强光下图像产生过曝问题，图 4 为该检测方法的图像处理框架<sup>[9]</sup>。此外，在智能交通领域中，先融合后检测方法可对行人或车辆进行有效避障<sup>[7]</sup>；或是构建先获取源图像的检测结果后对图像进行分割融合的鲁棒性检测系统<sup>[6]</sup>。

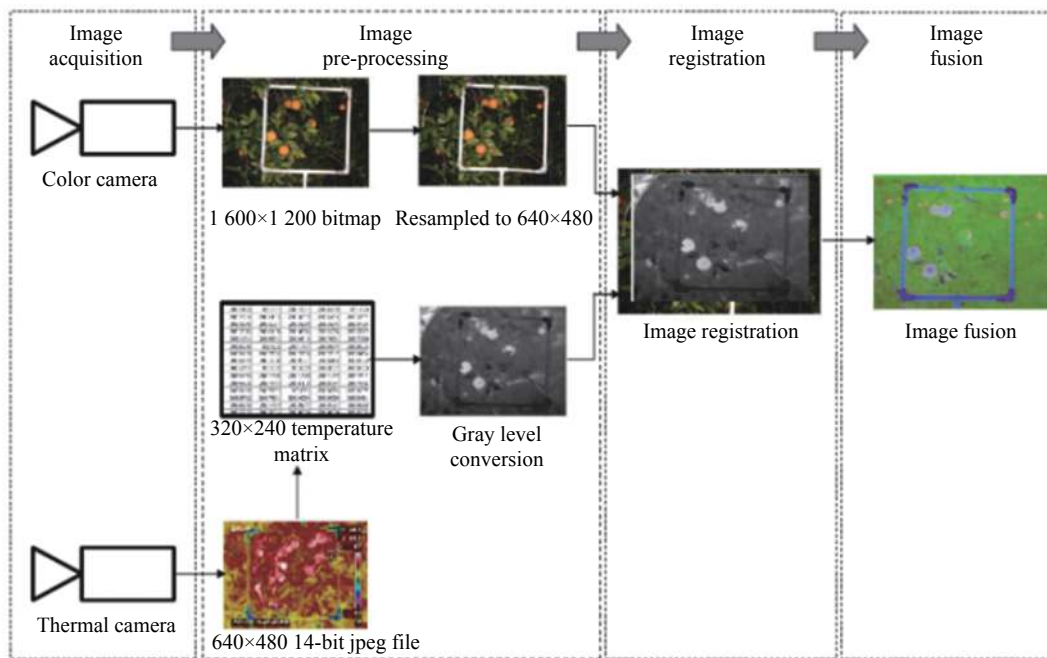


图 4 用于水果检测的图像融合处理框架<sup>[9]</sup>

Fig.4 Image fusion processing framework for fruit detection<sup>[9]</sup>

目标跟踪一般通过处理视频或时间序列图像实现。针对目标的移动特性，Han 等人<sup>[101]</sup>将目标的热辐射信息和移动信息作为显著性特征，实现目标的快速检测跟踪，增强监控应用中所需的可视化。目标跟踪与侦察行动中，通常需要具有成像传感器平台的机载设备，利用机载图像建立的红外与可见光图像融合系统，可实现全天候不间断工作，并实时突显目标信息<sup>[5]</sup>。与其他应用相比，目标跟踪更注重算法的实时性。融合质量高的复杂算法有时并不符合实际需求，所以在实际需求中，会采用较为简单的融合算法，例如利用比率低通金字塔降低计算成本<sup>[102]</sup>、均值滤波器和中值滤波器的两尺度图像分解方法<sup>[77]</sup>。但红外与可见光图像融合方法在目标检测与跟踪的实时性应用中还处于发展阶段，在硬件设施上的应用过少。

## 2.3 夜视监控

黑暗环境下的融合图像虽然可获得完整信息的图像，但人类视觉系统对彩色图像会更加敏感<sup>[103]</sup>。

夜视监控需要清晰且色彩丰富的图像或视频，而基于彩色图像的红外与可见光融合图像是进行夜视监控的理想融合源。所以将彩色图像的颜色特征转移至融合图像，可增强夜视监控的对比度，突出目标信息，提供良好的视觉感知<sup>[104-105]</sup>。

在城市安全监控防护中，Davis 等人<sup>[106]</sup>在图像融合方法中引入背景差分技术，构建的融合系统可呈现全天候彩色图像监控。在其他夜视监控应用中，常用的方法有利用 DWT 和目标区域分割建立的机载夜视监控系统<sup>[107]</sup>，基于 NSCT 建立的夜视导航和监视方法<sup>[108]</sup>，增强边缘信息的实时夜视监视系统<sup>[109]</sup>，基于 NMF 和颜色转移技术的夜间视觉摄像机系统<sup>[110]</sup>等。在军事目标应用中，Hogervorst 等人<sup>[111]</sup>利用颜色映射优化图像间的色彩匹配，生成更自然的融合图像。即使目标已进行伪装处理，该方法仍能在夜间将目标识别，因颜色映射过程简单快速，其已逐步应用于军事领域。如图 5 所示，图像 (f) 将伪装目标直观地表示



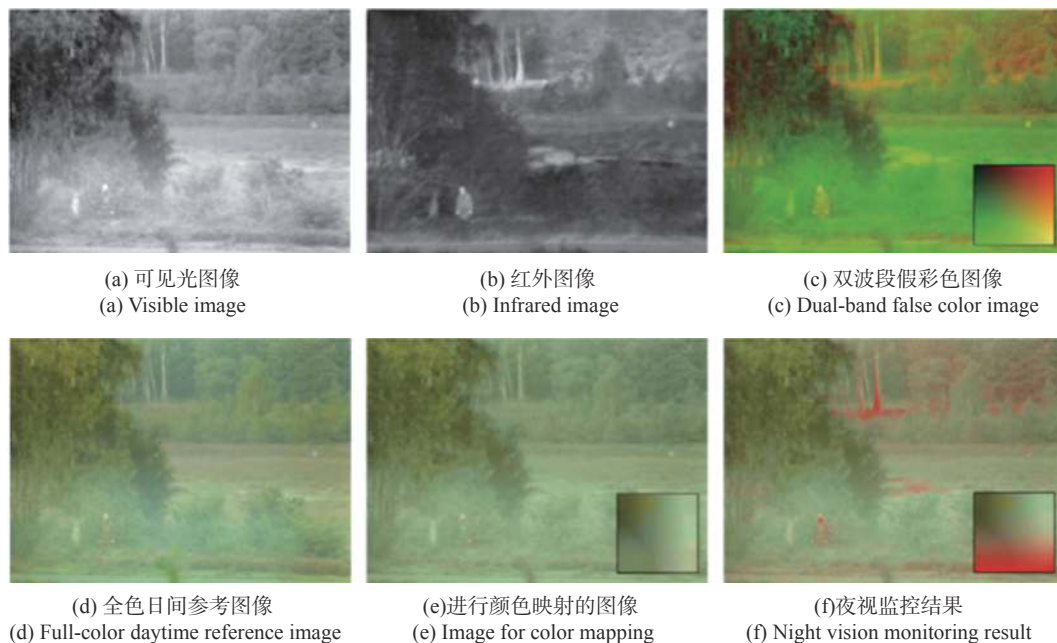


图 5 自然颜色映射在图像融合中的应用<sup>[111]</sup>

Fig.5 Application of natural color mapping in image fusion<sup>[111]</sup>

出来,有效地突出了潜在监测对象。

### 2.4 其他融合应用

红外和可见光图像融合技术的应用领域非常广泛,包括军事领域的彩色视觉增强、农业领域、医学成像领域、遥感领域等。农业领域,除上述对水果进行检测采摘外,还可以评估水果的硬度和可溶性固形物含量<sup>[112]</sup>、提高病态植物的早期检测准确性<sup>[10]</sup>。医学成像领域中,通过使用可调模式选择图像融合方法实时融合红外和可见光视频图像,有助于在术中评估关键胆管结构<sup>[113]</sup>。遥感领域,红外与可见光融合图像可应用于城市目标的空间信息<sup>[114]</sup>、气象预警<sup>[115]</sup>、高空无人机装备<sup>[116]</sup>及其他方面<sup>[117-118]</sup>。此外,红外与可见光的融合技术还可用于旧文档的视觉增强,提高旧文本的可读性<sup>[119-120]</sup>。目前大部分复杂算法注重提高融合质量,但在实际应用的缺乏一定的实用性,随着计算机成像技术、传感器技术的不断发展,一些低成本、高实时性的图像融合算法将会在更多领域应用<sup>[121-122]</sup>。

### 3 评价指标

图像融合质量的评价方法主要分为主观法和客观法。主观法是将图像划分为五个等级,分别是“特别好”、“好”、“一般”、“差”和“特别差”。主观法属于定性分析,具有较强的主观意识,对于两幅融合效果

较为相近的图像无法做出客观的判断,同时,相邻的评价级别没有明确的划分界限,存在着一定的不足。客观评价法是通过特定的公式计算图像的相关指标信息以对融合图像进行定量分析,主要分为无参考图像与有参考图像两类评价指标<sup>[123]</sup>。

常用的图像融合评价指标的定义及说明如表 2、表 3 所示。设源图像的尺寸大小为  $M \times N$ ,其中 A,B,S 表示源图像,F 和 S 表示融合图像和参考图像; $\mu$  为图像的灰度均值; $p_k$  表示像素值为  $k$  的概率 ( $k=0,1,2,\dots,255$ )。设  $Z=A, B, S, F, R; i=1,2,\dots,M; j=1,2,\dots,N$ 。其中,  $Z(i, j)$  表示图像  $Z$  的灰度值; $\Delta Z$  表示图像  $Z$  的差分;表示边缘信息量; $P_Z$  和  $P_{ZZ}$  分别表示图像的概率密度函数和图像间的联合概率密度;函数表示图像的边缘强度函数。

无参考图像的评价指标又可以分为基于单一图像的评价指标和基于源图像的评价指标。基于单一图像的图像评价方法是基于最终融合图像所进行的图像性能评价,包括信息熵 (Information Entropy, IE)、标准差 (Standard Deviation, SD)、平均梯度 (Average Gradient, AG)、空间频率等,其通过不同的方式度量融合图像本身的信息量、灰度值分布等。IE 和 SD 分别通过统计图像灰度分布和度量像素灰度值相较于灰度均值的偏离程度来反映融合图像的信息量。

表 2 无参考图像的评价指标

Tab.2 Evaluation index without reference image

Evaluation indicators	Definition	Explanation
IE <sup>[124]</sup>	$IE = - \sum_{i=0}^{L-1} p_i \log_2 p_i$	Amount of information contained in an image increases as IE improves
SD <sup>[125]</sup>	$SD = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (F(i, j) - \mu)^2}$	Deviation between pixels and pixel mean is evaluated by SD, which improves with the increase of SD, resulting in improvement in contrast of images
AG <sup>[126]</sup>	$AG = \frac{1}{(M-1)(N-1)} \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} \sqrt{\frac{(\Delta Z_i^2 + \Delta Z_j^2)}{2}}$	A wealth of detailed information is exhibited by a high value of AG which is used to reflect the gray variation of the image
Q <sup>AB/F</sup> <sup>[127]</sup>	$Q^{AB/F} = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (Q_{(i,j)}^{AF} w_{(i,j)}^A + Q_{(i,j)}^{BF} w_{(i,j)}^B)}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (w_{(i,j)}^A + w_{(i,j)}^B)}$	Fusion effect of image exhibits better as the value of Q <sup>AB/F</sup> which is used to evaluate the transfer of edge information, approaches 1
MI <sup>[2]</sup>	$I_{FA}(i, j) = \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} P_{FA}(i, j) \log_2 \frac{P_{FA}(i, j)}{P_F(i)P_B(j)}$ $MI_{AB}^F = I_{FA} + I_{FB}$	Amount of information preserved in an image increases with the improvement of MI which is utilized to characterize inheritance of image information
CC <sup>[128]</sup>	$CC = \frac{\sum_{i=1}^M \sum_{j=1}^N [(F(i, j) - \mu_F) \times (S(i, j) - \mu_S)]}{\sqrt{\sum_{i=1}^M \sum_{j=1}^N [(F(i, j) - \mu_F)^2] \sum_{i=1}^M \sum_{j=1}^N [(S(i, j) - \mu_S)^2]}}$	Similarity between images improves as CC increases, thereby preserving more image information

表 3 基于参考图像的评价指标

Tab.3 Evaluation index based on reference image

Evaluation indicators	Definition	Explanation
SSIM <sup>[129]</sup>	$SSIM_{RF} = \prod_{i=1}^3 \frac{2\mu_R\mu_F + c_i}{\mu_R^2 + \mu_F^2 + c_i}$	Similarity between source image and fusion image enhances with the increase of SSIM which is used to measure image luminance, contrast and structural distortion level
RMSE <sup>[2]</sup>	$RMSE = \sqrt{\frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N [R(i, j) - F(i, j)]^2}$	Performance indicators of images promote with the reduction of RMSE
PSNR <sup>[2]</sup>	$PSNR = 10 \cdot \lg \frac{(255^2 \times M \times N)}{\sum_{i=1}^M \sum_{j=1}^N [R(i, j) - F(i, j)]^2}$	The distortion of images decreases as the improvement of PSNR using to evaluate whether the image noise is suppressed

AG 和空间频率反映图像的灰度变化率和清晰度。基于源图像的评价指标通常只考虑图像的某一统计特征,与主观评价的结果有出入。无参考图像的评价指标还有一类是基于源图像进行衡量,主要从信息论的角度出发,度量融合图像从源图像处所提取的信息。常用的有互信息 (Mutual Information, MI)、相关系数 (Correlation Coefficient, CC) 以及边缘信息传递量的 Q<sup>AB/F</sup>。此外,还有从信息熵引申出来的交叉熵、联合熵,IE 反映的仅仅是融合图像的信息量,无法说明图像的整体融合效果,而交叉熵和联合熵可弥补该不足。

基于参考图像的评价指标是通过比较源图像与

标准参考图像间灰度值、噪声等的差异以评价其性能。主要包括结构相似度 (Structural Similarity, SSIM)、均方根误差 (Root-Mean-Square Error, RMSE)、峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR) 等。SSIM 是通过比较图像间的亮度、对比度、结构失真水平的差异性来评价图像的性能; RMSE、偏差指数和扭曲程度都是通过比较图像间像素的灰度值进行评估; PSNR 是通过度量融合图像的噪声是否得到抑制来评价图像的质量。在实际的图像融合过程中,往往没有参考图像作为一个标准,所以该评价方法目前还未大规模应用。

### 4 实 验

文中选取六个不同应用场景的红外与可见光图像, 分别为野外战营 (Nato-camp)、车辆 (Car)、坦克

(Tank)、烟雾场景 (Octec)、士兵行人 (People) 和街道 (Street), 九种较有代表性的融合方法以及六种评价指标进行实验对比分析。图 6 所选取的融合方法从上至下分别为: LP<sup>[11]</sup>、RP<sup>[38]</sup>、DTCWT<sup>[130]</sup>、曲波变换 (Cur-

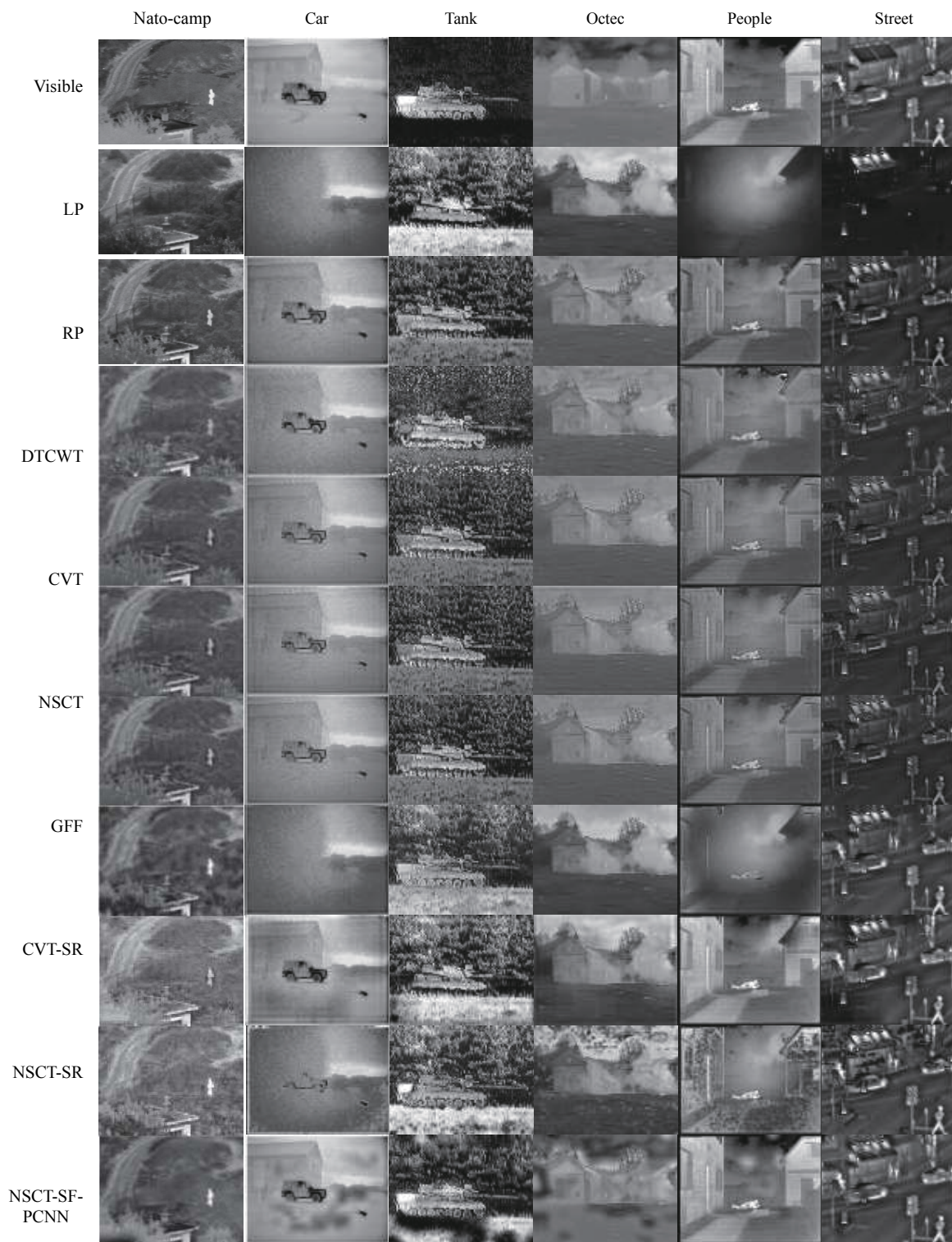


图 6 九种代表性方法对红外与可见光图像的融合效果

Fig.6 Effect of nine representative methods on the fusion of infrared and visible images

velet Transform, CVT)<sup>[130]</sup>、NSCT<sup>[130]</sup>、引导滤波器 (Guided Filtering-based Fusion, GFF)<sup>[38]</sup>、CVT-SR<sup>[38]</sup>、NSCT-SR<sup>[38]</sup>、NSCT-SF-PCNN<sup>[131]</sup>。评价指标分别是 IE、SD、MI、AG、边缘相似度 (Edge-association) 和  $Q^{AB/F}$  六种, 如图 7 所示。其中, 融合方法均采用公开代码, 所设参数与原文献一致。具体表现为: 所有多尺度变换方法的分解层数均为 4 层; 稀疏表示方法中滑动窗口的步长设置为 1, 字典大小为 256, 稀疏系数分解算法 K-SVD 的迭代次数固定为 180 次。除混合方法外, 所有算法都采用“高频绝对值取大, 低频加权平均”的融合规则。其余参数详见对应参考文献。

从图 6 的视觉效果来看, LP 方法在 Car 中出现房

子的细节丢失, Octec 图像的对比度、亮度低, Street 整体的融合视觉效果良好, 但部分细节出现块现象; RP 方法在六个场景中的融合视觉效果均不如 LP 方法; DTCWT、CVT、NSCT、CVT-SR 整体的融合效果良好; GFF 方法在 Car 和 People 中的融合效果较差, 车辆和人的轮廓信息大部分被烟雾遮挡; NSCT-SR 在两幅图像中出现图像块效应; NSCT-SF-PCNN 除在 Nato-camp 表现出良好的融合效果外, 在其余的图像对中没有表现出优越的融合效果。

从图 7 的客观评价指标来看, RP 方法在 AG 指标中数值最高, 在对图像清晰度有高要求的实际应用中, 可考虑应用此方法; DTCWT 方法整体的 SD 最低,

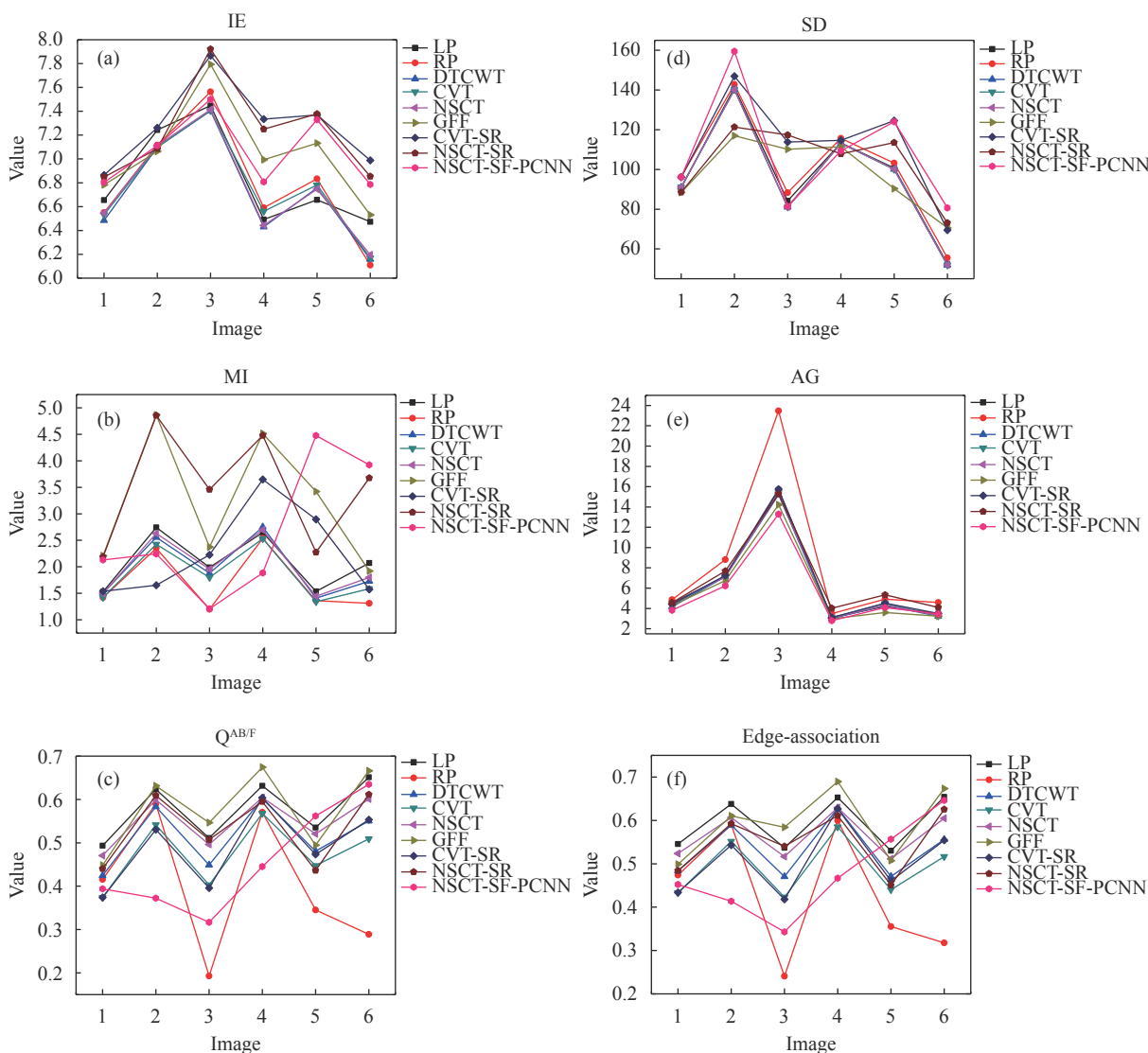


图 7 不同方法的客观评价指标

Fig.7 Objective evaluation metrics with different methods

$Q^{ABF}$  和边缘相似度的效果较好,说明基于 DTCWT 方法的图像对比度低,边缘信息保留度高;NSCT 以及它的混合方法 NSCT-SR 具有较高的 IE、MI,说明基于 NSCT 的方法及其改进方法有利于图像信息的保留,并可用于场景复杂的图像中;GFF 方法与 NSCT-SR 所获得的最佳评价指标值相似,但其不适用于烟雾场景中;CVT-SR 方法可获得良好的 IE 和 AG,由此观之,基于 SR 的融合方法可提高图像的 IE 和 AG;NSCT-SF-PCNN 方法在所选取的六个实验场景中,在六个评价指标中都没有突出效果,究其原因,可能是该方法的融合规则过于简单,或是所选取的图像场景不适用于此方法。

九种融合方法在六种典型场景中的实验结果表明,场景对图像融合效果的影响显著,根据实际场景选择合适的融合方法可以提高算法的实用性(表1)。金字塔变换和小波变换的融合算法简单,适用于光线充足、近距离拍摄的场景,如设备的无损检测。NSCT 信息保留度高,适用于背景复杂的场景,如人员搜救。NSST 实时性高,可用于机载夜视监控、农业自动化、智能交通等视频融合领域。具有高对比度的 GFF 融合方法,可应用于夜视环境下目标的侦察与识别,但不适用于烟雾场景。对烟尘透过能力强的 CVT-SR 在实际的森林火灾防控、火灾报警中具有一定的应用前景。基于神经网络的融合方法具有较强的自适应性,可用于复杂环境下的目标检测。

## 5 结论

文中综述了目前常用的红外与可见光图像融合方法的发展进程及应用研究,重点阐述基于多尺度变换、稀疏表示、神经网络等方法的核心思想、发展进程、优势不足,并对常用的融合方法进行总结对比;介绍了红外与可见光图像融合方法的应用领域,包括目标识别、目标检测与跟踪、夜视监控等;总结了九种目前较常使用的图像融合评价指标;最后针对六种典型场景,选择代表性的融合方法和六种评价指标进行实验对比。

针对当前红外与可见光图像融合算法存在的问题,有三方面改进建议:

(1) 基于多尺度变换的图像融合方法虽然已经成为图像融合领域的热门研究方向,但还可从固定基函

数和分解层数的自适应设计方面进一步改进算法;(2) 未来可根据不同场景的特点,对不同融合方法进行组合式创新。但在构建混合模型时,还需综合考虑算法的性能表现;(3) 基于深度学习的融合方法是未来的重点研究方向。近年来提出的端到端网络模型解决了卷积神经网络模型中存在的大部分问题,但还需根据红外与可见光图像成像原理来设计有针对性的损失函数,并在不同场景中采集大量的数据集,从而进一步提高该模型的泛化能力。

针对红外与可见光图像融合应用方面的挑战,可从以下方面进一步探索:

(1) 提高实际场景中红外与可见光图像的配准精度,将空间变换设为变量因素,实现配准和融合的同步,以减少伪影现象;(2) 减少红外图像的噪声,引入显著性检测算法,提取主要红外目标减少噪声干扰,或设计多孔径成像系统并对所得红外图像进行超分,在提高图像分辨率的同时扩大视场范围;(3) 提高融合算法的实时性,将并行运算应用于图像融合领域,实现算法时间和空间上的并行,从而提高运行效率。

## 参考文献:

- [1] 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.
- [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] Short N J, Yuffa A J, Videen G, et al. Effects of surface materials on polarimetric-thermal measurements: Applications to face recognition [J]. *Applied Optics*, 2016, 55(19): 5226-5233.
- [4] Heo J, Kong S G, Abidi B R. Fusion of visual and thermal signatures with eyeglass removal for robust face recognition[C]//Computer Vision and Pattern Recognition Workshop, 2004,19: 122-127.
- [5] Kumar K S, Kavitha G, Subramanian R, et al. MATLAB-A Ubiquitous Tool for the Practical Engineer[M]. Croatia: In Tech, 2011: 307-326.
- [6] Castillo J C, Fernandez-Caballero A, Serrano-Cuerda J, et al. Smart environment architecture for robust people detection by infrared and visible video fusion [J]. *Journal of Ambient Intelligence and Humanized Computing*, 2017, 8(2): 223-237.

- [7] Fendri E, Boukhriss R R, Hammami M. Fusion of thermal infrared and visible spectra for robust moving object detection [J]. *Pattern Analysis and Applications*, 2017, 20(4): 907-926.
- [8] Apatean A, Rogozan A, Bensrhair A. Visible-infrared fusion schemes for road obstacle classification [J]. *Transportation Research Part C-Emerging Technologies*, 2013, 35: 180-192.
- [9] Bulanon D M, Burks T F, Alchanatis V. Image fusion of visible and thermal images for fruit detection [J]. *Biosystems Engineering*, 2009, 103(1): 12-22.
- [10] Raza S E A, Sanchez V, Prince G, et al. Registration of thermal and visible light images of diseased plants using silhouette extraction in the wavelet domain [J]. *Pattern Recognition*, 2015, 48(7): 2119-2128.
- [11] Burt P, Adelson E. The laplacian pyramid as a compact image code [J]. *IEEE Transactions on Communications*, 1983, 31(4): 532-540.
- [12] Toet A. Image fusion by a ration of low-pass pyramid [J]. *Pattern Recognition Letters*, 1989, 9(4): 245-253.
- [13] Toet A, Vanruyven L J, Valetton J M. Merging thermal and visual images by a contrast pyramid [J]. *Optical Engineering*, 1989, 28(7): 789-792.
- [14] Toet A. A morphological pyramidal image decomposition [J]. *Pattern Recognition Letters*, 1989, 9(4): 255-261.
- [15] Freeman W T, Adelson E H, Intel M. The design and use of steerable filters [J]. *IEEE Transpattern Anal*, 1991, 13(9): 891-906.
- [16] Yu X L, Ren J L, Chen Q, et al. A false color image fusion method based on multi-resolution color transfer in normalization YCBCR space [J]. *Optik*, 2014, 125(20): 6010-6016.
- [17] Jin H Y, Jiao L C, Liu F, et al. Fusion of infrared and visual images based on contrast pyramid directional filter banks using clonal selection optimizing [J]. *Optical Engineering*, 2008, 47(2): 27002-27008.
- [18] He D X, Meng Y, Wang C Y. Contrast pyramid based image fusion scheme for infrared image and visible image[C]//2011 IEEE International Geoscience and Remote Sensing Symposium, 2011: 597-600.
- [19] Grossmann A, Morlet J. Decomposition of hardy functions into square integrable wavelets of constant shape [J]. *Siam Journal on Mathematical Analysis*, 1984, 15(4): 723-736.
- [20] Mallat S G. A theory for multiresolution signal decomposition-the wavelet representation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1989, 11(7): 674-693.
- [21] Niu Y, Xu S, Wu L, et al. Airborne infrared and visible image fusion for target perception based on target region segmentation and discrete wavelet transform [J]. *Mathematical Problems in Engineering*, 2012, 2012: 732-748.
- [22] Madheswari K, Venkateswaran N. Swarm intelligence based optimisation in thermal image fusion using dual tree discrete wavelet transform [J]. *Quantitative Infrared Thermography Journal*, 2017, 14(1): 24-43.
- [23] Zou Y, Liang X, Wang T. Visible and infrared image fusion using the lifting wavelet [J]. *Telkomnika Indonesian Journal of Electrical Engineering*, 2013, 11(11): 6290-6295.
- [24] Chai P F, Luo X Q, Zhang Z C. Image fusion using quaternion wavelet transform and multiple features [J]. *IEEE Access*, 2017, 5: 6724-6734.
- [25] Yan X, Qin H L, Li J, et al. Infrared and visible image fusion with spectral graph wavelet transform [J]. *Journal of the Optical Society of America a-Optics Image Science and Vision*, 2015, 32(9): 1643-1652.
- [26] Tao G Q, Li D P, Lu G H. On image fusion based on different fusion rules of wavelet transform [J]. *Acta Photonica Sinica*, 2004, 33(2): 221-224.
- [27] Selesnick I W, Baraniuk R G, Kingsbury N G. The dual-tree complex wavelet transform [J]. *IEEE Signal Processing Magazine*, 2005, 22(6): 123-151.
- [28] Da Cunha A L, Zhou J P, Do M N. The nonsubsampling contourlet transform: theory, design, and applications [J]. *IEEE Transactions on Image Processing*, 2006, 15(10): 3089-3101.
- [29] Yin S, Cao L, Tan Q, et al. Infrared and visible image fusion based on NSCT and fuzzy logic[C]//Proceedings of the 2010 IEEE International Conference on Mechatronics and Automation, 2010,5: 671-675.
- [30] Liu H X, Zhu T H, Zhao J J. Infrared and visible image fusion based on region of interest detection and nonsubsampling contourlet transform [J]. *Journal of Shanghai Jiaotong University (Science)*, 2013, 18(5): 526-534.
- [31] Guo K, Labate D. Optimally sparse multidimensional representation using shearlets [J]. *Siam Journal on Mathematical Analysis*, 2007, 39(1): 298-318.
- [32] 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.
- [33] Kong W W, Wang B H, Lei Y. Technique for infrared and visible image fusion based on non-subsampling shearlet transform and spiking cortical model [J]. *Infrared Physics & Technology*, 2015, 71: 87-98.

- [34] Kong W, Lei Y, Zhao H. Adaptive fusion method of visible light and infrared images based on non-subsampled shearlet transform and fast non-negative matrix factorization [J]. *Infrared Physics & Technology*, 2014, 67: 161-172.
- [35] Hu H M, Wu J W, Li B, et al. An adaptive fusion algorithm for visible and infrared videos based on entropy and the cumulative distribution of gray levels [J]. *IEEE Transactions on Multimedia*, 2017, 19(12): 2706-2719.
- [36] Zhang X Y, Ma Y, Fan F, et al. Infrared and visible image fusion via saliency analysis and local edge-preserving multi-scale decomposition [J]. *Journal of the Optical Society of America a-Optics Image Science and Vision*, 2017, 34(8): 1400-1410.
- [37] Yang B, Li S T. Multifocus image fusion and restoration with sparse representation [J]. *IEEE Transactions on Instrumentation and Measurement*, 2010, 59(4): 884-892.
- [38] 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.
- [39] Yin H T. Sparse representation with learned multiscale dictionary for image fusion [J]. *Neurocomputing*, 2015, 148: 600-610.
- [40] Yang B, Li S T. Pixel-level image fusion with simultaneous orthogonal matching pursuit [J]. *Information Fusion*, 2012, 13(1): 10-19.
- [41] Liu Y, Wang Z F. Simultaneous image fusion and denoising with adaptive sparse representation [J]. *Iet Image Processing*, 2015, 9(5): 347-357.
- [42] Yin H T, Li S T. Multimodal image fusion with joint sparsity model [J]. *Optical Engineering*, 2011, 50(6): 067007-067009.
- [43] Nejati M, Samavi S, Shirani S. Multi-focus image fusion using dictionary-based sparse representation [J]. *Information Fusion*, 2015, 25: 72-84.
- [44] Wang J, Peng J Y, Feng X Y, et al. Fusion method for infrared and visible images by using non-negative sparse representation [J]. *Infrared Physics & Technology*, 2014, 67: 477-489.
- [45] Zhang Q, Levine M D. Robust multi-focus image fusion using multi-task sparse representation and spatial context [J]. *IEEE Transactions on Image Processing*, 2016, 25(5): 2045-2058.
- [46] Zhang Q H, Fu Y L, Li H F, et al. Dictionary learning method for joint sparse representation-based image fusion [J]. *Optical Engineering*, 2013, 52(5): 1-11.
- [47] Yu N N, Qiu T S, Bi F, et al. Image features extraction and fusion based on joint sparse representation [J]. *IEEE Journal of Selected Topics in Signal Processing*, 2011, 5(5): 1074-1082.
- [48] Engan K, Aase S O, Husoy J H. Method of optimal directions for frame design[C]//1999 IEEE International Conference on Acoustics, Speech, and Signal Processing, 1999: 2443-2446.
- [49] Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation [J]. *IEEE Transactions on Signal Processing*, 2006, 54(11): 4311-4322.
- [50] Rubinstein R, Zibulevsky M, Elad M. Double sparsity: Learning sparse dictionaries for sparse signal approximation [J]. *IEEE Transactions on Signal Processing*, 2010, 58(3): 1553-1564.
- [51] Kim M, Han D K, Ko H. Joint patch clustering-based dictionary learning for multimodal image fusion [J]. *Information Fusion*, 2016, 27: 198-214.
- [52] Dong W S, Li X, Zhang L, et al. Sparsity-based image denoising via dictionary learning and structural clustering[C]//2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011: 457-464.
- [53] Chatterjee P, Milanfar P. Clustering-based denoising with locally learned dictionaries [J]. *IEEE Transactions on Image Processing*, 2009, 18(7): 1438-1451.
- [54] Yao Y, Guo P, Xin X, et al. Image fusion by hierarchical joint sparse representation [J]. *Cognitive Computation*, 2014, 6(3): 281-292.
- [55] Ophir B, Lustig M, Elad M. Multi-scale dictionary learning using wavelets [J]. *IEEE Journal of Selected Topics in Signal Processing*, 2011, 5(5): 1014-1024.
- [56] Lu X Q, Zhang B H, Zhao Y, et al. Theinfrared and visible image fusion algorithm based on target separation and sparse representation [J]. *Infrared Physics & Technology*, 2014, 67: 397-407.
- [57] Zhu Z Q, Yin H P, Chai Y, et al. A novel multi-modality image fusion method based on image decomposition and sparse representation [J]. *Information Sciences*, 2018, 432: 516-529.
- [58] Wang K P, Qi G Q, Zhu Z Q, et al. A novel geometric dictionary construction approach for sparse representation based image fusion [J]. *Entropy*, 2017, 19(7): 306.
- [59] Zhang Q, Liu Y, Blum R S, et al. Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: A review [J]. *Information Fusion*, 2018, 40: 57-75.
- [60] Kong W W, Zhang L J, Lei Y. Novel fusion method for visible light and infrared images based on NSST-SF-PCNN [J]. *Infrared Physics & Technology*, 2014, 65: 103-112.
- [61] Xiang T Z, Yan L, Gao R R. A fusion algorithm for infrared

- and visible images based on adaptive dual-channel unit-linking PCNN in NSCT domain [J]. *Infrared Physics & Technology*, 2015, 69: 53-61.
- [62] Ma L J, Zhao C H. An effective image fusion method based on nonsubsampling contourlet transform and pulse coupled neural network[C]//Proceedings of the 2nd International Conference on Computer and Information Applications (ICCIA 2012), 2012: 8-12.
- [63] Li Y, Song G-H, Yang S-C. Multi-sensor image fusion by NSCT-PCNN transform[C]//2011 IEEE International Conference on Computer Science and Automation Engineering, 2011: 638-642.
- [64] Kong W W, Lei Y J, Lei Y, et al. Image fusion technique based on non-subsampling contourlet transform and adaptive unit-fast-linking pulse-coupled neural network [J]. *Let Image Processing*, 2011, 5(2): 113-121.
- [65] Qu X B, Yan J W, Xiao H Z, et al. Image fusion algorithm based on spatio frequency-motivated pulse coupled neural networks in nonsubsampling contourlet transform Domain [J]. *Acta Automatica Sinica*, 2009, 12(34): 1508-1514.
- [66] El-taweel G S, Helmy A K. Image fusion scheme based on modified dual pulse coupled neural network [J]. *Let Image Processing*, 2013, 7(5): 407-414.
- [67] Yu Z, Yan L, Han N, et al. Image fusion algorithm based on contourlet transform and PCNN for detecting obstacles in forests [J]. *Cybernetics and Information Technologies*, 2015, 15(1): 116-125.
- [68] Liu S, Piao Y, Tahir M. Research on fusion technology based on low-light visible image and infrared image [J]. *Optical Engineering*, 2016, 55(12): 123104.
- [69] Li H, Wu X J, Kittler J. Infrared and visible image fusion using a deep learning framework[C]//2018 24th International Conference on Pattern Recognition, 2018: 2705-2710.
- [70] Ren X, Meng F, Hu T, et al. Infrared-visible image fusion based on convolutional neural networks (CNN)[C]//International Conference on Intelligent Science and Big Data Engineering, 2018: 301-307.
- [71] Ma J Y, Yu W, Liang P W, et al. FusionGAN: A generative adversarial network for infrared and visible image fusion [J]. *Information Fusion*, 2019, 48: 11-26.
- [72] Li H, Liu L, Huang W, et al. An improved fusion algorithm for infrared and visible images based on multi-scale transform [J]. *Infrared Physics & Technology*, 2016, 74: 28-37.
- [73] Fu Z Z, Wang X, Xu J, et al. Infrared and visible images fusion based on RPCA and NSCT [J]. *Infrared Physics & Technology*, 2016, 77: 114-123.
- [74] Cvejic N, Bull D, Canagarajah N. Region-based multimodal image fusion using ICA bases [J]. *IEEE Sensors Journal*, 2007, 7(5): 743-751.
- [75] Mou J, Gao W, Song Z. Image fusion based on non-negative matrix factorization and infrared feature extraction[C]// 2013 6th International Congress on Image and Signal Processing (CISP), 2013: 1046-1050.
- [76] 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.
- [77] Bavirisetti D P, Dhuli R. Two-scale image fusion of visible and infrared images using saliency detection [J]. *Infrared Physics & Technology*, 2016, 76: 52-64.
- [78] Gan W, Wu X H, Wu W, et al. Infrared and visible image fusion with the use of multi-scale edge-preserving decomposition and guided image filter [J]. *Infrared Physics & Technology*, 2015, 72: 37-51.
- [79] Cui G M, Feng H J, Xu Z H, et al. Detail preserved fusion of visible and infrared images using regional saliency extraction and multi-scale image decomposition [J]. *Optics Communications*, 2015, 341: 199-209.
- [80] Zhao J F, Zhou Q, Chen Y T, et al. Fusion of visible and infrared images using saliency analysis and detail preserving based image decomposition [J]. *Infrared Physics & Technology*, 2013, 56: 93-99.
- [81] Zhang B H, Lu X Q, Pei H Q, et al. A fusion algorithm for infrared and visible images based on saliency analysis and non-subsampling shearlet transform [J]. *Infrared Physics & Technology*, 2015, 73: 286-297.
- [82] Meng F, Song M, Guo B, et al. Image fusion based on object region detection and non-subsampling contourlet transform [J]. *Computers & Electrical Engineering*, 2017, 62: 375-383.
- [83] Cai J J, Cheng Q M, Peng M J, et al. Fusion of infrared and visible images based on nonsubsampling contourlet transform and sparse K-SVD dictionary learning [J]. *Infrared Physics & Technology*, 2017, 82: 85-95.
- [84] Yin M, Duan P H, Liu W, et al. A novel infrared and visible image fusion algorithm based on shift-invariant dual-tree complex shearlet transform and sparse representation [J]. *Neurocomputing*, 2017, 226: 182-191.
- [85] Chai Y, Li H F, Qu J F. Image fusion scheme using a novel dual-channel PCNN in lifting stationary wavelet domain [J].



- Optics Communications*, 2010, 283(19): 3591-3602.
- [86] Yang B, Li S T. Visual attention guided image fusion with sparse representation [J]. *Optik*, 2014, 125(17): 4881-4888.
- [87] 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.
- [88] Kong W W. Technique for gray-scale visual light and infrared image fusion based on non-subsampled shearlet transform [J]. *Infrared Physics & Technology*, 2014, 63: 110-118.
- [89] Adu J H, Gan J H, Wang Y, et al. Image fusion based on nonsubsampling contourlet transform for infrared and visible light image [J]. *Infrared Physics & Technology*, 2013, 61: 94-100.
- [90] Liu Z, Tsukada K, Hanasaki K, et al. Image fusion by using steerable pyramid [J]. *Pattern Recognition Letters*, 2001, 22(9): 929-939.
- [91] G Liu, Z L Jing, S Y Sun, et al. Image fusion based on expectation maximization algorithm and steerable pyramid [J]. *Chinese Optics Letters*, 2004, 2(7): 18-21.
- [92] Deng H, Ma Y. Image Fusion based on steerable pyramid and PCNN[C]//2009 Second International Conference on the Applications of Digital Information and Web Technologies, 2009: 569-573.
- [93] Zhan L, Zhuang Y, Huang L. Infrared and visible images fusion method based on discrete wavelet transform [J]. *Journal of Computers*, 2017, 28(2): 057-071.
- [94] Saeedi J, Faez K. Infrared and visible image fusion using fuzzy logic and population-based optimization [J]. *Applied Soft Computing*, 2012, 12(3): 1041-1054.
- [95] Chang L H, Feng X C, Zhang R, et al. Image decomposition fusion method based on sparse representation and neural network [J]. *Applied Optics*, 2017, 56(28): 7969-7977.
- [96] Omri F, Fofou S, Abidi M. NIR and visible image fusion for improving face recognition at long distance[C]//International Conference on Image and Signal Processing, 2014: 549-557.
- [97] Singh S, Gyaourova A, Bebis G, et al. Infrared and visible image fusion for face recognition[C]//Biometric Technology for Human Identification, International Society for Optics and Photonics, 2004: 585-596.
- [98] Heo J, Kong S G, Abidi B R, et al. Fusion of visual and thermal signatures with eyeglass removal for robust face recognition[C]//2004 Conference on Computer Vision and Pattern Recognition Workshop, 2004: 122-122.
- [99] Abaza A, Bourlai T. On ear-based human identification in the mid-wave infrared spectrum [J]. *Image Vision Computing*, 2013, 31(9): 640-648.
- [100] Uzair M, Mahmood A, Mian A, et al. Periocular region-based person identification in the visible, infrared and hyperspectral Imagery [J]. *Neurocomputing*, 2015, 149: 854-867.
- [101] Han J G, Pauwels E J, de Zeeuw P. Fast saliency-aware multi-modality image fusion [J]. *Neurocomputing*, 2013, 111: 70-80.
- [102] Schnelle S R, Chan A L. Enhanced target tracking through infrared-visible image fusion[C]//14th International Conference on Information Fusion, 2011: 1-8.
- [103] 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.
- [104] Toet A. Natural colour mapping for multiband nightvision imagery [J]. *Information Fusion*, 2003, 4(3): 155-166.
- [105] Toet A, Hogervorst M A. Progress in color night vision [J]. *Optical Engineering*, 2012, 51(1): 010901.
- [106] Davis J W, Sharma V. Background-subtraction using contour-based fusion of thermal and visible imagery [J]. *Computer Vision and Image Understanding*, 2007, 107(2-3): 162-182.
- [107] Niu Y F, Xu S T, Wu L Z, et al. Airborne infrared and visible image fusion for target perception based on target region segmentation and discrete wavelet transform [J]. *Mathematical Problems in Engineering*, 2012, 10: 732-748.
- [108] Bhatnagar G, Liu Z. A novel image fusion framework for night-vision navigation and surveillance [J]. *Signal Image and Video Processing*, 2015, 9: 165-175.
- [109] Paramanandham N, Rajendiran K. Multi sensor image fusion for surveillance applications using hybrid image fusion algorithm [J]. *Multimedia Tools and Applications*, 2018, 77(10): 12405-12436.
- [110] Tzagaris V, Anastassopoulos V. Fusion of visible and infrared imagery for night color vision [J]. *Displays*, 2005, 26(4): 191-196.
- [111] Hogervorst M A, Toet A. Fast natural color mapping for night-time imagery [J]. *Information Fusion*, 2010, 11(2): 69-77.
- [112] Mendoza F, Lu R F, Cen H Y. Comparison and fusion of four nondestructive sensors for predicting apple fruit firmness and soluble solids content [J]. *Postharvest Biology and Technology*, 2012, 73: 89-98.
- [113] Hanna B V, Gorbach A M, Gage F A, et al. Intraoperative assessment of critical biliary structures with visible range/infrared image fusion [J]. *Journal of the American College of Surgeons*, 2008, 206(6): 1227-1231.
- [114] Eslami M, Mohammadzadeh A. Developing a spectral-based strategy for urban object detection from airborne hyper-

- spectral TIR and visible data [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, 9(5): 1808-1816.
- [115] Han L, Wulie B, Yang Y L, et al. Direct fusion of geostationary meteorological satellite visible and infrared images based on thermal physical properties [J]. *Sensors*, 2015, 15(1): 703-714.
- [116] Li H G, Ding W R, Cao X B, et al. Image registration and fusion of visible and infrared integrated camera for medium-altitude unmanned aerial vehicle remote sensing [J]. *Remote Sensing*, 2017, 9(5): 441-469.
- [117] Chang X, Jiao L C, Liu F, et al. Multicontourlet-based adaptive fusion of infrared and visible remote sensing images [J]. *IEEE Geoscience and Remote Sensing Letters*, 2010, 7(3): 549-553.
- [118] Lu X C, Zhang J P, Li T, et al. Synergetic classification of long-wave infrared hyperspectral and visible images [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2015, 8(7): 3546-3557.
- [119] Gargano M, Bertani D, Greco M, et al. A perceptual approach to the fusion of visible and NIR images in the examination of ancient documents [J]. *Journal of Cultural Heritage*, 2015, 16(4): 518-525.
- [120] Kim S J, Deng F, Brown M S. Visual enhancement of old documents with hyperspectral imaging [J]. *Pattern Recognition*, 2011, 44(7): 1461-1469.
- [121] Feng Z J, Zhang X L, Yuan L Y, et al. Infrared target detection and location for visual surveillance using fusion scheme of visible and infrared images [J]. *Mathematical Problems in Engineering*, 2013, 2013(3): 831-842.
- [122] Zhao C, Guo Y, Wang Y. A fast fusion scheme for infrared and visible light images in NSCT domain [J]. *Infrared Physics & Technology*, 2015, 72: 266-275.
- [123] Zhang X L, Li X F, Li J. Validation and correlation analysis of metrics for evaluation performance of image fusion [J]. *Acta Automatica Sinica*, 2014, 40(2): 306-315. (in Chinese)
- [124] Aardt V, J an. Assessment of image fusion procedures using entropy, image quality, and multispectral classification [J]. *Journal of Applied Remote Sensing*, 2008, 2(1): 1-28.
- [125] Yun J, R ao. In-fibre Bragg grating sensors [J]. *Measurement Science Technology*, 1997, 8(4): 355.
- [126] Zhu X X, Bamler R. A sparse image fusion algorithm with application to pan-sharpening [J]. *IEEE Transactions on Geoscience Remote Sensing*, 2012, 51(5): 2827-2836.
- [127] Xydeas C S, V. P V. Objective image fusion performance measure [J]. *Military Technical Courier*, 2000, 56(4): 181-193.
- [128] Deshmukh M, Bhosale U. Image fusion and image quality assessment of fused images [J]. *International Journal of Image Processing*, 2010, 4(5): 484-508.
- [129] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: From error visibility to structural similarity [J]. *IEEE Trans Image Process*, 2004, 13(4): 600-612.
- [130] Li S, Yang B, Hu J. Performance comparison of different multi-resolution transforms for image fusion [J]. *Information Fusion*, 2011, 12(2): 74-84.
- [131] Qu X B, Yan J W, Xiao H Z, et al. Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampling contourlet transform domain [J]. *Acta Automatica Sinica*, 2008, 34(12): 1508-1514.