



计算机科学

COMPUTER SCIENCE

红外和可见光图像融合算法的研究进展

魏琦, 赵娟

引用本文

魏琦, 赵娟. 红外和可见光图像融合算法的研究进展[J]. 计算机科学, 2023, 50(2): 190-200.

WEI Qi, ZHAO Juan. Research Progress of Infrared and Visible Image Fusion Algorithms[J]. Computer Science, 2023, 50(2): 190-200.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于联合模型的端到端事件可信度识别](#)

End-to-End Event Factuality Identification with Joint Model

计算机科学, 2023, 50(2): 292-299. <https://doi.org/10.11896/jsjcx.211200108>

[基于图注意力的神经协同过滤社会推荐算法](#)

Neural Collaborative Filtering for Social Recommendation Algorithm Based on Graph Attention

计算机科学, 2023, 50(2): 115-122. <https://doi.org/10.11896/jsjcx.211200019>

[基于层级化数据记忆池的边缘侧半监督持续学习方法](#)

Hierarchical Memory Pool Based Edge Semi-supervised Continual Learning Method

计算机科学, 2023, 50(2): 23-31. <https://doi.org/10.11896/jsjcx.221100133>

[类脑计算脉冲神经网络模型及其学习算法研究进展](#)

Spiking Neural Network Model for Brain-like Computing and Progress of Its Learning Algorithm

计算机科学, 2023, 50(1): 229-242. <https://doi.org/10.11896/jsjcx.220100058>

[基于分割注意力与边界感知的中文嵌套命名实体识别算法](#)

Chinese Nested Named Entity Recognition Algorithm Based on Segmentation Attention and Boundary-aware

计算机科学, 2023, 50(1): 213-220. <https://doi.org/10.11896/jsjcx.211100257>

红外和可见光图像融合算法的研究进展

魏琦^{1,2} 赵娟¹

1 中国科学院深圳先进技术研究院 广东 深圳 518055

2 中国科学院大学 北京 100039

(qi.wei@sait.ac.cn)

摘要 红外图像便于识别热源目标,可见光图像包含丰富的纹理信息。红外和可见光的融合图像兼顾了两个波段传感器的优势,可以清楚地显示热源目标及其背景,在军事侦察、安防监控、遥感监测等领域有着广泛的应用,已成为图像融合领域的重点研究方向。近年来,国内外学者对红外和可见光图像融合算法开展了大量研究。文中首先对现有的图像融合算法进行了详细介绍,包括多尺度变换、稀疏表示的传统图像处理方法和基于CNN, GAN, AE这3种常见网络结构的深度学习图像融合算法。接着综述了融合图像的评价方法,对常见的多种客观评价指标进行了归类分析。然后开展对比实验,对各种方法进行了主观评价和定量分析,指出不同方法的优势和不足。最后,对红外和可见光图像融合技术的未来发展趋势进行展望。

关键词: 图像融合; 红外图像; 可见光图像; 神经网络; 评价准则

中图分类号 TP391

Research Progress of Infrared and Visible Image Fusion Algorithms

WEI Qi^{1,2} and ZHAO Juan¹

1 Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong 518055, China

2 University of Chinese Academy of Sciences, Beijing 100039, China

Abstract Infrared images are easy to identify thermal targets, and visible images have rich texture information. The fusion of infrared and visible images takes the advantages of both optical bands which can clearly show the targets and background. It has been widely used in many fields such as military reconnaissance, security monitoring, remote sensing measurement, and becomes a key research direction in the field of image fusion. In recent years, infrared and visible image fusion algorithms have attracted the attention of researchers around the world and have been studied abundantly. In this paper, the image fusion algorithms are introduced firstly, including traditional image processing methods of multi-scale transformation, sparse representation, and deep learning algorithms based on CNN, GAN, AE. Then, the evaluation methods of fusion images are summarized, and a variety of common objective evaluation indexes are classified. After that, comparative experiments are carried out to subjectively evaluate and quantitatively analyze the advantages and disadvantages of these algorithms. Finally, the development trend of infrared and visible image fusion methods is prospected.

Keywords Image fusion, Infrared image, Visible image, Neural network, Evaluation metric

1 引言

随着传感器技术的不断迭代,获取到的图像逐渐趋向于高信息量、高分辨率与多源性。单一的信息处理方法无法综合多源信息的特点,以满足多样化的分析处理需求,由此诞生了图像融合这一处理技术。图像融合是将多个传感器获取到的图像信息进行综合,得到一幅包含不同传感器成像特点、全面反映场景信息的融合图像^[1]。红外成像技术可以在不受光照、气候等的影响下得到清晰的热辐射目标,但图像的背景

模糊,纹理信息缺失;可见光成像技术通过反射可见光获得的图像细节丰富,但成像条件苛刻,易受自然气候影响^[2]。红外和可见光图像融合技术通过对红外和可见光图像进行合并处理,可得到包含清晰的热成像目标以及背景信息丰富的单张图像,已被广泛应用于各个领域,如军事侦察^[3-4]、目标识别追踪^[5-6]、安防监控^[7-8]、农业生产^[9-10]、遥感测量^[11-12]等。

自20世纪80年代起^[13-15],基于变换域的多源图像融合算法逐渐成为研究重点,这类算法将图像转换到变换域,构建融合策略对变换系数进行整合,进而重建得到融合图像。典型

到稿日期:2022-01-09 返修日期:2022-07-13

基金项目:国防科技创新特区项目(20-163-00-KX-001-002-02, 20-163-00-KX-001-003-02)

This work was supported by the National Defense Science and Technology Innovation Special Zone Project(20-163-00-KX-001-002-02, 20-163-00-KX-001-003-02).

通信作者:赵娟(juan.zhao@siat.ac.cn)

的方法包括基于金字塔、小波、轮廓波的多尺度变换^[14-17]和基于稀疏表示^[18-20]的方法。近年来,随着深度学习技术在计算机视觉和图像处理领域的发展,涌现出了一大批基于深度学习的红外和可见光图像融合算法,且取得了不俗的成就^[21-25]。

现有的红外与可见光图像融合的综述中,较少有文献从传统图像处理方法和深度学习理论两方面详细阐述,并且缺少对融合图像质量评价方法的统一归纳。本文首先综述了红外和可见光图像融合算法的研究进展,具体包括多尺度变换、稀疏表示和基于主流网络结构的深度学习方法。接着总结了图像融合结果的评价方式,针对4种传统图像融合算法和6种深度学习融合算法在相同数据集上开展实验,并利用8种评价准则对融合结果进行对比分析,指出了所选取的对比算法的优劣和应用场景。最后对红外与可见光图像融合技术的发展趋势进行了总结与展望。

2 红外和可见光图像融合算法

2.1 基于传统图像处理方法的融合算法

2.1.1 多尺度变换融合算法

在过去几十年里,图像融合领域对基于多尺度变换算法的研究从未停止^[1,14-17]。通过多尺度变换获得的融合图像具有良好的符合人眼特性的视觉观感。具体的融合过程如图1所示。

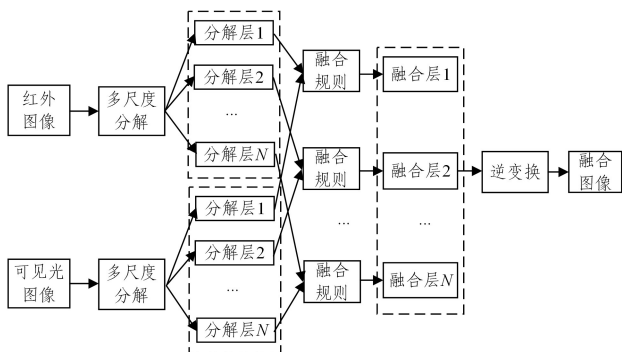


图1 基于多尺度变换的红外与可见光图像融合框架

Fig. 1 Infrared and visible image fusion framework based on multi-scale transformation

首先把源图像分解为不同尺度来提取特征,使用相应的融合策略对每一尺度的特征分别进行处理,然后通过对应的逆变换来获得融合结果图像。基于多尺度分解方法的关键是选择适当的分解方法和融合规则^[2],下文将从这两方面详细介绍。

(1) 金字塔变换

金字塔变换最早于20世纪80年代被提出^[13],通过对图像进行多次滤波,图像的分辨率随着分解层数的增加而降低,从而形成金字塔式的多层结构。在不同的分解层针对源图像特点,设计特定的融合规则对相应的图像金字塔进行合并,达到突出源图像特征与细节的目的。常见的基于金字塔变换的红外与可见光图像融合算法,按不同的金字塔形式可划分为拉普拉斯金字塔(Laplacian Pyramid, LP)变换^[13,16]、对比度金字塔(Contrast Pyramid, CP)变换^[14,26]、形态学金字塔变换^[27]等。

LP变换通过对源图像进行低通滤波和下采样的迭代操作得到高斯金字塔,再将高斯金字塔的每一层图像与其上一层上采样后得到的图像做差分运算得到拉普拉斯金字塔^[13]。通过插值运算,LP变换突出了图像的高频信息,可以有效保留可见光图像中的丰富细节,但融合图像的对比度较低,视觉观感不佳。CP变换与LP变换类似,其每一层图像通过计算高斯金字塔相邻两层图像的比率得到。与LP变换相比,其分解得到的子带对局部对比度更为敏感,改善了融合图像的视觉观感^[26]。

相比将多源图像自身或其局部作为特征,直接在原像素空间进行操作的空间域图像融合算法——金字塔变换开辟了多尺度变换图像融合研究的基本思路,在保留目标细节方面得到了显著提升。但其也具有一定的局限性,如金字塔相邻尺度的图像信息冗余度较高,且在分解的过程中因忽略了数据相关性而损失了图像特征等。随着小波变换理论的提出,基于小波变化的图像融合算法的研究得到了进一步扩展。

(2) 小波变换

小波变换的概念最早出现在文献^[28]中,其相比金字塔变换具备更合理的信息保留能力和更完善的图像重构能力,且图像在分解过程中没有信息损失和冗余信息堆积。基于小波变换的图像融合方法的研究主要包括两个方面:融合规则的设计和分解重构算法的选择^[29]。融合规则将在后文中详细介绍,而根据不同的分解重构算法出现了多种小波变换形式,包括离散小波变换^[29-30]、双树离散小波变换^[31-32]、提升小波变换^[33-34]等。

离散小波变换通过滤波器组将源图像分解为一系列低频和高频子带,但在分解重建的过程中,缺乏平移不变性和方向选择性,使得融合图像的边缘比较模糊^[29]。双树离散小波使用两个实小波来实现复小波变换,从而获得6个方向的细节信息,取得了比离散小波变换更好的方向选择性。同时采用的两组分解滤波器具有不同的延迟,可以消除平移引起的混叠现象,改善了平移不变性^[32]。提升小波变换被看作是第二代小波变换,它直接在稀疏域上计算,很大程度上降低了计算复杂度,节省了存储空间,基于提升机制的小波变换多被用于快速实时融合的场景中。

小波变换相比金字塔变换取得了很大的提升,但各向同性的特点使其在表示图像边缘、轮廓时效率不高,只能提取有限方向上的梯度信息,没有充分利用图像本身的几何特征。研究者们从人类视觉特性出发对小波分析理论进行了改进和丰富,提出了基于多尺度几何分析的图像融合方法^[35]。

(3) 非下采样多尺度变换

轮廓波变换最早是由Do等^[35]于2002年提出的,这种表示方法充分考虑到图像的边缘、轮廓信息,具有良好的频域、空域、方向性和各向异性的特点。但在变换过程中采样操作的加入导致平移不变性缺失,使图像出现吉布斯现象。为解决这一问题,出现了非下采样轮廓波变换(Non-Subsampled Contourlet Transform, NSCT)^[36]。NSCT由非下采样金字塔式滤波器组和非下采样方向滤波器组组成,在轮廓波变换的基础上,通过对分解和合成滤波器的上/下采样取代对图像的下/上采样操作,保证了平移不变性,同时减少了分解和重建

过程中信息的损失。

Meng 等^[37]将 NSCT 与显著性检测算法结合,整合红外图像提取到的显著目标到融合结果图上,使得融合结果保存了丰富的细节,且获得了显著的红外特征。但 NSCT 分解过程复杂,算法运行效率较低。在此基础上,出现了非下采样剪切波变换(Non-Subsampled Shearlet Transform, NSST)^[38-39],其借助剪切波更好的稀疏逼近性,在图像分解重建过程中不仅降低了计算复杂度,且获得的图像更加符合人眼特性,优化了视觉观感。

(4) 融合规则

经过多尺度变换得到的子带需要按照特定的融合规则处理,不同规则的优劣会直接影响融合图像的质量,包括基于像素和基于区域的两种规则。基于像素的规则直接将源图像像素点的灰度值作为融合权重的因子,过程简单、计算成本低、算法实时性好(包括加权平均法^[32]、绝对值取大法^[40]等),但缺乏局部区域内相邻像素(系数)之间的相关性,会损失一定的信息,且会造成图像边缘过于尖锐,影响视觉观感。基于区域的规则根据源图像窗口内的统计特性来决定融合权重,考虑了窗口内的结构特征(包括基于显著性区域^[39,37]以及基于能量^[29,41]等方法),但降低了对边缘的敏感性,容易引入虚假轮廓,且自适应性不强。

2.1.2 稀疏表示图像融合算法

稀疏表示方法将信号在一组过完备字典上进行分解,用尽可能少的原子来表示原始信号,从而得到一组稀疏向量^[18]。数学公式描述为:

$$X = \sum_{i=1}^K d_i \alpha_i = D\alpha \quad \text{s. t. } \min \|\alpha\| \quad (1)$$

其中, X 为单一原始信号, D 为过完备字典, α 为信号表示结果。为了让 α 满足稀疏性,加上 L_0 范数的稀疏性约束。

最早将稀疏表示用于图像融合领域是由 Yang 等^[42]提出的。首先从大量自然图片中学习得到一个过完备字典,然后在将源图像通过该字典稀疏表示的过程中增强显示有意义的特征^[20,43]。在多尺度变换的过程中,若分解层数过多,意外出现的配准误差和噪声会严重影响融合结果的图像质量。稀疏表示方法使用滑动窗口将源图像分解为多组重叠子块,在一定程度上避免了图像噪声引起的缺陷,提高了对非配准图像的鲁棒性。基于稀疏表示的红外与可见光图像融合框架如图 2 所示,其重点在于过完备字典的生成和稀疏分解算法^[40,42]。

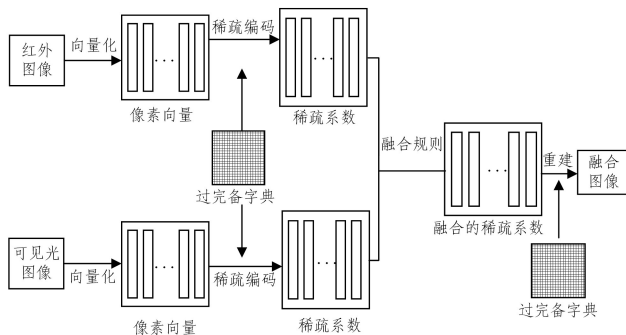


图 2 基于稀疏表示的红外与可见光图像融合框架

Fig. 2 Infrared and visible image fusion framework based on sparse representation

建造一个过完备字典的方法包括:基于数学模型的固定字典和基于样本学习的自适应字典。基于数学模型的方法采用数学分析的方式来构建,常见的模型包括离散余弦变换(Discrete Cosine Transform, DCT)^[44]和小波变换^[20]等。其中基于 DCT 的方法通过对离散余弦字典中的原子不断调制来获得过完备字典^[44],构建方式简单、高效,但字典一经生成,结构即固定,仅可对特定几何结构的图像进行稀疏表示。可学习的字典构建方法通过训练样本来获得有一定自适应性的更为灵活的字典。常用的方法有主成分分析(Principle Component Analysis, PCA)^[19]、基于奇异值分解的稀疏字典训练法(K -Singular Value Decomposition, K -SVD)^[45]等。

稀疏分解算法主要采用贪心策略,通过迭代从字典中逐步选择对应信号最优的基元素,利用系数和字典对输入信号进行表示。应用广泛的是正交匹配追踪(Orthogonal Matching Pursuit, OMP)算法^[46]、同步正交匹配追踪(Simultaneous OMP, SOMP)^[47]算法等。OMP 算法的思想是首先从过完备字典中选择与信号最匹配的原子,构建稀疏逼近,求出信号残差,然后继续选择与该残差最匹配的原子,不断迭代,直到差值在可以忽略的范围内。SOMP 在 OMP 基础上改进,考虑到了输入信号的内部一致性,简化了后续融合规则的设计,得到了更广泛的应用。

基于稀疏表示的融合方法减小了多尺度变换的配准误差影响,弥补了特征提取程度不足的问题,但仍具有局限性,在使用滑窗技术来分割图片时,会得到很多重叠子块,降低了算法运行效率。当源图像过于复杂时,过完备字典的表示能力有限,容易造成图像细节信息丢失,从而使稀疏表示的效果退化。

2.1.3 其他方法

在传统图像处理领域,还可以使用基于子空间的方法,包括主成分分析法^[48]、独立成分分析法^[49]、非负矩阵分解法^[50]等。这类方法把高维的源图像投射到低维空间中,因为在低维空间更容易捕捉到原始图像的内在结构,且处理低维子空间时的数据消耗的内存和空间减少,从而可以获得更高的处理效率。此外,可以结合多种方法的优点使用混合融合方法,如将多尺度变换和稀疏分解相结合^[51-53],将多尺度变换与显著性检测相结合^[54-55]等,在特殊应用场景中也获得了很好的融合效果。

2.2 基于深度学习方法的融合算法

对于图像融合任务,图像特征的提取效果会直接影响到融合性能。除了使用传统的滤波器、稀疏分解等方法来获得图像的高低频信息,更高效且充分地提取图像特征是改善融合图像性能的关键之一。近年来,深度学习技术在计算机视觉和图像处理领域飞速发展^[21-23],越来越多的研究人员将深度学习技术引入到图像融合领域,涌现出了一大批基于深度学习的图像融合算法。本节按常用网络结构,将这些算法分为卷积神经网络(Convolutional Neural Networks, CNN)、生成对抗网络(Generative Adversarial Networks, GAN)、自编码器(Auto-Encoder, AE)三大类别,分别进行阐述。

2.2.1 基于卷积神经网络的方法

在文献^[56]中, Liu 等提出了一种基于卷积稀疏表示

(Convolutional Sparse Representation, CSR) 的图像融合方法,CSR对整张图片进行最优提取,通过卷积操作,弥补了稀疏表示方法细节信息保存不足的缺点。随后,Liu等作者在文献[57]中引入CNN来解决多焦点图像融合,把融合任务看作图像分类问题,使用网络来产生决策图,用于决定每个清晰的图像块应来自于哪张源图像。

为了更充分地提取图像的深层信息,Li等设计了一种基于预训练网络的融合框架^[58],使用训练完成的VGG-19网络^[59]来提取图像高频信息。通过VGG网络,中间层的信息得到了有效提取并反映在最终的融合图像中,但因未对提取到的信息作进一步处理,其融合图像的观感提升有限。为了充分利用提取到的深度特征,Li等随后在文献[60]中提出了基于残差网络(Residual Neural Network, ResNet)的融合方法,使用训练完成的ResNet网络^[61]来获得深层特征,并对其做零阶成分分析(Zero-phase Component Analysis, ZCA)^[62]和L1-norm操作归一化处理得到初始权重,接着使用权重平均策略来得到融合图像。由于ResNet网络的特性,提取到的深度特征包含了多层信息,从而丰富了融合图像的细节信息。

借助CNN还可以建立更为通用的融合模型,在文献[63]中,整个图像融合框架(Image Fusion framework based on the Convolutional Neural Network, IFCNN)由特征提取、特征融合以及特征重建3个模块组成。采用两个卷积层来提取图像特征,特征融合模块对输入图像的多个卷积特征采取适当的融合策略进行处理,接着使用两个卷积层来重建融合特征得到融合结果图。在AAAI 2020上,Xu等^[64]提出了另一种无监督的网络框架,将原始图像的权重信息输入到稠密连接网络(DenseNet)^[65]的损失函数中,通过训练DenseNet根据不同图片的类型来提取特征并通过损失函数来重建图像。由于源图像权重是基于数据驱动的,通过衡量不同图像中的信息数量和质量获得,因此可以有选择性地保留源图像中有用的信息,应用于多种图像融合任务。此外,文献[66]提出了端到端通用模型,通过特征提取和信息测量,自动估计相应源图像的重要性,并提出自适应信息保留度,在多种图像融合问题中均取得了不错的性能。

2.2.2 基于生成对抗网络的方法

在一般的网络模型中,需要为模型提供参考图像(Ground Truth),在图像融合领域,很难定义结果图像的标准水平。现有的网络模型通过忽略参考图像、借助神经网络来确定每个来自源图像的图像块的模糊程度,通过计算权重来确定是否保留。使用端到端的生成式对抗网络模型,也可以一定程度避免上述问题^[67]。

Ma等在文献[22]中首次提出了融合生成网络框架(fusionGAN),文献中GAN网络首次被用于红外和可见光图像融合领域。整个网络的核心思想是使融合图像不断保存来自红外图像的热辐射信息和可见光图像的梯度信息。通过将两种图像在通道方向级联起来输入到生成器中,得到初始融合图。接着使用判别器来不断区分融合图像和可见光图像。通过生成器和判别器的博弈过程,融合图像将逐渐增加来自可见光图像的细节信息。但该方法生成的融合图像清晰度较低,保存的细节信息不足,且在有些情况下会引入预期之外的

噪声。随后Ma等改进了损失函数^[68],以提升细节信息的质量和锐化热辐射目标的边缘,使得最终的融合结果不仅含有可见光图像中丰富的纹理信息,还保留了红外图像中的对边缘锐化处理过的热辐射目标。

fusionGAN首次将生成对抗网络引入红外和可见光图像融合领域,在融合结果的视觉效果上取得了一定的突破,但其网络训练困难,计算效率不高。有研究者在其基础上,使用可学习的分组卷积(Group Convolution)来代替原生成器中的普通卷积层^[69],显著减少了计算量和内存消耗。同时利用多级残差网络作为基本网络构建单元,降低了网络训练难度,使融合结果的亮度和对比度得到增强,融合效果提升显著。

一般GAN网络中,生成器的输入均为红外或可见光图像,文献[70]创新性地将预融合好的图像作为标签数据,从而使生成器以预融合图像作为基准,调整生成器参数。其核心网络为相对耦合对抗神经网络(Relativistic Coupled GAN),通过该网络模型同时学习各个源图像的特征信息,提升了融合图像的纹理细节和热辐射信息的保留效果。此外,文献[71]提出了一种基于特征分解引导和GAN的混合模型,通过GAN网络合并低频分量获得基础层图像,将高频分量通过加权最小二乘法融合得到细节层图像,最终将基础层和细节层线性组合重建获得融合图像。

2.2.3 基于自编码器网络的方法

自编码器网络结构如图3所示,主要由3部分组成:编码器、融合层、解码器。其中编码器用于提取源图像的特征,经融合层处理后输入解码器中从而获得最终融合图像。

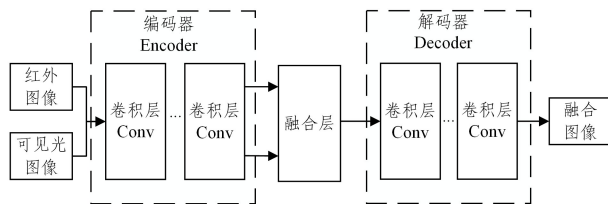


图3 基于自编码器的红外和可见光图像融合示意图

Fig. 3 Schematic diagram of infrared and visible image fusion based on auto-encoder network

在ICCV 2017^[23],Prabhakar等开启了图像融合领域使用AE网络的相关研究。其网络模型由5个卷积层组成,前两层作为编码器,使用孪生网络框架,固定权重,叠加提取到的特征输入到后三层作为解码器来重建得到融合图像。该方法取得了不错的融合效果,但网络框架相对简单,对深度特征的提取不够充分,不同网络层之间的连接没有做特殊处理,导致部分实用信息在迭代过程中损失。在此基础上,有研究者在文献[72]中将dense block^[65]整合到编码器中,使网络尽可能多地保存信息,利用每一层提取到的特征来构建融合特征图,分别使用叠加和l1-norm融合策略,将特征输入到解码器中来获得结果图像。

对于解码器也可以提出改进,在文献[73]中,Li等使用基于nest connection^[74]的网络框架,在网络层之间加入了上采样操作和跳转连接,从而在解码阶段更充分地利用深度特征来重建图像。文章还提出了基于空间注意力和通道注意力模型的融合策略,改善了图像的整体观感。文献[75]中提出

了另一种 AE 结构,其编码器含有细节和语义两个分支结构,细节分支使用稠密连接来提取浅层和边缘特征信息,语义分支经快速下采样 3 次,上采样为原始尺寸大小,来提取全局结构特征信息。最终通过通道融合策略,根据红外和可见光图像在各自通道的重要性来得到改善的增强特征。

上述方法均采用手动设计的融合规则来整合深层特征,一定程度上限制了融合性能。在文献[76]中,Xu 等提出基于显著性像素分类的 AE 网络融合框架,使用分类器计算特征图像素的显著性来得到融合权重,提高了融合效率。

表 1 按本文介绍的顺序列出了各种红外与可见光图像融合算法,并总结了每种方法的优缺点。

表 1 红外与可见光图像融合算法对比
Table 1 Comparison of infrared and visible image fusion methods

融合算法	具体方法	优势	劣势		
传统处理方法	金字塔变换	拉普拉斯金字塔	突出高频信息,计算效率高	图像细节信息保留不足	
		对比度金字塔	图像对比度高,视觉观感不错	分解过程信息冗余,计算效率低	
	多尺度变换	离散小波变换	纹理信息丰富,图像信噪比高	易出现图像振铃效应,导致轮廓模糊	
		小波变换	双树离散小波变换	更好的方向选择性,具备平移不变性	计算复杂,效率低
			提升小波变换	计算效率高,可实时处理	会损失一定的细节信息
	非下采样	非下采样轮廓波变换	消除了吉布斯效应,丰富了图像细节信息	分解过程复杂,运行效率低	
	多尺度变换	非下采样剪切波变换	降低了计算复杂度,视觉观感更好	配准要求高	
	稀疏表示	稀疏分解方法	算法稳定性好,配准误差的影响小,更好的亮度表现	计算复杂,轮廓及边缘信息模糊	
	其他方法	基于子空间	减少了内存消耗,计算效率高	对图像噪声敏感,配准要求高	
		混合方法	保留了各种方法的优点,适用场景丰富	设计复杂,难以做到实时处理	
基于卷积神经网络	卷积稀疏表示(CSR)[56]	弥补稀疏表示信息保留不足的问题,将卷积网络引入图像融合	仅用于多焦点图像融合问题,中间层信息损失严重		
	借助预训练完成的网络[58,60,77]	能够有效提取图像的多尺度深度特征,原理相对简单	缺乏训练过程,网络参数固定,模型性能有限		
	IFCNN ^[63] , FusionDN ^[64] U2Fusion ^[66]	网络泛化能力强,适用多种图像融合场景	并未针对红外和可见光融合任务优化融合规则,在特殊场景下表现不佳		
	fusionGAN ^[22,68]	首次引入 GAN 网络,避免了手工设计权重参数及融合规则	算法性能不够稳定,部分场景会损失信息		
基于深度学习	基于生成对抗网络	FLGC-fusionGAN ^[69]	减少了计算量和内存消耗,融合图像亮度和对比度增强	融合图像锐度不足,边缘模糊	
		相对耦合对抗网络 ^[70]	降低了网络训练难度,且融合图像细节信息丰富	需要手动标记预融合图像	
基于自编码器网络	deepFuse ^[23]	模型设计简单,计算效率高,启发了自编码网络的研究	对源图像的显著性信息提取不够完善,且中间层信息在迭代过程中损失		
	denseFuse ^[72]	尽可能多地保留了原始图像的特征信息,且网络易于训练	未优化融合策略,在特殊场景中红外目标不够显著		
	nestFuse ^[73] , rfnNest ^[25]	改进了融合策略和损失函数,结果图像细节丰富,显著性特征充分	计算复杂,实时性差,图像整体亮度偏低		

3 融合结果评价

在图像融合领域,对于生成的结果图像,缺乏参考标准(Ground Truth),如何准确评估融合结果图像的质量,一直是研究难点之一^[78]。评价融合图像质量的方式有两种,即主观评价方式和客观评价方式。

3.1 主观评价方式

主观评价方式根据观察者的主观感受来对图像质量做出评价,一般情况下为基于打分的方式。但具体实施起来需要

考虑参试者的人数设置,以及是否具有专业背景知识、屏幕分辨率、亮度设置等细节信息。由于依赖人类视觉系统主观感受得到的评价结果不够科学、全面,仅可以作为辅助评价准则,因此需要借助客观评价方式来做出进一步评判。

3.2 客观评价方式

客观评价方式是指采用定量指标对图像质量进行评价,本文参考了文献[79],将这些指标分为三大类别:基于信息论、基于图像特征以及基于结构相似性进行综述,其归纳总结如表 2 所列。

表 2 客观评价准则
Table 2 Objective evaluation criteria

评价准则	解释	
基于信息论	信息熵(EN)	统计图像中所包含的信息数量
	交叉熵(CE)	用来衡量源图像和融合图像之间差异程度
	互信息 MI	评价从源图像到融合图像所保留的信息数量
基于图像特征	平均梯度(AG)	通过计算梯度信息来表征图像清晰程度,反映了融合图像的细节和纹理信息
	基于边缘的相似性度量(Q ^{AB/F}) 空间频率(SF)	统计融合图像保存的源图像的边缘信息数量 用来衡量融合图像的梯度分布情况,从而反映细节和纹理信息
基于结构相似性	结构相似性(SSIM)	由相关度损失、亮度损失以及对比度失真程度组成,衡量融合图像和源图像的相似程度
	均方根误差(RMSE)	计算融合图像和源图像之间的误差来比较差异性

3.2.1 基于信息论的评价准则

(1) 信息熵

信息熵(Entropy, EN)是用来衡量图片中所包含的信息数量的评价指标^[80]。该数值越大,表示图片中所包含的信息越丰富,融合性能越好,其计算式为:

$$EN = - \sum_{l=0}^{N-1} p_l \log_2 p_l \quad (2)$$

其中, l 表示图像的灰度等级, p_l 表示融合图像中相应灰度级的归一化直方图。

(2) 交叉熵

交叉熵(Cross Entropy, CE)是用来衡量源图像和融合图像之间差异的评价指标^[9]。该数值越小,表示融合图像和源图像之间的差异越小,融合性能越好,其计算式为:

$$CE = \frac{Ce(A, F) + Ce(B, F)}{2} \quad (3)$$

其中, $Ce(A, F)$, $Ce(B, F)$ 分别表示源图像 A , B 和融合图像 F 的交叉熵,具体计算式为:

$$Ce(X, F) = \sum_{i=0}^{N-1} p_X(i) \log_2 \frac{p_X(i)}{p_F(i)} \quad (4)$$

其中, $p_X(i)$ 表示源图像中像素灰度值为 i 的像素数占整个图像像素数量的比例, N 为图像的灰度等级。

(3) 互信息

互信息(Mutual Information, MI)是用于衡量从源图像到融合图像所保留的信息数量的评价指标^[81]。该数值越大,即融合图像从源图像获得的信息越多,融合性能越好。其计算式为:

$$MI = MI_{A,F} + MI_{B,F} \quad (5)$$

其中, $MI_{A,F}$, $MI_{B,F}$ 分别表示融合图像从源图像中保留的信息数量,通过下式计算得到:

$$MI_{X,F} = \sum_{x,f} p_{X,F}(x, f) \log \frac{p_{X,F}(x, f)}{p_X(x) p_F(f)} \quad (6)$$

其中, $p_X(x)$ 和 $p_F(f)$ 分别为源图像和融合图像的边缘直方图,而 $p_{X,F}(x, f)$ 为源图像 x 和融合图像 f 的联合直方图。

3.2.2 基于图像特征的评价准则

(1) 平均梯度

平均梯度(Average Gradient, AG)是通过计算梯度信息来表征图像清晰程度的评价指标,其反映了融合图像的细节和纹理信息^[82]。该数值越大,融合图像信息越丰富,融合性能越好。其计算式为:

$$AG = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \sqrt{\frac{1}{2} (F(i, j) - F(i+1, j))^2 + (F(i, j) - F(i, j+1))^2} \quad (7)$$

其中, M 和 N 分别代表融合图像的高度和宽度, $F(i, j)$ 表示图像第 i 行第 j 列像素的值。

(2) 基于边缘的相似性度量 $Q^{AB/F}$

$Q^{AB/F}$ 是衡量融合图像保存的源图像的边缘信息数量的评价指标^[83]。计算得到的 $Q^{AB/F}$ 的取值范围为 $[0, 1]$, 其值越接近 0, 表示损失的边缘信息越多; 该值越大, 表示融合图像

中保存了越多的边缘信息, 融合性能越好。其计算式为:

$$Q^{AB/F} = \frac{\sum_{i=1}^M \sum_{j=1}^N Q^{AF}(i, j) \omega^A(i, j) + Q^{BF}(i, j) \omega^B(i, j)}{\sum_{i=1}^M \sum_{j=1}^N (\omega^A(i, j) + \omega^B(i, j))} \quad (8)$$

其中, $Q^{XF}(i, j) = Q_g^{XF}(i, j) Q_a^{XF}(i, j)$, $Q_g^{XF}(i, j)$ 和 $Q_a^{XF}(i, j)$ 分别表示在位置 (i, j) 处的边缘强度和方向值; ω^x 为权重值, 用于表示源图像在融合图像中所占的比例。

(3) 空间频率

空间频率(Spatial frequency, SF)是基于梯度的图像质量指标, 用于衡量融合图像的梯度分布情况, 从而反映细节和纹理信息^[84]。空间频率越大, 融合图像包含的边缘和纹理信息越丰富, 融合性能也就越好。其计算式为:

$$SF = \sqrt{RF^2 + CF^2} \quad (9)$$

其中, RF , CF 分别表示空间行频率和空间列频率, 可通过下式进一步得到:

$$RF = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=2}^N ((p(i, j) - p(i, j-1)))^2} \quad (10)$$

$$CF = \sqrt{\frac{1}{MN} \sum_{i=2}^M \sum_{j=1}^N (p(i, j) - p(i-1, j))^2} \quad (11)$$

其中, M 和 N 分别表示融合图像的高度和宽度, $p(i, j)$ 表示图像第 i 行第 j 列像素的值。

3.2.3 基于结构相似性的评价准则

(1) 结构相似性

结构相似性(Structural Similarity Index Measure, SSIM)是用来衡量融合图像结构损失和失真程度的评价指标^[85]。该数值越大, 表示融合性能越好。

结构相似性由 3 部分组成: 相关度损失、亮度损失以及对比度失真程度。计算式如下:

$$SSIM = SSIM_{A,F} + SSIM_{B,F} \quad (12)$$

$$SSIM_{X,F} = \sum_{x,f} \frac{2\mu_x\mu_f + C_1}{\mu_x^2 + \mu_f^2 + C_1} \cdot \frac{2\sigma_x\sigma_f + C_2}{\sigma_x^2 + \sigma_f^2 + C_2} \cdot \frac{\sigma_{xf} + C_3}{\sigma_x\sigma_f + C_3} \quad (13)$$

其中, $SSIM_{X,F}$ 表示源图像 X 和融合图像 F 的结构相似性值, 其中 μ_x , μ_f 和 σ_x , σ_f 分别代表源图像和融合图像的均值和标准差, $\sigma_{x,f}$ 表示源图像和融合图像的协方差, C_1 , C_2 , C_3 为正常数, 用于避免分母为 0。

(2) 均方根误差

均方根误差(Root Mean Squared Error, RMSE)是通过计算融合图像和源图像之间的误差来衡量差异性的评价指标^[79]。该数值越小, 表明融合过程产生的误差越小, 融合性能越好, 其计算式为:

$$RMSE = \frac{RMSE_{A,F} + RMSE_{B,F}}{2} \quad (14)$$

$$RMSE_{X,F} = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (X(i, j) - F(i, j))^2} \quad (15)$$

其中, $RMSE_{X,F}$ 表示源图像 X 和融合图像 F 的均方误差值, M 和 N 分别代表融合图像的高度和宽度, $X(i, j)$ 和 $F(i, j)$ 表示图像第 i 行第 j 列像素的值。

4 实验分析

本节从 TNO 数据库¹⁾ 中选取了 5 组红外和可见光图像,

¹⁾ <https://doi.org/10.6084/m9.figshare.1008029.v1>

如图 4 所示,分别为丛林、城市、海洋、烟雾和复杂光照 5 种不同的应用场景。使用具有代表性的 10 种算法得到融合图像,基于客观评价准则和计算定量指标,综合比较并分析。选取的融合方法包括:LP^[13]、NSCT_SR^[36]、引导滤波融合(Guided Filtering Fusion, GFF)^[17]、联合稀疏表示(Joint Sparse Representation, JSR)^[43]、CNN^[24]、ResNet^[60]、IFC-NN^[63]、fusionGAN^[22]、denseFuse^[72]、基于嵌套连接的残差融合网络(residual fusion network Nest connection, rfnNest)^[25]。其中 LP, NSCT_SR, GFF 为基于多尺度变换方法,JSR 为基于稀疏表示方法,CNN, ResNet, IFCNN 为基于卷积神经网络的方法,fusionGAN 为基于生成对抗网络的方法,denseFuse 和 rfnNest 为基于自编码器的方法。算法的融合效果如图 5 所示,图像分辨率均设置为 360×270 。图 6 给出了不同方法基于客观评价标准的统计结果。其中,融合方法均采用公开代码,所设参数与原文献一致,实验环境为 CPU Intel i7-10700k, GPU NVIDIA RTX 3070。

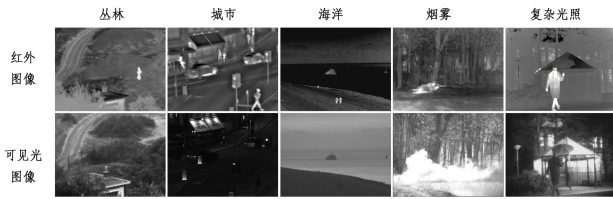


图 4 5 组实验原始图像

Fig. 4 Five groups of experimental original images

从图 5 主观视觉效果来整体评价,LP 方法观感良好,但细节信息保留不足,红外热辐射目标不够显著;NSCT_SR 方法突出了热辐射目标,但同样缺少细节,图像轮廓边缘模糊;GFF 方法丰富了图像细节,在传统图像处理方法中获得了最佳的视觉效果;JSR 方法提高了图像整体亮度,但在城市场景中引入了缺陷信息,在烟雾场景中忽略了热辐射目标。从 CNN, ResNet, IFCNN, denseFuse, rfnNest 方法的融合结果来看,图像整体对比度良好,纹理信息丰富,具有不错的视觉观感。fusionGAN 在所有深度学习方法中,对于细节信息的保留最少,且红外目标轮廓模糊,观感不佳。

选取图像中的复杂区域,按不同场景做进一步分析。对于丛林场景的热辐射目标,LP, ResNet, denseFuse 方法的人物亮度较暗,与背景区分不够显著;fusionGAN 和 rfnNest 方法目标边缘模糊,缺乏细节信息。在围栏的处理上,NSCT_SR, GFF, fusionGAN 方法普遍存在涂抹痕迹;LP 和 CNN 方法的处理相对比较自然。在城市场景中,选取的重点区域需要准确处理灯光边缘与广告牌上的信息。NSCT_SR 方法图像的文字无法辨认,画面模糊;JSR 方法引入了缺陷信息,降低了图像质量;CNN 和 IFCNN 方法对光源信息保留充分,图像亮度适宜;GFF 和 fusionGAN 方法物体边缘模糊,视觉观感不佳。对于海洋场景,选取的重点区域包括了红外目标与自然背景,可以看出,NSCT_SR 和 fusionGAN 方法的海水颜色呈现为黑色,难以与海岸区分;GFF 和 JSR 方法对天空的处理保留了无用的亮度信息,影响了图像完整性。在烟雾场景中,JSR 和 CNN 方法错误计算了红外图像权重,使得烟雾

中的士兵被完全隐藏;NSCT_SR 和 fusionGAN 方法则完全消除了烟雾,损失了可见光图像重要的场景信息。对于复杂光照场景,需要识别出人物和雨伞、房屋与天空的不同层级关系,NSCT_SR 方法最模糊,整幅图像均缺乏细节信息;GFF 方法人物不够显著,无法区分和雨伞的交界;JSR 方法偏亮;fusionGAN 方法偏暗,对比度不佳;CNN 方法整体观感最佳,最大程度地还原了场景中的丰富信息。

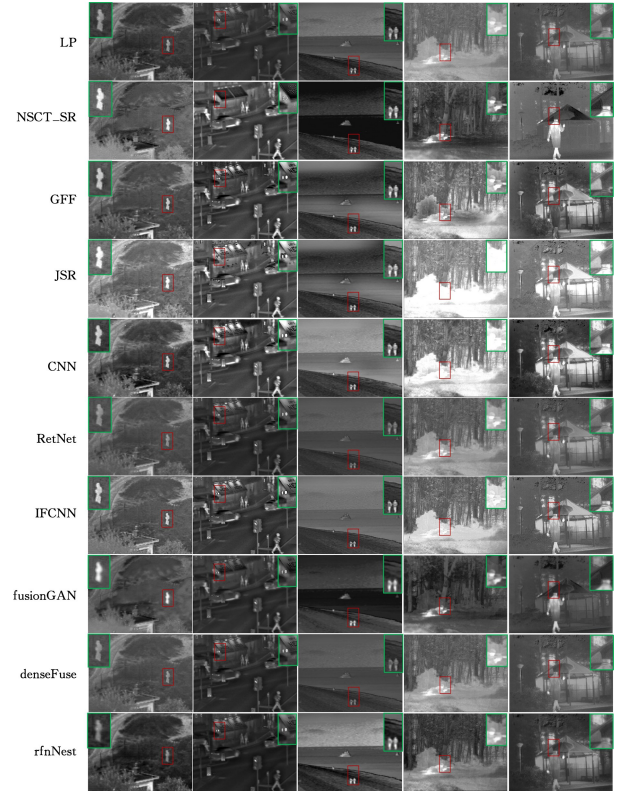


图 5 10 种代表性方法的红外与可见光图像的融合结果

Fig. 5 Fusion results of infrared and visible images of 10 representative methods

从定量角度继续比较各种方法,如图 6 所示,不同方法在不同的评价准则上差异很大。GFF, CNN, IFCNN 方法在 EN, AG, $Q^{AB/F}$ 指标上相对比较突出,适合需要对源图像细节信息和边缘信息保存程度高的场合。NSCT_SR 方法在 CE 和 MI 指标上表现优秀,对细节信息保留充分,但在部分特殊场景中表现欠佳,如海洋场景中,其 EN 值在各种算法中最低。JSR 方法的 AG 和 SF 指标数值突出,在 SSIM 和 RMSE 指标上数值最低,这与图 5 相对应,其融合结果图像更为清晰,梯度分布均匀,但对比度和亮度表现较差,适合更在意图像细节和轮廓边缘完整性的场合。而 ResNet, denseFuse 和 rfnNest 方法在 SSIM 和 RMSE 指标上表现优秀,在 AG 和 SF 指标上数值较低,其融合图像整体结构完整性不错,视觉观感良好,但在图像纹理细节和边缘信息方面存在一定程度的缺失。FusionGAN 方法在 AG, $Q^{AB/F}$, SF 指标中表现均最低,这与其损失函数的设计有关,未能有效保存到图像的边缘、轮廓信息。LP 方法在所选取的场景的各个指标中表现均不突出,原因是该方法的融合原理及规则相对简单,不适合实际应用。

综上,对于算法的适用场景,如果需要保存更多的来自源图像的信息,可以选择 GFF, NSCT_SR, CNN 的方法;若更在意图像的清晰程度、轮廓及边缘完整性,可以选择

JSR, CNN, IFCNN 方法。而在强调图像整体观感,对亮度、对比度有更高要求的场景中,建议选择 denseFuse 和 ResNet 方法。

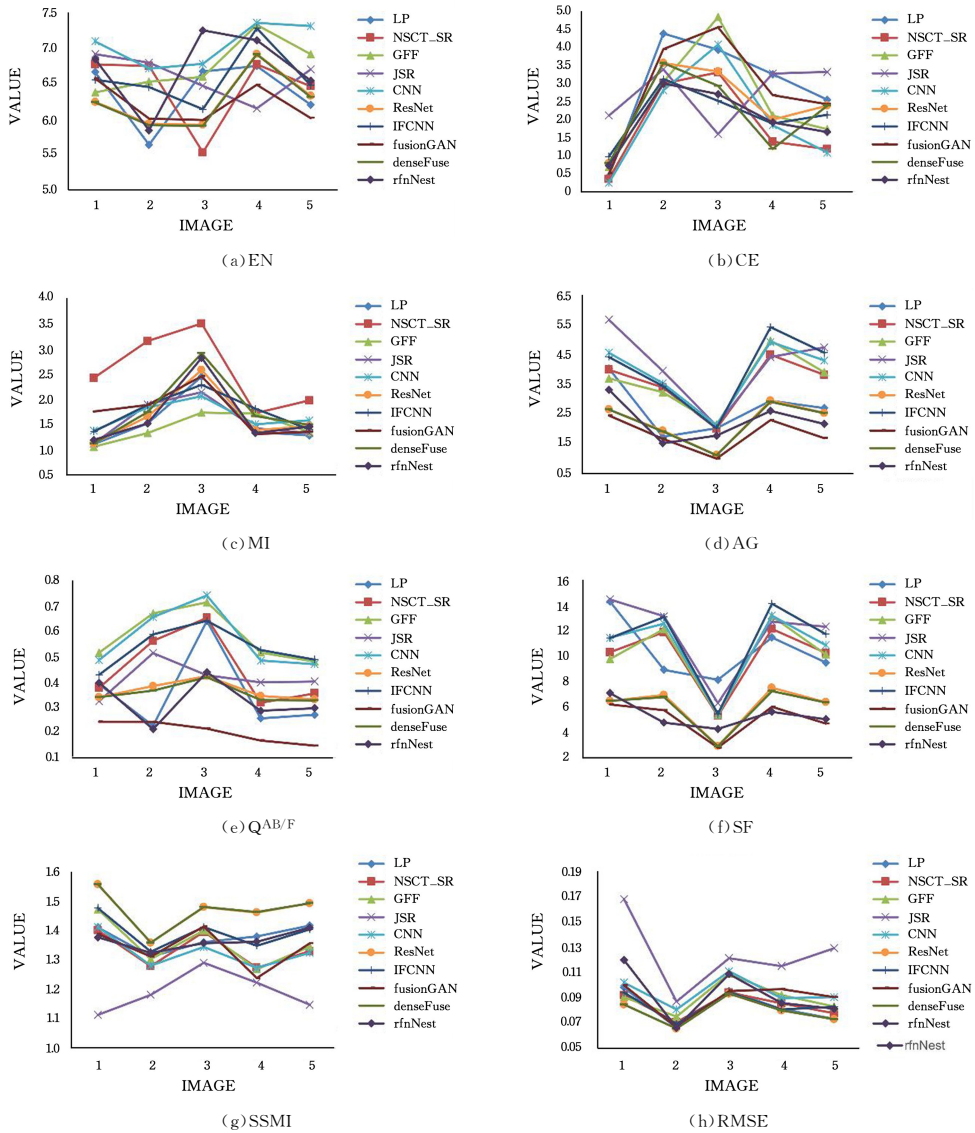


图6 不同方法的客观评价准则

Fig. 6 Objective evaluation metrics of different methods

表3列出了所选取的10种算法在5组图像上运行时间的比较,其中图像的分辨率均设置为 360×270 ,统计值为单张图像的处理时间平均值。可以看出,LP算法原理简单,使用叠加融合规则,其运行效率在传统方法中最高。NSCT_SR算法不仅分解和重建难度增加,其融合策略的设计中加入了显著性检测,相对比较耗时。GFF算法配合多种滤波器的使用,计算量相对较小,处理效率高。JSR算法在使用字典对图像分解的过程中,最为耗时。值得注意的是,JSR, CNN, ResNet算法的运行平台为MATLAB,它们仅使用CPU运算,而其他深度学习算法在Pytorch平台上运行,借助GPU加速,运行时间得到了极大优化,其中rfnNest方法使用两阶段的融合策略,计算量大,效率略低。IFCNN方法处理速度最快,可用于需要实时处理的场景中。

表3 10种融合算法的平均运行时间

Table 3 Average running time of 10 fusion algorithms

Method	Time/s
LP	0.051
NSCT_SR	7.430
GFF	0.505
JSR	128.401
CNN	40.461
ResNet	2.532
IFCNN	0.039
fusionGAN	0.041
denseFuse	0.760
rfnNest	1.035

注:其中IFCNN, fusionGAN, denseFuse, rfnNest为在GPU上运行所得的时间,其他算法为在CPU上运行所得的时间

结束语 本文对红外和可见光图像的融合算法进行了详细介绍,分别从基于传统图像处理理论和深度学习理论两个

角度介绍了具体的算法原理及研究进展。接着总结了对图像融合结果的评价方式,除了主观评价方式外,还从基于信息论、图像特征以及结构相似性 3 个方面使用客观准则对融合图像结果进行定量分析。最后选择了基于传统图像处理理论的比较成熟的以及最新的基于深度学习理论的多种算法,在相同数据集上进行实验,综合比较分析,指出了不同方法的优势和不足以及应用场景。

针对当前红外与可见光图像融合算法存在的问题,以及未来的研究方向,总结出以下 4 点建议:

(1) 基于深度学习的算法的研究比较活跃,但本文通过比较发现其所带来的提升有限,并未展现出对传统图像处理方法的压倒性优势。在后续研究中,应尝试将传统方法和深度学习的优势结合起来,从而带来视觉效果和定量分析层面的综合提升。

(2) 由于不同研究人员在评价图像质量时使用的方法不一致,已出现了数十种客观评判标准。在未来短时间内很难规范图像的评价方式。建议在应用算法时可以参考本文总结的方法,使用尽可能多的客观准则来综合评价算法性能。

(3) 基于嵌入式平台的红外/可见光融合技术是本领域的重要发展方向之一,如何在低算力的条件下实现双光图像的实时融合,是仍待解决的问题。在未来,需要提出更高效的融合模型,在不降低图像质量的情况下,精简模型复杂度,完成实时处理。

(4) 图像的信噪比仍然是影响融合效果的重要因素之一,对于采集到的原始红外和可见光图像,可以结合显著性检测技术和图像去噪技术,统一进行图像增强后再做进一步融合处理,提高融合效果。

参 考 文 献

- [1] LI S, KANG X, FANG L, et al. Pixel-level image fusion: A survey of the state of the art[J]. *Information Fusion*, 2017, 33: 100-112.
- [2] MA J, MA Y, LI C. Infrared and visible image fusion methods and applications: A survey[J]. *Information Fusion*, 2019, 45: 153-178.
- [3] HOGERVORST M, TOET A. Fast natural color mapping for night-time imagery[J]. *Information Fusion*, 2010, 11(2): 69-77.
- [4] TOET A. Natural colour mapping for multiband nightvision imagery[J]. *Information Fusion*, 2003, 4(3): 155-166.
- [5] HAN J, BHANU B. Fusion of color and infrared video for moving human detection[J]. *Pattern Recognition*, 2007, 40(6): 1771-1784.
- [6] SINGH R, VATSA M, NOORE A. Integrated multilevel image fusion and match score fusion of visible and infrared face images for robust face recognition[J]. *Pattern Recognition*, 2008, 41(3): 880-893.
- [7] BHATNAGAR G, LIU Z. A novel image fusion framework for night-vision navigation and surveillance[J]. *Signal, Image and Video Processing*, 2015, 9(1): 165-175.
- [8] 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.
- [9] BULANON D, BURKS T, ALCHANATIS V. Image fusion of visible and thermal images for fruit detection[J]. *Biosystems engineering*, 2009, 103(1): 12-22.
- [10] SANCHEZ V, PRINCE G, CLARKSON JP, 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] SIMONE G, FARINA A, MORABITO F C, et al. Image fusion techniques for remote sensing applications[J]. *Information fusion*, 2002, 3(1): 3-15.
- [12] LI H, DING W, CAO X, 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.
- [13] BURT P J, ADELSON E H. The Laplacian Pyramid as a Compact Image Code[J]. *IEEE Transactions on Communications*, 1983, 31(4): 532-540.
- [14] TOET A, VAN RUYVEN L J, VALETON J M. Merging thermal and visual images by a contrast pyramid[J]. *Optical Engineering*, 1989, 28(7): 2877-89.
- [15] ZHANG Z, BLUM R S. A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application[C]// *Proceedings of the IEEE*. 1999: 1315-1326.
- [16] CHEN J, LI X, LUO L, et al. Infrared and visible image fusion based on target-enhanced multiscale transform decomposition[J]. *Information Sciences*, 2020, 508: 64-78.
- [17] LI S, KANG X, HU J. Image fusion with guided filtering[J]. *IEEE Trans Image Process*, 2013, 22(7): 2864-2875.
- [18] 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.
- [19] WANG K, QI G, ZHU Z, et al. A novel geometric dictionary construction approach for sparse representation based image fusion[J]. *Entropy*, 2017, 19(7): 306.
- [20] YIN H. Sparse representation with learned multiscale dictionary for image fusion[J]. *Neurocomputing*, 2015, 148: 600-610.
- [21] SUN C, ZHANG C, XIONG N. Infrared and Visible Image Fusion Techniques Based on Deep Learning: A Review[J]. *Electronics*, 2020, 9(12): 2162.
- [22] MA J, YU W, LIANG P, et al. FusionGAN: A generative adversarial network for infrared and visible image fusion[J]. *Information Fusion*, 2019, 48: 11-26.
- [23] PRABHAKAR K R, SRIKAR V S, BABU R V. Deepfuse: A deep unsupervised approach for exposure fusion with extreme exposure image pairs[C]// *Proceedings of the IEEE International Conference on Computer Vision*. 2017: 4714-4722.
- [24] LIU Y, CHEN X, CHENG J, et al. Infrared and visible image fusion with convolutional neural networks[J]. *International Journal of Wavelets, Multiresolution and Information Processing*, 2018, 16(3): 1850018.
- [25] LI H, WU X J, KITTLER J. RFN-Nest: An end-to-end residual fusion network for infrared and visible images[J]. *Information*

- Fusion,2021,73:72-86.
- [26] 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.
- [27] TOET A. A morphological pyramidal image decomposition[J]. Pattern Recognition Letters,1989,9(4):255-261.
- [28] 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.
- [29] ZHAN L, ZHUANG Y, HUANG L. Infrared and visible images fusion method based on discrete wavelet transform[J]. Journal of Computers,2017,28(2):57-71.
- [30] HAN X, LI Z L, YAO D L, et al. Fusion of infrared and visible images based on discrete wavelet transform[C]//The Photoelectronic Technology Committee Conferences. 2015:387-392.
- [31] SELESNICK I W, BARANIUK R G, KINGSBURY N C. The dual-tree complex wavelet transform[J]. IEEE Signal Processing Magazine,2005,22(6):123-151.
- [32] ZUO Y, LIU J, BAI G, et al. Airborne infrared and visible image fusion combined with region segmentation[J]. Sensors,2017,17(5):1127.
- [33] MINGHUI S, LU L, YUANXI P, et al. Infrared & visible images fusion based on redundant directional lifting-based wavelet and saliency detection [J]. Infrared Physics & Technology,2019,101:45-55.
- [34] 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.
- [35] DO M N, VETTERLI M. Contourlets: a directional multiresolution image representation [C] // Proceedings of International Conference on Image Processing. 2002.
- [36] 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.
- [37] MENG F, SONG M, GUO B, et al. Image fusion based on object region detection and Non-Subsampled Contourlet Transform [J]. Computers & Electrical Engineering,2017,62:375-383.
- [38] KONG W, WANG B, 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.
- [39] CAO Y, LI X, CHEN X, et al. Research on Fusion Method of Infrared and Visible Image Based on NSST and Guided Filtering [J]. Journal of Changchun University of Science and Technology (Natural Science Edition),2022,45(3):28-34.
- [40] LIU Y, LIU S, WANG Z. A general framework for image fusion based on multi-scale transform and sparse representation[J]. Information Fusion,2015,24:147-164.
- [41] LIU G T, ZHOU J H, LI T, et al. Infrared and visible image fusion through hybrid curvature filtering image decomposition[J]. Infrared Physics & Technology,2022,120:103938.
- [42] YANG B, LI S. Multifocus image fusion and restoration with sparse representation[J]. IEEE Transactions on Instrumentation and Measurement,2009,59(4):884-892.
- [43] 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.
- [44] BIN Y, CHAO Y, GUO Y H. Efficient image fusion with approximate sparse representation [J]. International Journal of Wavelets, Multiresolution and Information Processing,2016,14(4):1650024.
- [45] 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.
- [46] PATI Y C, REZAIIFAR R, KRISHNAPRASAD P S. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition[C]// Proceedings of 27th Asilomar Conference on Signals, Systems and Computers. 1993:40-44.
- [47] YANG B, LI S. Pixel-level image fusion with simultaneous orthogonal matching pursuit[J]. Information Fusion,2012,13(1):10-19.
- [48] 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.
- [49] MITIANOUDIS N, ANTONOPOULOS S A, STATHAKI T. Region-based ICA image fusion using textural information[C]//2013 18th International Conference on Digital Signal Processing (DSP). 2013:1-6.
- [50] KONG W, LEI Y, ZHAO H. Adaptive fusion method of visible light and infrared images based on non-subsampling shearlet transform and fast non-negative matrix factorization[J]. Infrared Physics & Technology,2014,67:161-172.
- [51] 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.
- [52] CHEN G, CHEN Y, LI J, et al. Infrared and visible image fusion based on discrete nonseparable shearlet transform and convolutional sparse representation[J]. Journal of Jilin University (Engineering and Technology Edition),2021,51(3):996-1010.
- [53] 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.
- [54] 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.
- [55] HOU R C, ZHOU D M, NIE R C, et al. Infrared and Visible Images Fusion Using Visual Saliency and Dual-PCNN[J]. Computer Science,2018,45(6A):162-166.
- [56] 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.
- [57] 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.
- [58] 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(ICPR). 2018;2705-2710.
- [59] SIMONYAN K,ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv:1409.1556, 2014.
- [60] LI H,WU X J,DURRANI T S. Infrared and visible image fusion with ResNet and zero-phase component analysis[J]. *Infrared Physics & Technology*,2019,102;103039.
- [61] HE K,ZHANG X,REN S,et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016;770-778.
- [62] KESSY A,LEWIN A,STRIMMER K. Optimal whitening and decorrelation[J]. *The American Statistician*, 2018, 72(4):309-314.
- [63] ZHANG Y,LIU Y,SUN P,et al. IFCNN:A general image fusion framework based on convolutional neural network[J]. *Information Fusion*,2020,54;99-118.
- [64] XU H,MA J,LE Z,et al. FusionDn:A unified densely connected network for image fusion[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2020;12484-12491.
- [65] HUANG G,LIU Z,VAN DER MAATEN L,et al. Densely connected convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017:4700-4708.
- [66] XU H,MA J,JIANG J,et al. U2Fusion:A Unified Unsupervised Image Fusion Network[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*,2022,44(1);502-518.
- [67] LI C,HOU J,LI J,et al. Infrared and Visible Image Fusion Method Based on Attention and Residual Concatenation[J]. *Computer Engineering*,2022,48(7);234-240.
- [68] MA J,LIANG P,YU W,et al. Infrared and visible image fusion via detail preserving adversarial learning[J]. *Information Fusion*,2020,54;85-98.
- [69] YUAN C,SUN C,TANG X,et al. FLGC-Fusion GAN:An Enhanced Fusion GAN Model by Importing Fully Learnable Group Convolution[J]. *Mathematical Problems in Engineering*, 2020, 2020;6384831.
- [70] LI Q,LU L,LI Z,et al. Coupled GAN With Relativistic Discriminators for Infrared and Visible Images Fusion[J]. *IEEE Sensors Journal*,2021,21(6);7458-7467.
- [71] TANG L L,LIU G,XIAO G,et al. p Infrared and visible image fusion based on guided hybrid model and generative adversarial network[J]. *Infrared Physics & Technology*,2022,120;103914.
- [72] LI H,WU X J. DenseFuse:A fusion approach to infrared and visible images[J]. *IEEE Transactions on Image Processing*, 2018, 28(5):2614-2623.
- [73] LI H,WU X J,DURRANI T. NestFuse: An Infrared and Visible Image Fusion Architecture Based on Nest Connection and Spatial/Channel Attention Models[J]. *IEEE Transactions on Instrumentation and Measurement*,2020,69(12);9645-9656.
- [74] ZHOU Z,SIDDIQUEE M M R,TAJBAKHSH N,et al. UNet++:A Nested U-Net Architecture for Medical Image Segmentation[C]//Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, 2018;3-11.
- [75] FU Y,WU X J. A dual-branch network for infrared and visible image fusion[C]//2020 25th International Conference on Pattern Recognition(ICPR). 2021;10675-10680.
- [76] XU H,ZHANG H,MA J. Classification Saliency-Based Rule for Visible and Infrared Image Fusion[J]. *IEEE Transactions on Computational Imaging*,2021,7;824-836.
- [77] ZHANG L,LI H,ZHU R,et al. An infrared and visible image fusion algorithm based on ResNet-152[J]. *Multimedia Tools and Applications*,2022,81(7);9277-9287.
- [78] ZHANG X,YE P,XIAO G. VIFB:a visible and infrared image fusion benchmark[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2020;104-105.
- [79] LIU Z,BLASCH E,XUE Z,et al. Objective Assessment of Multiresolution Image Fusion Algorithms for Context Enhancement in Night Vision: A Comparative Study[J]. *IEEE Trans. Pattern. Anal. Mach. Intell*,2012,34(1);94-109.
- [80] ROBERTS J W,VAN AARDT J A,AHMED F B. Assessment of image fusion procedures using entropy, image quality, and multispectral classification[J]. *Journal of Applied Remote Sensing*,2008,2(1);023522.
- [81] QU G,ZHANG D,YAN P. Information measure for performance of image fusion[J]. *Electronics Letters*,2002,38(7);313-315.
- [82] CUI G,FENG H,XU Z,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.
- [83] XYDEAS C A,PETROVIC V. Objective image fusion performance measure[J]. *Electronics Letters*,2000,36(4);308-309.
- [84] ESKICIOGLU A M,FISHER P S. Image quality measures and their performance[J]. *IEEE Transactions on Communications*, 1995, 43(12);2959-2965.
- [85] WANG Z,BOVIK A C,SHEIKH H R,et al. Image quality assessment:from error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*,2004,13(4);600-612.



WEI Qi, born in 1997, postgraduate. His main research interests include image fusion and so on.



ZHAO Juan, born in 1985, Ph.D, senior engineer, master supervisor. Her main research interests include optical signal processing and image processing.