

红外与可见光图像融合研究

毛义坪, 马小艳

重庆对外经贸学院数学与计算机学院, 重庆

收稿日期: 2023年7月10日; 录用日期: 2023年8月10日; 发布日期: 2023年8月17日

摘要

红外与可见光图像融合是指将同一场景下的相机获取的图像与红外传感器获得的图像融合为一张图像。融合图像具备源图像的重要信息, 尽量无冗余信息。融合图像在广泛应用在计算机视觉、农业、遥感、医学等领域。本文主要对其研究进展做整理, 使感兴趣学者快速掌握红外与可见光图像融合研究脉络。首先将研究算法主要分为: 多尺度变换、稀疏表示、神经网络以及其他方法等几大类。接着对每类进行详细探讨, 然后对目前主流算法做一些优劣评价与总结。最后对本研究领域未来值得突破点进行展望。

关键词

图像融合, 红外图像, 可见光图像

Research on Infrared and Visible Image Fusion

Yiping Mao, Xiaoyan Ma

School of Mathematics and Computer Science, Chongqing College of International Business and Economics, Chongqing

Received: Jul. 10th, 2023; accepted: Aug. 10th, 2023; published: Aug. 17th, 2023

Abstract

Visible and infrared image fusion refers to the fusion of images obtained by cameras and infrared sensors in the same scene into one image. The fused image has important information from the source image, and there should be as little redundant information as possible. Fusion images are widely used in many fields such as medical, computer vision, agriculture, and remote sensing. This article mainly summarizes the research progress of visible and infrared image fusion, enabling interested scholars to quickly grasp the research context. Firstly, the algorithms are mainly divided into several categories: multi-scale transformation, sparse representation, neural networks,

and other methods. Secondly, a detailed discussion will be conducted on each category, followed by some evaluation and summary of the current mainstream algorithms. Finally, prospects for future breakthroughs in this research field.

Keywords

Image Fusion, Infrared Images, Visible Image

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着科技不断发展, 传感器技术不断进步。可获得大量不同种类的信息与数据。同一类型的传感器只能从一个方面获取信息, 不能提供下游计算需要的所有信息。因此, 在现代应用技术中, 将多传感器信息融合为一体发挥着越来越重要的作用。信息融合技术是将不同传感器获得的信息融合为一体的技术。融合后的信息具有多个源信息的互补特征。

可见光图像一般由普通相机获取, 和人类视觉几乎一致。获得的信息是具有高清细节、高分辨率的图像。但在下雨、夜晚、雾霾等情况下获得的图像几乎不能使用。红外图像是指红外传感器捕获信息源发出的热辐射能而形成的图像。其特点通常是分辨率低、细节纹理较模糊等。但不会受到来自恶劣天气、普通遮挡的影响。

由于可见光和红外图像优缺点具有互补性, 互相结合后具有更大的利用价值。于是可见光和红外图像融合技术顺应而生。可见光与红外图像融合是指将可见光图像与红外图像融合为一张图像, 融合后的图像具有两种类型图像互补信息, 表达力更加丰富。

融合后的信息广泛应用于计算机视觉领域[1], 如: 目标监控、目标跟踪、目标识别。农业领域, 如水果采摘、水果质量评估、农作物特征检测。医学领域、遥感领域等[2]。

融合算法始于上个世纪, 距今已有几十年的历史。根据研究者提出的融合算法架构, 大致可分为多尺度分解法、稀疏表示法、神经网络法、其他混合法等。接下来章节分别对其种类进行概述, 最后一小节对可见光与红外图像融合算法进行总结与展望。

2. 多尺度分解法

本多尺度分解法是深度学习算法提出之前最火的方法之一。其主要思想是首先将源图像分解为若干层, 然后利用相应的融合规则对每一层进行融合, 最后对每层融合后的图像进行逆变换, 最终的得到融合图像。根据其融合框架可得知, 分解算法与融合规则决定最终融合质量。

2.1. 金字塔分解

最早基于多尺度分解的可见光和红外图像融合算法是基于金字塔分解法。文献[3]产生拉普拉斯金字塔的多尺度变换包括四个步骤: 低通滤波、子采样、插值和差分, 所有这些都是不同的尺度上完成的。图像的融合是通过组合源图像生成的两个拉普拉斯金字塔的差图像, 然后应用逆变换得到最终融合图像。还有学者对传统金字塔进行改进, 提出可操纵金字塔, 具有多尺度、多方向和自反转的图像分解优点。

活动水平测量和权重分配是图像融合的两个关键部分。Li [4]提出了一种新的基于主成分分析网络(PCANet)和图像金字塔的红外和可见光融合方法。首先,使用轻量级的学习网络 PCANet 来获得源图像的活动水平测量和权重分配。然后,通过图像金字塔将权重和源图像分解为多个尺度,并在每个尺度上应用加权平均融合规则。最后,通过重构得到融合后的图像。

2.2. 小波变换

第二种是基于离散小波变换法。文献[5]通过离散小波变换获得源图像的稀疏系数。然后通过改进的熵加权融合规则和基于绝对值最大的融合规则对红外和可见光图像的低系数和高系数进行融合。在使用随机高斯矩阵的局部线性投影之后,通过压缩采样匹配追踪算法重建融合图像。取得了不错的融合效果。

普通离散小波本身具有其局限性,如不具有平移不变性、提取多方向信息不丰富等缺点。为了克服这些缺点,学者们提出双树小波小波、四元数小波、提升平稳小波等,有效的克服普通小波的缺点。

还有其小波算法的变种,如 Chen 提出了一种基于深度小波密集网络(WT DenseNet) [6]的红外和可见光图像融合方法。WT DenseNet 包括三个网络层,即混合特征提取层、融合层和图像重建层。最后,通过图像重建层输出融合图像。

2.3. 多尺度几何分析

第三种多尺度分解法是多尺度几何分析(MGA)法。其方法主要思想来源也是克服小波表示图像的缺点。比如小波在二维空间中捕获方向信息的能力有限。大多数小波变换都无法从信号稀疏性的角度得到图像的最优表示等。多尺度几何分析法应用到可见光和红外图像融合中最初是轮廓波法。

Shen 提出了一种基于非下采样轮廓波变换(NSCT)和改进的双边滤波器的红外和可见光图像融合算法[7],该算法利用 NSCT 将图像分解为低频分量和高频分量。采用改进的双边滤波方法对两幅图像的高频分量进行处理,对图像的噪声进行滤波,计算红外图像高频分量的图像细节。对于低频系数的融合规则,采用了局部标准方差系数法。最后,根据 NSCT 的逆变换,得到融合图像。

多尺度几何分析除了轮廓波还有曲波变换法、剪切波法等等。

3. 稀疏表示法

由于多尺度分解法的融合规则几乎都是高频系数取绝对值最大,低频系数取平均。在一定程度上会造成融合结果模糊。于是学者将稀疏表示理论引入图像融合领域中。

稀疏表示的理论指任何一组信号都可以基于一个过完备字典去线性表示。即使用尽量少的字典中原子线性的去描述一组信号称为稀疏表示。基于稀疏表示[8]的可见光和红外图像融合算法基本步骤主要有四步。首先,使用滑动窗口策略将每个源图像分解为步长为 1 的重叠的小块。第二,学习过完备字典,过完备字典一般有离线学习和在线学习两种学习方式。并对每个块进行稀疏编码,利用学习到的过完备字典获取稀疏表示系数。第三,对稀疏表示系数进行融合。最后,利用过完备字典与融合后的系数重构出最终图像。过完备字典、稀疏编码、融合规则是融合质量的关键。

3.1. 离线字典学习

过完备字典一般有离线学习和在线学习两种学习方式。离线学习字典主要是采用固定基构造或通过自然图像训练得到。文献[9]采用了三种离线字典方式进行可见光和红外图像融合。第一种对不同频率的余弦波进行采样,构造过完备离散余弦变换(DCT)字典。第二种是混合字典,它由 DCT 基、小波“db1”基、Gabor 基和脊波基组成。第三种是使用迭代 K-SVD 算法从自然样本中获得训练字典。训练数据从 50

幅自然图像的数据库中随机挑选, 然后分割成 50,000 个 8×8 个图像块进行训练。

3.2. 在线字典学习

过完备字典在线学习法是指采用融合前的源图像进行学习构造字典。文献[10]提出了一种基于联合块聚类的高效字典学习方法。根据结构相似性, 聚集来自不同传感器模态源图像的所有块, 并且只选择几个有效描述每个联合块的主要部分, 以形成过完备的字典。构造信息丰富且紧凑字典是研究目标。为了在字典构建中获得足够的稀疏表示信息, 文献[11]提出一种新的几何字典构建方法。首先基于形态学相似性将源图像中的图像块分类为不同的组。然后根据随机坐标编码提取相应的图像块信息, 最后构建字典。

对于使用稀疏表示的图像融合方法, 自适应字典和融合规则对多模态图像融合有很大影响, 最大 L1 范数融合规则可能会造成融合结果的灰色不一致的现象。为了解决这个问题, Wang 根据联合补丁聚类的自适应字典与稀疏表示相关理论, 提出了一种改进的图像融合方法[12]。首先, 使用高斯滤波器来分离高频和低频信息。第二, 采用局部能量加权策略来完成低频融合。第三, 使用联合补丁聚类算法重建过完备自适应学习字典。根据设计混合融合规则, 融合高频部分。最后, 通过将频域转换为空间域获得融合结果。

4. 神经网络法

近年来, 深度学习算法受到学者们青睐。正是由于神经网络通常由许多神经元组成, 可以模仿人脑的感知行为机制。神经网络还具备强大的特征提取和重构能力。深度学习的发展也促进了图像融合的巨大进步, 使得可见光和红外图像融合很有前景[13]。基于神经网络的可见光和红外图像融合算法一般分为基于预训练网络法、基于自编码法、基于端对端网络法 3 类。

4.1. 基于预训练网络法

基于预训练网络法主要思想是将提前训练好的神经网络引入到可见光和红外图像融合。算法框架是首先利用通过固定网络获取图像深度特征, 然后采用合适策略进行特征融合。最后用恰当方法来重建融合图像。如文献 Xu [14]提出将解耦表示应用于可见光和红外图像融合。首先根据不同源图像中的信息源进行分解表示。然后应用不同的策略来融合这些不同类型的表示。最后将融合的代表输入到预训练的生成器中, 以生成融合结果。

4.2. 基于自编码网络法

自编码网络可见光与红外图像融合算法一般由编码器、融合层和解码器三部分组成。其中编码器从源图像中提取充分的深度特征, 融合层根据相应融合策略用来融合深层特征。解码器根据编码的特征自适应地重建原始数据。通过解码器网络对融合后的图像进行重构。文献[15]的编码器包含 C1 与密集块 (DenseBlock) 两个部分。第一部分 C1 包含 3×3 的滤波器来提取粗略特征, 第二部分 DenseBlock 包含三个卷积层(每个层的输出为下一层的输入), 其中滤波器大小也是 3×3 。融合层选择了加法策略和 L1 范数策略两种融合策略来组合编码器获得的显著特征图。融合层的输出将作为解码器的输入。解码器包含四个卷积层, 每层也是 3×3 大小的滤波器。此方法取得了不错的融合效果。但此方法在深层特征方向上比较冗余。

Ren [16]提出了一种自编码网络法并命名为基于变分自动编码器法(VAE), 该所提出的网络分为图像融合网络和红外特征补偿网络。首先, 在图像融合网络中, 创建图像融合网络的编码器, 从源图像中生

成隐藏层中的潜在向量。然后根据高斯概率密度的乘积,把两个不同的潜在向量合并为一个。随着损失函数值的下降,解码器开始重建融合图像。为了提高网络训练的效率,在网络中加入了残差块和对称跳跃连接方法。并且为了克服融合网络中损失函数设置的缺陷,还设计了一个网络来补偿红外图像的关键辐射特征。此方法在一定程度上改善了特征中方向信息比较冗余问题,使融合后的特征更精细。

4.3. 基于端对端网络法

端对端网络法结构比较简单,输入源图像与输出最终融合结果之间就只有一个设计好的融合网络。关键优势在于,直接通过训练学习得到源图像和融合图像之间的映射,因此不需要过多的处理步骤。避免了像传统方法那样手动设计复杂的活动水平测量和融合规则。

为了克服手动且复杂设计融合规则等缺点, Ma 等人提出一种生成对抗性端到端网络 FusionGAN 算法,并设计了一个专门用于红外和可见光图像融合的损失函数。可以从输入源图像自动生成融合图像,而无需手动设计活动水平测量或融合规则。该方法在生成器和鉴别器之间建立了一个对抗性游戏,其中生成器主要生成的融合图像的红外强度信息与可见光图像梯度信息,鉴别器旨在使融合图像具有可见光图像中更多的细节信息。相对来说,融合结果不太清晰,细节较丰富。

更大程度为了使融合结果目标更清晰,细节更丰富。Ma [17]继续改进 FusionGAN 提出了一种基于细节保留对抗性学习的端到端模型。它能够克服传统融合方法中活动水平测量和融合规则手工设计和复杂设计的局限性。考虑到红外和可见光图像的特定信息,设计了两个损失函数。包括细节损失和目标边缘增强损失,以提高细节信息的质量,并在生成对抗性网络的框架下锐化红外目标的边缘。此方法使融合图像能够保留热辐射的同时,还可以锐化红外图像中的目标边界和可见光图像中丰富的纹理细节。其融合结果取得了优异成绩。

为了进一步避免手工设计的复杂且不准确的融合规则,与充分利用红外和可见光图像中的信息。Li 提出了一种实用的融合方法 RCGAN [18]。将耦合生成对抗性网络开创性引入到图像融合领域且提出了使用预融合图像作为标签的策略。实验表明 RCGAN 法融合结果优秀,能够有效地保持可见光图像中丰富的纹理和红外图像中的热辐射信息。

值得注意的是,一个好的模型应该应用相对广泛。Xu 提出了一种统一且无监督的端到端图像融合网络 U2Fusion [19],它根据特征提取和信息测量,能够自动估计被融合图像的重要性。能对多模态、多曝光和多聚焦图像进行融合。通过实验,验证了 U2Fusion 在三个典型图像融合任务上的有效性和通用性。

5. 其他方法

除了上述总结的三类方法以外,学者还提出了基于其他理论法以及混合方法等。比如刘等人提出多尺度变换与稀疏表示结合的方法[20]。其融合框架中,首先对每个源图像执行多尺度变换,获得它们的低频系数和高频系数。然后,使用基于稀疏表示法合并低频系数,高频系数的合并采用绝对值最大法。最后对合并后的系数执行多尺度逆变换,最终获得融合后的图像。还有基于主成分分析理论(PCA)法、独立成分分析法等。

为了解决红外和可见光图像融合中亮度和细节信息丢失的问题, Li [21]提出了一种利用滚动引导滤波和梯度显著性图进行红外和可见光图像融合的有效方法。滚动引导滤波用于将输入图像分解为近似层和残差层;能量属性融合模型用于融合近似层;通过重构融合的近似层子图像和残差层子图像来生成融合图像。

然而,现有的大多数融合方法都是用于灰度图像融合任务,不能应用于真实的融合成像系统。为了解决这些问题, Mo [22]提出了一种属性引导的红外和可见光图像融合算法,首先,对红外图像进行属性

滤波, 得到显著物体的初始权重图。然后, 利用边缘保持滤波器来细化初始权重图。最后, 采用拉普拉斯金字塔融合策略将源图像与所获得的权重图进行融合。与以往的方法相比, 该方法利用了红外图像的固有特性, 设计了一种基于属性滤波器的特征提取规则, 无需训练步骤或先验信息, 即可有效地提取红外图像中的突出目标。

6. 总结与展望

本文主要对红外与可见光图像融合算法研究进展做梳理。使新学者快速掌握此研究方向脉络, 并为本领域做出贡献。对主流算法与结果的主要总结如下:

1) 融合算法大致可分为多尺度变换、稀疏表示、神经网络以及其他方法混合法等几大类。并对每个大类进行细分类与探讨。

2) 神经网络深度学习算法目前最火, 但其融合结果并不是完全碾压传统算法[23]。

3) 主观上看, 多尺度分解法融合结果比较稳定, 但细节保留不足, 尤其是对红外目标淡化。

4) 在纹理复杂的图像中, 稀疏表示法有时细节模糊, 可能原因是字典的表达能力不足, 且其最大的缺点就是计算代价太大, 不适用于实时任务。

近几十年的发展, 红外和可见光图像融合方面已经取得了长足的进步, 但仍有一些问题需要进一步研究解决。对本课题展望如下:

1) 大多数基于多尺度变换法的融合规则都比较简单(如: 高频绝对值最大, 低频取平均), 如何设计有价值融合规则值得进一步探讨。

2) 一般稀疏表示法采用滑动窗口技术对图像分块, 然后融合每个块, 最终融合图像由融合后的块拼接而成。其并未考虑块与块之间的关联性, 且此算法计算开销很大, 不适合实时处理的系统。如何考虑关联性, 值得关注。

3) 神经网络算法中端到端算法较火, 但设计损失函数非常主观。如何有效的考虑融合与后续应用之间的相关性来设计设计损失函数。是一个有趣的研究点。

4) 由于深度学习算法不是完全碾压传统算法, 学者大可不必对传统算法保持悲观。开发有效的传统特征提取算法或者传统算法与深度学习算法结合是一个突破点。如文献[24]提出的显著性检测与多层潜在低秩表示的融合方法。

5) 研究理论的目标是最终落实到实时应用。然而应用到真实的成像系统较少, 开发计算代价小与实时处理的平台也是一个重要的发展点。

基金项目

重庆对外经贸学院科学研究项目(KYKJ202002)。

参考文献

- [1] Ma, W., Wang, K., Li, J., Yang, S.X., Li, J., Song, L. and Li, Q. (2023) Infrared and Visible Image Fusion Technology and Application: A Review. *Sensors*, **23**, Article No. 599. <https://doi.org/10.3390/s23020599>
- [2] 沈英, 黄春红, 黄峰, 等. 红外与可见光图像融合技术的研究进展[J]. 红外与激光工程, 2021, 50(9): 152-169.
- [3] Bulanon, D.M., Burks, T.F. and Alchanatis, V. (2009) Image Fusion of Visible and Thermal Images for Fruit Detection. *Biosystems Engineering*, **103**, 12-22. <https://doi.org/10.1016/j.biosystemseng.2009.02.009>
- [4] Li, S., Zou, Y., Wang, G. and Lin, C. (2023) Infrared and Visible Image Fusion Method Based on a Principal Component Analysis Network and Image Pyramid. *Remote Sensing*, **15**, Article No. 685. <https://doi.org/10.3390/rs15030685>
- [5] Liu, Z., Yin, H., Fang, B., et al. (2015) A Novel Fusion Scheme for Visible and Infrared Images Based on Compressive Sensing. *Optics Communications*, **335**, 168-177. <https://doi.org/10.1016/j.optcom.2014.07.093>

-
- [6] Chen, Y.L., *et al.* (2023) Infrared and Visible Image Fusion with Deep Wavelet-Dense Network. *Optica Applicata*, **53**, 49-64. <https://doi.org/10.37190/oa230104>
- [7] Shen, Y., Xiang, K.Y., *et al.* (2021) A Noisy Infrared and Visible Light Image Fusion Algorithm. *Journal of Information Processing Systems*, **17**, 1004-1019.
- [8] Zhang, Q., Liu, Y., Blum, R.S., *et al.* (2018) Sparse Representation Based Multi-Sensor Image Fusion for Multi-Focus and Multi-Modality Images: A Review. *Information Fusion*, **40**, 57-75. <https://doi.org/10.1016/j.inffus.2017.05.006>
- [9] Yang, B. and Li, S. (2012) Pixel-Level Image Fusion with Simultaneous Orthogonal Matching Pursuit. *Information Fusion*, **13**, 10-19. <https://doi.org/10.1016/j.inffus.2010.04.001>
- [10] Kim, M., Han, D.K. and Ko, H. (2016) Joint Patch Clustering-Based Dictionary Learning for Multimodal Image Fusion. *Information Fusion*, **27**, 198-214. <https://doi.org/10.1016/j.inffus.2015.03.003>
- [11] Wang, K., Qi, G., Zhu, Z., *et al.* (2017) A Novel Geometric Dictionary Construction Approach for Sparse Representation Based Image Fusion. *Entropy*, **19**, 306. <https://doi.org/10.3390/e19070306>
- [12] Wang, C., Wu, Y., Yu, Y., *et al.* (2022) Joint Patch Clustering-Based Adaptive Dictionary and Sparse Representation for Multi-Modality Image Fusion. *Machine Vision and Applications*, **33**, Article No. 69. <https://doi.org/10.1007/s00138-022-01322-w>
- [13] Zhang, H., Xu, H., Tian, X., *et al.* (2021) Image Fusion Meets Deep Learning: A Survey and Perspective. *Information Fusion*, **76**, 323-336. <https://doi.org/10.1016/j.inffus.2021.06.008>
- [14] Xu, H., Wang, X. and Ma, J. (2021) DRF: Disentangled Representation for Visible and Infrared Image Fusion. *IEEE Transactions on Instrumentation and Measurement*, **70**, Article ID: 5006713. <https://doi.org/10.1109/TIM.2021.3056645>
- [15] Li, H. and Wu, X.J. (2019) DenseFuse: A Fusion Approach to Infrared and Visible Images. *IEEE Transactions on Image Processing*, **28**, 2614-2623. <https://doi.org/10.1109/TIP.2018.2887342>
- [16] Ren, L., Pan, Z.B., Cao, J.Z., *et al.* (2021) Infrared and Visible Image Fusion Based on Variational Auto-Encoder and Infrared Feature Compensation. *Infrared Physics & Technology*, **117**, Article ID: 103839. <https://doi.org/10.1016/j.infrared.2021.103839>
- [17] Ma, J.Y., Liang, P.W., Yu W., *et al.* (2020) Infrared and Visible Image Fusion via Detail Preserving Adversarial Learning. *Information Fusion*, **54**, 85-98. <https://doi.org/10.1016/j.inffus.2019.07.005>
- [18] Li, Q., *et al.* (2021) Coupled GAN with Relativistic Discriminators for Infrared and Visible Images Fusion. *IEEE Sensors Journal*, **21**, 7458-7467. <https://doi.org/10.1109/JSEN.2019.2921803>
- [19] Xu, H., Ma, J., Jiang, J., Guo, X. and Ling, H. (2022) U2Fusion: A Unified Unsupervised Image Fusion Network. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **44**, 502-518. <https://doi.org/10.1109/TPAMI.2020.3012548>
- [20] Liu, Y., Liu, S.P. and Wang, Z.F. (2015) A General Framework for Image Fusion Based on Multi-Scale Transform and Sparse Representation. *Information Fusion*, **24**, 147-164. <https://doi.org/10.1016/j.inffus.2014.09.004>
- [21] Li, L., Lv, M., Jia, Z., Jin, Q., Liu, M., Chen, L. and Ma, H. (2023) An Effective Infrared and Visible Image Fusion Approach via Rolling Guidance Filtering and Gradient Saliency Map. *Remote Sensing*, **15**, Article No. 2486. <https://doi.org/10.3390/rs15102486>
- [22] Mo, Y., Kang, X.D., *et al.* (2021) Attribute Filter Based Infrared and Visible Image Fusion. *Information Fusion*, **75**, 41-54. <https://doi.org/10.1016/j.inffus.2021.04.005>
- [23] 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.
- [24] 龙志亮, 邓月明, 王润民, 等. 基于显著性检测与MDLatLRR分解的红外与可见光图像融合[J]. 红外技术, 2023, 45(7): 705-713.