

文章编号 1004-924X(2023)10-1548-15

基于快速交替引导滤波和 CNN 的红外与 可见光图像融合

杨艳春*, 李永萍, 党建武, 王阳萍

(兰州交通大学 电子与信息工程学院, 甘肃 兰州 730070)

摘要:为了解决红外与可见光图像融合中出现细节信息丢失, 边缘模糊以及伪影的问题, 本文提出一种快速交替引导滤波, 在保证融合图像质量的前提下有效提高运行效率, 结合 CNN(卷积神经网络)以及红外特征提取进行有效的融合。首先, 对源图像利用四叉树分解和贝塞尔插值来提取红外亮度特征结合可见光图像得到初始融合图像。其次, 通过快速交替引导滤波获取源图像的基础层与细节层信息, 基础层通过 CNN 与拉普拉斯变换得到融合后的基础图像, 细节层通过显著性测量的方法得到融合后的细节图像。最后, 将初始融合图、基础融合图以及细节融合图进行相加得到最终融合结果。本算法涉及到的快速交替引导滤波以及特征提取性能使得最终融合结果中包含丰富的纹理细节信息, 边缘清晰。经实验表明, 本算法所得融合结果在视觉方面具有较好的保真度, 客观评价指标较对比方法其信息熵、标准差、空间频率、小波特征互信息、视觉保真度以及平均梯度分别平均提高了 9.9%, 6.8%, 43.6%, 11.3%, 32.3%, 47.1%。

关键词:快速交替引导滤波; 红外特征提取; 卷积神经网络; 红外与可见光图像融合

中图分类号: TP391 **文献标识码:** A **doi:** 10.37188/OPE.20233110.1548

Infrared and visible image fusion based on fast alternating guided filtering and CNN

YANG Yanchun*, LI Yongping, DANG Jianwu, WANG Yangping

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

* Corresponding author, E-mail: yangyanchun102@sina.com

Abstract: In order to solve the problems of the loss of detail information, blurred edges, and artifacts in infrared and visible image fusion, this paper proposes a fast alternating guided filter, which significantly increases the operation efficiency while ensuring the quality of the fused image. The proposed filter combines a convolutional neural network (CNN) and infrared feature extraction effective fusion. First, quadtree decomposition and Bessel interpolation are used to extract the infrared brightness features of the source images, and the initial fusion image is obtained by combining the visible image. Second, the information of the base layer and the detail layer of the source images is obtained through fast alternating guided filtering.

收稿日期: 2022-08-17; **修订日期:** 2022-09-26.

基金项目: 长江学者和创新团队发展计划资助(No. IRT_16R36); 国家自然科学基金(No. 62067006); 甘肃省科技计划项目(No. 18JR3RA104); 甘肃省高等学校产业支撑计划项目(No. 2020C-19); 兰州市科技计划项目(No. 2019-4-49); 甘肃省教育厅: 青年博士基金项目 2022QB-067; 甘肃省自然科学基金项目(No. 21JR7RA300); 兰州交通大学天佑创新团队(No. TY202003); 兰州交通大学-天津大学联合创新基金项目(No. 2021052)

The base layer obtains the fused base image through the CNN and Laplace transform, and the detail layer obtains the fused detail image through the saliency measurement method. Finally, the initial fusion map, basic fusion map, and detail fusion map are added to obtain the final fusion result. Because of the fast alternating guided filtering and feature extraction performance of this algorithm, the final fusion result contains rich texture details and clear edges. The experimental results indicate that the fusion results obtained by the algorithm have good fidelity in vision, and its objective evaluation indicators are compared with those of other methods. The information entropy, standard deviation, spatial frequency, wavelet feature mutual information, visual fidelity, and average gradient show improvements by 9.9%, 6.8%, 43.6%, 11.3%, 32.3%, and 47.1%, respectively, on average.

Key words: fast alternating guided filtering; infrared feature extraction; convolutional neural network; infrared and visible image fusion

1 引言

红外与可见光图像融合大多需要在弱光环境下进行监控,在这种环境下,可见光图像捕捉到被监控环境中大量的可见信息,而红外图像通常捕捉到不可见目标对象信息。红外图像是通过“测量”物体向外辐射的热量获得,描述的是目标与背景所保持温度的差异,而可见光图像的优点在于不容易受到温度等的影响,主要呈现场景中的细节以及纹理信息。由此在同一场景下,为了实现信息的互补,可通过红外与可见光图像融合技术来获取一张具有更全面场景目标信息的图像。到目前为止红外与可见光图像融合技术已在电力设备监测、资源勘探、生物医学、人脸识别和空间探索等领域得到广泛和深入的研究^[1-4]。

近年来,深度学习在图像融合等各个领域迅速发展并得到了广泛应用^[5-8]。文献[9]通过训练使得网络具有分类能力,利用二值化和两个一致性策略得到决策映射,并由该映射作为权重重构融合图像;文献[10]提出一种基于端到端的在生成器和判别器之间通过对抗性博弈产生融合图像的网络 FusionGAN;文献[11]为了更好地进行特征提取,提出了一种由编码层、融合层、解码层和输出层组成的基于监督卷积神经网络图像融合方法;文献[12]中采用完全卷积模型,以端到端的方式进行训练,利用两个卷积层从多个输入图像中提取特征,并根据输入图像类型选择适当的融合规则进行融合得到融合图像;文献[13]建立权重独立的双通道残差卷积神经网络,对可见光与红外频段下的目标图像进行特征提取,生

成多尺度复合频段特征图组的融合方法。目前学者们利用深度学习方法在红外与可见光图像融合上取得较好的融合效果,但是在融合过程中没有结合图像的尺度信息,往往影响融合图像的视觉效果,容易出现细节丢失、边缘模糊以及出现伪影等问题。

为解决以上问题以及提高算法运算速率,更好地提取源图像的低频信息,有效保留源图像的高频信息,从而进一步提升融合图像质量。本文提出一种改进的快速交替引导滤波(Fast Alternating Guided Filtering, FAGF)与深度学习网络 CNN 结合的红外与可见光图像融合算法,将源图像利用快速交替引导滤波进行分解,将分解后的细节层信息通过显著性测量来获得融合后的细节信息;基础层信息通过 CNN 与拉普拉斯结合来获得融合后的基础图像;与此同时将源图像通过红外特征的提取以及可见光信息的补充,来获得一个基础融合图像;最后将得到的三部分融合图像相加得到最终融合结果。实验结果表明,本文方法保留了更多源图像中的主要能量信息,提取到更多源图像的细节与纹理信息、边缘清晰,使融合图像具有更真实的视觉效果。

2 基本原理

2.1 交替引导滤波

滚动引导滤波器(Rolling Guidance Filter, RGF)旨在平滑图像中有灰度波动的纹理以及修复图像有噪声像素点,先去除双边滤波器中小于高斯核的细节,然后通过迭代更新引导图像来恢

复剩余细节的边缘^[14];与滚动引导滤波器 RGF 相似,平滑迭代恢复滤波器(Smooth and Iteratively Restore Filter, SIR)通过高斯滤波去除小尺度的细节,并使用边缘感知滤波器迭代恢复大尺度的边缘。RGF 迭代地过滤输入图像,同时迭代地恢复最初模糊的引导图像,SIR 则相反,它使用原始输入图像作为引导,逐渐恢复最初模糊的图像。

交替引导滤波(Alternating Guided Filtering, AGF)结合了 RGF 的大规模边缘和局部强度保持特性^[15],具有 SIR 的边缘恢复特性,即通过 RGF 进行边缘曲率平滑和局部强度降低以及通过 SIR 恢复大尺度边缘附近的小尺度细节。AGF 步骤如下:

(1)对输入图像应用联合双边滤波器去除小结构。在公式(1)中, I 和 O 分别表示输入和输出图像, i 和 j 表示图像中像素坐标的索引。 α 表示归一化因子, f 和 r 分别表示空间滤波器核范围和强度滤波器核, Ω 表示核的空间支持。通过使用先前迭代步骤对输入图像应用联合滤波器获得第 t 次迭代滤波结果 O^{t+1} :

$$O_i^{t+1} = \frac{1}{a} \sum_{j \in \Omega} I_j \cdot f(\|i - j\|) \cdot r(\|O_i^t - O_j^t\|). \quad (1)$$

(2)使用源图像作为引导图像,将联合双边滤波器应用于上一步获得的结果。详细过程可以描述如式(2)所示:

$$O_i^{t+1} = \frac{1}{a_i} \sum_{j \in \Omega} Q_j \cdot f(\|i - j\|) \cdot r(\|I_i - I_j\|). \quad (2)$$

(3)对(2)的结果进行中值滤波。然后,过滤

后的结果保证了图像强度和大尺寸边缘曲率的保留。中值滤波器的应用可防止在大尺度边缘附近重新引入过滤后的小尺度细节。

2.2 改进的快速交替引导滤波

为了在保证融合图像质量的前提下提高交替引导滤波的运行速度,本文在 AGF 的基础上进行改进,使得运行时间更快,运行效率更高。本文所提出的改进快速交替引导滤波步骤以及结构模型如下所示:

(1)下采样处理:对采用最近邻域插值算法对输入图 p 和引导图像 I 尺寸为 $M \times N$,进行比率为 s 的下采样操作,使得图像成为 M, N 分别缩小 s 倍的图像,其中每一个像素点的值即该 $s \times s$ 尺寸窗口内所有像素的均值。

(2)RGF 操作过程:利用上一次迭代的结果 O' 作为引导图像,对子采样后的图像应用联合双边滤波器去除小结构。

(3)SIR 操作过程:使用源图像作为引导图像,将联合双边滤波器应用于前一步获得的结果。

(4)对第三步中得到的结果进行小核(例如 3×3)中值滤波处理。

(5)上采样处理:采用双线性内插算法用四个最近邻估计给定的灰度做了二次一维的线性插值,在原有图像像素的基础上利用四个邻域像素的灰度在两个方向上进行内插插入新的元素。使得新图像的像素点对应落在原始图像四个像素点中间,通过系数映射进行双线性上采样到原始大小。

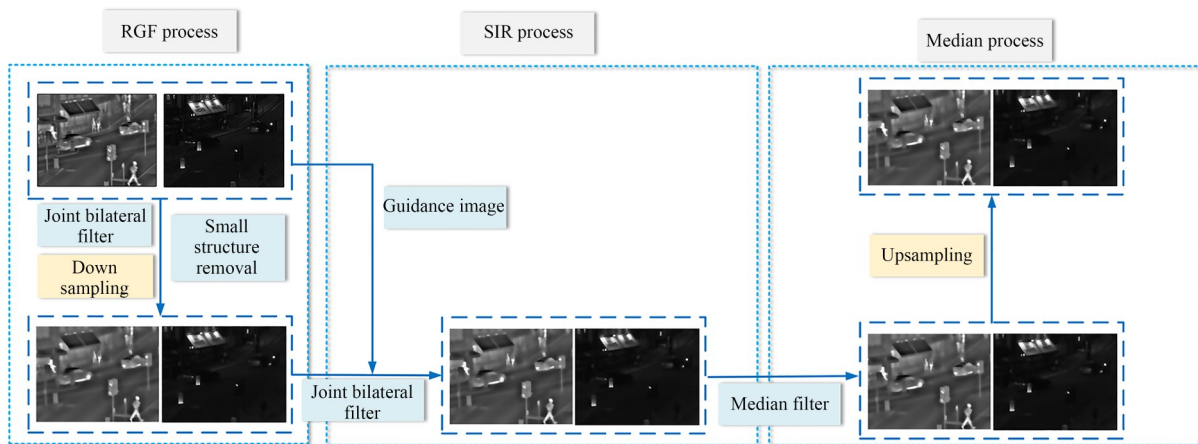


图1 FAGF 结构模型

Fig.1 FAGF structural model

为了验证快速交替引导滤波器的有效性,本文选择图 8 中的 6 组图像进行仿真实验。快速交替引导滤波这种加速策略即对滤波输入图像和引导图像进行子采样,计算局部线性系数,并对这些系数进行上采样,在原始的引导图像上采用上采样系数来产生输出。经实验表明,对于子采样率 s ,该方法将时间复杂度 $O(N)$ 降低到 $O(N/s^2)$ 。该方法用于本算法中不仅提升了原有交替引导滤波的运行速率,同时对融合效果得到了保持(如图 6、图 7 所示),快速交替引导滤波器能够更快更精确地分解出包含轮廓信息基础层与细节纹理信息细节层,有助于提升最终融合效果。以下是 AGF 与 FAGF (Fast Alternating Guided Filtering, FAGF) 对分解图像的运行时间对比,运算时间平均减少 81.4%:

表 1 FAGF 运行时间对比(单位:秒)

Tab. 1 FAGF running time comparison (unit:s)

	AGF	FAGF
Img1	0.296 764	0.055 198
Img2	0.772 967	0.143 772
Img3	1.051 652	0.189 297
Img4	0.756 234	0.140 659
Img5	0.345 330	0.055 252
Img6	0.663 786	0.146 032

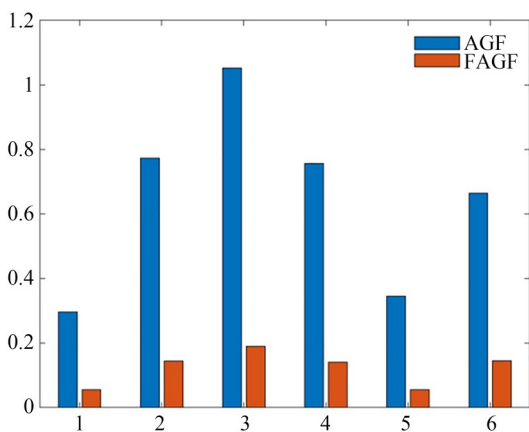


图 2 FAGF 运行时间对比柱状图

Fig. 2 FAGF runtime comparison histogram

2.3 卷积神经网络(CNN)框架

卷积神经网络 CNN 是一种较强的图像特征表示方法,在计算机视觉和图像处理中取得了巨大的成功。通常在图像融合算法中使用的卷积神经网络是一种孪生网络^[9]。CNN 在图像融合领域中卷积层和全连接层可以分别被视为图像融合中的活动水平测量和权重分配部分,在特征提取过程中,能够较好地提取源图像更多的细节以及纹理信息,应用于图像融合领域的网络模型如图 3 所示。可以看出,网络中的每个分支都有三个卷积层和一个最大池化层。将每个分支得到的特征图串联起来,然后用与特征图相同大小的特征向量进行全连接后被馈送到 softmax 层。

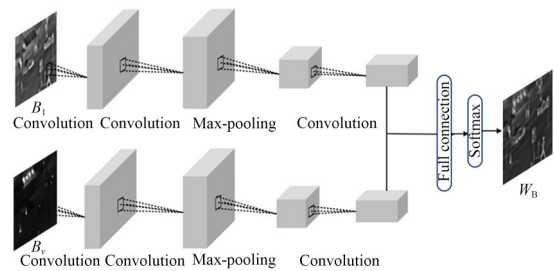


图 3 CNN 网络模型

Fig. 3 CNN network model

3 本文算法

本文提出了一种基于快速交替引导滤波 (FAGF) 的红外与可见光图像融合方法。具体步骤如下:

(1) 将源图像通过红外特征的提取以及可见光信息的补充,来获得一个基础融合图像;

(2) 源图像通过 FAGF 分解为包含大尺度强度变化信息的基础层与包含小尺度细节信息的细节层;

(3) 基础层通过 CNN 框架与拉普拉斯算子进行特征提取及融合,得到融合后的基础层融合图;

(4) 细节层信息通过显著性测量来获得融合后的细节层融合图;

(5) 将得到的三部分融合图像相加得到最终融合结果。以下是本算法融合框架:

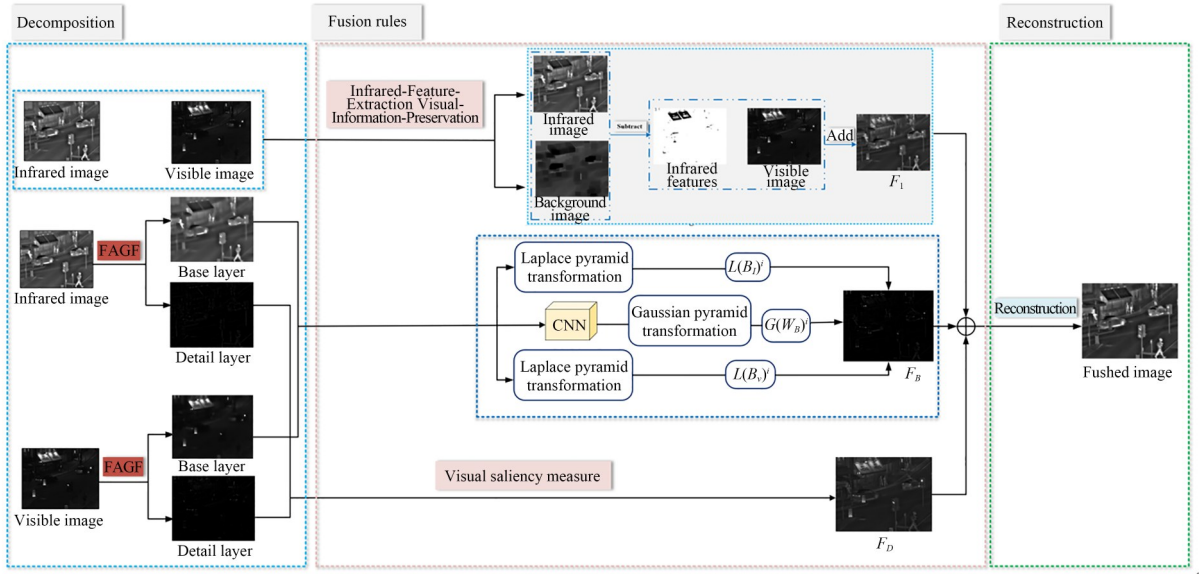


图4 本文算法思路框图

Fig. 4 Algorithm idea block diagram of this paper

3.1 红外特征提取

首先,为了对红外图像中所含的目标信息进行估计,本文采用四叉树结构将红外图像进行分解,其分解步骤如下所示:

(1)以红外图像初始化四叉树结构的根,计算当前根子块中的最大灰度值与最小灰度值的差值进行判断,若该值大于预先设定的阈值 t ,则将其分割为四个更小的子块,直到差值小于或等于 t 为止。用于判断四叉树中的块是否需要进一步分解,通常选择一个较小的阈值,通过噪声像素来抑制一个平滑块的波动。而允许的最小尺寸应该大到足以覆盖整个或大部分红外物体,因此本文将红外物体的最大尺寸设置为参数。取红外图像前5%灰度的平均值作为阈值,以此来估计出更多的红外信息。

(2)从四叉树结构的每个子块中均匀采样16个点,并设置这些点的灰度值为对应位置的局部最小灰度值。然后根据对应的16个点的纵横坐标和灰度值进行插值,对每个子块进行贝塞尔曲面的重构,如下式所示:

$$H(q, v) = QMPM^T V^T, \quad (3)$$

其中: (q, v) 为插值点,插值点的位置用0~1的插补比表示。 (Q, V) 为变量插值系数, M 为常插值系数矩阵。 P 表示由16个点组成的 4×4 矩阵。其中, Q, V, M 和 P 定义如下:

$$Q = [q^3, q^2, q^1, q^0], \quad (4)$$

$$V = [v^3, v^2, v^1, v^0]$$

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 3 & 0 & 0 \\ 3 & -6 & 3 & 0 \\ -1 & 3 & -3 & 1 \end{bmatrix}, \quad (5)$$

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{bmatrix}, \quad (6)$$

其中: p_{ij} 表示16个点坐标的维度特征。贝塞尔通过均匀设置 i 和 j ,可以重构出每个子块的贝塞尔曲面。由于采用不同的点,可能会导致贝塞尔曲面存在伪影,因此对红外背景采用高斯滤波器进行平滑操作,如式(7)所示:

$$I_{FB} = I_B * g(s, \sigma), \quad (7)$$

其中: s 表示高斯滤波器大小, σ 为滤波器参数。由于高斯滤波器起到平滑图像的作用,其方差越大,模糊效果越大,通过不同参数的实验对比。在本文中,将高斯滤波器尺寸设置为 $s=11$ 和 $\sigma=5$ 。通过上式可以得到一个平滑后的红外背景图像 I_{FB} ,然后,通过从红外图像 I 中减去红外背景图像 I_{FB} ,即可提取出红外亮度特征。

$$F_1' = \max(I - I_{FB}, 0). \quad (8)$$

最后将红外亮度特征直接加到可见光图像中得到初始融合图像 F_1 :

$$F_1 = F_1' + V. \quad (9)$$

3.2 FAGF 分解

为保留更多的细节以及纹理信息,并提高算法运行效率,本文采用 FAGF 方法进行源图像双尺度的分解,具体步骤如式(10)所示:

应用 FAGF 滤波将源图像 I_1, I_2 分解为基础层 B_1, B_V 和细节层 D_1, D_V 。基础层如下所示:

$$B = FAGF(I, \delta_r, \delta_s, n, s), \quad (10)$$

其中: I 表示作为引导图像的源图像, δ_r 表示双边滤波器的空间权重。 δ_s 表示双边滤波器的范围权重, n 表示滚动引导的迭代次数, s 表示变化率,本文通过多次实验对比,当 $\delta_s=5, \delta_r=0.05$ 时图像边缘平滑效果最好,因此选择该参数。

(2)细节层是从源图像中减去基础层,定义如下:

$$D = I - B. \quad (11)$$

3.3 基础层融合

(1)由于 CNN 网络具有网络结构局部连接和权重共享的特点,具有复杂特征表达能力,通过 CNN 网络获得图像的权重图能够较好地提取到源图像中的边缘细节纹理特征,本文将分解得到的基础图像 B_I 和 B_V 通过预训练的 CNN 模型进行处理,得到低频系数的最终决策图 W_B 。本文中,如图 5 所示通过一个稳定的活跃水平测量和权重分配策略,通过训练卷积神经网络 CNN 来获取从源图像到权重图的直接映射,每个分支由三个卷积层和一个最大池化层组成,卷积核采用

3×3 大小,步长为 2,在全连接层将全连接层转换为两个卷积层,转换后的网络在处理图像时生成一个包含图像对应位置的二维向量,并对该二维向量的总和进行归一化,最后,将该值分配为对应位置的像素值,并对重叠像素进行平均得到基础图像的权重图。

$$W_B = CNN(B_I, B_V). \quad (12)$$

(2)其次,分别用拉普拉斯金字塔 $Lap(\cdot)$ 分解 B_I 和 B_V ,并通过高斯金字塔 $Gauss(\cdot)$ 分解基础决策图 W_B 。其中每个金字塔的分解级数由源图像大小设定,源图像空间大小为 $H \times W$,则 $i = \lfloor \log_2 \min(H, W) \rfloor$ 。公式如式(13)所示:

$$\begin{cases} L(B_I)^i = Lap(B_I, i) \\ L(B_V)^i = Lap(B_V, i) \\ G(W_B)^i = Gauss(W_B, i) \end{cases} \quad (13)$$

(3)采用局部相似度计算 B_I 和 B_V 的能量系数。

$$E(B_I)^i(x, y) = \sum_m \sum_n L(B_I)^i(x + m, y + n)^2, \quad (14)$$

$$E(B_V)^i(x, y) = \sum_m \sum_n L(B_V)^i(x + m, y + n)^2. \quad (15)$$

用于融合模式矩阵行列式值的相似性度量方法如式(16)所示:

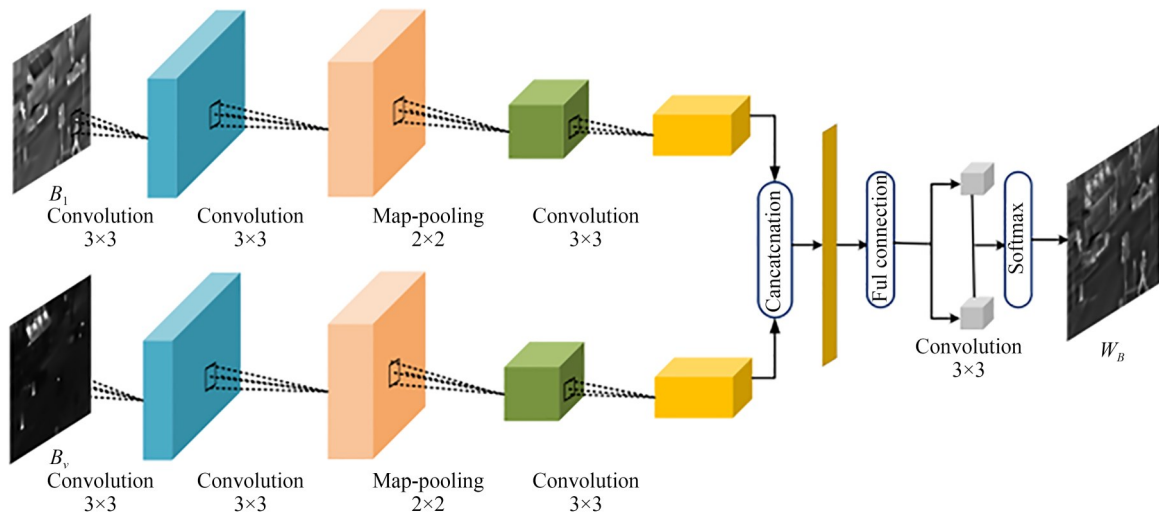


图 5 本文 CNN 网络模型

Fig. 5 CNN network model in this paper

$$M_{(B)}^i(x, y) = \frac{2 \sum_m \sum_n L(B_I)^i(x+m, y+n) L(B_V)^i(x+m, y+n)}{E(B_I)^i(x, y) + E(B_V)^i(x, y)}. \quad (16)$$

(4)其中相似性度量算法中匹配值的取值范围为[-1 1],本文设置一个阈值 $t=0.6$ 来确定融合模式,若上式中的匹配值大于等于阈值 t ,则采用平均权重方式;若匹配值小于阈值 t ,则采用最大像素值作为中心像素。

(5)对融合后的低频子带 $L(F_B)^i$ 进行拉普拉斯逆变换得到 F_B :

$$L(F_B)^i(x, y) = G(W_B)^i(x, y) \cdot L(B_I)^i(x, y) + (1 - G(W_B)^i(x, y)) \cdot L(B_V)^i(x, y), \quad (17)$$

$$L(F_B)^i(x, y) = \begin{cases} L(B_I)^i(x, y), & E(B_I)^i(x, y) \geq E(B_V)^i(x, y), \\ L(B_V)^i(x, y), & E(B_I)^i(x, y) < E(B_V)^i(x, y), \end{cases} \quad (18)$$

$$F_B = R_lap(L(F_B)). \quad (19)$$

3.4 细节层融合

VSM (Visual Saliency Measure, VSM) 可以反映显著的图像特征,并根据图像的显著值分配不同的融合权重,有良好的融合效果,同时增强

图像对比度,使得图像具有更好的人眼视觉效果, VSM 方法具有简单性和有效性^[16]。令 V_1 和 V_2 分别表示输入红外图像和可见光图像的显著性值。可通过下式获得融合细节层 F_D :

$$F_D = W_d D_I + (1 - W_d) D_V, \quad (20)$$

其中权重 W_d 定义为:

$$W_d = 0.5 + \frac{V_1 - V_2}{2}. \quad (21)$$

根据加权平均融合规则,如在某些位置 $V_1 = V_2$,则权重 W_d 为 0.5。如 $V_1 > V_2$,则权重 W_d 将大于 0.5,融合图像将融合来自细节层 D_I 更多信息。 $V_1 < V_2$,融合图像将融合来自细节层 D_V 的更多信息。该方法保证了融合后的图像都是由两幅源图像中显著性较高的信息组成。经过以上的分析和运算,最终的融合结果 F 可由式(22)得到:

$$F = F_1 + F_B + F_D. \quad (22)$$

4 实验结果及分析

4.1 快速交替引导滤波性能分析

如图 6 所示为 RGF, SIR, AGF 和 FAGF 四

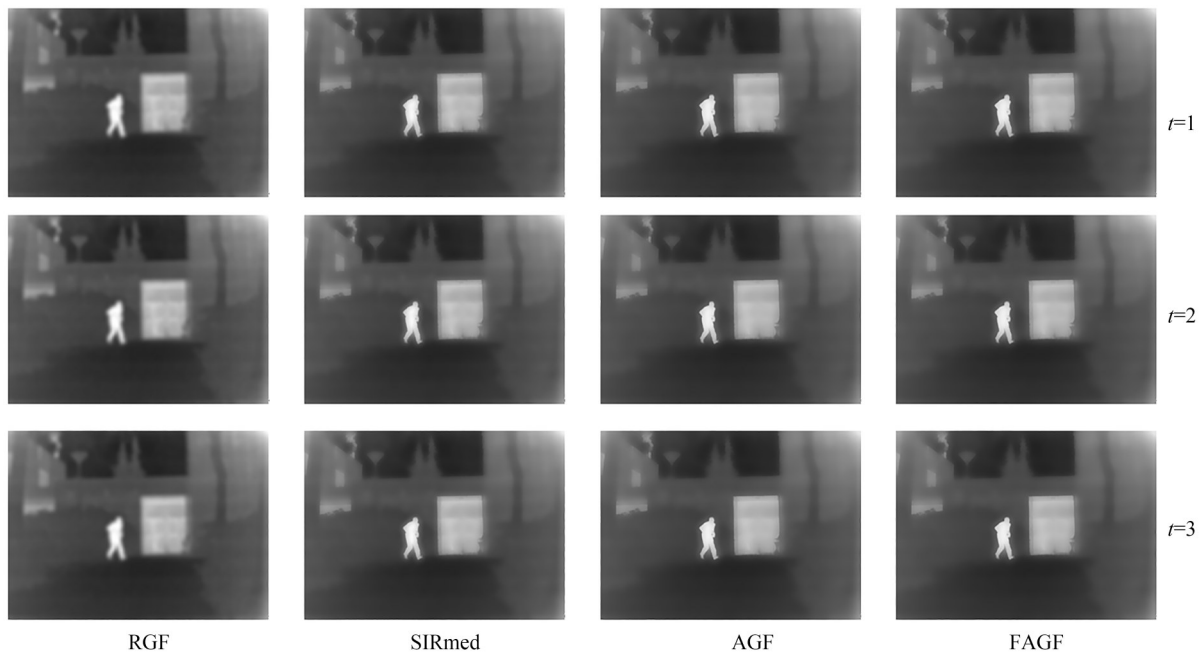


图 6 RGF, SIR, AGF 和 FAGF 的滤波结果

Fig. 6 Filtering results of RGF, SIR, AGF and FAGF

种滤波器当双边滤波器的 $\delta_s=5, \delta_r=0.05$ 时,前 3 次迭代的滤波结果(在图中记为 t)。在此过程中滤波器会迭代地大规模恢复图像边缘,RGF 逐渐平滑边缘曲率,SIR 则会平滑图像整体强度信息,从而降低图像全局对比度。AGF 克服了 RGF 曲率平滑和 SIR 局部对比度低的问题,即在不平滑曲率和保持图像亮度的情况下恢复大尺度边缘,同时克服图像边缘模糊的问题。FAGF 在保持 AGF 恢复大尺度边缘的同时,提高了 AGF 算法的运算效率。

图 7 选用具有代表性的像素区域路灯作为滤波性能分析对象,当 $t=3$ 时 RGF, SIR, AGF, FAGF 的滤波结果、滤波结果放大图以及各自对应的三维截面图。三维图像可以看到像素的分布情况以及像素值的大小,选取 Iteration=3 的滤波结果图进行分析,可以看出由于 RGF 滤波器具有小结构消除以及边缘保持、边缘曲率

平滑的特点,因此 RGF 的等高线像素分布均匀,而且 RGF 会逐渐平滑大尺度边缘的曲率边缘。而 SIR 弥补了 RGF 的缺点,同时较好地表示产生的边缘的曲率特性。从三维图可以看出由于 SIR 在平滑过程中,会平滑掉图像整体图像强度,导致图像降低全局对比度,像素值整体较低,与 RGF 相比大尺度峰值总体降低,AGF 在不平滑曲率和保持图像亮度的情况下恢复大尺度边缘,同时克服图像边缘模糊的问题。相比之下,FAGF 结合了 RGF 与 SIR 的优点,在防止平滑曲率和保持整体图像亮度以及对比度的同时恢复了大尺度边缘的同时相比 AGF 提高了算法的运算速率。从三维截面图可以看出,FAGF 在平滑曲率和保持局部图像对比度的情况下恢复了大尺度边缘,保持了图像更多的边缘区域。

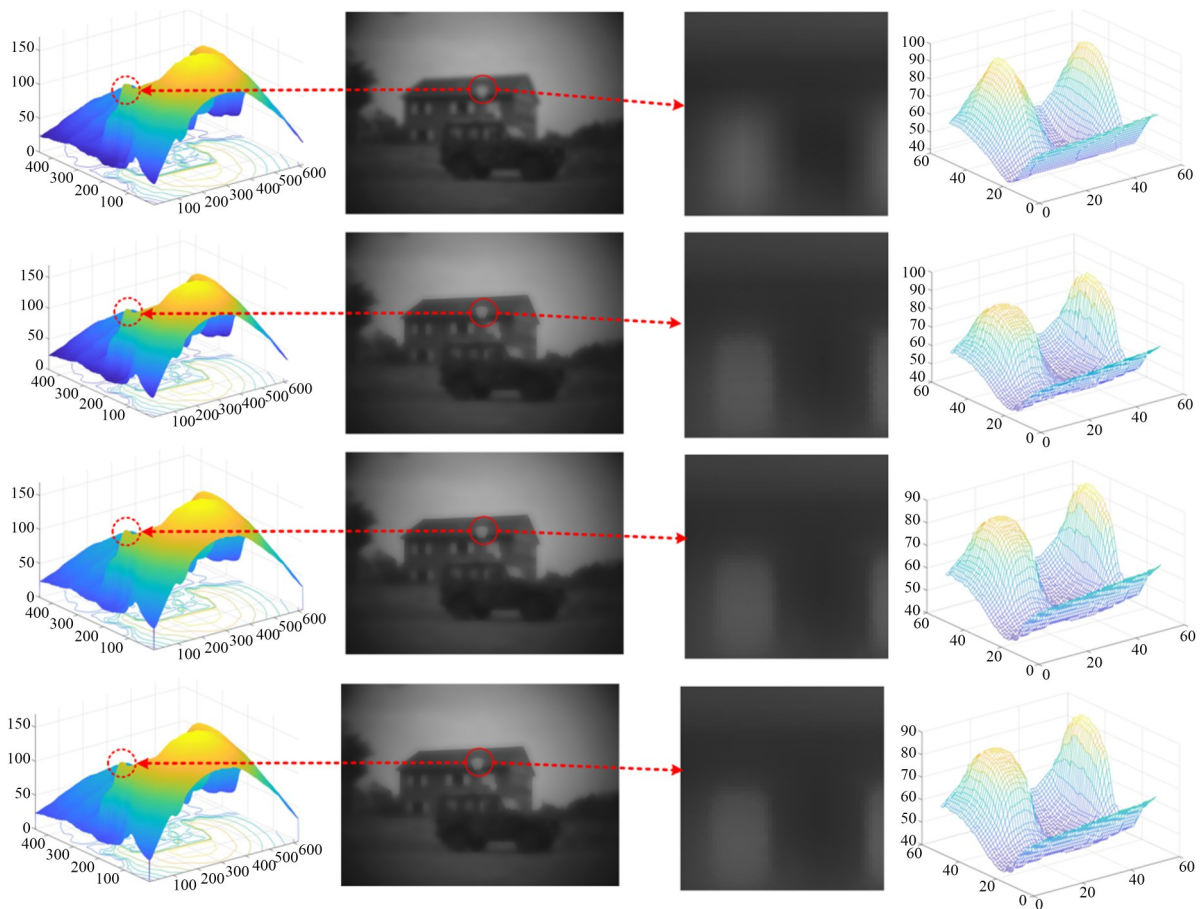


图 7 RGF, SIR, AGF, FAGF 滤波结果分析
Fig. 7 Analysis of RGF, SIR, AGF, FAGF filtering result

4.2 实验说明

本文红外与可见光源图像选取公共数据集 TNO^[17]中的 21 组图像进行实验,实验仿真平台采用 Intel 酷睿 i7,编程环境为 Matlab2019b。为验证算法性能,本文选定 6 组已配准图像进行实验分析,其大小分别为 360×270 pixel, 632×496 pixel, 505×510 pixel, 640×480 pixel, 360×270 pixel, 595×328 pixel。为验证本文算法的可行性与有效性,将本文方法(简称 FAGF)与 WLS 方法^[18]、GF 方法^[19]、VGG 方法^[20]、RES 方法^[21]与 PIAF 方法^[22]进行实验对比。

其中 WLS 方法是一种基于视觉显着性图(VSM)和加权最小二乘(WLS)优化的新型多尺度融合方法;GF 方法是一种新的基于模糊梯度阈值函数和全局优化的滤波器,克服了传统基于局部加权平均的滤波器在分解过程中由于亮度的色散而导致的边缘模糊问题,保证图像的整体亮度和边缘信息;VGG 方法是利用深度学习框架 VGGNet-19 网络的红外与可见光图像融合方法;RES 方法一种基于 ResNet152 深度学习模型的红外与可见光图像融合算法;PIAF 方法通过一个光照感知子网络来估计光照分布并计算光照概率,利用光照概率来构建光照感知损失来指导融合网络的训练,是一种基于光照感知的渐进式红外和可见光图像融合网络。

本文算法中 CNN 网络的训练集来自 ImageNet 数据集的 50 000 张高质量图像。首先对每幅图像进行灰度转换与高斯滤波模糊,其次,随机采样 20 对 16×16 的块,根据输入和输出神经元的数量自适应地决定每个卷积层的权值初始化的规模。softmax 损失函数作为优化目标,通过随机梯度下降(SGD)算法使其最小化。训练过程在深度学习框架 Caffe 上进行操作。通过 200 万个实例训练,经过 10 个迭代,最终得到了训练后的网络。

当高斯滤波器尺寸设置为 $s=11$ 和 $\sigma=5$, FAGF 参数 $\delta_s=5$, $\delta_r=0.05$, $n=3$ 时效果最好,因此将它们作为本文方法的默认参数设置。

4.3 主观评价

通过 6 组实验的融合结果如图 8 所示,其中红框内容为每组图像的局部细节放大图。Img1 的融合图像中 WLS 方法、GF 方法、VGG 方法、

PIAF 方法融合效果整体模糊,目标对象周围出现伪影,且细节信息丢失树木边缘有虚假信息的掺入,RES 方法融合整体效果好,目标对象显著,但是和本文方法相比纹理细节信息不够清晰,相比之下,本文方法目标清晰,能够清楚看到树木的分支结构。Img2 的融合图像中 GF 方法、VGG 方法和 RES 方法相较于 FAGF 方法由于边缘不够清晰,导致目标信息不够显著,其中 GF 方法屋顶、车辆以及人物产生了虚影,VGG 方法整体融合效果色调偏暗,丢失路面的纹理信息,路灯边缘模糊,VGG 方法融合图像左上方橱窗里的信息不如 FAGF 方法丰富;WLS 方法对比度高,目标对象显著,但是目标边缘较为模糊,PIAF 方法整体融合效果好,但是色调偏亮,丢失一部分可见光信息,导致地面以及车辆的纹理信息不够清晰,对比度不高。而 FAGF 方法在目标清晰的前提下,人物以及建筑物的边缘轮廓更加清晰,人物身上有较对比方法更多的纹理信息。Img3 的融合图像中 GF 方法、VGG 方法和 RES 方法从融合色调来看,融合了更多的红外图像信息,但是整体色调偏暗,且船只轮廓不够清晰,而且丢失了水面波纹,WLS 方法和 PIAF 方法融合效果较好,WLS 方法,水面以及天空的纹理较明显,但不如 FAGF 方法色调丰富,PIAF 方法船只的纹理信息较为清晰,但是船只的影子泛白较为严重,相比之下,FAGF 方法水面及天空的纹理信息丰富,饱和感更高,船只上面保留了较多可见光信息,融合效果相较于其余对比方法具有更高的保真度。Img4 的融合图像中 WLS 方法和 RES 方法融合效果较好,但是缺乏对比度低,地面信息较为模糊,缺乏人物信息,VGG 方法整体融合效果偏暗,图像边缘处轮廓不清晰,地面处有较为明显的块效应,从而影响整体融合效果;GF 方法和 PIAF 方法对比度不高,且 GF 方法在天空以及房屋边缘处有大面积的晕影导致目标对象不显著,FAGF 方法在房屋轮廓有较为明显的优势,除了能清楚看见人物信息之外,同时房屋边缘清晰可见,草地细节信息丰富。Img5 的融合图像中从整体融合效果看,WLS 方法和 PIAF 方法泛白较为严重导致图像中的人物信息不够明确,树木分支模糊不清晰,GF 方法在树木范围

产生了大量虚影导致人物目标轮廓模糊,融合结果色调偏暗,VGG方法和RES方法融合效果较好,但是对比度不高,目标对象显著,人物身上可以看到较为丰富的纹理信息,但是人物轮廓不清晰,边缘模糊,而FAGF方法除了能够清楚的看到人物目标之外,还可以看到有较为清晰的树木的分支,包含较对比方法更多的可见光信息。Img6的融合图像中从树木边缘来看,WLS方法、GF方法、VGG方法与PIAF方法均出

现边缘晕染,目标不够显著的特征。而FAGF方法,有较为清晰的树木边缘信息,并且FAGF方法在飞机头部位包含更丰富地细节信息,能够更清晰地看出飞机头上面地纹理,具有更高的保真度。

综上,通过上述实验对比可以看出,本文所提方法其融合结果均有更加清晰的边缘细节信息,晕影和块效应不明显,整体融合效果图的场景信息丰富具有更高的对比度。

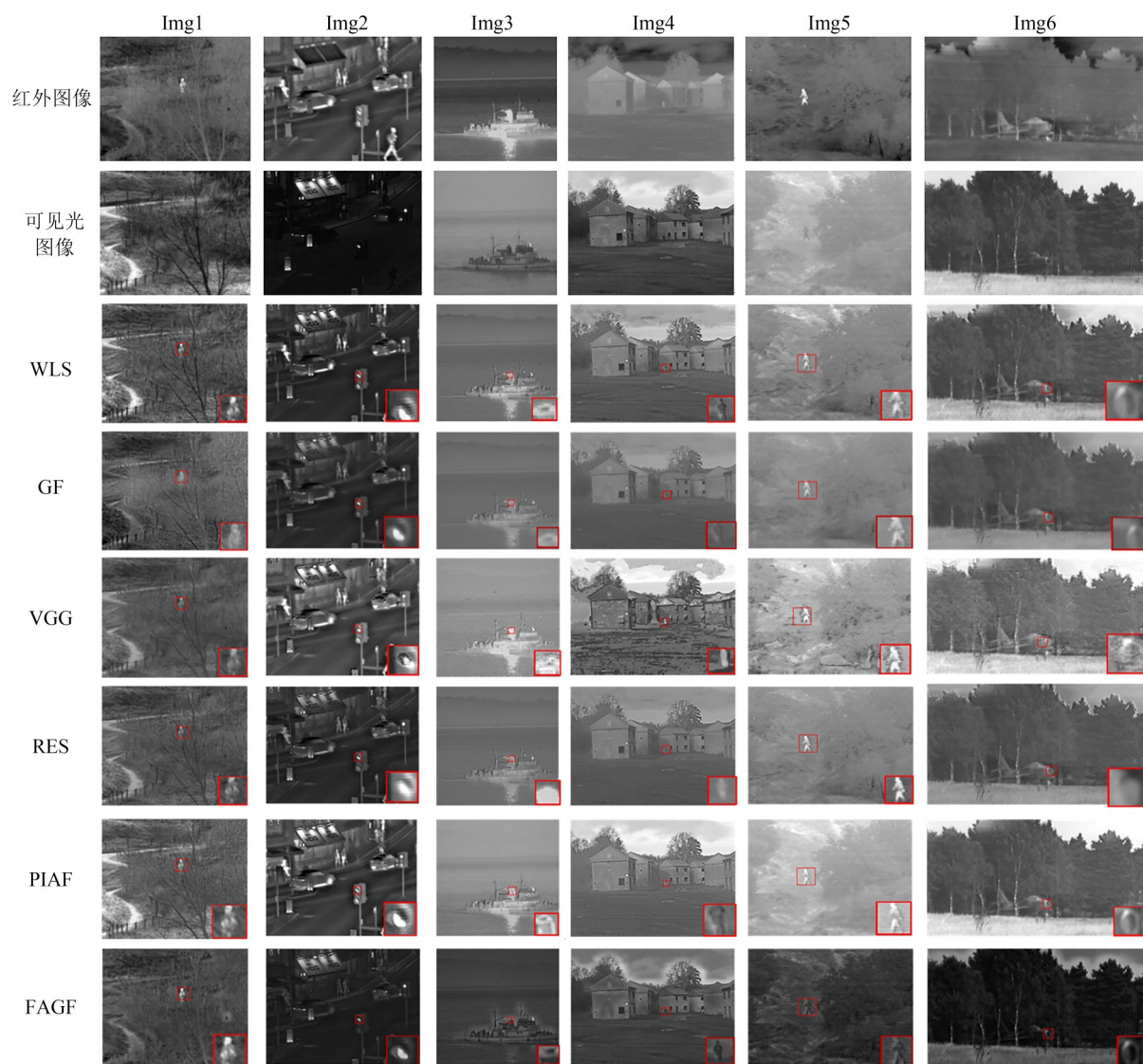


图 8 实验结果图

Fig. 8 experimental results

三维分析对比图可以较好地表示融合结果图中细节与纹理信息的分布情况。如图 9 所示,以 Img2 为例,可以看出同一场景下红外图像和

可见光图像在三维分析中所表现出来的特征信息,红外图像可根据温度的高低区分图像,可见光图像则根据光照的亮度来区分场景信息。相

比 FAGF 方法, WLS 方法与 RES 方法虽分布较为均匀, 但细节边缘线条不及本文算法清晰; GF 方法缺少可见光信息, 整体亮度较低; VGG 方法从图像看均有少量噪声伪影出现, 缺少红外信息在三维分布中可以看出轮廓信息有过拟合现象;

PIAF 方法信息分布均匀, 但相比 FAGF 方法, 提取到的可见光信息较少, 即图像目标显著但是, 边缘模糊; 因此, 从三维分析图可以看出, 本文算法具有细节纹理信息丰富, 边缘信息清晰的优势。

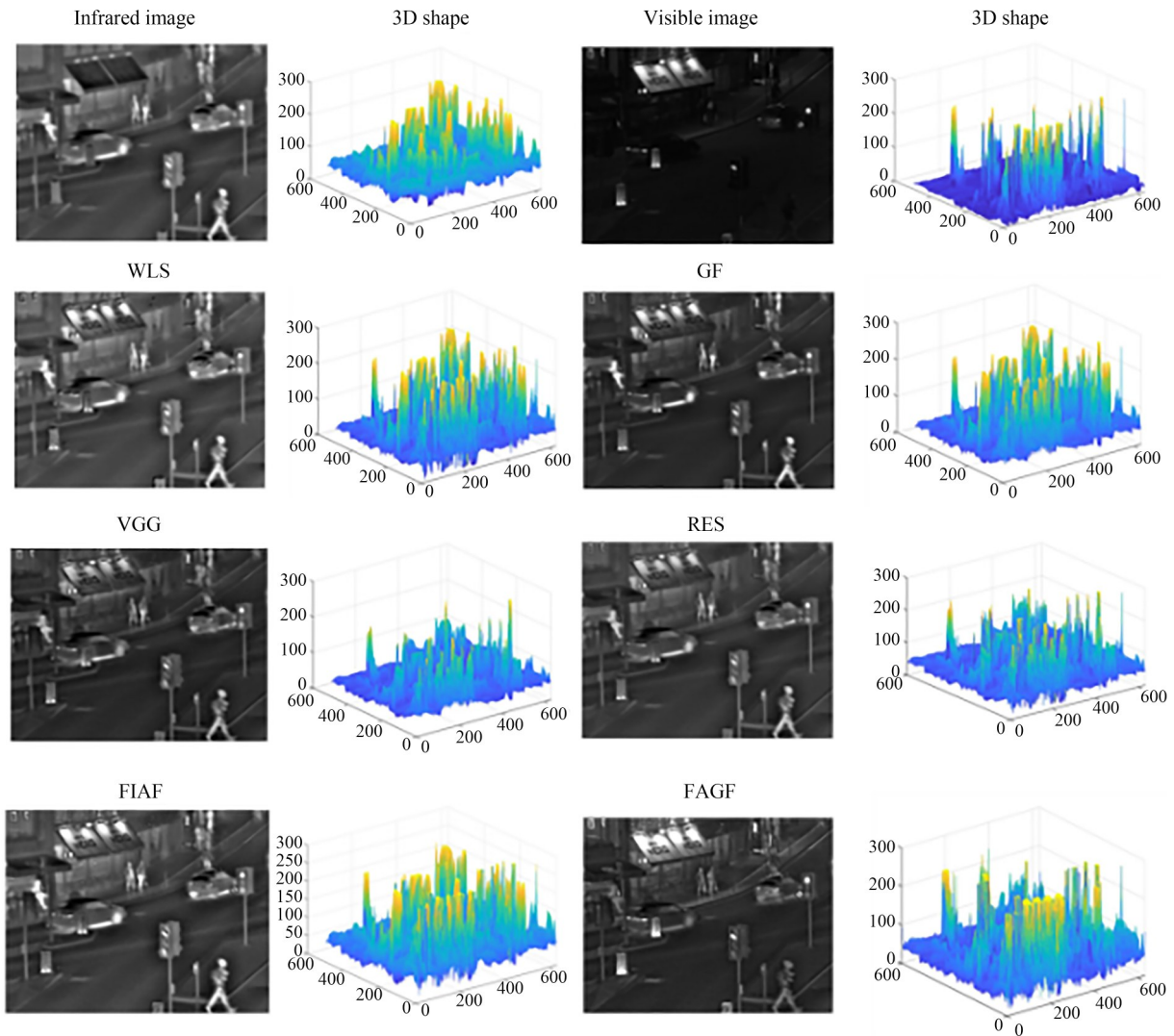


图 9 融合结果三维对比分析图

Fig. 9 Three-dimensional comparative analysis chart of fusion results

4.4 客观评价

为验证算法性能及图像融合质量, 本文对融合结果的定量分析主要采用 6 种在图像融合方面广泛使用的评价指标, 依次为信息熵(EN)^[23]、标准差(SD)、空间频率(SF)^[24]、小波特征互信息(FIM_W)^[25]、视觉保真度(VIF)^[26]、平均梯度(AG)。以上六种指标值越大, 则说明融合图像与源图像具有较高的信息相关性, 即在融合过程

中损失的细节信息越少, 反映其相似程度。

6 组实验图像的客观评价指标值如表 2 所示。通过实验指标值可以看出本文方法的 EN, SF, AG 指标均优于其余 5 种对比方法, 指标 SD, FIM_W, VIF 指标值有绝大多数处于最优。表明本文方法能够从源图像中提取到更多的细节边缘信息, 包含更丰富的纹理结构, 具有较高的保真度。

表 2 客观评价指标对比
Table 2 Comparison of objective evaluation indicators

	Method	EN	SD	SF	FMI_W	VIF	AG
Img1	WLS	6.643 5	7.908 8	0.056 8	0.323 7	0.651 3	6.367 8
	GF	6.510 4	7.149 2	0.040 7	0.490 9	0.444 6	4.791 5
	VGG	6.099 6	7.382 6	0.026 9	0.382 9	0.487 8	3.161 7
	RES	6.237 5	7.573 3	0.039 2	0.398 8	0.513 3	4.550 7
	PIAF	6.709 8	7.998 9	0.041 5	0.427 6	0.756 4	5.116 5
	FAGF	6.915 8	8.211 4	0.063 4	0.577 3	0.758 3	6.478 9
Img2	WLS	6.138 7	7.047 1	0.052 4	0.357 5	0.761 6	3.397 3
	GF	6.620 5	7.994 7	0.037 0	0.353 5	0.735 8	2.448 3
	VGG	5.921 6	7.146 2	0.028 7	0.372 5	0.726 4	1.957 8
	RES	6.031 9	7.120 7	0.044 7	0.408 9	0.736 1	2.787 5
	PIAF	6.745 2	7.882 0	0.051 3	0.387 5	0.990 9	3.839 4
	FAGF	6.830 5	7.183 1	0.063 9	0.409 2	0.789 2	4.108 7
Img3	WLS	5.319 3	5.275 6	0.029 8	0.363 6	0.577 3	2.184 9
	GF	6.439 0	7.655 1	0.012 3	0.345 0	0.564 4	1.086 6
	VGG	4.888 8	5.075 9	0.015 3	0.356 5	0.440 6	1.157 7
	RES	5.021 2	5.131 1	0.024 3	0.402 6	0.485 3	1.767 4
	PIAF	5.747 0	7.900 9	0.018 8	0.362 7	0.729 0	1.408 6
	FAGF	6.818 1	6.604 1	0.030 8	0.404 1	0.903 4	2.812 8
Img4	WLS	6.849 9	9.366 2	0.057 7	0.397 0	0.847 9	4.387 9
	GF	6.480 2	8.688 3	0.048 7	0.434 4	0.625 5	3.422 7
	VGG	6.205 1	9.429 1	0.030 0	0.387 3	0.607 7	2.175 2
	RES	6.367 5	9.537 3	0.049 3	0.424 6	0.639 2	3.485 9
	PIAF	6.824 9	8.482 3	0.057 0	0.419 7	0.773 5	4.242 4
	FAGF	7.186 3	9.805 6	0.094 1	0.459 5	1.088 6	7.191 4
Img5	WLS	6.315 2	7.754 2	0.040 8	0.379 5	0.994 6	3.852 7
	GF	6.020 5	8.634 9	0.020 4	0.434 5	0.663 5	2.172 7
	VGG	5.904 5	7.547 3	0.019 6	0.403 4	0.742 6	1.899 1
	RES	5.979 0	7.560 0	0.030 6	0.429 2	0.800 3	2.843 4
	PIAF	6.304 6	8.097 4	0.024 6	0.412 7	1.156 0	2.617 2
	FAGF	6.554 7	8.496 4	0.050 2	0.427 1	1.228 2	5.467 9
Img6	WLS	6.609 7	8.354 9	0.025 8	0.297 1	0.929 6	2.585 7
	GF	5.856 4	7.114 0	0.017 3	0.331 3	0.643 5	1.721 4
	VGG	6.187 3	8.440 0	0.013 4	0.302 6	0.746 7	1.401 2
	RES	6.256 7	8.441 4	0.019 3	0.334 9	0.783 6	1.952 8
	PIAF	6.663 5	8.436 0	0.026 1	0.334 9	0.976 6	2.497 5
	FAGF	6.711 3	8.622 6	0.041 5	0.380 0	1.135 9	4.308 0

综合主客观评价指标,本文所提算法大多数指标均优于对比融合方法,较好地保留了可见光图像中的边缘及纹理信息,同时突出了红外图像的目标,具有更优的清晰度与对比度。

如图 10 所示,分别为 5 种对比方法在 6 种不同指标下的指标数值折线图,从上图表可以清晰地看出本文方法所存在指标方面的优越性。从 EN、SF、AG 的指标对比折线图来看,本文方法的融合效果均处于最优。

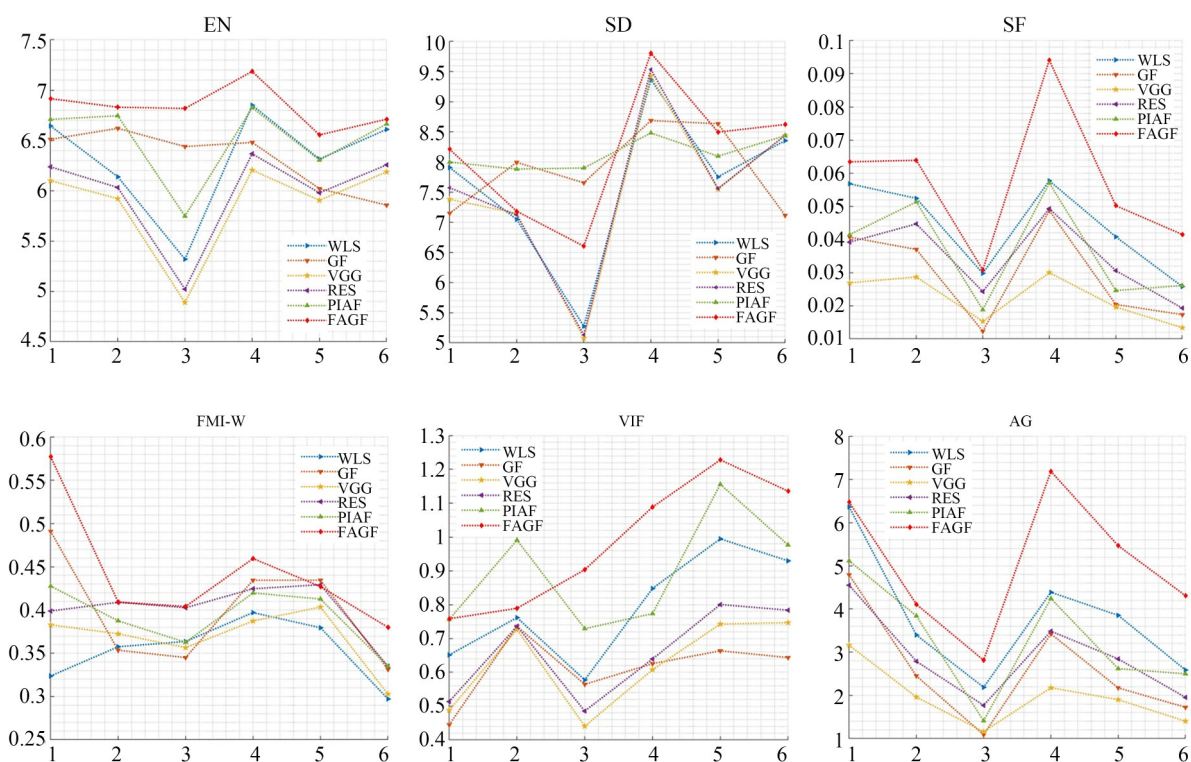


图 10 指标对比折线图

Fig. 10 Indicator comparison line chart

5 结 论

本文结合快速交替引导滤波以及 CNN 的红外与可见光图像的融合算法,在保证融合图像质量的前提下有效提高运行过程中交替引导滤波的运算效率,并改善了已有方法中存在的融合结果图细节信息丢失,边缘模糊以及产生伪影的问题。本算法涉及到快速交替引导滤波的分解,卷积神经网络的特征提取性能,以及四叉树分解和贝塞尔插值来提取红外亮度特征结合可见光图像信息得到融合图像。实验结果表明,本文方法融合结果图包含丰富的纹理细节信息,边缘清晰。通过计算每幅图像各项指标的平均增长率,

可以得出 FAGF 方法与 WLS 方法、GF 方法、VGG 方法、RES 方法以及 PIAF 方法五种对比方法相比其信息熵、标准差、空间频率、小波特征互信息、视觉保真度以及平均梯度指标分别平均提高了 9.9%, 6.8%, 43.6%, 11.3%, 32.3%, 47.1%。

然而基于深度学习的红外与可见光图像的融合仍然还需要很多改进,例如网络的具体架构,生成训练样本的方法,融合方案的总体设计(与传统融合技术更有效结合)以及其他图像融合问题的应用(例如,医疗图像融合)等等。上述许多问题相互关联,它们相互影响。今后,将集中讨论这些问题,希望进一步提高图像融合效果。

参考文献:

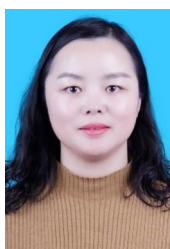
- [1] 刘先红, 陈志斌, 秦梦泽. 结合引导滤波和卷积稀疏表示的红外与可见光图像融合[J]. 光学精密工程, 2018, 26(5): 1242-1253.
LIU X H, CHEN Z B, QIN M Z. Infrared and visible image fusion using guided filter and convolutional

- sparse representation [J]. *Opt. Precision Eng.*, 2018, 26(5): 1242-1253. (in Chinese)
[2] HU W R, YANG Y H, ZHANG W S, et al. Moving object detection using tensor-based low-rank and saliently fused-sparse decomposition [J]. *IEEE Transactions on Image Processing*, 2017, 26(2): 724-737.

- [3] 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.
- [4] MA C, MIAO Z J, ZHANG X P, *et al.* A saliency prior context model for real-time object tracking[J]. *IEEE Transactions on Multimedia*, 2017, 19(11): 2415-2424.
- [5] 王昕, 吉桐伯, 刘富. 结合目标提取和压缩感知的红外与可见光图像融合[J]. *光学精密工程*, 2016, 24(7): 1743-1753.
- WANG X, JI T B, LIU F. Fusion of infrared and visible images based on target segmentation and compressed sensing [J]. *Opt. Precision Eng.*, 2016, 24(7): 1743-1753. (in Chinese)
- [6] 张蕾, 金龙旭, 韩双丽, 等. 采用非采样Contourlet变换与区域分类的红外和可见光图像融合[J]. *光学精密工程*, 2015, 23(3): 810-818.
- ZHANG L, JIN L X, HAN S L, *et al.* Fusion of infrared and visual images based on non-sampled Contourlet transform and region classification [J]. *Opt. Precision Eng.*, 2015, 23(3): 810-818. (in Chinese)
- [7] AZARANG A, MANOOCHEHRI H E, KEHTARNAVAZ N. Convolutional autoencoder-based multispectral image fusion[J]. *IEEE Access*, 2019, 7: 35673-35683.
- [8] HOU R C, ZHOU D M, NIE R C, *et al.* VIF-net: an unsupervised framework for infrared and visible image fusion[J]. *IEEE Transactions on Computational Imaging*, 2020, 6: 640-651.
- [9] 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.
- [10] 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.
- [11] AN W B, WANG H M. Infrared and visible image fusion with supervised convolutional neural network[J]. *Optik*, 2020, 219: 165120.
- [12] 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.
- [13] 郝永平, 曹昭睿, 白帆, 等. 基于兴趣区域掩码卷积神经网络的红外-可见光图像融合与目标识别算法研究[J]. *光子学报*, 2021, 50(2): 0210002.
- HAO Y P, CAO Z R, BAI F, *et al.* Research on infrared visible image fusion and target recognition algorithm based on region of interest mask convolution neural network [J]. *Acta Photonica Sinica*, 2021, 50(2): 0210002. (in Chinese)
- [14] ZHANG Q, SHEN X, XU L, *et al.* Rolling Guidance Filter [C]. *Computer Vision-ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part III* 13. Springer International Publishing, 2014: 815-830.
- [15] TOET A. Alternating guided image filtering [J]. *PeerJ Computer Science*, 2016, 2: e72.
- [16] ZHAI Y, SHAH M. Visual Attention Detection in Video Sequences Using Spatiotemporal Cues [C]. *Proceedings of the 14th ACM international conference on Multimedia. October 23 - 27, 2006, Santa Barbara, CA, USA*. New York: ACM, 2006: 815-824.
- [17] TOET A. TNO image fusion dataset [DB/OL]. (2014). https://figshare.com/articles/dataset/TNO_Image_Fusion_Dataset/1008029
- [18] MA J, ZHOU Z, WANG B, *et al.* Infrared and visible image fusion based on visual saliency map and weighted least square optimization [J]. *Infrared Physics & Technology*, 2017, 82: 8-17.
- [19] MA J Y, ZHOU Y. Infrared and visible image fusion via gradientlet filter [J]. *Computer Vision and Image Understanding*, 2020, 197/198: 103016.
- [20] 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)*. 20-24, 2018, Beijing, China. IEEE, 2018: 2705-2710.
- [21] 李恒, 张黎明, 蒋美容, 等. 一种基于ResNet152的红外与可见光图像融合算法[J]. *激光与光电子学进展*, 2020, 57(8): 081013.
- LI H, ZHANG L M, JIANG M R, *et al.* An infrared and visible image fusion algorithm based on ResNet152 [J]. *Laser & Optoelectronics Progress*, 2020, 57(8): 081013. (in Chinese)
- [22] TANG L, YUAN J, ZHANG H, *et al.* PIAFusion: a progressive infrared and visible image fusion network based on illumination aware [J]. *Information Fusion*, 2022, 83/84: 79-92.
- [23] 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.
- [24] ESKICIOGLU A M, FISHER P S. Image quality measures and their performance [J]. *IEEE Transactions on Communications*, 1995, 43(12) : 2959-2965.
- [25] HAGIGHAT M, RAZIAN M A. Fast-FMI: non-reference image fusion metric[C]. 2014 8th International Conference on Application of Information and Communication Technologies(AICT). Oct. 15-17, 2014, Astana, Kazakhstan. New York: IEEE, 2014: 1-3.
- [26] HAN Y, CAI Y, CAO Y, *et al.* A new image fusion performance metric based on visual information fidelity, *Inf. Fusion* 14 (2) (2013) 127-135.

作者简介:



杨艳春(1979—),女,新疆五家渠人,副教授,博士,硕士生导师,2002年、2007年、2014年于兰州交通大学分别获得学士、硕士和博士学位,主要从事图像融合和图像处理的研究。Email: yangyanchun102@sina.com



李永萍(1996—),女,甘肃省武威人,硕士研究生,2019年于宁德师范学院获得学士学位,主要从事图像融合和图像处理的研究。E-mail: 2647336295@qq.com