

红外与可见光图像多特征自适应融合方法

王君尧, 王志社, 武圆圆, 陈彦林, 邵文禹

(太原科技大学 应用科学学院, 山西 太原 030024)

摘要: 由于成像机理不同, 红外图像以像素分布表征典型目标, 而可见光图像以边缘和梯度描述纹理细节, 现有的融合方法不能依据源图像特征自适应变化, 造成融合结果不能同时保留红外目标特征与可见光纹理细节。为此, 本文提出红外与可见光图像多特征自适应融合方法。首先, 构建了多尺度密集连接网络, 可以有效聚合所有不同尺度不同层级的中间特征, 利于增强特征提取和特征重构能力。其次, 设计了多特征自适应损失函数, 采用 VGG-16 网络提取源图像的多尺度特征, 以像素强度和梯度为测量准则, 以特征保留度计算特征权重系数。多特征自适应损失函数监督网络训练, 可以均衡提取源图像各自的特征信息, 从而获得更优的融合效果。公开数据集的实验结果表明, 该方法在主、客观评价方面均优于其他典型方法。

关键词: 图像融合; 密集连接; 自适应损失函数; 可见光图像; 红外图像

中图分类号: TP391.4 文献标识码: A 文章编号: 1001-8891(2022)06-0571-09

Multi-Feature Adaptive Fusion Method for Infrared and Visible Images

WANG Junyao, WANG Zhishe, WU Yuanyuan, CHEN Yanlin, SHAO Wenyu

(School of Applied Science, Taiyuan University of Science and Technology, Taiyuan 030024, China)

Abstract: Owing to different imaging mechanisms, infrared images represent typical targets by pixel distribution, whereas visible images describe texture details through edges and gradients. Existing fusion methods fail to adaptively change according to the characteristics of the source images, thereby resulting in fusion results that do not retain infrared target features and visible texture details simultaneously. Therefore, a multi-feature adaptive fusion method for infrared and visible images is proposed in this study. First, a multi-scale dense connection network that can effectively reuse all intermediate features of different scales and levels, and further enhance the ability of feature extraction and reconstruction is constructed. Second, a multi-feature adaptive loss function is designed. Using the pixel intensity and gradient as measurement criteria, the multi-scale features of the source image are extracted by the VGG-16 network and the feature weight coefficients are calculated by the degree of information preservation. The multi-feature adaptive loss function can supervise network training and evenly extract the respective feature information of the source image to obtain a better fusion effect. The experimental results of public datasets demonstrate that the proposed method is superior to other typical methods in terms of subjective and objective evaluations.

Key words: image fusion, dense skip connection, adaptive loss function, visible image, infrared image

0 引言

红外传感器具备全天时全天候工作能力, 能够获取捕捉目标的热辐射信息, 但所获取的红外图像存在分辨率低、对比度差、边缘模糊等缺点。可见光传感器通过光反射成像, 能够获取场景、纹理和细

节等信息, 但容易受恶劣天气、光照等环境因素影响。红外与可见光传感器成像机理不同, 成像频段不同, 表征前景目标和背景细节的特征信息差异大, 成像条件和成像特性都具有很强的互补性。因此, 图像融合技术可以整合红外与可见光图像的互补特征信息, 从而合成一幅有利于人眼观察或其他视觉

收稿日期: 2022-02-22; 修订日期: 2022-03-14.

作者简介: 王君尧(1997-)女, 硕士研究生, 研究方向为图像融合, 深度学习。E-mail: jywang0119@163.com.

通信作者: 王志社(1982-)男, 副教授, 博士, 研究方向为智能信息处理, 机器视觉和机器学习。E-mail: wangzs@tyust.edu.cn.

基金项目: 山西省基础研究计划资助项目(No.201901D111260), 信息探测与处理山西省重点实验室开放基金(No.ISPT2020-4)。

任务的单一图像，能够广泛应用于实时监控^[1]、目标跟踪^[2]、目标检测^[3]、行人识别^[4]等领域。

传统图像融合方法通常依据目标特性或成像机制等先验知识，构建相应特征提取模型，再设计相应的融合规则，通过反变换得到最终的融合图像。常见的融合方法有多尺度变换^[5]、稀疏表示^[6]、显著性融合^[7]和混合方法^[8-9]等。传统的融合方法依赖固定的数学模型，仅能对某一类特征敏感，且无差别提取红外与可见光图像特征，特征提取能力不足。此外，需要人工设计相应的融合规则，计算复杂度高，且规则单一，缺乏对图像特征自适应的提取能力。因此，传统方法的融合图像容易导致细节丢失、图像扭曲和人造纹理等问题，无法达到理想效果。

近年来，由于卷积运算具有较强特征提取能力，基于深度学习融合方法^[10-13]利用大量图像数据，不断训练网络模型，获取最优模型参数，有效地提高了图像融合性能。文献[14]提出了 DenseFuse，在编码器中引入密集连接模块，利用中间层特征增强特征表征能力，但其网络结构简单，且不能提取图像的多尺度特征。文献[15]提出了 IFCNN，选择不同的特征融合策略以适应不同的融合要求，实现多任务图像融合。文献[16]提出了 UNFusion，构建了多尺度编码器与解码器，且采用密集连接结构，可以有效地提取和重建多尺度深度特征。尽管这些方法取得了较好的融合效果，但仍需人工设计相应的融合规则，是非端到端方法。为解决上述问题，文献[17]提出 FusionGAN，首次将生成对抗网络引入图像融合领域中，将图像融合问题视为特征对抗问题。接着，又提出了 GANMcC^[18]，进一步将图像融合问题转化为多分类限定问题。文献[19]设计了 RFN-Nest，采用两阶段训练，先训练编码器和解码器，再训练融合规则。这些虽然规避了融合规则的人为限制，但无法依据源图像特征自适应变化，造成图像特征

提取不充分，导致融合结果不能同时保留红外目标特征与可见光纹理细节。

针对上述问题，本文提出了红外与可见光图像多特征自适应融合方法。首先，构建了多尺度密集连接网络，在解码和编码器中建立密集连接，可以聚合所有不同尺度不同层的中间特征，有效地提高了网络模型的特征提取和重构能力。其次，构造了多特征自适应损失函数，依据 VGG-16 网络提取源图像各自的特征信息，自适应监督融合网络训练，从而获得更加平衡的融合结果。

1 融合方法

1.1 网络结构

多特征自适应融合方法的原理框架如图 1 所示，其中，左半部分为编码器（encoder），普通卷积层 Conv1-4 采用 3×3 卷积核，用于提取 4 个不同尺度的浅层特征。ECB 表示编码卷积模块，由卷积核为 1×1 、 3×3 的两个卷积层组成，用于提取图像深层特征。网络的右半部分为解码器（decoder），DCB 表示解码卷积模块，包含了两个 3×3 卷积层，用于重构图像特征。卷积层 Conv5 用于输出最终结果，采用 Sigmoid 作为激活函数，而其他卷积层都采用 PReLU 作为激活函数。此外，向下箭头表示下采样操作，采用步长卷积运算，而向上箭头表征上采样操作，采用最近邻插值法。可以看出，垂直方向表明了网络的尺度变化，水平方向表征了网络的深度变化，利用密集跳转连接，可以聚合所有不同尺度不同层级的中间特征，进一步提高网络的特征提取与重构能力。

在训练时，将红外与可见光图像分别输入至 VGG-16 网络的 3 个通道中，并于最大池化前获取 5 个不同尺度的特征图，最后通过计算，得到多特征自适应权重 ω_r 、 ω_{vis} ，以控制损失函数中的比例参数，

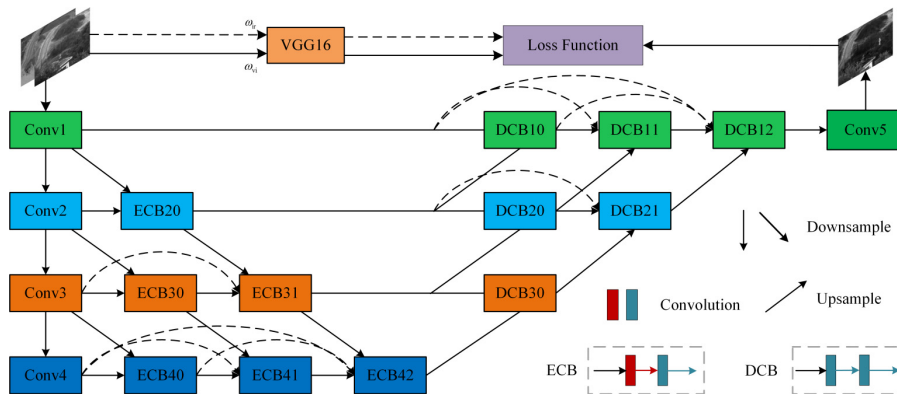


图 1 多特征自适应融合方法原理

Fig.1 The principle of multi-feature adaptive fusion method

自适应调整不同源图像与融合图像之间的特征误差, 获得信息更加均衡的结果。在测试时, 网络无需加载 VGG-16 网络与多特征自适应损失函数, 仅将红外与可见光图像进行通道合并 (concatenate), 输入至编码器, 最后, 将获取的不同尺度的特征图通过解码器, 重构得到融合结果。

1.2 多特征自适应损失函数

红外与可见光图像融合应同时保留红外图像与可见光图像的特征信息, 即红外目标与可见光纹理细节。若某一源图像包含丰富的信息, 那么其特征保留的比例应该增加, 使融合结果与该源图像具有较高的相似性。而现有的融合方法等同看待源图像的固有特征, 无法针对源图像的固有特征的差异性, 自适应调节融合图像与源图像之间的特征保留比例, 造成信息损失。为此, 本文方法综合考虑不同类型源图像的特征差异, 设计了统一的信息度量。通过计算源图像与融合结果之间的相似关系, 自适应调整损失函数的参数, 以监督网络训练过程, 有效解决了此类问题。本文结合红外与可见光源图像的各自的固有特征, 以像素结构与像素强度为度量准则, 构造了多特征自适应损失函数。依据 VGG-16 网络提取源图像各自的特征信息, 自适应监督融合网络训练, 均衡地提取可见光图像的纹理细节与红外图像的目标亮度, 从而获得更加平衡的融合结果。

在训练时, 采用 VGG-16 网络提取源图像 I 的多尺度特征 $\phi_j^k(I)$, 其中 j 表示尺度, $j=1, 2, \dots, 5$, k 表示通道索引, 取值区间为 $[1, C_j]$ 。通过计算特征图的梯度特征来获取源图像的结构信息保留度 δ_s , 权重越高代表结构信息保留程度越高, 如公式(1)所示:

$$\delta_s = \frac{1}{5} \sum_{j=1}^5 \frac{1}{H_j W_j C_j} \sum_{k=1}^{C_j} \|\nabla \phi_j^k(I)\|_F^2 \quad (1)$$

式中: ∇ 代表梯度算子; $\|\cdot\|_F$ 代表 Frobenius 范数; H_j 、 W_j 为 j 尺度特征图的高度、宽度。

计算结构信息保留度可以评估边缘和细节等高频信息, 但容易引起图像像素强度以及目标特征信息损失, 因此需要获取源图像的像素强度信息保留度 δ_l , 以均衡测量源图像与融合图像之间的特征相似性, 如公式(2)所示:

$$\delta_l = \frac{1}{5} \sum_{j=1}^5 \frac{1}{H_j W_j C_j} \sum_{k=1}^{C_j} \|\phi_j^k(I)\|_F^2 \quad (2)$$

综合上述的结构特征保留度和像素强度保留度, 总的特征信息保留度如公式(3)所示:

$$\delta = \delta_l + \alpha \cdot \delta_s \quad (3)$$

式中: α 为超参数, 用于权衡 δ_s 与 δ_l 的比例关系。

每组源图像 I_{ir} 、 I_{vis} 可由信息保留度 δ_{ir} 、 δ_{vis} 计算得到 ω_{ir} 、 ω_{vis} , 如公式(4)所示:

$$[\omega_{ir}, \omega_{vis}] = \text{softmax} \left(\left[\frac{\delta_{ir}}{c}, \frac{\delta_{vis}}{c} \right] \right) \quad (4)$$

式中: 超参数 c 用于缩放数值, 以获得更好的权重分配。softmax 函数将多特征自适应权重 ω_{ir} 、 ω_{vis} 映射到区间 $[0,1]$, 且 $\omega_{ir} + \omega_{vis} = 1$ 。

将多特征自适应权重应用于损失函数, 以监督网络训练。多特征自适应损失函数从结构相似性和强度分布两个方面实现相似约束。结构相似性指数测度 (SSIM) 通过在亮度、对比度和结构 3 个方面衡量两幅图像相似度, 以该指标来测量融合结果与源图像之间的信息损失, 如公式(5)所示:

$$L_{MS_SSIM} = \omega_{ir} L_{MS_SSIM}(I_f, I_{ir}) + \omega_{vis} L_{MS_SSIM}(I_f, I_{vis}) \quad (5)$$

由于结构相似性指数测度容易引起亮度变化和颜色偏差。因此, 本文采用 L1 损失函数以保留图像目标亮度信息, 弥补 SSIM 损失函数的不足, 如公式(6)所示:

$$L_{L1} = \omega_{ir} L_{L1}(I_f, I_{ir}) + \omega_{vis} L_{L1}(I_f, I_{vis}) \quad (6)$$

总损失函数如公式(7)所示, 其中 β 为超参数, 以均衡两类损失函数差异。

$$L_{total} = \beta \cdot L_{MS_SSIM} + L_{L1} \quad (7)$$

2 实验

2.1 实验参数设定

在训练阶段, 选取 TNO 数据集中的红外与可见光图像为训练集。为了扩大训练数据集, 以滑动步长为 12, 将其裁剪为大小为 256×256 的红外与可见光图像对, 共 10563 对, 灰度值取值范围为 $[0,255]$ 。此外, 设置迭代数 (epoch) 为 1, 优化器为 Adam, 学习率 (learning rate) 为 0.001, 每批训练量 (batch size) 为 8。超参数设置为 $\alpha=1$, $c=2$, $\beta=0.87$ 。测试及训练配置为 Intel i9-9900、GeForce RTX 2080Ti 和 32 GB 内存。选定的编译环境为 pytorch。

在测试阶段, 选取 TNO^[20]、Roadscene^[21] 与 OTCBVS^[22] 为验证数据集, 以验证方法的有效性。将本文方法与 MDLatLRR (multi-level decomposition of LatLRR) ^[6]、DenseFuse^[14]、IFCNN^[15]、UNFusion^[16]、GANMcC^[18]、U2Fusion^[23]、RFN-Nest^[19] 这 7 种典型方法进行比较。在定性比较中, 为了便于观察, 将融合图像中的红外目标与可见光细节分别标注, 并将局部细节放大。定量评价指标选择信息熵 (entropy, EN) ^[24]、非线性相关信息熵 (nonlinear correlation

information entropy, NCIE) [25]、标准差 (standard deviation, SD) [26]、相关差异和 (the sum of the correlation differences, SCD) [27]、视觉信息融合保真度 (the visual information fidelity for fusion, VIFF) [28]、多尺度结构相似性 (multi-scale structural similarity index measure, MS_SSIM) [29], 可从图像特征信息、结构相似性、视觉感知等方面评估融合性能。

2.2 消融实验

2.2.1 密集连接有效性验证

为验证密集连接的有效性, 采用3个对比验证模型, 分别是编码器和解码器都无密集连接, 记作 No_Dense, 仅编码器有密集连接, 记作 Only_En, 仅解码器有密集连接, 记作 Only_De。实验选择 TNO 数据集中的 Kaptein_1654、House、Helicopter 三组典型的红外与可见光图像进行定性比较, 如图2所示。可以看出, No_Dense、Only_En 模型融合结果丢失了可见光图像的纹理细节, 降低了目标与背景的对比度, 且图像的背景出现噪点, 部分边缘细节处产生伪光晕; Only_De 模型结果在一定程度上提高了图像对比度, 但仍缺乏纹理结构信息。例如, 在 Helicopter 结果中, 天空云彩的结构与灰度分布有所缺失。相比之下, 本文方法在提高目标亮度的同时突出了图像细节与对比度, 图像较为清晰, 视觉效果优于对比模型。

此外, 为了避免视觉感知偏差, 本文采用 TNO 数据集的25组图像以及6个指标, 在客观方面量化比较模型有效性, 表1给出了密集连接的定量比较结果。观察表中数据, 本文模型获得的融合结果在 EN、SCD、VIFF 与 MS_SSIM 指标上获得最优值, 且 NCIE, SD 指标仅次于 Only_De。从结果可以看出, 密集连接可以实现特征复用, 有效增强图像特

征表示能力。且在解码和编码器中建立密集连接在主客观评价上均优于其他模型, 即 No_Dense、Only_En 以及 Only_De, 表明本文模型可以有效聚合所有不同尺度不同层的中间特征, 提高网络模型的特征提取和重构能力。

表1 密集连接的定量比较结果

Table 1 Quantitative comparison results of dense skip connection

Metrics	No_dense	Only_En	Only_De	Ours
EN	6.90676	<u>7.00322</u>	6.98832	7.00507
NCIE	0.80460	0.80443	0.80519	<u>0.80464</u>
SD	35.68939	38.39500	43.99837	<u>41.44601</u>
SCD	1.79833	<u>1.84037</u>	1.78040	1.84590
VIFF	0.56981	<u>0.60081</u>	0.58553	0.62035
MS_SSIM	<u>0.92346</u>	0.92305	0.91745	0.93027

注: 将最优值以黑体加粗标注, 次优值以下划线标注。客观评价指标越高, 表明融合性能越好。

Note: The best average value is marked in bold, and the second value is marked with an underline. The objective evaluation index is positively correlated with the fusion performance.

2.2.2 多特征自适应损失函数有效性验证

为验证多特征自适应损失函数的有效性, 采用3个对比验证模型, 分别是仅保留像素亮度测量准则的融合模型, 记作 F_{inten} , 仅保留梯度测量准则的融合模型, 记作 F_{grad} , 直接采用固定权重系数, 即 ω_{ir} 、 ω_{vis} 取值为 0.5, 记作 F_{mean} 。图3给出了多特征自适应损失函数的定性比较结果, 由图观察得知, F_{inten} 融合结果边缘细节不清晰, 出现伪光晕, 且图像对比度不高。 F_{grad} 、 F_{mean} 突出了可见光细节, 但目标

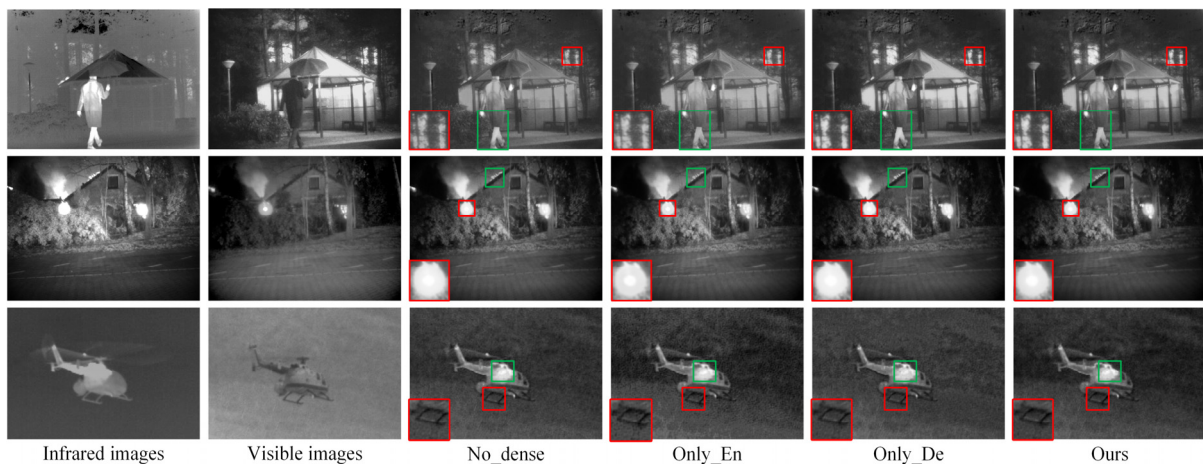


图2 密集连接的定性比较结果

Fig.2 Qualitative comparison results of dense skip connections

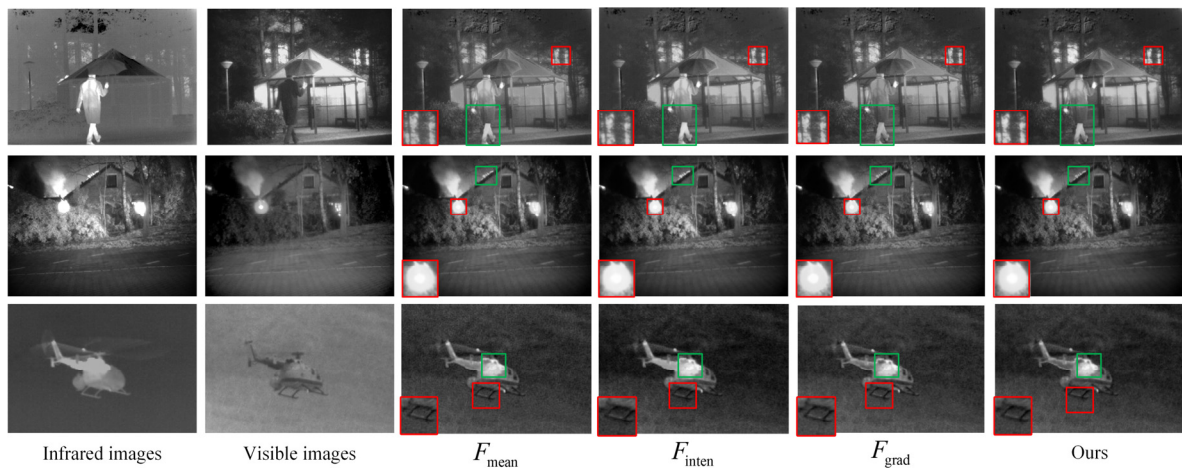


图3 多特征自适应模块的定性比较结果

Fig.3 Qualitative comparison results of multi-feature adaptive modules

亮度与对比度依旧不高。相比于对比验证模型，本文设计的多特征自适应模块既保留了红外图像的目标亮度，又突出了可见光细节，取得更好的主观融合效果。通过定性比较，证明了多特征自适应损失函数的优越性。

此外，表2给出了多特征自适应损失函数的定量比较结果，进一步验证本文模型设计的合理性与有效性。观察表中数据，本文方法在所有指标上均获得最大值，与主观评价结果一致，表明多特征自适应损失函数可以有效获取源图像纹理结构以及像素强度的信息保留度，实现源图像与融合图像的特征相似性约束，包括边缘细节、亮度变化以及灰度分布等，并平衡了两类图像的差异信息，获得良好的融合结果。

表2 多特征自适应损失函数的定量比较结果

Table 2 Quantitative comparison results of multi-feature adaptive loss function

Metrics	F_{mean}	F_{inten}	F_{grad}	Ours
EN	6.82654	<u>6.95075</u>	6.92680	7.00507
NCIE	0.80432	<u>0.80460</u>	0.80455	0.80464
SD	33.71341	<u>38.96415</u>	36.28725	41.44601
SCD	1.77467	<u>1.83212</u>	1.82509	1.84590
VIFF	0.55015	0.59584	<u>0.59834</u>	0.62035
MS_SSIM	0.91588	0.92522	<u>0.92957</u>	0.93027

2.3 TNO 数据集

为了验证本文方法的优越性，从TNO数据集中选取25组图像进行实验验证，其中选取7组典型的红外与可见光图像进行主观评价，包括Kaptein_1654、Sandpath、Nato_camp、Movie_18、Helicopter、Movie_01和Marne_04。TNO数据集的定性比较结果如图4所示。观察结果得知，MDLatLRR、DenseFuse以及IFCNN融合结果保留了可见光图像纹理细节，但是目标亮度严重缺乏，且图像背景较暗，导致其对比度偏低，部分细节不明显。GANMcC获得的融合结果倾向于保留红外典型目标，而损失了大量的可见光结构信息，例如Sandpath结果中，路边围栏模糊不清。U2Fusion和RFN-Nest获得较为良好的结果。图像清晰，同时保留了丰富的细节，提高了目标与背景的对比度，但是其结果中典型目标亮度依旧不高。相比之下，UNFusion和本文方法可以同时具有两类图像的差异信息，在不损失目标亮度的同时，突出细节，保留丰富的场景信息。主观评价结果表明，本文方法能够获得较好的融合效果。

表3给出了TNO数据集的定量比较结果，从客观角度验证融合方法的有效性，由表可知，本文方法在指标EN、SD、SCD以及MS_SSIM上获得最优值，并在指标NCIE和VIFF上仅次于UNFusion以及U2Fusion。表明本文方法融合结果良好，具有较高的对比度以及图像质量。

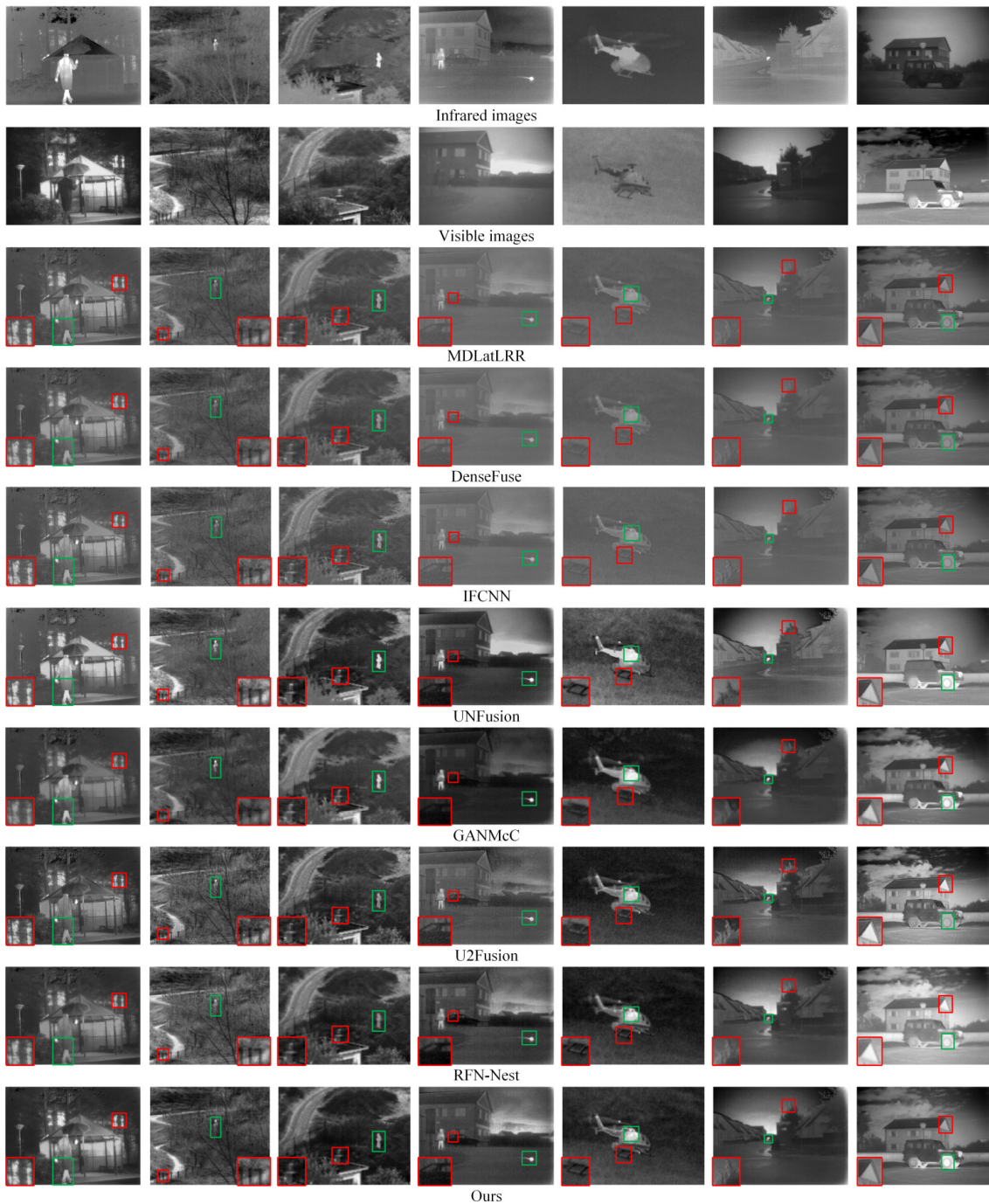


图4 TNO数据集的定性比较结果

Fig.4 Qualitative comparison results of TNO datasets

表3 TNO数据集的定量比较结果

Table 3 Quantitative comparison results of TNO datasets

Metrics	MDLatLRR	DenseFuse	IFCNN	UNFusion	GANMcC	U2Fusion	RFN-Nest	Ours
EN	6.29523	6.25275	6.33795	<u>6.98828</u>	6.57763	6.84306	6.89803	7.00507
NCIE	0.80435	0.80451	0.80404	0.80912	0.80452	0.80392	0.80428	<u>0.80464</u>
SD	23.70282	22.85769	24.06712	<u>40.93903</u>	29.92973	33.59608	34.85373	41.44601
SCD	1.59002	1.57018	1.59052	1.70351	1.67191	1.76712	<u>1.78875</u>	1.84590
VIFF	0.30902	0.28295	0.34396	0.45033	0.40468	0.62469	0.53121	<u>0.62035</u>
MS_SSIM	0.90133	0.87696	0.90474	0.87404	0.85915	0.87283	<u>0.91217</u>	0.93027

2.4 Roadscene 数据集

为进一步验证该融合方法的有效性,选取 Roadscene 数据集中的 221 组图像进行实验验证,图 5 为 Roadscene 数据集的定性比较结果,观察可知,MDLatLRR、DenseFuse、IFCNN 的融合图像倾向于可见光图像,提取红外图像的亮度信息不足,且图像对比度低。GANMcC、RFN-Nest 结果倾向于红外图像,保留了目标亮度,但损失大量纹理细节,且目标边缘模糊,RFN-Nest 存在较为严重的伪影。UNFusion 虽可均衡获得红外目标与可见光细节,但丢失了天空云彩的边缘纹理。U2Fusion 及本文方法获得优于其他方法的融合结果,根据定性分析,两者结果差异不明显,故从定量分析进一步比较融合性能。

Roadscene 数据集的定量比较如表 4 所示,本文方法取得了 EN、SD、SCD 与 VIFF 指标的最大值。EN、SD 值最高,证明多特征自适应融合方法整合不同尺度不同层级的中间层特征后,提取到更多的特

征信息。SCD、VIFF 取得最优值,则说明在图像质量与视觉效果比较中,本文方法优于其他典型方法。虽然 NCIE 指标略次于 UNFusion 和 U2Fusion,MS_SSIM 指标仅次于 U2Fusion,但综合分析,多特征自适应融合方法的融合效果综合性能更优。

2.5 OTCBVS 数据集

为进一步验证该融合方法的鲁棒性,选取了 OTCBVS 基准数据集的 40 组红外与可见光图像对进行实验。图 6 给出了 OTCBVS 数据集的定性比较结果,其中,MDLatLRR、DenseFuse、IFCNN 结果保留了可见光纹理信息,但对比度低,导致部分细节不清晰。GANMcC 目标亮度较高,但缺乏大量结构信息,例如,图像中几乎没有地缝,且目标边缘模糊。RFN-Nest 目标亮度不高,可见光边缘细节模糊,且存在伪影。在定性比较中,U2Fusion、UNFusion 以及本文方法的融合结果优于其他方法。而从主观上来看,三者结果差异不明显,因此从客观上比较他们的融合性能。

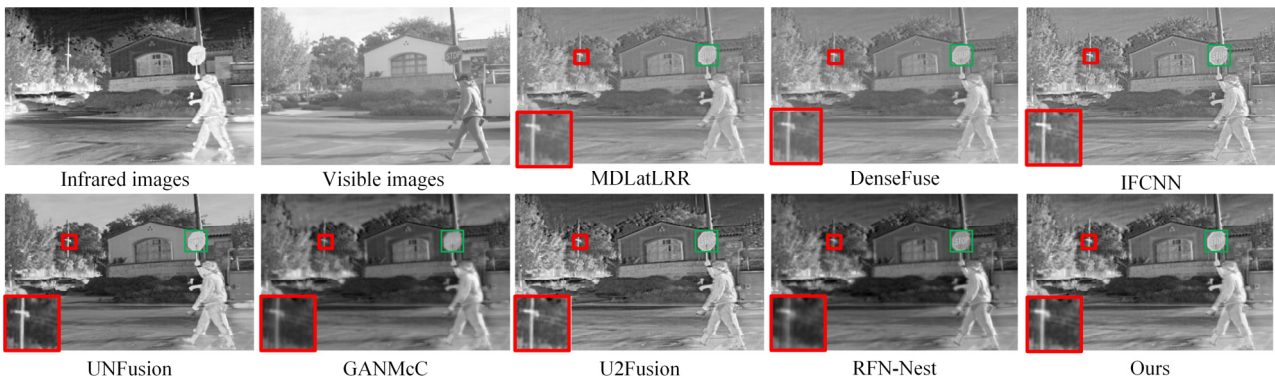


图 5 Roadscene 数据集的定性比较结果

Fig.5 Qualitative comparison results of Roadscene datasets

表 4 Roadscene 数据集的定量比较结果

Table 4 Quantitative comparison results of Roadscene datasets

Metrics	MDLatLRR	DenseFuse	IFCNN	UNFusion	GANMcC	U2Fusion	RFN-Nest	Ours
EN	6.86973	6.82473	6.91635	7.37792	7.23683	<u>7.41910</u>	7.33677	7.47617
NCIE	0.80701	0.80718	0.80629	0.80998	0.80670	<u>0.80771</u>	0.80649	0.80658
SD	33.05830	32.22811	33.64622	<u>50.26212</u>	43.81169	49.15315	46.03340	51.43262
SCD	1.38030	1.36329	1.38949	1.63715	1.60983	<u>1.76111</u>	1.72851	1.84427
VIFF	0.39051	0.35676	0.41529	0.48462	0.45465	<u>0.58028</u>	0.47894	0.61286
MS_SSIM	0.88881	0.85857	0.89583	0.85420	0.85744	0.93322	0.87277	<u>0.92484</u>

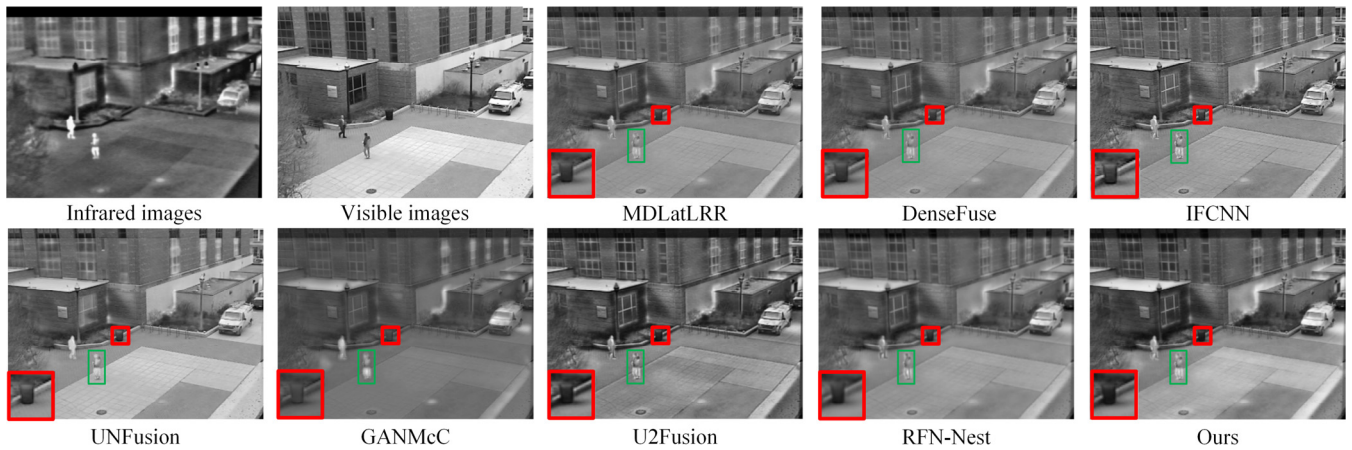


图6 OTCBVS数据集的定性比较结果

Fig.6 Qualitative comparison results of OTCBVS datasets

表5为OTCBVS数据集的定量比较结果,对比可得,本文方法在EN、SD、MS_SSIM、VIFF、SCD指标均优于其他方法,仅NCIE指标略次于UNFusion。进一步验证了本文的融合性能。

本文采用TNO、Roadscene与OTCBVS三个公开数据集进行实验验证,综合视觉感知以及多个指

标进行评价。相比于其他典型方法,主观上更均衡地保留了两类图像的典型特征,包括红外图像目标特征以及可见光图像的纹理特征,客观上随着数据集的变化,评价指标趋势基本一致,表明本文方法具有较强的鲁棒性与优越性。

表5 OTCBVS数据集的定量比较结果

Table 5 Quantitative comparison results of OTCBVS datasets

Metrics	MDLatLRR	DenseFuse	IFCNN	UNFusion	GANMcC	U2Fusion	RFN-Nest	Ours
EN	7.12984	7.07264	7.23707	7.28585	6.60437	<u>7.35554</u>	7.19973	7.57158
NCIE	0.80557	0.80569	0.80505	0.81011	0.80560	0.80501	0.80528	<u>0.80578</u>
SD	36.43416	35.31958	38.72566	<u>44.91661</u>	27.50886	43.12793	39.04909	50.45001
SCD	1.44604	1.41625	1.45256	1.42248	1.09337	<u>1.62165</u>	1.48793	1.75565
VIFF	0.27563	0.27216	0.29275	0.22372	0.16750	<u>0.32515</u>	0.26596	0.35130
MS_SSIM	0.83939	0.81697	0.86125	0.79352	0.67759	<u>0.87722</u>	0.80530	0.88162

3 结论

本文提出了多特征自适应红外与可见光图像融合方法。首先,采用了多尺度密集连接的网络结构,充分利用不同尺度的浅层及深层特征;其次,设计多特征自适应损失函数,引入VGG-16网络提取特征并用于计算特征的信息保留度,进而约束损失函数。特征自适应损失函数针对性地提取了红外图像的目标信息和可见光图像的纹理细节信息,监督网络模型获得更平衡的融合结果。消融实验验证了密集连接与多特征自适应损失函数的有效性。此外,在TNO、Roadscene和OTCBVS三种不同的数据集上,与7种典型方法进行了主客观实验,结果表明本文方法的优越性和鲁棒性。

参考文献:

- [1] 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.
- [2] ZHANG Xingchen, YE Ping, QIAO Dan, et al. Object fusion tracking based on visible and infrared images: a comprehensive review[J]. *Information Fusion*, 2020, 63: 166-187.
- [3] TU Zhengzheng, LI Zhun, LI Chenglong, et al. Multi-interactive dual-decoder for RGB-thermal salient object detection[J]. *IEEE Transactions on Image Processing*, 2021, 30: 5678-5691.
- [4] FENG Zhanxiang, LAI Jianhuang, XIE Xiaohua. Learning modality-specific representations for visible-infrared person re-identification[J]. *IEEE Transactions on Image Processing*, 2020, 29: 579-590.
- [5] MO Yang, KANG Xudong, DUAN Puhong, et al. Attribute filter based

- infrared and visible image fusion[J]. *Information Fusion*, 2021, **75**: 41-54.
- [6] LI Hui, WU Xiaojun, Kittle J. MDLatLRR: a novel decomposition method for infrared and visible image fusion[J]. *IEEE Transactions on Image Processing*, 2020, **29**: 4733-4746.
- [7] 李辰阳, 丁坤, 翁帅, 等. 基于改进谱残差显著性图的红外与可见光图像融合[J]. *红外技术*, 2020, **42**(11): 1042-1047.
LI Chenyang, DING Kun, WENG Shuai, et al. Image fusion of infrared and visible images based on residual significance[J]. *Infrared Technology*, 2020, **42**(11): 1042-1047.
- [8] WANG Zhishe, YANG Fengbao, PENG Zhihao, et al. Multi-sensor image enhanced fusion algorithm based on NSST and top-hat transformation[J]. *Optik-International Journal for Light and Electron Optics*, 2015, **126**(23): 4184-4190.
- [9] LIU Yu, CHEN Xun, PENG Hu, et al. Multi-focus image fusion with a deep convolutional neural network[J]. *Information Fusion*, 2017, **36**: 191-207.
- [10] WANG Zhishe; WU Yuanyuan; WANG Junyao, et al. Res2Fusion: infrared and visible image fusion based on dense Res2net and double non-local attention models[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, **71**: 1-12.
- [11] MA Jiayi, MA Yong, LI Chang. Infrared and visible image fusion methods and applications: a survey[J]. *Information Fusion*, 2019, **45**: 153-178.
- [12] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation[C]//*Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015: 234-241.
- [13] Toet A. Computational versus psychophysical bottom-up image saliency: a comparative evaluation study[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2011, **33**(11): 2131-2146.
- [14] LI Hui, WU Xiaojun. DenseFuse: a fusion approach to infrared and visible images[J]. *IEEE Transactions on Image Processing*, 2019, **28**(5): 2614-2623.
- [15] ZHANG Yu, LIU Yu, SUN Peng, et al. IFCNN: a general image fusion framework based on convolutional neural network[J]. *Information Fusion*, 2020, **54**: 99-118.
- [16] WANG Zhishe, WANG Junyao, WU Yuanyuan, et al. UNFusion: a unified multi-scale densely connected network for infrared and visible image fusion[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022, **32**(6): 3360-3374.
- [17] MA Jiayi, YU Wei, LIANG Pengwei, et al. FusionGAN: a generative adversarial network for infrared and visible image fusion[J]. *Information Fusion*, 2019, **48**: 11-26.
- [18] MA Jiayi, ZHANG Hao, SHAO Zhenfeng, et al. GANMcC: a generative adversarial network with multiclassification constraints for infrared and visible image fusion[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, **70**: 1-14.
- [19] LI Hui, WU Xiaojun, Josef Kittler. RFN-Nest: an end-to-end residual fusion network for infrared and visible images[J]. *Information Fusion*, 2021, **73**: 72-86.
- [20] TOET A. TNO Image Fusion Dataset[DB/OL]. [2014-04-26]. [https://figshare.com/articles/TN Image Fusion Dataset/1008029](https://figshare.com/articles/TN_Image_Fusion_Dataset/1008029).
- [21] XU Han. Roadscene Database[DB/OL]. [2020-08-07]. <https://github.com/hanna-xu/RoadScene>.
- [22] Ariffin S. OTCBVS Database[DB/OL]. [2007-06]. <http://vcip-okstate.org/pbvs/bench/>.
- [23] XU Han, MA Jiayi, JIANG Junjun, et al. U2Fusion: a unified unsupervised image fusion network[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, **44**(1): 502-518.
- [24] Aslantas V, Bendes E. Assessment of image fusion procedures using entropy, image quality, and multispectral classification[J]. *Journal of Applied Remote Sensing*, 2008(2): 1-28.
- [25] LIU Zheng, Blasch E, XUE Zhiyun, et al. Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: A comparative study[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **34**: 94-109.
- [26] RAO Yunjiang. In-fibre bragg grating sensors[J]. *Measurement Science and Technology*, 1997(8): 355-375.
- [27] Aslantas V, Bendes E. A new image quality metric for image fusion: The sum of the correlations of differences[J]. *AEU-Int. J. Electron. C.*, 2015, **69**: 1890-1896.
- [28] HAN Yu, CAI Yunze, CAO Yin, et al. A new image fusion performance metric based on visual information fidelity[J]. *Information Fusion*, 2013(14): 127-135.
- [29] MA Kede, ZENG Kai, WANG Zhou. Perceptual quality assessment for multi-exposure image fusion[J]. *IEEE Trans. Image Process*, 2015, **24**: 3345-3356.