

基于 ResNet50 与卷积稀疏表达的红外与可见光图像融合算法

邵大光¹ 邵现振² 刘鹏² 赵闯² 陶青川^{1*}

¹(四川大学电子信息学院 四川 成都 610000)

²(中国石油化工股份有限公司胜利油田分公司河口采油厂 山东 东营 257000)

摘要 提出一种基于 ResNet50 神经网络与卷积稀疏表达的红外与可见光图像融合算法。通过低通滤波将红外与可见光图像分解成基础层和细节层;运用卷积稀疏表达对基础层进行处理得到新的基础层,使用 ResNet50 神经网络对细节层进行特征提取,对得到的特征图进行 L1 正则化和最大选择策略得到最大权重层,经过权重分配得到新的细节层;对新的基础层和细节层进行重建,得到融合图像。该算法针对基础层和细节层提出了新的融合策略,并且能较好地保留细节信息和结构信息。实验结果表明,该算法在主观和客观指标证明上优于对比算法。

关键词 图像融合 ResNet50 卷积稀疏表达 红外图像 可见光图像

中图分类号 TP391.4

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2024.05.030

THE INFRARED AND VISIBLE IMAGE FUSION ALGORITHM BASED ON RESNET50 AND CONVOLUTION SPARSE REPRESENTATION

Shao Daguang¹ Shao Xianzhen² Liu Peng² Zhao Chuang² Tao Qingchuan^{1*}

¹(The College of Electronics Information and Engineering, Sichuan University, Chengdu 610000, Sichuan, China)

²(Hekou Oil Production Plant of Shengli Oilfield Branch of Sinopec, Dongying 257000, Shandong, China)

Abstract This paper proposes an infrared and visible image fusion algorithm based on ResNet50 neural network and convolution sparse representation. The infrared and visible images were decomposed into the base layers and the detail layers through low-pass filtering. The convolution sparse representation was utilized to process the base layer to obtain the fused base layer, and the ResNet50 neural network was applied to extract the features of the detail layer. For the feature map, we performed L1-regularization and choose-max strategy to obtain the maximum weight layer. And the fused detail layer could be generated through weight distribution. We generated the new base layer and detail layer to obtain the fused image. We proposed new fusion strategies for the base layer and the detail layer, and the fused image retained more detail information and structural information. The experimental results show that our method is superior to the comparison algorithm in subjective and objective metrics.

Keywords Image fusion ResNet50 Convolution sparse representation Infrared image Visible image

0 引言

图像融合是一种图像处理技术,可以将不同信道针对同一目标所得到的图像数据中的互补信息、有利信息等信息综合生成更高质量的图像。融合后的高质

量图像综合了不同信道的有利信息,提高了其他图像处理技术如:目标检测、目标跟踪的精度和可靠性^[1]。红外图像是使用红外热成像设备捕获目标和场景的热红外辐射而来,图像中各像素点的亮度值取决于目标和场景的热辐射能量强弱,在弱光照场景下也可以很好地显示目标和场景的结构特征,但热成像设备普遍

分辨率低,无法详细呈现场景细节信息。可见光图像是根据物体对光线反射特征的成像,成像清晰度和光照强度直接相关,在光照充分的情况下,图像分辨率较高并包含丰富的细节信息,但是在弱光场景下分辨率会明显下降,目标与场景几乎无法区分。由此,红外与可见光图像融合技术可应用在各行各业,例如:监控、军事监测、目标检测等领域。在实际的场景中,红外图像和可见光图像有各自的优缺点,同时两种图像包含重要的互补信息,因此将红外与可见光图像进行融合,可以综合利用二者的有利信息,从而提高在复杂环境下的图像质量。

目前,主流的红外与可见光图像融合的方法可以根据其原理分为如下几大类^[2]:基于多尺度分解的方法、基于稀疏表达的方法、基于神经网络的方法、复合方法等。基于多尺度的方法又可分为:金字塔变换^[3-4]、小波变换^[5-6]、曲线波变换(CVT)^[7-8]、双树复小波变换(DTCWT)^[9-11]、非下采样轮廓变换(NSCT)^[12-13];基于稀疏表达的方法可分为:稀疏表达(SR)^[14-15],还有在此基础上进行改变的自适应稀疏表达(ASR)^[16]、卷积稀疏表达(CSR)^[17-19];基于神经网络的方法可分为:脉冲耦合神经网络和卷积神经网络^[20-21];除此以外,VGG、ResNet 等网络模型也可用于图像融合的处理,例如李辉提出的 VGGML^[22]、ResNet-ZCA^[23]方法。另外还有一些复合融合方法及其他方法^[24-26]。

考虑到现有图像融合方法的优势和不足。例如多尺度方法中,曲波变换 CVT 可以很好地保存细节和边缘信息,但它的采样原理导致了最终图片容易失真和频谱混叠;基于稀疏表达的方法可以减少块伪影,但是纹理细节的保存效果欠佳;而基于神经网络的方法有很强的适应性,也就是高鲁棒性、高容错能力,但是训练过程复杂,处理时间较长。每个方法有各自的特点,我们在进行图像融合的过程中需要考虑到这些情况。考虑到上述方法存在的问题,提出了一种基于卷积稀疏表达和 ResNet50 的红外与可见光图像融合方法。本文所选用的数据集多为弱光的场景,红外图像可以较好地呈现热源特征,但对背景信息不敏感,而可见光呈现的内容比较丰富,为了结合红外和可见光图像各自的特点,首先讲红外和可见光图片分解为基础层(Base Layer)和细节层(Detail Layer),然后通过卷积稀疏表达(CSR)提取基础层特征,通过 ResNet50 提取细节层的特征,融合重建后的图片既有丰富的结构信息,又能突出目标特征,可以对弱光场景下的目标检测、跟踪等图像处理提供方法提供帮助。

1 相关工作

1.1 CSR

卷积稀疏表达就是一种增强的稀疏表达,稀疏表达技术可以很好地作用于图像块,但是它不能优化整幅图片,同时稀疏表达在图像融合的细节保留能力是有限的。因此,提出了卷积稀疏表达技术。卷积稀疏表达技术将原始图像建模为卷积图和字典的卷积和,卷积图由式(1)可得。

$$\operatorname{argmin}_{\{X_m\}} \frac{1}{2} \left\| \sum_m d_m * X_m - S \right\|_2^2 + \lambda \sum_m \|X_m\|_1 \quad (1)$$

式中: m 代表第 m 个图像块样本; $*$ 代表卷积操作; S 代表输入图片; d_m 代表一组字典滤波器; X_m 代表卷积图。字典滤波器从字典学习算法中学习得到。CSR 可以解决 SR 存在的问题,并且由于单值的结果而优化了整个图像。同时,卷积形式的原理是基于整个图像的,这样 CSR 可以在图像融合过程中保留更多的纹理细节。在本文中,CSR 对基础层进行处理并融合,很好地保留了基础层的结构特征和基础信息。

1.2 ResNet50

深度学习网络的对图像特征提取的能力与网络深度密切相关,然而随着网络层数的增加会出现梯度消失导致网络性能下降。ResNet 解决了这个梯度下降的问题,通过带有段链接的残差块,使网络有了无限加深的可能,并且相对于其他同深度的网络模型,训练速度更快。残差块的结构如图 1 所示。

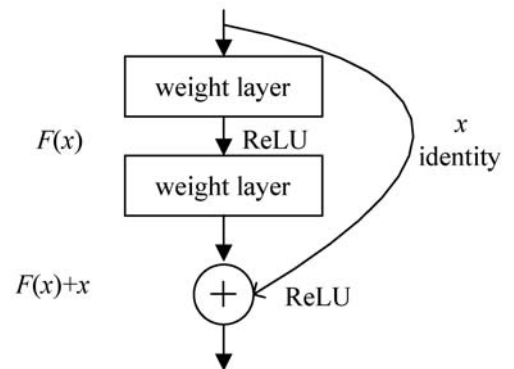


图1 残差块结构

其中 X 表示残差块的输入, $F(x)$ 表示 x 经过权重后的映射, ReLU 为修正线性单元, 残差块的输入为 $F(x) + x$ 。该网络结构更有效地避免了梯度消失。

本文的细节层采用了基于 ResNet50 的融合策略,将细节信息更好地保留了下来,其网络结构见表 1。

表 1 ResNet50 网络结构

序号	layer name	output size	ResNet50
1-3	conv1	112 × 112	7 × 7, 64, stride = 2
4-36	conv2_x	56 × 56	3 × 3, max pool stride = 2
			$\begin{bmatrix} 1 \times 1 & 64 \\ 3 \times 3 & 64 \\ 1 \times 1 & 256 \end{bmatrix} \times 3$
37-78	conv3_x	28 × 28	$\begin{bmatrix} 1 \times 1 & 128 \\ 3 \times 3 & 128 \\ 1 \times 1 & 512 \end{bmatrix} \times 4$
79-140	conv4_x	14 × 14	$\begin{bmatrix} 1 \times 1 & 256 \\ 3 \times 3 & 256 \\ 1 \times 1 & 1024 \end{bmatrix} \times 6$
141-172	conv5_x	7 × 7	$\begin{bmatrix} 1 \times 1 & 512 \\ 3 \times 3 & 512 \\ 1 \times 1 & 2048 \end{bmatrix} \times 3$
173-174		1 × 1	average pool, fc-1000, Softmax
175	FLOPs		3.8 × 109

2 图像融合方法

本文将 I_1 记为可见光源图片,将 I_2 记为红外源图片, F_b 记为基础层, F_d 记为细节层。融合方法总体架构如图 2 所示。

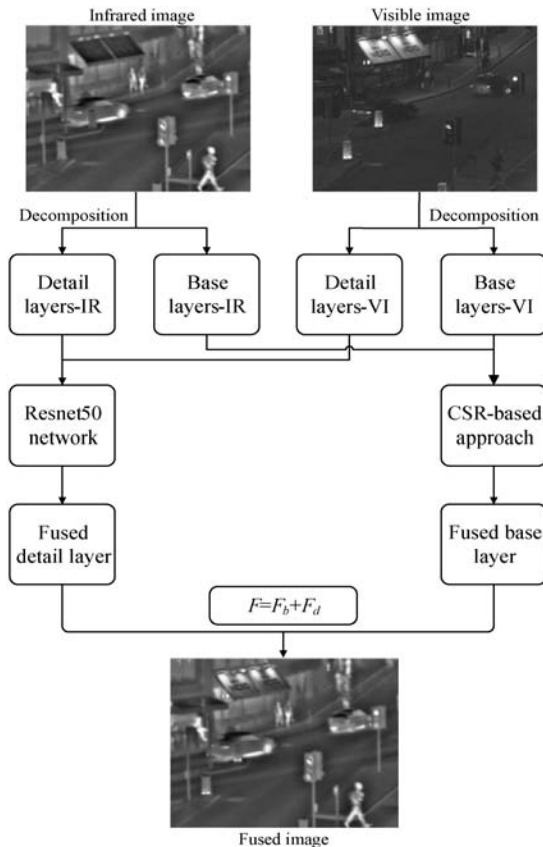


图 2 算法总架构

低通滤波是频域滤波的一种,可以去掉图片的高频信息,即细节信息。通过低通滤波对 $I_k, k \in \{1, 2\}$ 进行分解,逆运算得到图片的基础层 I_k^b 和细节层 I_k^d , 表示为:

$$I_k^b = \text{lowpass}(I_k, f_{\text{pass}}) \quad (2)$$

$$I_k^d = I_k - I_k^b \quad (3)$$

式中: f_{pass} 为低通截止频率。通过 CSR 提取基础层特征得到 F_b , 通过 ResNet50 提取细节层特征得到 F_d 。最后将基础层部分 F_b 、细节层部分 F_d 重建得到融合图片。

2.1 基础层处理

源图像的基础层段 I_k^b 采用 CSR 进行表达,可以通过计算下面的 CSR 模型得到。

$$\underset{\{C_{k,m}\}}{\text{argmin}} \frac{1}{2} \left\| \sum_{m=1}^M d_m \times C_{k,m} - I_k^b \right\|_2^2 + \lambda \sum_{m=1}^M \|C_{k,m}\|_1 \quad (4)$$

式中:稀疏系数映射 $C_{k,m}$ 是 M 维向量; M 是图像像素的数目; I_k^b 是源图像的低频频带。根据 SR 方法, $C_{k,m}$ 的 L1 范数被认为是源图像的活动层级测度,因此,活动层级图如式(5)所示。

$$A_k(x, y) = \|C_{k,1:M}(x, y)\|_1 \quad (5)$$

另外,基于窗口的 $A_k(x, y)$ 平均策略使该方法节省了更多的结构信息,从而可以得到优化的活动层级图:

$$A_k(x, y) = \frac{\sum_{p=-r'}^{r'} \sum_{q=-r'}^{r'} A_k(x+p, y+q)}{(2r'+1)^2} \quad (6)$$

式中: r' 表示窗口的大小;变量 $\{p, q\}$ 在 $(-r, r)$ 中。当 r' 太小时,它可能无法表示图像的特征。然而,当 r' 太大时,会导致复杂性的大幅度增加。因此,采用合适的 r' 进行图像融合具有重要意义。本文通过大量实验验证,该方法在 r' 等于 3 时效果最好。然后,通过执行选择最大策略,需要融合系数映射,如式(7)所示。

$$C_{f,1:M}(x, y) = C_{k^*,1:M}(x, y), k^* = \underset{k}{\text{argmin}}(\bar{A}_k(x, y)) \quad (7)$$

式中: $C_{f,1:M}(x, y)$ 表示融合稀疏系数映射。通过学习字典和卷积稀疏向量重构融合后得到新的基础层,计算式如式(8)所示。

$$F_b = \sum_{m=1}^M d_m \times C_{f,m} \quad (8)$$

2.2 细节层处理

对于细节层,使用 ResNet50 做特征提取,对提取的特征 L1 正则化,特征层通过平均选择策略得到权重层 $W_k, k \in \{1, 2\}$, 将权重作为特征图 I_k^d 的权重系数得到细节层融合图片 F_d , 具体的细节层融合框架如图 3 所示。

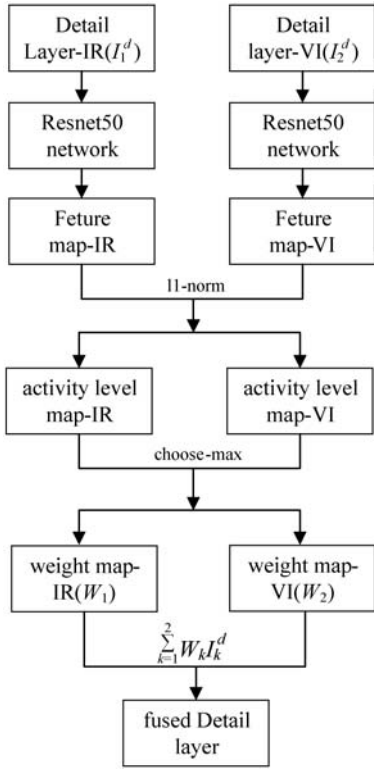


图3 细节层融合框架

细节层融合的具体步骤如下:

1) 特征提取:为了更好地利用图片的细节信息,将红外与可见光图片细节层 I_k^d 输入神经网络 ResNet50 提取特征。分别选取了 resnet 神经网络的 conv2_1、conv3_1、conv4_1 和 conv5_1 作为特征提取层 $L_k^{i,m}$,其中, $i \in (2,5)$, m 表示第 i 个特征层的通道数 $m = 64 \times 2^{i-1}$ 。

2) L1 范数正则化处理:对特征图进行 L1 范数处理,得到活动层级图 A_k^i 计算过程如式(9)所示。

$$A_k^i(x, y) = \|L_k^{i,m}(x, y)\|_1 \quad (9)$$

3) 权重计算:所得到的活动层级图采用最大选择的策略,得到权重图 $W_k^i(x, y)$ 。

$$W_k^i(x, y) = \begin{cases} 1 & k = \operatorname{argmax}_k(A_k(x, y)) \\ 0 & k = \operatorname{argmin}_k(A_k(x, y)) \end{cases} \quad (10)$$

4) 融合过程:经过权重分配,得到最终融合后的细节层,内部实验证明,将 conv5_1 输出作为特征图效果最佳。过程如式(11)所示。

$$F_d = \sum_{k=1}^2 W_k I_k^d \quad (11)$$

2.3 融合重建

获取细节层的 F_d 和基础层的 F_b 后,对图像进行重建可以得到最终的融合图片。

$$F(x, y) = F_b(x, y) + F_d(x, y) \quad (12)$$

式中: $F_b(x, y)$ 表示基础层在处的值; $F_d(x, y)$ 表示细

节层在 (x, y) 处的值; $F(x, y)$ 表示融合后图片在 (x, y) 处的值。

3 实验

3.1 实验设置

本文实验数据源于网上开源数据集 TNO,选取 4 种场景分别为“helicopter”“stroll”“tree”“street”,红外与可见光源如图 4 所示。内部实验对比了 resnet50 不同层的特征提取能力,因为 conv1 中没有残差块,选取 conv2_1、conv3_1、conv4_1、conv5_1 做比较。外部选取 6 种常用融合方法进行比较,分别是 CVT^[7]、DTCWT^[9]、NSCT^[12]、CSR^[19]、VGGML^[22]、ResNet-ZCA^[23],上述 6 种方法的参数取值均与其论文保持一致。为了定量且全面地比较本文方法和对比方法,选用以下四种评价指标:信息熵(EN)^[27]、互信息(MI)^[28]、标准差(SD)^[29]、视觉信息保真度(VIF)^[30]。实验环境为 Windows 10, MATLAB R2019a, Inter Core i7-7700HQ, 8 GB。

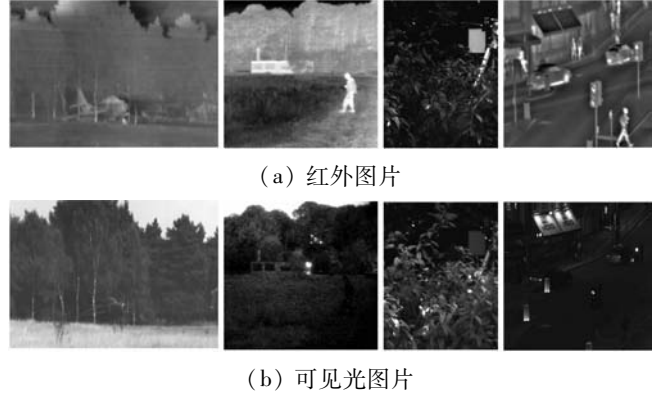


图4 四种场景所对应红外图片与可见光图片

3.2 内部实验

本文算法内部实验选取 conv2_1、conv3_1、conv4_1、conv5_1 进行特征提取并比较,由客观指标 EN、MI、VIF、SD 评判,如表 2 - 表 5 所示。从客观指标上看,conv5_1 在这四种场景下,基本达到了最好的效果。只有“stroll”场景下的 VIF 指标略小于 conv4_1。综上,我们所选取的特征层,在这四种场景,有很好的融合效果和很强的细节信息保留的能力。实验结果图如图 5 所示。

表2 场景 helicopter 内部实验指标

Layers	EN	MI	SD	VIF
conv2_1	6.651 0	4.042 0	47.350 4	0.402 9
conv3_1	6.651 0	4.042 1	47.350 3	0.402 9
conv4_1	6.650 9	4.042 4	47.350 9	0.402 9
conv5_1	6.651 0	4.042 4	47.351 1	0.402 9

表 3 场景 stroll 内部实验指标

Layers	EN	MI	SD	VIF
conv2_1	7.508 2	4.078 0	54.845 1	0.319 9
conv3_1	7.508 2	4.078 2	54.845 2	0.319 9
conv4_1	7.508 2	4.077 4	54.847 0	0.320 1
conv5_1	7.508 2	4.078 8	54.848 1	0.320 0

表 4 场景 tree 内部实验指标

Layers	EN	MI	SD	VIF
conv2_1	6.826 7	4.299 4	42.326 6	0.816 0
conv3_1	6.826 7	4.299 5	42.327 7	0.816 1
conv4_1	6.826 7	4.300 3	42.330 5	0.816 2
conv5_1	6.826 7	4.301 0	42.332 2	0.816 2

表 5 场景 street 内部实验指标

Layers	EN	MI	SD	VIF
conv2_1	6.687 5	3.518 6	33.380 5	0.514 7
conv3_1	6.687 3	3.518 5	33.381 6	0.514 7
conv4_1	6.687 3	3.515 9	33.388 2	0.515 1
conv5_1	6.687 7	3.522 6	33.402 9	0.515 1

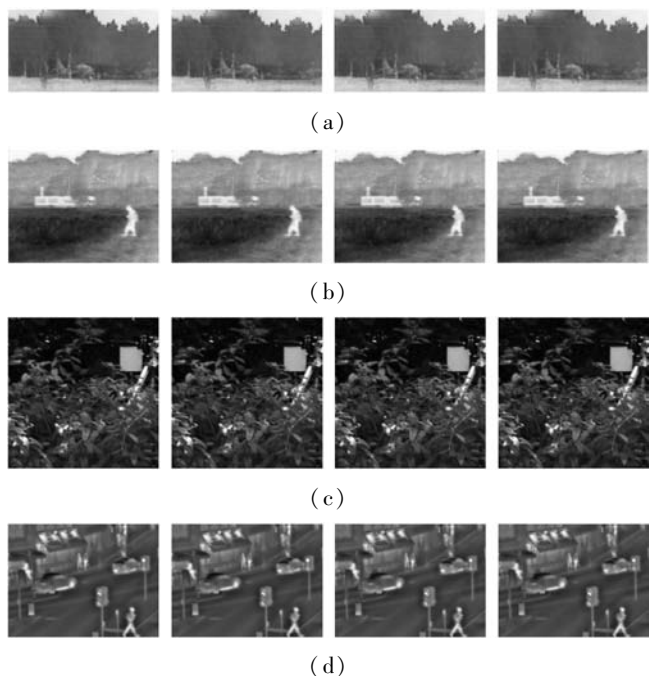


图 5 四个场景由不同卷积层提取特征的融合结果

3.3 外部实验

3.3.1 融合图片主观评价

1) 对于场景 helicopter。

各个方法的融合结果如图 6 所示。对方框选中的部分进行重点对比。将选中部分放大两倍置于左下角,可以看到对于飞机的前轮亮度,本文方法保存是最亮的。除此以外,方法融合后的图片中,树干的亮度

与细节也被较好地保存,CVT、CSR 和 DTCWT 在树干融合上丢失了较多的细节。整幅图像中,在树林与天空的分界上,本文方法融合得十分清晰,并且得到了较为真实天空的画面。总的来说,本文方法融合了红外图像中所展示的更多的细节,并在亮度保存上有较好的效果,所以在主观评价上表现得比其他方法要好。

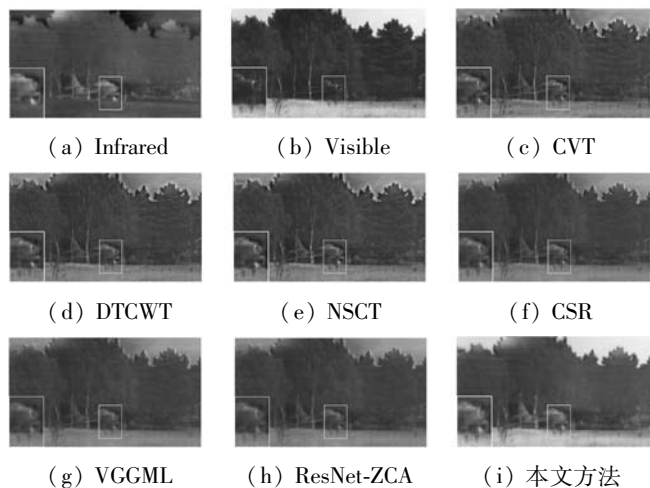


图 6 场景 helicopter 融合结果

2) 对于场景 stroll。

各个方法的融合结果如图 7 所示。可以看到本文方法在车头与背景的交界部分,边缘界限比较清晰,但是 CSR、CVT、VGGML 方法稍微有些模糊,这意味着,本文方法在边缘保存的效果上较好,车的部分亮度也是最高的。另外,方框选中部分,树干的保存较完整。除了被重点框出的车头部分,人的融合是所有方法融合得到的图片中亮度保存最好的,同时,在马路与低处分界的地方,本文方法的融合效果是最好的,其余方法丢失了马路的亮度与细节。所以,主观上本文方法比所其他方法效果更好。

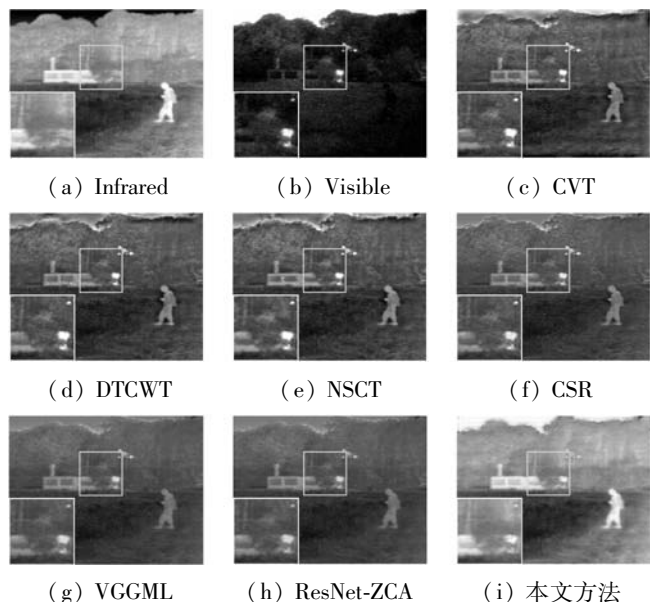


图 7 场景 stroll 融合结果

3) 对于场景 tree。

各个方法的融合结果如图 8 所示。可以看到花朵的亮度和边缘保留得很好, CVT 和 VGGML 方法融合后降低了花朵的亮度。本文方法所融合的图片中, 对白色木杆的融合效果很好, 边缘清晰不模糊, 完好地保存了红外图像中的亮度信息与边缘细节。同时, 对于木杆两侧的叶片, 本文方法相比较于 DTCWT、CSR 保存了更多的叶片细节。从整体图片来看, 本文方法所保存的叶片亮度与边缘信息优于其他方法。所以主观上, 方法较其他方法更有优越性。

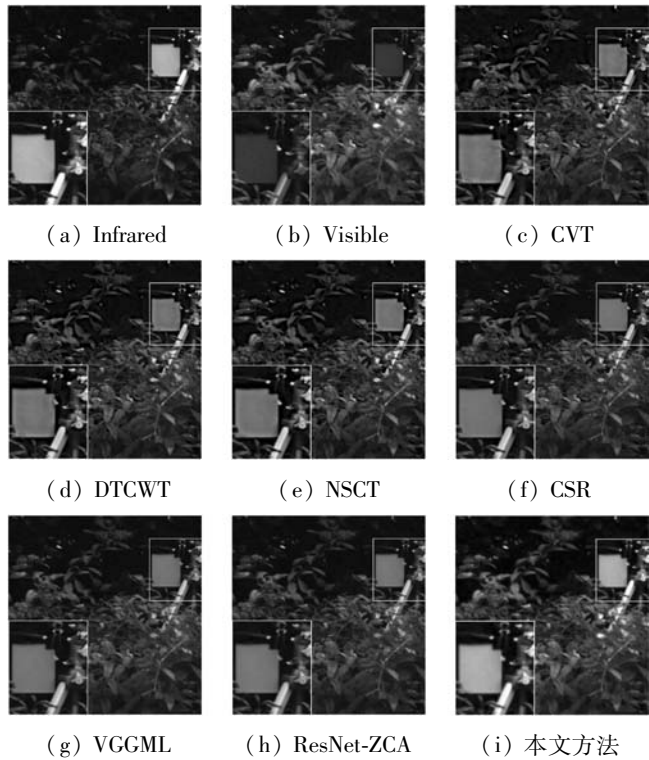


图 8 场景 tree 融合结果

4) 对于场景 street。

各个方法的融合结果如图 9 所示。可以看到, 人物的亮度信息和细节信息被保存得很好, 亮度远高于其余方法, 另外, 路沿部分边缘十分清晰, 而 ResNet-ZCA、VGGML、CSR 等方法稍微有些模糊并且损失了亮度。同时, 对于小的指示灯, 本文方法比其他方法更亮, 细节更丰富。从整体图片上来看, 对于广告牌上的英文字体, 本文方法所融合得到的也十分清晰, 但是 ResNet-ZCA、VGGML 方法较为模糊。总的来说, 方法在边缘保存和细节保留上优于其余方法的。

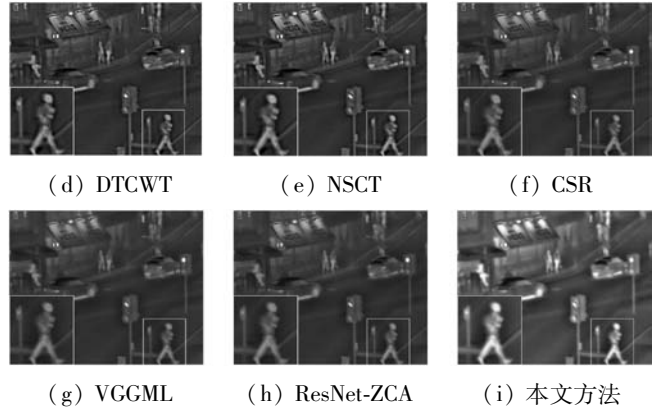
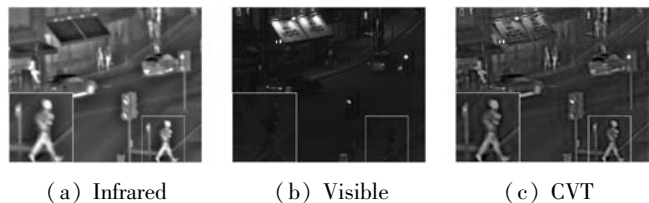


图 9 场景 street 融合结果

3.3.2 融合图片客观评价

由本文算法融合后的图片与典型的六种方法进行比较, 由客观指标 EN、MI、VIF、SD 评判, 如表 6 - 表 9 所示。从客观指标上看, 我们所提出的方法这四种场景下, 基本达到了最好的效果。对于第一组图片融合, 可以看到, 图片的互信息和细节信息都保存得最好, 指标的值远大于其他所比较的方法。在第二组图片中, 虽然视觉信息评价指标 (VIF) 小于其他方法, 但是在其余指标 EN、MI、SD 都远远大于所比较的方法, 说明本文方法所得到的融合图片视觉效果虽然稍差, 但是在细节保留和互补信息融合上是远远好于其他方法的。在第三组和第四组图片的上, 这四个评价指标也好于其余方法。综上, 本文方法在这四种场景中有很好的融合效果和很强的细节信息保留的能力。

表 6 场景 helicopter 外部实验指标

算法	EN	MI	SD	VIF
CVT	6.287 0	1.982 3	29.977 7	0.356 7
DTCWT	6.253 2	2.055 7	30.054 2	0.336 2
NSCT	6.306 8	2.142 5	30.539 9	0.417 0
CSR	6.132 3	2.468 1	29.166 0	0.342 1
VGGML	6.091 3	2.573 8	28.580 0	0.309 1
ResNet-ZCA	6.091 0	2.555 7	28.735 6	0.307 8
Proposed	6.651 0	4.042 4	47.351 1	0.402 9

表 7 场景 stroll 外部实验指标

算法	EN	MI	SD	VIF
CVT	7.019 2	2.133 7	33.098 8	0.390 5
DTCWT	7.046 3	2.332 8	34.193 8	0.413 1
NSCT	7.065 1	2.388 6	34.816 3	0.468 5
CSR	6.946 8	2.722 7	31.704 7	0.357 4
VGGML	6.910 5	3.030 0	30.868 8	0.355 4
ResNet-ZCA	6.904 7	3.036 6	30.678 5	0.349 0
Proposed	7.508 2	4.078 8	54.848 1	0.320 0

表 8 场景 tree 外部实验指标

算法	EN	MI	SD	VIF
CVT	6.595 3	3.204 7	35.171 9	0.738 2
DTCWT	6.587 6	3.336 6	35.352 8	0.749 8
NSCT	6.587 7	3.454 3	36.207 6	0.807 3
CSR	6.504 3	3.796 9	33.958 2	0.698 9
VGGML	6.425 1	4.252 8	32.168 6	0.620 5
ResNet-ZCA	6.424 9	4.279 3	32.112 5	0.612 1
Proposed	6.826 7	4.301 0	42.332 2	0.816 2

表 9 场景 street 外部实验指标

算法	EN	MI	SD	VIF
CVT	6.165 8	1.604 1	25.198 9	0.452 0
DTCWT	6.159 6	1.744 1	25.347 7	0.456 8
NSCT	6.201 4	1.808 2	26.424 8	0.514 1
CSR	6.021 2	2.313 2	23.389 8	0.415 9
VGGML	5.933 2	2.540 9	21.442 2	0.368 6
ResNet-ZCA	5.953 1	2.361 3	21.661 2	0.363 0
Proposed	6.687 7	3.522 6	33.402 9	0.515 1

3.3.3 算法运行时间对比

各个算法在四个场景进行图像融合的运行时间的对比如表 10 所示,对比算法包括 CVT、DTCWT、NSCT、CSR、VGGML 和 ResNet-ZCA,在同样的计算环境下运行。

表 10 各个算法在各场景中图像融合的运行时间对比

算法	helicopter	stroll	tree	street
CVT	0.016	0.012	0.015	0.017
DTCWT	0.017	0.007	0.015	0.018
NSCT	0.015	0.008	0.016	0.022
CSR	115.880	37.190	147.400	180.080
VGGML	8.470	3.780	8.760	10.560
ResNet-ZCA	3.180	1.450	4.290	4.830
Proposed	69.920	15.740	45.300	98.640

对比可得,传统算法(CVT、DTCWT、NSCT)运行时间短于本文算法,但是融合图像的客观与主观评价上,本文方法远远超过这几种传统方法。本文算法基于 CSR 和 ResNet 的算法,但运行速度处于该两种算法之间,比 CSR 算法效率有明显提升,图像效果同样提升显著。在牺牲了一定效率的情况下,本文算法大大提升了融合图像的质量。

4 结 语

本文使用 ResNet50 与 CSR 提取特征来融合红外

线和可见光图像。首先,将图像分解成包含大尺度特征的基础层和包含纹理信息的细节层。基础层部分使用 CSR 提取特征,平均权重进行融合。使用 ResNet50 网络对细节层进行特征提取,得到最大权重层,之后细节层和最大权重层相乘得到融合后的细节层。最后由融合后的基础层和细节层对图像进行重建。实验结果表明,本文方法不仅很好地保留了源图像的纹理特征,且很大程度上降低了融合图像的伪影和噪声,在主观评价和客观评价两个方面均取得了较好的融合效果,因此本文方法具有一定的实用价值与理论意义。

参 考 文 献

[1] 丁斋生,周冬明,聂仁灿,等. 基于视觉显著性与残差网络的红外-可见光图像融合方法[J]. 云南大学学报(自然科学版),2019,41(6):1108-1117.

[2] Ma J Y, Ma Y, Li C. Infrared and visible image fusion methods and applications: A survey[J]. Information Fusion, 2019, 45:153-178.

[3] Bulanon D M, Burks T F, Alchanatis V. Image fusion of visible and thermal images for fruit detection[J]. Biosystems Engineering, 2009, 103(1):12-22.

[4] Vanmali, A, Gadre, V. Visible and NIR image fusion using weight-map-guided Laplacian-Gaussian pyramid for improving scene visibility[J]. Sādhanā,2017,42, 1063-1082.

[5] Li H, Manjunath B, Mitra S. Multisensor image fusion using the wavelet transform[J]. Graphical Models and Image Processing, 1995, 57(3):235-245.

[6] Fan D Y. An advanced algorithm of multi-focus images fusion based on wavelet transform[J]. Advanced Materials Research 2011, 1198:775-779.

[7] Nencini F, Garzelli A, Baronti S, et al. Remote sensing image fusion using the curvelet transform[J]. Information Fusion,2007, 8(2):143-156.

[8] He T, Shang H. Direct-wave denoising of low-frequency ground-penetrating radar in open pits based on empirical curvelet transform[J]. Near Surface Geophysics,2020, 18(3): 295-305.

[9] 张贵仓,苏金凤,拓明秀. DTCWT 域的红外与可见光图像融合算法[J]. 计算机工程与科学,2020,42(7):1226-1233.

[10] Zuo Y, Liu J, Bai G, Wang X, et al. Airborne infrared and visible image fusion combined with region segmentation[J]. Sensors,2017,17(5).

[11] 齐海生,荣传振,肖力铭,等. 基于双树复小波变换与引导滤波的红外与可见光图像融合算法[J]. 通信技术,2019, 52(2):330-336.

[12] Fu Z, Wang X, Xu J, et al. Infrared and visible images fusion based on RPCA and NSCT[J]. Infrared Physics &

- Technology, 2016, 77:114 – 123.
- [13] 邓立暖, 尧新峰. 基于 NSST 的红外与可见光图像融合算法[J]. 电子学报, 2017, 45(12):2965 – 2970.
- [14] Wright J, Ma Y, Mairal J, et al. Sparse representation for computer vision and pattern recognition[J]. Proceedings of the IEEE, 98 (6):1031 – 1044.
- [15] Zhang Q, Liu Y, Blum R, 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.
- [16] Liu Y, Wang Z. Simultaneous image fusion and denoising with adaptive sparse representation[J]. IET Image Processing, 2015, 9(5):347 – 357.
- [17] Liu F, Chen L, Lu L, et al. Medical image fusion method by using Laplacian pyramid and convolutional sparse representation[J]. Concurrency and Computation: Practice and Experience, 2020, 32(17).
- [18] 董安勇, 苏斌, 赵文博, 等. 基于卷积稀疏表示的红外与可见光图像融合[J]. 激光与红外, 2018, 48(12):1547 – 1553.
- [19] Wohlberg B. Efficient algorithms for convolutional sparse representations[J]. IEEE Transactions on Image Processing, 2016, 25(1):301 – 315.
- [20] Zhao C, Wang T, Lei B. Medical image fusion method based on dense block and deep convolutional generative adversarial network[J]. Neural Computing and Applications, 2021, 33(12):6595 – 6610.
- [21] 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.
- [22] 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. IEEE, 2018:2705 – 2710.
- [23] Li H, Wu X J, Durrani T S. Infrared and visible image fusion with ResNet and zero-phase component analysis[J]. Infrared Physics & Technology, 2018, 102:103039.
- [24] Ma J, Yu W, Liang P, et al. FusionGAN: A generative adversarial network for infrared and visible image fusion[J]. Information Fusion, 2018, 48:11 – 26.
- [25] 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.
- [26] Gao C, Liu F, Yan H. Infrared and visible image fusion using dual-tree complex wavelet transform and convolutional sparse representation[J]. Journal of Intelligent and Fuzzy Systems, 2020(5):1 – 13.
- [27] Roberts J, Aardt J, Ahmed F. Assessment of image fusion procedures using entropy, image quality, and multispectral classification[J]. Journal of Applied Remote Sensing, 2008, 2(1):1 – 28.
- [28] Liu J, Shao Z. Feature-based remote sensing image fusion quality metrics using structure similarity[J]. Acta Photonica Sinica, 2011, 40(1):126 – 131.
- [29] Ma J, Ma Y, Li C. Infrared and visible image fusion methods and applications: A survey[J]. Information Fusion, 2019:153 – 178.
- [30] Han Y, Cai Y, Cao Y, et al. A new image fusion performance metric based on visual information fidelity[J]. Information Fusion, 2013, 14(2):127 – 135.
- ~~~~~
- (上接第 170 页)
- [9] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//IEEE International Conference on Computer Vision and Pattern Recognition, 2016:770 – 778.
- [10] Hou Q B, Cheng M, Hu X W, et al. Deeply supervised salient object detection with short connections[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2015:815 – 828.
- [11] Luo Z M, Mishra A, Achkar A, et al. Non-local deep features for salient object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2017:6593 – 6601.
- [12] Movahedi V, Elder J H. Design and perceptual validation of performance measures for salient object segmentation[C]//IEEE International Conference on Computer Vision and Pattern Recognition, 2010:49 – 56.
- [13] Yan Q, Xu L, Shi J P, et al. Hierarchical saliency detection[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2013:1155 – 1162.
- [14] Li Y, Hou X D, Koch C, et al. The secrets of salient object segmentation[C]//IEEE International Conference on Computer Vision and Pattern Recognition, 2014:280 – 287.
- [15] Wang L J, Lu H C, Wang Y F, et al. Learning to detect salient objects with image-level supervision[C]//IEEE International Conference on Computer Vision and Pattern Recognition, 2017:3796 – 3805.
- [16] Li G B, Yu Y Z. Visual saliency based on multiscale deep features[C]//IEEE International Conference on Computer Vision and Pattern Recognition, 2015:5455 – 5463.
- [17] Feng M Y, Lu H C, Ding E. Attentive feedback network for boundary aware salient object detection[C]//IEEE International Conference on Computer Vision and Pattern Recognition, 2019:1623 – 1632.
- [18] Wang L Z, Lu H J, Zhang P, et al. Saliency detection with recurrent fully convolutional networks[C]//European Conference on Computer Vision, 2016:825 – 841.