

DOI:10.16136/j.joel.2022.03.0372

基于多尺度局部极值分解与 ResNet152 的红外与可见光图像融合

陈广秋*, 王 帅, 黄丹丹, 段 锦

(长春理工大学 电子信息工程学院, 吉林 长春 130022)

摘要: 为了进一步提升红外与可见光图像融合方法的性能, 本文提出了一种基于多尺度局部极值分解与深度学习网络 ResNet152 的红外与可见光图像融合方法。首先, 利用多尺度局部极值分解 (multiscale local extrema decomposition, MLED) 方法将源图像分解为近似图像和细节图像, 分离出源图像中重叠的重要特征信息。然后采用残差网络 ResNet152 深度提取源图像的多维显著特征, 以 l_1 -范数作为活性测度生成显著特征图, 对近似图像进行加权平均融合, 以保持能量和残留细节信息不丢失。在细节图像中, 利用“系数绝对值取大”规则获得初始决策图, 源图像作为引导图像, 初始决策图作为输入图像进行引导滤波处理, 得到优化决策图, 计算加权局部能量得到能量显著图, 对细节图像进行加权平均融合, 使融合图像具有丰富的纹理细节和良好的视觉边缘感知。最后, 对近似融合图像和细节融合图像进行重构, 得到融合图像。实验结果表明, 与现有的典型融合方法相比, 本文所提出的融合方法在客观评价和视觉感受方面都取得了最好的效果。

关键词: 数字图像处理; 图像融合; 多尺度局部极值分解 (multiscale local extrema decomposition, MLED); 残差网络; 显著特征图; 能量显著图

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1005-0086(2022)03-0283-13

Infrared and visible image fusion based on multiscale local extrema decomposition and ResNet152

CHEN Guangqiu, WANG Shuai, HUANG Dandan, DUAN Jin

(School of Electronic and Information Engineering, Changchun University of Science and Technology, Changchun, Jilin 130022, China)

Abstract: In order to further improving the performance of infrared and visible image fusion method, an infrared and visible image fusion method based on multiscale local extrema decomposition (MLED) and deep learning network ResNet152 is proposed in this paper. Firstly, the source images are decomposed into approximate images and many detail images using MLED, which can separate out the overlapped important feature information. Secondly, the residual network ResNet152 is used to extract the multi-dimensional deep features of the source images, and the l_1 -norm is used as the activity level measure to generate the salient feature maps, the weighted average fusion algorithm is carried out for the approximate images, which can keep the energy and residual details not lost. For the detail images, the initial decision map is obtained by the rule “coefficient absolute max”. The source images are used as the guided images, and the initial decision maps are used as the input images for guided filtering. So the optimized decision maps are obtained, the weighted local energy is calculated to get the energy saliency maps. The weighted average algorithm is carried out for the detail images, which can make the fusion image having rich texture details and good visual edge perception. Finally, the fusion image is obtained by reconstructing the fused approximate image and detail images. The experimental results show that, compared with

* E-mail: guangqiu_chen@126.com

收稿日期: 2021-06-01 修订日期: 2021-07-13

基金项目: 吉林省科技发展规划项目 (20200201167JC) 和吉林省教育厅“十三五”科学技术项目 (JJKH20200785KJ) 资助项目

the existing typical fusion methods, the proposed method can achieve state-of-the-art results in terms of both objective evaluation and visual quality.

Key words: digital image processing; image fusion; multiscale local extrema decomposition (MLEED); residual network; salient feature map; energy saliency map

1 引 言

图像融合是一种将多个具有互补性的特定场景图像信息合并到一个图像中,融合后的图像比任何单个源图像所提供的信息更丰富,便于视觉观察或其他一些计算机视觉任务处理。红外与可见光图像融合由于应用领域广泛而成为多年来持续研究的热点。可见光传感器是通过捕获场景中物体的反射光而成像的,图像中包含丰富的外观细节信息,有利于建立判别模型,但成像质量易受环境因素影响,如低照度光线、烟、雾、霾或恶劣天气等。红外传感器主要捕获物体发出的热辐射,受光照变化或伪装的影响较小,因此能够克服环境等不利因素而捕获目标。因此,准确地将红外与可见光图像进行融合,有利于目标的自动检测和精确定位,拓展图像处理的应用领域^[1]。

图像融合的关键问题是如何从源图像中提取显著特征,并将其合并生成融合图像,过去几十年里已经提出了多种红外与可见光图像融合的方法,其中基于固定基函数的多尺度分解和多尺度边缘保持平滑滤波器 2 种融合方法应用最为广泛,前者主要包括离散小波变换(discrete wavelet transform, DWT)^[2]、曲波变换^[3]、轮廓波变换^[4]、剪切波变换^[5]等,这类融合方法首先对图像进行多尺度与多方向分解,然后采用预先设计的融合规则对分解后的系数进行合并,最后利用逆变换重构生成融合图像,但这种预先定义的基函数并不一定能够准确地表达图像的边缘和纹理,从而影响融合效果,在边缘周围容易出现“晕”现象。后者是利用边缘保持平滑滤波器(edge preserving smoothing filter, EPSF)对图像进行迭代滤波,生成一个基层和多个细节层图像^[6],主要包括双边滤波器(bilateral filter, BF)^[7]、引导滤波器(guided filter, GF)^[8]、 l_0 梯度最小化滤波器(l_0 gradient minimization filter, l_0 GMF)^[9]和滚动滤波器(rolling guidance filter, RGF)^[10]等,但传统的 EPSF 是根据图像的对比度去除细节信息,从而达到边缘保持的目的,较少考虑它们的空间尺度,在图像分解过程中并不能对空间重叠特征进行多尺度分离,在图像边缘处容易出现梯度反转现象,从而在融合图像中引入伪影或虚假信息。多尺度图像融合方法中,近似图像和细节图像所采用的“等权值

平均”和“系数绝对值取大”融合规则削弱了融合图像的亮度和对比度,容易将源图像的噪声转移到融合图像中,造成融合图像质量的下降^[11]。

随着压缩感知理论的发展,基于稀疏表示和低秩表示的融合方法也出现在许多文献中,并取得了较好的融合效果,这类融合方法主要采用分块稀疏或低秩表示,利用 l_1 范数取大规则合并系数,采用块平均的方法重构图像,由于这种块稀疏表示是多值的且不相等,块平均重构算法容易平滑掉边缘信息且易出现块效应^[12]。

近年,基于深度学习的融合方法成为研究热点,相比较于非学习类的融合方法,深度学习(deep learning, DL)方法以其强大的特征提取和数据表示能力,取得了更好的融合性能和时间效率^[13]。文献[14]提出了一种基于卷积神经网络的融合方法,该方法采用孪生卷积神经网络获得权重图,根据拉普拉斯金字塔分解后子带图像的能量匹配度进行选择融合或加权性融合,取得了较好的融合效果,但该方法中起主导作用的是选择性融合,在融合过程中权重图的参与度非常有限。文献[15]使用了一种基于 VGG 网络框架的融合方法,该方法中基层图像采用等权值加权平均,利用 VGG19 网络提取细节层图像的深层特征并生成权重图,对细节图像进行加权平均融合,该方法中只是利用 VGG19 网络提取了细节图像的特征信息,并没有充分利用源图像的内容信息,基层图像融合方法过于简单,融合图像中丢失了大量有用的信息。文献[16]提出了一种基于卷积稀疏表示的融合方法,该方法首先对源图像进行 2 层分解,然后采用卷积稀疏表示提取细节图像的特征信息,利用特征值取大规则选取稀疏系数,近似图像采用等权值加权平均融合规则,该方法中特征值取大规则容易误选噪声作为融合系数,近似图像的融合规则仍然过于简单,降低了图像的亮度和对比度。

为了克服上述缺陷,本文提出了一种多尺度局部极值分解与 ResNet152 的红外与可见光图像融合方法,利用不同尺度模板的局部极值分解方法,将源图像分解为包含不同尺度空间特征信息的近似图像和细节图像,采用残差网络 ResNet152 深度提取源图像的多维显著特征,以 l_1 范数作为活性测度生成显著特征图,对近似图像进

行加权平均融合,以保持能量和剩余细节信息不丢失,使融合图像的整体外观和对比度更加符合人的视觉观察习惯。在细节图像中,利用“系数绝对值取大”规则获得初始决策图,以源图像作为引导图像,初始决策图作为输入图像进行引导滤波,得到优化决策图,对优化决策图进行局部能量计算得到能量显著图,对细节图像进行加权平均融合,使融合图像具有丰富的纹理信息和良好的视觉边缘感知,避免传统“系数绝对值取大”规则在图像中融入过多的不相干红外细节信息和噪声。

2 多尺度局部极值分解

局部极值分解能够区分图像中高对比度纹理和强边缘成分,在平滑图像的同时能够有效地保留边缘信息,在不考虑对比度信息的情况下提取小尺度细节信息^[17]。假设输入图像为 I , $I(p)$ 为 p 点像素的强度,局部极值分解算法具体过程如下:

1) 局部极值识别

利用大小为 $e \times e$ 的尺度模板扫描输入图像 I ,如果在模板范围内最多有 $e-1$ 个像素值大于(或小于) $I(p)$ 时,则表示 p 点为极大值(或极小值)点, $e=2j+1$, j 为尺度值,取值为 $1 \leq j \leq J$ 的整数。

2) 构造极值包络

对于输入图像 I 和局部极值像素集合 S ,采用 LEVIN 等^[18]提出的插值算法计算图像的极值包络 E ,依据图像局域性原理,相邻像素值具有一定的相似性,如果 $I(r)$ 和 $I(s)$ 具有相似性,那么相邻像素的 $E(r)$ 和 $E(s)$ 也具有相似性,这样就可以通过最小化式(1)所表示的能量函数获得极值包络。

$$\sum_r \left(E(r) - \sum_{s \in N(r)} \omega_{rs} E(s) \right)^2, \quad \forall p \in S, E(p) = I(p), \quad (1)$$

式中, $E(r)$ 表示 r 的邻域,权重 ω_{rs} 可以利用 r 的邻域方差 σ_r^2 计算得到:

$$\omega_{rs} \propto \exp \left(- \frac{(I(r) - I(s))^2}{2\sigma_r^2} \right). \quad (2)$$

采用加权最小二乘法最小化式(1)所表示的二次函数,将其简化为求解一个 $E(r)$ 为 3×3 局部邻域的稀疏线性系统,得到极值包络的输出 $E_{\max}$ 和 $E_{\min}$ 。

3) 获得基层图像

对两个极值包络进行平均运算处理,得到基层图像,即:

$$I^B = \frac{E_{\max} + E_{\min}}{2}. \quad (3)$$

4) 获得细节图像

将原图像与基层图像进行差分运算得到细节图像,即:

$$I^d = I - I^B. \quad (4)$$

不断的改变模板尺寸,重复步骤(1)–(4),能够得到多个细节图像 I_j^d 和多个基层图像 I_j^B ,对基层图像进行平均运算处理,得到最后的近似图像 I^A ,即 $I^A = \frac{1}{J} \sum_{j=1}^J I_j^B$,这样就形成了多尺度局部极值分解。图1所示为图像“Barbara”的2个尺度局部极值分解,尺度模板大小分别为 3×3 和 5×5 。

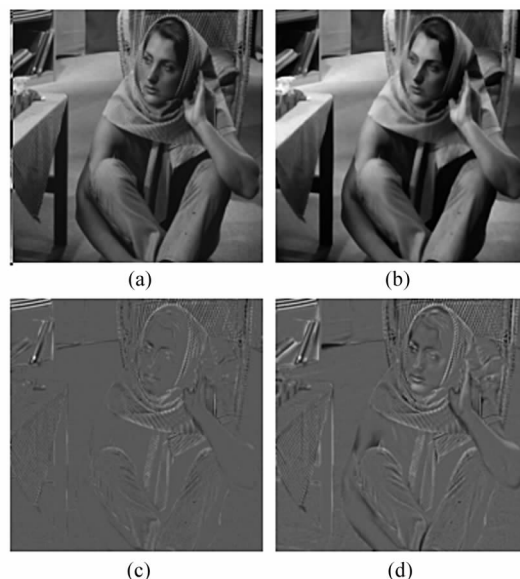


图1 图像“Barbara”的2个尺度局部极值分解:

(a) 原图像; (b) 近似图像;

(c) 细节图像1; (d) 细节图像2

Fig. 1 Two-scale local extrema decomposition for image “Barbara”: (a) Original image; (b) Approximate image; (c) Detail image 1; (d) Detail image 2

由上述分解过程可以得出,局部极值分解实质上是一种边缘保持平滑算子,输入图像 I 可以被分解成一个近似图像 I^A 和多个细节图像 I_j^d ,近似图像 I^A 包含源图像的主要成分,细节图像 I_j^d 包含不同尺度上的纹理和边缘信息,重构图像时,只需将 I_j^A 和多个 I_j^d 的均值求和即可,其表达式为:

$$I = I^A + \frac{1}{J} \sum_{j=1}^J I_j^d. \quad (5)$$

3 残差网络

深度卷积网络能够以端到端的多层方式提取低/中/高层次特征,并且堆叠层数越多,特征层次越丰富,但随着网络深度的不断增加,出现梯度消失和梯度爆炸问题,从而阻碍整个网络的收敛,这一问题通过标准初始化层和中间归一化层得到了初步解决,但随着网络深度的进一步增加,在训练集上出现了精度饱和后迅速下降的退化问题,而且这种退化

并不是由于过度拟合引起的,文献[19]推断这种退化是由于深度网络难以训练,求解器不能用多个非线性层拟合逼近恒等映射引起的。

为了解决上述问题,HE等^[19]提出了一种深度残差网络学习框架,该框架利用少数卷积层和一个短连接形成一个残差单元,然后通过堆叠残差单元形成残差网络,在训练过程中,只需学习残差函数,不会增加额外的参数和计算复杂度,使该网络更容易被优化和具有更高的精度。两种不同的残差网络单元如图2所示。原残差网络单元如图2(a)所示,计算式可表示为:

$$y_l = h(x_l) + F(x_l, W_l), x_{l+1} = f(y_l), \quad (6)$$

式中, x_l 和 x_{l+1} 分别为第 l 个残差单元的输入和输出特征, $W_l = \{M_{l,k} \mid 1 \leq k \leq K\}$ 为与第 l 个残差单元相关的一系列权重, K 为一个残差单元内的层数。 F 表示要学习的残差映射, f 为激活函数,是逐像素相加运算,函数被设置成恒等映射 $h(x_l) = x_l$ 。

图2中ReLU为线性修正单元,BN为批规范化。

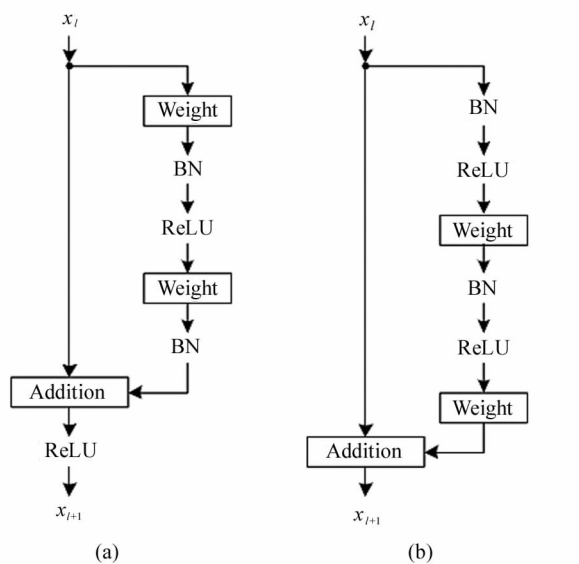


图2 两种不同的残差网络单元:

(a) 原残差网络单元; (b) 改进的残差网络单元

Fig. 2 Two different residual network units:

(a) Original residual network unit;

(b) Improved residual network unit

为了进一步提升训练精度和收敛速度,HE等^[20]在文献[19]的基础上对原残差单元进行了改进,如图2(b)所示,图中有2个恒等映射,分别为恒等跳跃连接 $h(x_l) = x_l$ 和恒等映射 $f \equiv y_l$,则式(6)可改写成下式:

$$x_{l+1} = x_l + F(x_l, W_l). \quad (7)$$

对上式进行递归运算:

$$x_{l+2} = x_{l+1} + F(x_{l+1}, W_{l+1}) = x_l + F(x_l, W_l) + F(x_{l+1}, W_{l+1}), \quad (8)$$

有下式成立为:

$$x_L = x_l + \sum_{i=l}^{L-1} F(x_i, W_i). \quad (9)$$

对于任何深单元 L 和浅单元 l ,式(9)具有下列性质:(1)任意较深单元 L 的特征可以表示为较浅单元 l 的特征 x_l 加上一个残差函数 $\sum_{i=l}^{L-1} F(x_i, W_i)$,这表明该模型在任何单元 L 和 l 之间以残差方式存在;(2)任意单元深度特征 $x_L = x_0 + \sum_{i=l}^{L-1} F(x_i, W_{i,k})$ 可以表示成 L 前所有残差函数的输出和再加上 x_0 ,而普通网络的特征是一系列矩阵向量的乘积。

同时式(9) also 具有良好的反向传播特性,假设用反向传播链规则表示的损失函数 ϵ ,其表达式为:

$$\frac{\partial \epsilon}{\partial x_l} = \frac{\partial \epsilon}{\partial x_L} \frac{\partial x_L}{\partial x_l} = \frac{\partial \epsilon}{\partial x_L} \left(1 + \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, W_i) \right). \quad (10)$$

上式表明梯度 $\frac{\partial \epsilon}{\partial x_l}$ 可以分解为2项的和,其中 $\frac{\partial \epsilon}{\partial x_L}$ 项可以直接传播信息而不涉及任何权重层,确保信息直接传回任意浅层单元 l ,另一项 $\frac{\partial \epsilon}{\partial x_L} \cdot \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, W_i)$ 通过权重层传播信息。当残差梯度 $\frac{\partial F(x_l, W_l)}{\partial x_l}$ 很小时,层的梯度也不会消失。式(9)和式(10)表明,信号可以从任何一个单元直接传播到另一个单元,包括前向传播和后向传播。

HE等^[20]构造了深度达到152层的ResNet152网络,并与VGG/GoogLeNet/PReLU-net/Plain Net进行了实验对比,在训练的复杂度和精度方面均具有一定的优势,并且网络深度越深,训练精度也越高,完美地解决了退化问题。ResNet152网络采用3层卷积结构代替图2(b)中的2层卷积结构,形成1个残差网络单元,共有4个特征提取层,分别为Conv2_x、Conv3_x、Conv4_x、Conv5_x,对应的残差网络单元数量分别为3、8、36、3,再加上输入卷积层和全连接卷积层,共有152个卷积层,ResNet152网络的结构如图3所示,其中Conv1为输入卷积层,FC为全连接卷积层。ResNet152网络是VGG网络深度的8倍,在具备更强的图像特征提取能力的同时,具有更低的复杂度和训练误差。因此本文采用Res-

Net152网络作为红外与可见光图像的特征提取网络。

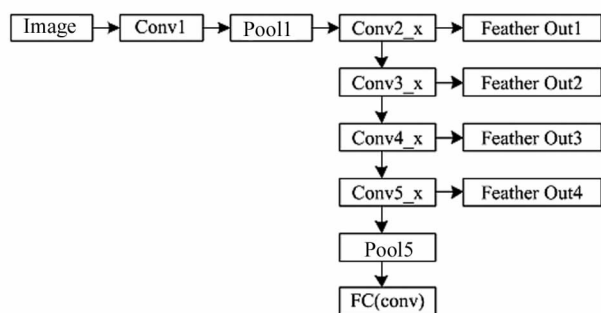


图3 ResNet152网络的结构

Fig. 3 Structure of ResNet152 network

4 融合规则

本文提出的融合方法主要包括3个部分,分别

为图像分解(或图像分析)、系数合并和图像重构(或图像综合),首先通过多尺度局部极值分解算法将输入图像分解为近似图像和细节图像,分离出不同尺度空间内的边缘和纹理特征,然后应用残差网络ResNet152对源图像进行深度特征提取,生成显著特征图,对近似图像进行加权平均融合,以保持能量和剩余细节信息不丢失。利用“系数绝对值取大”规则获得细节图像系数的初始决策图,将源图像作为引导图像,对决策图进行引导滤波器处理,获得边缘保持而纹理增强的优化决策图,计算图像系数的加权局部能量值得到能量显著图,对细节图像进行加权平均融合,以保留图像中的视觉结构和边缘,降低噪声的影响。最后,对融合后的近似图像和细节图像进行重构,获得融合图像,本文融合方法的总体框图如图4所示。

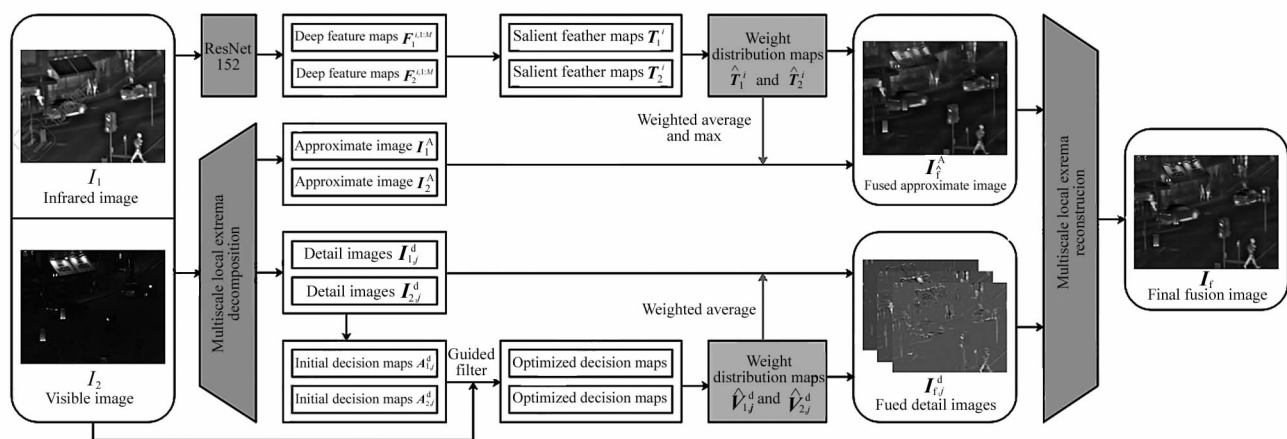


图4 本文融合方法的总体框图

Fig. 4 The overall block diagram of the proposed fusion method

4.1 近似图像融合规则

近似图像集中了源图像的主要能量和部分细节信息,由于红外与可见光图像的成像机制不同,在同一空间位置上的2类图像的像素强度差别较大,传统的等权值加权平均融合规则没有充分考虑这些信息特征,造成融合图像中能量损失严重,部分细节缺失,整体亮度下降,图像的视觉感知能力降低,严重地影响了融合图像的质量。

为了解决上述问题,本文提出一种基于残差网络ResNet152的加权平均近似图像融合规则,ResNet152网络能够深度提取图像的特征信息,如目标结构、纹理等,并以多维深度特征图像表现出来,利

用 l_1 -范数整合这些特征层图像,生成一幅包含所有特征信息的显著特征图,作为近似图像的活性测度,对近似图像进行加权平均融合,具体实现过程如下所述。

利用ResNet152提取源图像 I_n 的深层特征,表示为 $\mathbf{F}_n^{i,1:M}$, $n \in \{1, 2\}$, $i \in \{1, 2, 3, 4\}$ 表示特征提取层, M 为第 i 层输出的通道数, $M = 256 \times 2^{i-1}$, $\mathbf{F}_n^{i,m}(u, v)$ 表示第 i 层第 m 个深度特征图在空间位置 (u, v) 的灰度值, $m \in \{1, \dots, M\}$, $\mathbf{F}_n^{i,1:M}(u, v)$ 表示一个 M 维的向量。受文献[16]启发, $\mathbf{F}_n^{i,1:M}(u, v)$ 的 l_1 范数能够作为源图像信息内容的活性测度,活性测度图可由下式计算得到:

$$\hat{\mathbf{F}}_n^i(u, v) = \|\mathbf{F}_n^{i,1,M}(u, v)\|_1. \quad (11)$$

为了使融合规则对噪声和失配更具有鲁棒性,对 $\hat{\mathbf{F}}_n^i(u, v)$ 进行滑动分块平均处理,得到初始显著特征图,如下式所示:

$$\mathbf{P}_n^i(u, v) = \frac{\sum_{\alpha=-\theta}^{\theta} \sum_{\beta=-\theta}^{\theta} \hat{\mathbf{F}}_n^i(u + \alpha, v + \beta)}{(2\theta + 1)^2}, \quad (12)$$

式中, θ 为块尺寸半径, θ 值过小,鲁棒性差, θ 值过大,会丢失细节信息,文中 $\theta=1$ 。

由于网络中存在池化层,需要对初始显著特征图进行上采样运算处理,使其与源图像尺寸一致,本文利用双三次插值算法对初始显著特征图进行调整,得到尺寸大小与源图像一致的显著特征图 $\mathbf{T}_n^i$:

$$\mathbf{T}_n^i = \text{BiCubic}(\mathbf{P}_n^i), \quad (13)$$

式中, $\text{BiCubic}(\cdot)$ 表示双三次插值运算。对 $\mathbf{T}_n^i$ 进行归一化处理,得到最终的权重分配图表示为:

$$\begin{aligned} \hat{\mathbf{T}}_1^i(u, v) &= \frac{\mathbf{T}_1^i(u, v)}{\mathbf{T}_1^i(u, v) + \mathbf{T}_2^i(u, v)}, \\ \hat{\mathbf{T}}_2^i(u, v) &= 1 - \hat{\mathbf{T}}_1^i(u, v), \end{aligned} \quad (14)$$

对近似图像进行加权平均表示为:

$$\begin{aligned} \mathbf{I}_i^{\wedge}(u, v) &= \hat{\mathbf{T}}_1^i(u, v) \cdot \mathbf{I}_1^{\wedge}(u, v) + \\ &\quad \hat{\mathbf{T}}_2^i(u, v) \cdot \mathbf{I}_2^{\wedge}(u, v), \end{aligned} \quad (15)$$

最后的近似融合图像由下式得到:

$$\mathbf{I}_i^{\wedge}(u, v) = \max[\mathbf{I}_i^{\wedge}(u, v) \mid i \in \{1, 2, 3, 4\}]. \quad (16)$$

4.2 细节图像的融合规则

细节图像中包含了源图像的边缘和部分纹理细节信息,大的系数绝对值意味着更显著的边缘和纹理信息,所以许多文献中采用“系数绝对值取大”规则合并细节图像,并取得了较好的效果,但对于红外与可见光图像的融合,红外图像一般的目标亮度较高,背景噪声严重,传统的“系数绝对值取大”规则容易融入过多的红外图像信息和噪声,从而缺少可见光背景细节信息,造成融合图像视觉感受较差。为了克服上述缺陷,本文提出了一种基于能量显著图的加权融合规则,在传统的“系数绝对值取大”规则的基础上,通过引导滤波器对决策图进行引导滤波^[8],得到边缘保留和纹理信息增强的优化决策图,为了有效地提取细节信息,利用加权局部能量作为活性测度生成能量显著图,对细节图像进行加权平均融合,优化红外图像和可见光图像在融合图像中的比例,使融合图像更适合视觉观察。具体实现过程如下:

首先,利用“系数绝对值取大”规则获得初始决

策图:

$$\begin{aligned} \mathbf{A}_{1,j}^d(u, v) &= \begin{cases} 1, & |\mathbf{I}_{1,j}^d(u, v)| \geq |\mathbf{I}_{2,j}^d(u, v)| \\ 0, & \text{otherwise} \end{cases}, \\ \mathbf{A}_{2,j}^d(u, v) &= 1 - \mathbf{A}_{1,j}^d(u, v). \end{aligned} \quad (17)$$

然后,源图像 $\mathbf{I}_n$ 为引导图像,初始决策图为输入图像进行引导滤波处理,得到优化决策图:

$$\hat{\mathbf{A}}_{n,j}^d = GF(\mathbf{A}_{n,j}^d, \mathbf{I}_n, z, \lambda), n \in \{1, 2\}, \quad (18)$$

式中, $\hat{\mathbf{A}}_{n,j}^d$ 为引导滤波器的输出, $GF(\cdot)$ 表示引导滤波器, z 和 λ 分别为滤波器尺寸半径和正则化参数,滤波器尺寸大小为 $(2z+1) \times (2z+1)$ 。

经过引导滤波器优化后,决策图的边缘部分与输入图像 $\mathbf{A}_{n,j}^d$ 相似,纹理部分与引导图像 $\mathbf{I}_n$ 相似,这样就将源图像中部分纹理细节转移到决策图中。图像中局部能量越大对应的像素变化越剧烈,图像信息越丰富,可见局部能量表示图像的清晰度,能够反映灰度变化的绝对强度,所以本文通过计算优化决策图的加权局部能量生成能量显著图,对细节图像进行加权平均融合。图像的加权局部能量^[21]可由下式计算得到:

$$\mathbf{V}_{n,j}^d(u, v) = \sum_{t=-q}^q \sum_{t'=-q}^q \mathbf{W} \cdot \hat{\mathbf{A}}_{n,j}^d(u + t, v + t')^2, \quad (19)$$

式中, $\mathbf{W}$ 为权重矩阵,半径为 q ,尺寸为 $(2q+1) \times (2q+1)$ 。

$\mathbf{W}$ 中的元素值为 2^{2q-c} , c 为到中心点的4-邻域距离,为了不丢失边缘细节信息,本文中 $q=1$,此时3

$$\times 3 \text{ 的权重矩阵为 } \mathbf{W} = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}.$$

对 $\mathbf{V}_{1,j}^d$ 和 $\mathbf{V}_{2,j}^d$ 进行归一化处理,得到细节图像的权重分配图:

$$\begin{aligned} \hat{\mathbf{V}}_{1,j}^d(u, v) &= \frac{\mathbf{V}_{1,j}^d(u, v)}{\mathbf{V}_{1,j}^d(u, v) + \mathbf{V}_{2,j}^d(u, v)}, \\ \hat{\mathbf{V}}_{2,j}^d(u, v) &= 1 - \hat{\mathbf{V}}_{1,j}^d(u, v). \end{aligned} \quad (20)$$

最后,细节融合图像可由下式计算得到为:

$$\begin{aligned} \mathbf{I}_{f,j}^d(u, v) &= \hat{\mathbf{V}}_{1,j}^d(u, v) \cdot \mathbf{I}_{1,j}^d(u, v) + \\ &\quad \hat{\mathbf{V}}_{2,j}^d(u, v) \cdot \mathbf{I}_{2,j}^d(u, v), \end{aligned} \quad (21)$$

重构后的融合图像为:

$$\mathbf{I}_i = \mathbf{I}_i^{\wedge} + \frac{1}{J} \sum_{j=1}^J \mathbf{I}_{f,j}^d. \quad (22)$$

5 实验与分析

实验由2部分组成,分别为自由参数的分析确定和不同典型融合方法的对比分析,所用图像均来

自 https://figshare.com/articles/TNO_Image_Fusion_Dataset/1008029 的 TNO_Image_Fusion_Dataset 数据集。用于特征提取的残差网络 ResNet152 是一种通过 ImageNet 训练的固定网络,可在 <https://www.vlfeat.org/matconvnet/pretrained/#pretrained-models> 网站下载得到,实验结果的评价方式采用主观视觉观察和客观定量指标值,本文所用的客观评价指标包括结构信息相似度 (structural similarity, SSIM)^[22],人类视觉系统感知质量评价系数 Q_{HVS} ^[22],特征互信息 (feature mutual information, FMI_w)^[22]、边缘融合质量因子 Q_E ^[22] 和视觉信息保真度 (visual information fidelity, VIFF)^[22]。SSIM 是利用融合图像与源图像在亮度、对比度和结构方面的相似程度来评价融合图像的质量。 Q_{HVS} 是基于人眼视觉特性模型,通过计算融合图像与源图像感知对比度的保持程度对融合图像质量进行评价。 FMI_w 是通过计算融合图像与源图像特征信息间的互信息来评价融合图像的质量。 Q_E 是通过显著特征加权边缘图像的结构信息相似度对融合图像进行质量评价。VIFF 是结合自然图像统计模型、图像失真模型和人眼视觉系统模型的图像质量评价指标。上述评价指标值越大表明融合图像质量越好,反之越差。本文所用对比方法的实验结果均为参考文献作者提供的源代码仿真获得,参数均为作者默认参数。实验环境为: Intel(R) Core (TM) i5-4210M CPU @2.6 GHz 2.60 GHz, 安装内

存为 8.00 GB, 64 位操作系统, Windows7 旗舰版, 软件为 Matlab2018a 和 Visual Studio 2015。

5.1 自由参数的分析确定

本文融合方法中含有 3 个自由参数, 分别为源图像分解尺度数 j 、引导滤波器的窗口尺寸半径 z 和正则化参数 λ , 不同的参数值会对融合性能产生影响, 为了确定最优的参数值, 采用一组独立的图像集进行融合性能分析, 通过 5 种客观评价指标对融合结果进行客观评价, 每种指标的平均值作为最优参数值的确定依据。在分析不同参数对融合性能的影响时, 采用变量控制法进行研究, 即固定其他变量, 只研究一个变量对融合性能的影响, 得到该变量的最优值, 重复上述过程, 能够得到所有变量的最优值, 这也是图像融合领域在分析自由参数时常用的方法^[23]。在分析过程中, 分解尺度值 $j=1, 2, 3, 4, 5, 6, 7, 8$, 引导滤波器的窗口尺寸半径 $z=1, 2, 3, 4, 5, 6, 7, 8$, 对应的窗口大小为 $3 \times 3, 5 \times 5, 7 \times 7, 9 \times 9, 11 \times 11, 13 \times 13, 15 \times 15, 17 \times 17$, 正则化参数 $\lambda=10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}$ 。考虑到文章篇幅的限制, 这里只给出每种参数的一组分析结果, 此时其他 2 种参数被设置为最优值。图 5 所示为测试图像集合, 图 6 为不同参数下本文融合方法的融合性能。

图 6(a) 为分解尺度参数 j 的影响, 图中 5 种客观评价指标值随着 j 值的增加而增长, 在 $j \geq 5$ 时, 又开始逐渐下降, 这主要是因为值过小, 不同尺度上的



图 5 测试图像集合

Fig. 5 Test image set

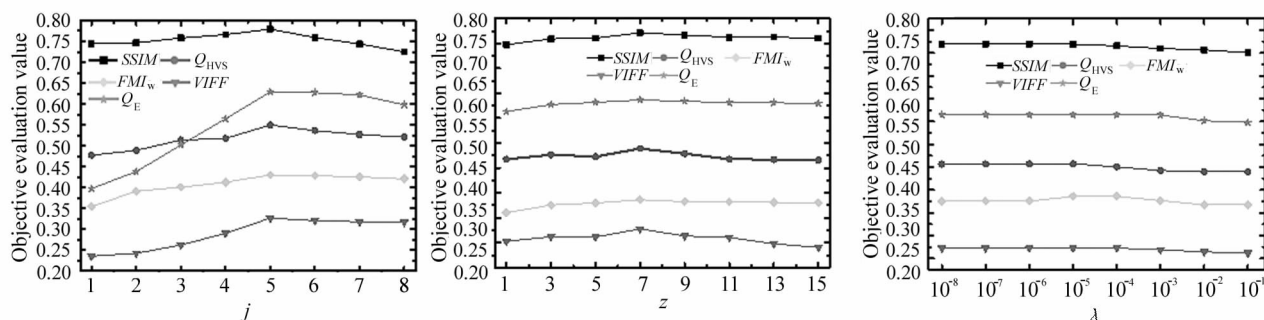


图 6 不同参数下本文融合方法的融合性能:(a) 参数 j ; (b) 参数 z ; (c) 参数

Fig. 6 Performance of the proposed method with different parameter:(a) Parameter j ; (b) Parameter z ; (c) Parameter λ

显著特征不能有效地分离出来, j 值过大, 近似图像中包含的结构细节信息过少, 利用源图像显著特征图构造的权重图与近似图像不能作最优匹配, 所以 j 值取 5 为最合适。图 6(b) 为引导滤波器窗口半径参数 z 的影响, 图中 5 种客观评价指标值在 $z \geq 7$ 时均呈现逐渐增长趋势, 在 $z > 7$ 时, $VIFF$ 和 Q_{HVS} 2 种评价指标值为逐渐减小趋势, $SSIM$ 、 FMI_w 和 Q_E 3 种评价指标值基本保持不变, 这主要是由于 z 较小时, 融合算法对噪声过于敏感, 但能保留更多的细节信息, z 较大时, 算法的鲁棒性较强, 但会平滑更多的边缘细节信息, 所以 z 值取 7 为最合适。图 6(c) 为引导滤波器正则化参数的影响, 图中 5 种客观评价指标值变化较小, 只有 Q_{HVS} 和 FIM_w 值在 $\lambda \geq 10^{-4}$ 时, 下降较为明显, 所以 λ 值取 10^{-4} 较为合适, 基于上述讨论, 最后的实验参数设定为 $j = 5$ 、 $z = 7$ 和 $\lambda = 10^{-4}$ 。

5.2 不同融合方法的对比与分析

为了验证本文所提出的融合方法的有效性, 将本文融合方法 (记为 LeResNet) 与 7 种典型融合方法进行对比实验。对比的融合方法包括基于多尺度滚动滤波器 (multi-scale rolling guidance filter, MRGF) 图像融合方法^[24]、基于带有非凸惩罚的稀疏正则化 (sparse regularization with non-convex penalties, SRNCP) 图像融合方法^[25]、基于红外特征提取与视觉信息保留 (infrared feature extraction and visual information preservation, IFEVIP) 的图像融合方法^[26]、基于迭代联合双边滤波器的加权 (iterative joint bilateral filter, IJBF) 融合方法^[27]、基于孪生卷积神经网络 (siamese convolutional neural network, SCNN) 的融合方法^[14]、基于 VGG19 的深度学习框架的融合方法^[15] 和基于卷积稀疏表示 (convolutional sparse representation, ConvSR) 的融合方法^[16]。其中 MRGF、SRNCP 和 IJBF 为最新的 3 种融合方法, SCNN、IFEVIP、VGG19 和 ConvSR 为 4 种标杆式的融合方法, 经常被其他文献引用。图 7 和图 8 分别为图像 “Kaptein_1123” 和图像 “Soldier_behind_smoke”, 图 9 和图 10 为 8 种不同融合方法对图像 “Kaptein_1123” 和图像 “Soldier_behind_smoke” 的融合结果, 表 1 为客观评价指标数据, 为了方便观察比较, 融合图像的局部区域用红色方框作了标记。

图 7(a) 和 7(b) 所示为 “Kaptein_1123” 红外与可见光图像, 场景中主要包括行人、甬路及两旁的灌木丛、未点亮的路灯、门廊及灯光、门廊后的松树、甬路旁的树干和屋顶的烟囱等目标和背景信息。图 9 所

示的 8 种融合方法都能够较好地提取红外目标信息, 如行人和烟囱等, 但 MRGF 和 IJBF 结果中出现了明显的可见光背景信息缺失和空间错乱现象, 视觉效果极差, 局部区域灰度扭曲, 容易误导观察者, 如图 9(a)、(d) 所示中门廊后侧的松树及天空、门廊前侧的甬路等。SRNCP 和 SCNN 结果整体亮度较好, 目标较清晰全面, 但局部区域融入了过多的红外信息, 丢失了许多细节信息, 如图 9(b)、(e) 所示中甬路上的方砖、灌木丛等。IFEVIP、VGG19 和 ConvSR 结果整体亮度偏低, 在松树周围出现了 “晕” 现象, 不利于视觉观察, 如图 9(c)、(f)、(g) 所示中天空和甬路部分, 其中 VGG19 和 ConvSR 结果中能量损失较严重, 视觉效果较差。本文提出的融合方法能够很好地提取红外热目标, 融入合理比例的可见光背景细节信息, 结果图像清晰自然, 有较高的亮度和对比度, 视觉感受良好, 如图 9(h) 中的行人、烟囱、松树和甬路等信息。



图 7 图像 “Kaptein_1123”:

(a) 红外图像; (b) 可见光图像

Fig. 7 Image “Kaptein_1123”:

(a) Infrared image; (b) Visible image

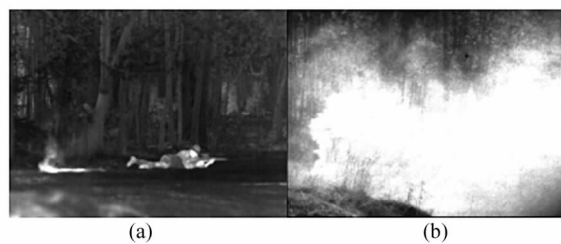


图 8 图像 “Soldier_behind_smoke”:

(a) 红外图像; (b) 可见光图像

Fig. 8 Image “Soldier_behind_smoke”:

(a) Infrared image; (b) Visible image

图 8(a) 和 8(b) 所示为 “Soldier_behind_smoke” 红外与可见光图像, 场景中主要包括卧在地上拿着枪的士兵、树木、荒草和烟雾等目标和背景信息。图 10 所示为 8 种融合方法得到的融合结果, MRGF、SRNCP、IFEVIP 和 SCNN 结果中都没有准确地提取士兵目标信息, MRGF 和 SCNN 结果中士兵细节

信息丢失严重,只表现出一个士兵形状的白色亮斑,如图 10(a)、(e)所示,IFEVIP 结果中几乎没有融入任何有效的红外图像信息,如图 10(c)所示,SRNCP 结果中树木和士兵信息缺失严重,整体亮度偏低,丢失了许多能量,不便于视觉观察,如图 10(b)所示,IJBF 结果中,士兵信息较完整,但烟雾和士兵的相对位置信息表达不明确,容易误导观察者,如图 10(d)所示;VGG19 和 ConvSR 的融合结果中,整体亮度改善较明显,基本包含了主要目标信息,但烟雾和荒草

的细节信息相对较少,如图 10(f)、(g)所示。本文提出的融合方法能够合理地分配红外与可见光图像信息的比例,融合结果图像亮度和对比度适中,目标和背景位置信息表达明确,视觉感受良好,如图 10(h)所示。表 1 中不同融合方法融合结果的客观评价数据基本与上述视觉主观评价一致,数值的大小反映了图像质量的差别。数值越大,表明融合方法的性能越好,结果图像中目标与细节信息越丰富,视觉感知越自然,反之,越差。

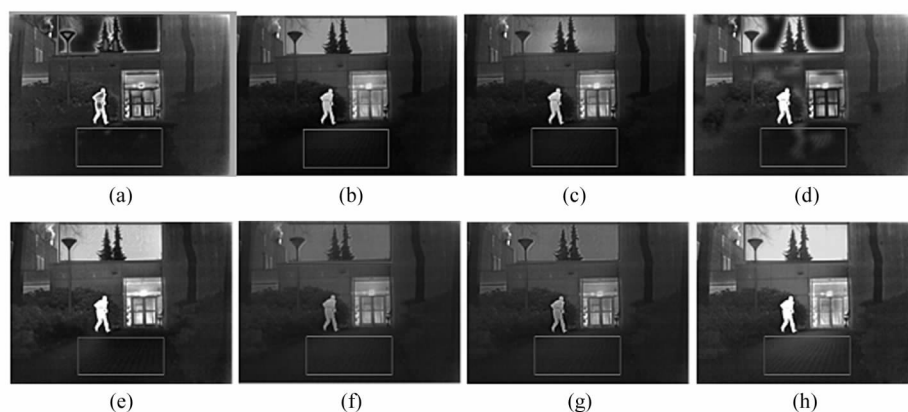


图 9 不同融合方法对图像“Kaptein_1123”的融合结果:(a) MRGF 方法;(b) SRNCP 方法;(c) IFEVIP 方法;(d) IJBF 方法;(e) SCNN 方法;(f) VGG19 方法;(g) ConSR 方法;(h) LeResNet 方法

Fig. 9 Fusion results based on different fusion methods on “Kaptein_1123” image: (a) MRGF method; (b) SRNCP method; (c) IFEVIP method; (d) IJBFF method; (e) SCNN method; (f) VGG19 method; (g) ConSR method; (h) LeResNet method

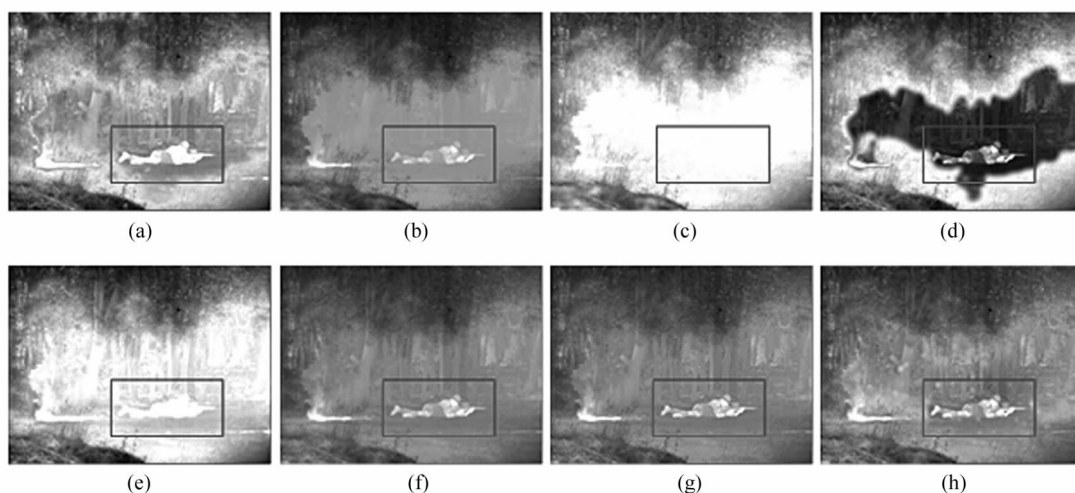


图 10 不同融合方法对图像“Soldier_behind_smoke”的融合结果:
(a) MRGF 方法;(b) SRNCP 方法;(c) IFEVIP 方法;

(d) IJBF 方法;(e) SCNN 方法;(f) VGG19 方法;(g) ConvSR 方法;(h) LeResNet 方法

Fig. 10 Fusion results based on different fusion methods on “Soldier_behind_smoke” image:

(a) MRGF method; (b) SRNCP method; (c) IFEVIP method; (d) IJBF method;

(e) SCNN method; (f) VGG19 method; (g) ConvSR method; (h) LeResNet method

表 1 不同融合方法融合结果的客观评价

Tab. 1 Objective evaluation of fusion results based on different fusion methods

Image	Metric	MRGF	SRNCP	IFEVIP	IJBF	SCNN	VGG19	ConvSR	LeResNet
Kaptein_1123	<i>SSIM</i>	0.672 53	0.711 42	0.725 19	0.665 75	0.723 03	0.707 47	0.714 15	0.73 57
	Q_{HVS}	0.446 25	0.490 30	0.475 07	0.466 72	0.501 53	0.503 18	0.469 61	0.518 19
	FMI_w	0.392 65	0.404 58	0.444 17	0.410 37	0.391 84	0.403 10	0.378 91	0.465 25
	Q_E	0.705 01	0.616 58	0.595 83	0.710 68	0.717 08	0.447 87	0.634 32	0.728 51
	<i>VIFF</i>	0.253 07	0.421 02	0.393 57	0.271 30	0.429 68	0.301 57	0.307 53	0.446 10
Soldier_behind_smoke	<i>SSIM</i>	0.582 40	0.680 93	0.532 18	0.549 77	0.617 43	0.624 09	0.635 42	0.695 42
	Q_{HVS}	0.447 82	0.352 91	0.316 41	0.403 41	0.383 29	0.384 20	0.387 33	0.457 33
	FMI_w	0.485 10	0.404 92	0.398 58	0.522 73	0.432 14	0.433 38	0.407 47	0.537 47
	Q_E	0.571 82	0.473 92	0.381 95	0.533 04	0.465 14	0.455 28	0.517 95	0.597 95
	<i>VIFF</i>	0.278 20	0.220 32	0.116 30	0.269 37	0.353 99	0.300 47	0.286 45	0.386 45

从上述分析可知,本文融合方法在主观和客观上均取得了最优的融合结果,其他 7 种方法得到的融合结果从视觉感知和客观评价数据上都存在一些明显的缺陷,这主要是因为多尺度局部极值分解能够很好的将源图像中重要的尺度特征信息分离,在近似图像和细节图像中,ResNet152 网络能够深度地提取源图像的多层特征,利用生成的显著特征图作为近似图像融合的权重,能够准确地将显著的红外与可见光信息转移到融合图像中,保持源图像中的能量不损失,在细节图像融合过程中,采用引导滤波器对初始决策图进行优化,得到边缘保留和纹理信息增强的决策图,能量显著图作为权重对细节图像进行加权平均,充分利用邻域像素间的强相关性最优地分配权重,保留更多的纹理信息和边缘细节,防

止局部区域亮度过高,造成过度曝光效果。

为了说明本文融合方法具有一定的稳定性,对来自 TNO_Image_Fusion_Dataset 数据集中图 11 所示 6 对红外与可见光图像进行对比融合实验,不同融合方法对 6 对红外与可见光图像融合的结果如图 12 所示。不同融合方法融合结果的平均定量客观评价如表 2 所示,表中数据为各评价指标的平均值,图 13 所示方为不同融合方法的 5 种客观评价指标对比曲线图。

从图 12、表 2 和图 13 中可得出,本文融合方法得到结果图像在主观视觉感知和客观评价数据方面均优于其他 7 种方法,这说明本文融合方法能够稳定地得到高质量的融合结果。



图 11 6 对红外与可见光图像

Fig. 11 Six pairs of infrared and visible images



图 12 不同融合方法对 6 对红外与可见光图像融合的结果:(a) MRGF 方法;(b) SRNCP 方法;(c) IFEVIP 方法;
(d) IJBF 方法;(e) SCNN 方法;(f) VGG19 方法;(g) ConvSR 方法;(h) LeResNet 方法

Fig. 12 Fusion results based on different fusion methods on six pairs of infrared and visible images:(a) MRGF method;
(b) SRNCP method;(c) IFEVIP method;(d) IJBF method;(e) SCNN method;
(f) VGG19 method;(g) ConvSR method;(h) LeResNet method

表 2 不同融合方法融合结果的平均定量客观评价

Tab. 2 Average quantitative evaluation of fusion results based on different fusion methods

Metric	MRGF	SRNCP	IFEVIP	IJBF	SCNN	VGG19	ConvSR	LeResNet
SSIM	0.586 53	0.621 02	0.626 08	0.581 46	0.631 21	0.593 26	0.616 36	0.641 53
Q_{HVS}	0.512 42	0.441 51	0.468 30	0.504 83	0.548 34	0.443 153	0.442 25	0.550 42
FMI_w	0.412 75	0.409 54	0.425 06	0.414 38	0.424 66	0.434 386	0.407 39	0.449 30
Q_E	0.713 88	0.586 01	0.558 53	0.693 52	0.722 62	0.441 08	0.629 44	0.747 38
VIFF	0.374 46	0.370 29	0.392 52	0.366 09	0.502 26	0.275 74	0.284 28	0.523 48

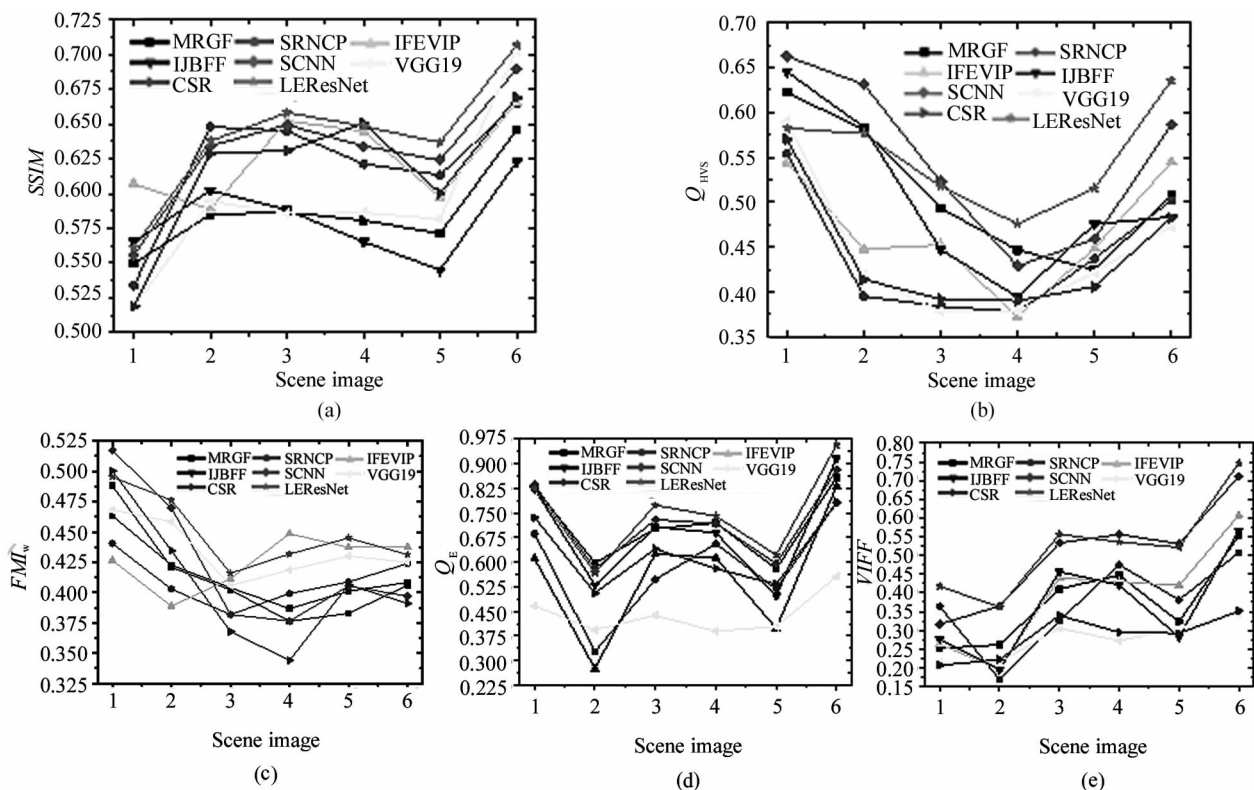


图 13 不同融合方法的 5 种客观评价指标对比曲线图: (a) SSIM 值; (b) Q_{IHS} 值; (c) FMI_w 值; (d) QE 值; (e) VIFF 值

Fig. 13 Five objective evaluation comparison curves of different fusion methods: (a) SSIM value; (b) Q_{IHS} value; (c) FMI_w value; (d) QE value; (e) VIFF value

6 结 论

本文提出了一种基于多尺度局部极值分解与 ResNet152 的红外与可见光图像融合方法,多尺度局部极值分解能够将源图像中不同尺度的重要特征信息分离出来,得到一个近似图像和多个细节图像。为了消除传统加权平均融合规则不能充分利用近似图像信息的缺陷,本文提出了一种基于显著特征图加权的近似图像融合规则,利用残差网络 ResNet152 深度提取源图像的显著特征,通过计算 l_1 范数和局部窗口均值构造显著特征图,最优地分配近似图像各系数的融合比例,保证融合后的图像具有良好的整体外观和对比度,为了克服“系数绝对值取大”规则对噪声敏感的缺陷,在细节图像融合过程中引入引导滤波器,提出了一种能量显著图的加权融合规则,更好地将源图像中显著细节信息转移到融合图像中,削弱噪声的影响。实验结果表明,本文提出的融合方法在能量保持、整体亮度和局部对比度等方面均能得到稳定的最优结果,在主观视觉评价和客观指标评价方面均优于已有文献中具有代表性的融合方法。同时本文所用的多尺度分解方法可用于其

他图像处理领域,如目标识别和特征提取等,融合规则可以用于其他多尺度融合方法。

参考文献:

- [1] CHEN G Q, CHEN Y C, LI J Y, 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.
陈广秋, 陈昱存, 李佳悦, 等. 基于 DNST 和卷积稀疏表示的红外与可见光图像融合 [J]. 吉林大学学报(工学版), 2021, 51(3): 996-1010.
- [2] KAUR A, SANDHYA S, GANJOO P, et al. A REVIEW: image fusion using DCT and DWT [J]. International Journal of Scientific and Engineering Research, 2019, 10(2): 702-707.
- [3] SAI W, ZHANG S W, HU Y P. A remote sensing image fusion algorithm based on second generation curvelet transform and signification judgment mechanism [J]. Optical Technique, 2020, 46(6): 721-727.
赛伟, 张善文, 胡予璞. 基于二代 Curvelet 变换耦合显著判定机制的遥感图像融合算法 [J]. 光学技术, 2020, 46(6): 721-727.
- [4] DAI W Z, WANG Q. Research on fusion method of visible and infrared image based on PCNN and IFS [J]. Journal of Optoelectronic · Laser, 2020, 31(7): 738-744.
戴文战, 王琪. 基于 PCNN 与 IFS 的可见光与红外图像

- 融合方法[J], 光电子·激光, 2020, 31(7):738-744.
- [5] WANG Z, DU Q Z, DONG A Y, et al. Infrared and visible image fusion based on nonsubsampling shearlet transform[J]. Journal of Optoelectronic · Laser, 2020, 31(10):1062-1073.
王昭, 杜庆治, 董安勇, 等. 基于非下采样剪切波变换的红外与可见光图像融合[J]. 光电子·激光, 2020, 31(10):1062-1073.
- [6] LIU Z J, JIA P, XIA Y H, et al. Development and performance evaluation of infrared and visual image fusion technology[J]. Laser & Infrared, 2019, 45(5):633-640.
刘智嘉, 贾鹏, 夏寅辉, 等. 基于红外与可见光图像融合技术发展性能评价[J]. 激光与红外, 2019, 45(5):633-640.
- [7] DAS A, SHYLAJA S S. Video deraining for mutual motion by fast bilateral filtering on spatiotemporal features[J]. International Journal of Innovative Technology and Exploring Engineering, 2020, 9(3):1772-1782.
- [8] LIU X X, YAN C W, JIA X F, et al. Stereo matching based on secondary weighted guided filtering fusion of double color models[J]. Journal of Optoelectronic · Laser, 2020, 31(8):842-849.
刘徐秀, 言传为, 贾晓芬, 等. 二次加权引导滤波融合双色彩空间模型的立体匹配[J]. 光电子·激光, 2020, 31(8):842-849.
- [9] CHEN L G, FU G. Structure-preserving image smoothing with semantic cues[J]. The Visual Computer, 2020, 36(10):2017-2027.
- [10] YU Q Z, ZHOU S, JIANG Y X, et al. High-performance SAR image matching using improved SIFT framework based on rolling guidance filter and ROEWA-powered feature[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019, 12(3):920-933.
- [11] HERMESSI H, MOURALI O, ZAGROUBA E. Multimodal medical image fusion review: Theoretical background and recent advances[J]. Signal Processing, 2021, 183:108036.
- [12] HE G Q, XING S Y, HE X J, et al. Image fusion method based on simultaneous sparse representation with non-subsampling contourlet transform[J]. IET Computer Vision, 2019, 13(2):240-248.
- [13] SUN C Q, ZHANG C, XIONG N X. Infrared and visible image fusion techniques based on deep learning: a review[J]. Electronics, 2020, 9(12):1-24.
- [14] 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):1850018-1-1850018-20.
- [15] LI H, WU X J, KITTLER J. Infrared and visible image fusion using a deep learning framework[C]//24th International Conference on Pattern Recognition, Aug. 20-24, 2018, Beijing, China. New York: IEEE, 2018, 2705-2710.
- [16] 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.
- [17] DING Z S, ZHOU D M, NIE R C, et al. Medical image fusion combining multi-scale local extrema decomposition and SCM model[J]. Computer Applications and Software, 2020, 37(7):142-147.
丁斋生, 周冬明, 聂仁灿, 等. 结合多尺度局部极值分解和SCM模型的医学图像融合方法[J]. 计算机应用与软件, 2020, 37(7):142-147.
- [18] LEVIN A, LISCHINSKI D, WEISS Y. Colorization using optimization[J]. ACM Transactions on Graphics, 2004, 23(3):689-694.
- [19] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, Jun. 27-30, 2016, Las Vegas, Nevada, USA. New York: IEEE, 2016, 770-778.
- [20] HE K M, ZHANG X Y, REN S Q, et al. Identity mappings in deep residual networks[C]//14th European Conference on Computer Vision, Oct. 11-14, 2016, Amsterdam, The Netherlands. Berlin: Springer, 2016, 630-645.
- [21] YIN M, LIU X N, LIU Y. Medical image fusion with parameter-adaptive pulse coupled neural network in non-subsampling shearlet transform domain[J]. IEEE Transactions on Instrumentation and Measurement, 2019, 68(1):49-64.
- [22] YANG Y C, LI J, WANG Y P. Review of image fusion quality evaluation methods[J]. Journal of Frontiers of Computer Science and Technology, 2018, 12(7):1021-1035.
杨艳春, 李娇, 王阳萍. 图像融合质量评价方法研究综述[J]. 计算机科学与探索, 2018, 12(7):1021-1035.
- [23] DURGA P B, RAVINDRA D. Two-scale image fusion of visible and infrared images using saliency detection[J]. Infrared Physics & Technology, 2016, 76(5):52-64.
- [24] JIAN L, YANG X, ZHOU Z, et al. Multi-scale image fusion through rolling guidance filter[J]. Future Generation Computer Systems, 2018, 83(6):310-325.
- [25] ANANTRASIRICHAI N, ZHENG R C, SELESNICK I, et al. Image fusion via sparse regularization with non-convex penalties[J]. Pattern Recognition Letters, 2020, 131(3):355-360.
- [26] ZHANG Y, ZHANG L J, BAI X Z, et al. Infrared and visible image fusion through infrared feature extraction and visual information preservation[J]. Infrared Physics & Technology, 2017, 83(6):227-237.
- [27] ZHAN K, KONG L W, LIU B, et al. Multimodal image seamless fusion[J]. Journal of Electronic Imaging, 2019, 28(2):023027-1-023027-9.

作者简介:

陈广秋 (1977—), 男, 博士, 副教授, 硕士生导师, 研究方向为图像处理与机器视觉。