文章编号:2096-1472(2023)11-0052-07 DOI:10.19644/j.cnki.issn2096-1472.2023.011.012

基于深度特征提取的红外与可见光图像融合

张 薇[1],汪雷丹[1],刘晓亮[1],李坤煌[2]

(1.广东粤港供水有限公司,广东 深圳 518021;
2.深圳高度创新技术有限公司,广东 深圳 518000)

✉ 506303235@qq.com; wangld@163.com; liuxliang@163.com; kunhli@163.com

摘 要: 针对红外与可见光图像融合算法中细节模糊、对比度低等问题,提出一种基于深度特征提取的融合算法。该算法模型具有较好的泛化性,通过迁移学习可应用于水利工程无人机自动巡检系统的安全监测。首先,利用滚动引导滤波对源图像进行多级分解。其次,针对基础层引入PCANet网络指导融合,针对细节层引入显著测量方法和引导滤波指导细节层融合。最后,将各融合部分进行叠加重构,得到融合图像。实验结果表明,该算法在主观视觉上具有细节信息丰富、高对比度的特点,在客观评价指标上也有较好的结果,尤其是标准差指标,较以往大部分方法领先幅度超过40%。

关键词: 红外与可见光图像;滚动引导滤波;PCANet;无人机;特征提取

中图分类号: TP391 **文献标志码:** A

Infrared and Visible Image Fusion Based on Depth Feature Extraction

ZHANG Wei[1], WANG Leidan[1], LIU Xiaoliang[1], LI Kunhuang[2]

(1.*Guangdong Yue Gang Water Supply Co., Ltd., Shenzhen 518021, China*;
2.*Shenzhen Highly Innovative Technology Co., Ltd., Shenzhen 518000, China*)

✉ 506303235@qq.com; wangld@163.com; liuxliang@163.com; kunhli@163.com

Abstract: Aiming at the problems of blurred details and low contrast in the fusion algorithm of infrared and visible images, a fusion algorithm based on deep feature extraction is proposed. The algorithm model has good generalization and can be applied to the safety monitoring of water conservancy engineering UAV (Unmanned Aerial Vehicle) automatic inspection system through transfer learning. Firstly, the source image is decomposed at multiple levels using rolling guided filtering. Then, PCANet network is introduced to guide fusion in the basic layer, and significant measurements and guided filtering are introduced to guide fusion in the detail layer. Finally, the fusion parts are overlaid and reconstructed to obtain the fused image. The experimental results show that the algorithm has rich detail information and high contrast in subjective vision, and also has better results in objective evaluation index, especially on the standard deviation indicator, which is more than 40% ahead of most previous methods.

Key words: infrared and visible image; rolling guided filtering; PCANet; UAV; feature extraction

0 引言(Introduction)

水利工程建设旨在加强水资源保护,为人们的安全用水和社会的可持续发展提供保障。东江-深圳供水工程(简称“东深供水工程”)是党中央为解决香港地区同胞饮水困难而兴建的跨流域大型调水工程。在工程建设过程中,需要管理的设备多、涉及范围广,为保证工程的顺利实施,需要对这些设备进行安全监管。随着无人机领域的讯速发展,配备无人机自动巡检系统的无人机飞行器可自动进行巡查工作,通过配套的AI(人工智能)项目,可以实现图像的目标识别,如识别水面悬浮物等,但是无人机拍摄的照片会受到环境等因素的干扰,影响对

收稿日期:2023-03-10

目标物的准确识别。图像融合是图像处理和计算机视觉应用中的一项重要技术,已在目标识别、视频监督和图像增强等方面有较好的应用[1-3]。本文提出基于深度特征提取的红外与可见光图像融合算法。首先,引入滚动引导滤波对源红外与可见光图像进行多级分解,以提取更多的细节边缘信息。其次,针对基础层引入 PCANet 网络进行特征提取,指导基础层融合;针对细节层采用拉普拉斯滤波、高斯滤波和引导滤波指导细节层融合。最后,将融合之后的基础层与细节层叠加重构,得到最后的融合结果。该算法模型具有较好的泛化性,通过迁移学习可应用于水利工程无人机自动巡检系统的安全监测。

1　相关研究(Related research)

通过图像融合,同一场景的不同模态图像可以合并为单一的融合图像。融合后的图像可以提供更全面的场景信息,更有利于人类和机器的感知。基于多尺度变换(Multi-Scale Transformation,MST)的融合算法是目前最主要的红外与可见光图像融合算法[4]。比较常见的该类型算法有基于曲波变换(Curvelet Transform,CVT)算法[5]、基于双树复小波变换(Dual-Tree Complex Wavelet Transform, DTCWT)算法[6]、基于Tetrolet 变换算法[7]及基于非下采样轮廓波变换(Non-Sampled Contourlet Transform,NSCT)算法[8]等。由于上述算法都是利用预先定义的基函数提取图像特征,所以对源图像的细节纹理信息等重要特征的提取效果不好,并且 MST 算法是将源图像投影到频率域,增大了计算复杂性。JIAN 等[9]将滚动引导滤波器(Rolling Guided Filtering,RGF)[10]用于图像融合,充分利用该滤波器平滑小尺度结构和大尺度结构边缘恢复特性分解源图像,很好地保留了源图像细节边缘信息。基于深度神经网络(Deep Neural Network,DNN)的算法在红外与可见光图像融合方面具有很强的特征提取能力。ZHOU 等[11]使用预训练网络 VGG-19 提取源图像的特征信息,从而使得融合图像保留更多的细节特征和结构信息。MA 等[12]将生成式对抗网络(Generative Adversarial Network,GAN)用于红外与可见光图像融合,用生成网络生成融合图像,而对抗网络使融合图像具有更多的细节信息。CHAN[13]等提出一种更简单、高效的深度学习网络——PCANet,该网络在图像处理方面具有很好的特征信息提取性能[14]。

2　分解方法(Decomposition method)

RGF 同时具有尺度感知和边缘保护的特性,并且其迭代收敛速度比较快,用其对源图像进行分解,能够有效保留目标的细节边缘信息[10]。该滤波器包括两个主要步骤:平滑小尺度结构和大尺度结构边缘恢复。

第一步使用高斯滤波器获取原始引导图像,可以有效去除源图像 I 中的小尺度结构,源图像 I 在中心像素 a 处的高斯滤波可表示如下:

$$G(a)=\frac{1}{\sum_{b\in N}\exp\left(-\frac{a-b}{2\sigma_s^2}\right)}\sum_{b\in N}\exp\left(-\frac{\|a-b\|^2}{2{\sigma_s}^2}\right)I(b) \tag{1}$$

其中:a 和 b 分别表示中心像素和相邻像素,N 表示中心像素 a 处相邻像素 b 的集合。根据尺度空间理论,结构尺度参数 σ_s 可以定义为高斯核的最小标准偏差,G 表示原始引导图像。

第二步使用联合双边滤波器进行迭代式边缘恢复,该滤波器具有计算高效和边缘保留较好等特性。这一步是一个迭代过程,恢复的图像 J^t 被迭代更新,初始图像 J^1 是高斯平滑的图像 G。第 t 次迭代可表示如下:

$$J^{t+1}(a)=\frac{1}{M}\sum_{b\in N}\exp\left[-\frac{a-b^2}{2\sigma_s^2}-\frac{J^t(a)-(J^t(b))^2}{2\sigma_r^2}\right]I(b) \tag{2}$$

其中:$M=\sum_{b\in N}\left(-\frac{\|a-b\|}{2\sigma_s^2}\right)$用于归一化,$\sigma_s$ 和σ_r 分别控制空间范围权重和强度差范围权重,J^{t+1} 表示第 t 次迭代的结果,默认参数 $t=4$。

迭代操作在保留大尺度结构边缘的同时,能很好地平滑不相关的细节。滚动引导滤波操作可定义如下:

$$U=RGF(I,\sigma_s,\sigma_r) \tag{3}$$

其中:$RGF(\cdot)$ 表示滚动引导滤波函数,通过改变参数 σ_s 和σ_r,可以对源图像进行平滑处理,实现一系列不同尺度的分层,U 是滚动滤波器的输出图像,I 是输入图像。

本文引入 RGF 将源红外与可见光图像分别分解为基础层和多个细节层,与其他多尺度分解方法相比,该方法能够最大限度地利用源图像中的相关细节信息。假设有两张已经预配准源红外与可见光图像,其表示为 I_k,$k\in\{1,2\}$。对于每一张源图像 I_k,基础层可表示如下:

$$I_k^{b,i}=RGF(I_k,\sigma_s^i,\sigma_r^i),i=1,2,\cdots,N-1 \tag{4}$$

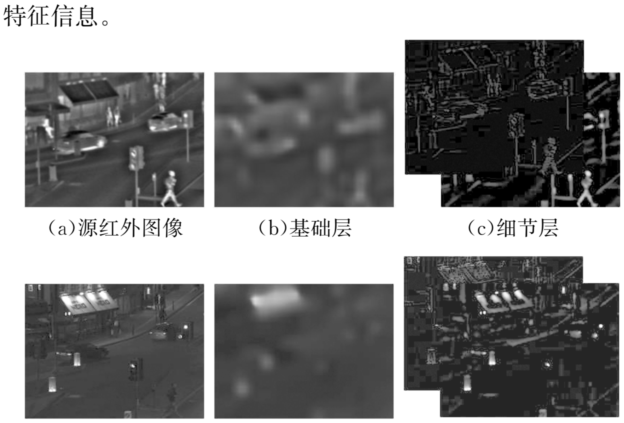
其中:$I_k^{b,i}$ 是对应第 i 层源图像的基础层,I_k 表示源图像,$RGF(\cdot)$表示滚动引导滤波函数,σ_s^i 表示第 i 层空间范围权重,σ_r^i 表示第 i 层强度差范围权重。

在给定相同 $\sigma_r(\sigma_r^{i+1}=\sigma_r^i)$条件下,参数 σ_s 决定了图像模糊程度,例如假设 $\sigma_s^{i+1}>\sigma_s^i$,那么平滑程度 $I_k^{b,i+1}>I_k^{b,i}$。因此,$I_k^{b,i}$ 比 $I_k^{b,i+1}$ 包含更多的结构信息。令 $I_k^{b,0}=I_k$,可以通过不等式关系得到细节层,其表达如下:

$$I_k^{d,i}=I_k^{b,i-1}-I_k^{b,i},i=1,2,\cdots,N-1 \tag{5}$$

$I_k^{d,i}$ 是第 i 层源图像的细节层;$I_k^{b,N-1}$ 是源图像最平滑图像,作为源图像的基础层。

源图像采用 RGF 的分解结果如图 1 所示。其中,I_1 是源红外图像,$I_1^{b,i}$ 是 I_1 经 RGF 分解后得到的基础层;$I_1^{d,i}$ 是经 RGF 分解后得到的细节层。I_2 是源可见光图像,同理可得,$I_2^{b,i}$ 和 $I_2^{d,i}$ 分别是 I_2 经 RGF 分解得到的基础层和细节层。从图 1 中可以看出,基础层 $I_1^{b,i}$ 和 $I_2^{b,i}$ 含有源图像的亮度和轮廓信息,而细节层 $I_1^{d,i}$ 和 $I_2^{d,i}$ 含有源图像的边缘信息和细节显著特征信息。



(a)源红外图像　(b)基础层　(c)细节层

(d)源可见光图像　(e)基础层　(f)细节层

图 1　RGF 分解结果

Fig. 1 RGF decomposition results

3 融合规则(Fusion rule)

3.1 基于PCANet特征提取的基础层融合

基础层中还残余有用的低频信息,卷积神经网络能够有效提取图像中的信息,但考虑到卷积神经网络训练时间较长、调参比较复杂等,CHAN等[13]提出了一种简单、高效的能够适应不同任务和数据的深度学习网络PCANet。该网络采用主成分分析的方法学习多级滤波器组,不需要调节复杂参数和解决复杂数值优化问题,因此其网络结构简单且性能高效。PCANet框架图如图2所示。

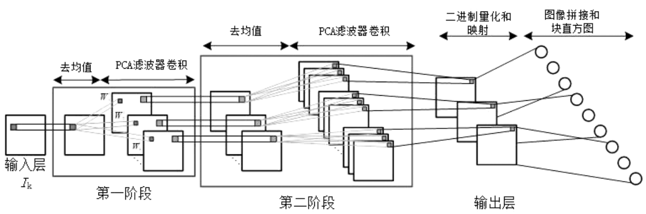


图2 PCANet框架图

Fig. 2 PCANet framework chart

对于每张 $M\times N$ 大小的输入图像,PCANet通过一个 $K_1\times K_2$ 大小的滑动窗口(一般选取窗口大小为3、5、7个像素的正方形)获取输入图像的局部特征,并将获取的这些块组合在一起,即 $\boldsymbol{x}_{i,1},\boldsymbol{x}_{i,2},\cdots,\boldsymbol{x}_{i,MN}\in\mathbb{R}^{K_1K_2}$,其中 $\boldsymbol{x}_{i,j}$ 是输入图像 I_i 里面的第 j 个向量块。将获取的每个块减去其平均值,实现去均值操作,处理之后得到 $\boldsymbol{X}_i=[\bar{\boldsymbol{x}}_{i,1},\bar{\boldsymbol{x}}_{i,2},\cdots,\bar{\boldsymbol{x}}_{i,MN}]$,其中 $\bar{\boldsymbol{x}}_{i,j}$ 是输入图像 I_i 中的第 j 个去均值块。将全部输入源图像 $\{I_i\}_i^K$(K 表示输入源图像数量)执行同样的处理,并将得到的结果放在一起,即

$$\boldsymbol{X}=[\bar{\boldsymbol{x}}_1,\bar{\boldsymbol{x}}_2,\cdots,\bar{\boldsymbol{x}}_K]\in\mathbb{R}^{K_1K_2\times K_{MN}} \tag{6}$$

假设第 i 层使用滤波器个数为 S_i,则第一层滤波器个数为 S_1。根据主成分分析法,可得最小重构误差:

$$\min_{\boldsymbol{U}\in\mathbb{R}^{K_1K_2\times S_1}}\|\boldsymbol{X}-\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{X}\|_F^2,\ \text{s.t.}\ \boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}=\boldsymbol{I}_{S_1} \tag{7}$$

其中:$\boldsymbol{U}$ 表示 $\boldsymbol{X}$ 通过第一层滤波器 S_1 特征值所对应的特征向量组合而成的特征矩阵,$\boldsymbol{I}_{S_1}$ 表示 $S_1\times S_1$ 大小的单位矩阵。由此可得PCA滤波器如下:

$$\boldsymbol{W}_l^1=\underset{k_1,k_2}{map}(p_l(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}))\in\mathbb{R}^{K_1K_2},l=1,2,\cdots,S_1 \tag{8}$$

其中:$p_l(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}})$ 表示 $\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}$ 里面第 l 个主特征向量,$\underset{k_1,k_2}{map}(p_l(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}))$ 表示可以把 $p_l(\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}})$ 映射到 $\boldsymbol{W}\in\mathbb{R}^{K_1K_2}$ 上的函数。能够采用PCA滤波器提取图像特征,主要是由于主特征向量可以获取全部均值训练块的变化。

本研究仅采用PCA滤波器提取红外与可见光基础层图像特征,而没有采用块直方图和二进制哈希。在PCANet的两个阶段中,在第一阶段将第一层滤波器数量 S_1 设置为8,在第二阶段将第二层滤波器数量 S_2 也设置为8。因此,在第一阶段能够得到8个基础层图像特征,在第二阶段能够得到64个基础层图像特征。受数据的非冗余性和时间效率等因素的影响,本文仅使用第一阶段提取到的图像特征融合基础层图像。图3是采用PCANet对基础层图像进行特征提取,获取基础层权值图 W_k^b 的过程。

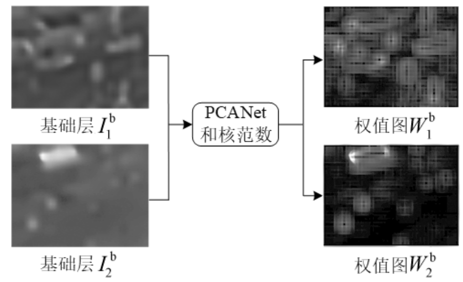


图3 PCANet处理过程

Fig. 3 Processing procedures of PCANet

3.2 基于显著图引导滤波的细节层融合

图像融合需要将两幅源图像中具有视觉意义的信息整合到一副图像中,可以通过给源图像的细节层分配适当的权重(不重要信息的像素分配低权重,重要信息的像素分配高权重)实现。人类视觉系统识别图像的细节层信息比其基础层敏感,本文采用基于显著图的权值图融合细节层。如图4所示,将拉普拉斯滤波应用于每张源图像,以获得高通图像 H_k:

$$H_k=I_k\times L \tag{9}$$

其中:L 是一个大小为 3×3 拉普拉斯滤波;H_k 绝对值的局部平均值,用于构建显著图 S_k:

$$S_k=|H_k|\times g_{r_g,\sigma_g} \tag{10}$$

其中:g 是大小为 $(2r_g+1)(2\sigma_g+1)$ 的高斯低通滤波,参数 r_g 和 σ_g 设置为5。

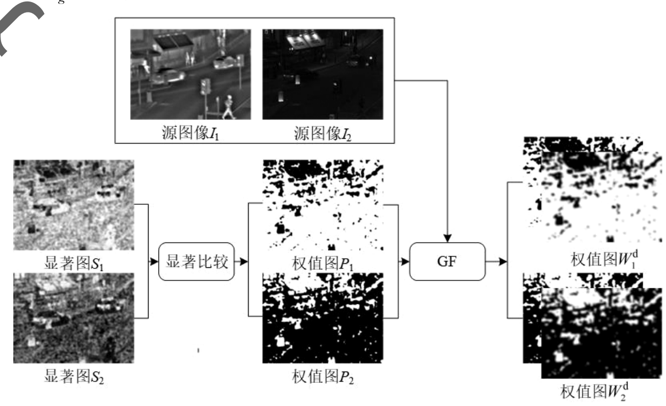


图4 GF处理结果

Fig. 4 GF processing results

显著图可以很好地提供源图像中的细节信息。接下来,比较显著图以确定权值图,如公式(11)所示:

$$P_k^i=\begin{cases}1 & \text{if } S_k^i=\max(S_1^i,S_2^i,\cdots,S_K^i)\\ 0 & \text{otherwise}\end{cases} \tag{11}$$

其中:K 是源图像的数量,S_K^i 是第 K 张图像中像素 i 的显著图值,P_k^i 是显著图的初始权值图。然而,按照上述方法获取的初始权值图通常是有噪声的,并且没有与目标边界对齐,这可能会对融合图像产生伪影。使用空间一致性[15]的方法能够有效解决上述问题。空间一致性方法表明,如果两个邻近像素具

有相似的颜色或亮度值，它们的权重也具有相似性。本文利用引导滤波可以让亮度相近的相邻像素具有相同的权重值。

假设引导滤波的输出图像 V 是引导图像 G 在以像素 p 为中心的局部窗口 w_p 中的线性变换模型，即

$$V_q = \mathrm{a_p} G_q + \mathrm{b_p}, \forall q \in w_p \tag{12}$$

其中：在窗口 w_p 内，线性系数 $\mathrm{a_p}$ 和 $\mathrm{b_p}$ 都是常数，并且可以通过最小化输出图像 V 和输入图像 I 之间的平方差进行估计：

$$E(\mathrm{a_p}, \mathrm{b_p}) = \sum_{i \in w_p} [(\mathrm{a_p} G_q + \mathrm{b_p} - I_q)^2 + \varepsilon \mathrm{a_p}^2] \tag{13}$$

其中，ε 表示正则化参数。利用线性回归[16]求解公式(13)，得到线性系数 $\mathrm{a_p}$ 和 $\mathrm{b_p}$ 的结果如下：

$$\mathrm{a_p} = \frac{\frac{1}{|w|} \sum_{q \in w_p} G_q I_q - \mu_p \bar{I}_q}{\delta_p + \varepsilon} \tag{14}$$

$$\mathrm{b_p} = \bar{I}_q - \mathrm{a_p} \mu_p \tag{15}$$

其中：δ_p 和 μ_p 分别表示在窗口 w_p 内引导图像 G 的方差和均值，$|w|$ 表示窗口 w_p 内的像素数，$\bar{I}_q$ 表示窗口 w_p 内输入图像的均值。在本文中，$GF(G, I, r, \varepsilon)$ 表示引导滤波算子，r 和 ε 分别决定了引导滤波的滤波器大小和模糊程度的参数，G 和 I 分别指的是引导图像和输入图像。根据公式(12)、公式(13)、公式(14)可以看出，如果一个位置 i 的局部方差很小，表明该像素处于引导图像的平坦区域，$\mathrm{a_p}$ 接近于 0，引导滤波输出为相邻输入像素的平均值；相反，如果像素 i 的局部方差很大，表明该像素处于引导图像的边缘区域，$\mathrm{a_p}$ 远离于 0，此时 $V_q \approx \bar{a} V_q$，意味着只有边缘一侧的权重会被平均。在这两种情况下，具有相似亮度和颜色的像素同样具有相似的权重。这与空间一致性的原理相同。

因此，为了让对应的权重图具有与源图像相似的平滑度和锐化度(图 4)，以源图像 I_k 作为引导图像，对每个权值图 P_k 进行引导滤波处理，即

$$W_k^{d_i} = GF(I_k, P_k, r_i, \varepsilon_i), i = 1, 2, \cdots, N-1 \tag{16}$$

其中：$GF(\cdot)$ 表示引导滤波函数，r_i 和 ε_i 是引导滤波的参数，$W_k^{d_i}$ 是显著图的最终权值图，P_k 是源图像 I_k 的初始权值图，I_k 是引导图像。

3.3　融合结果重构

最终目标是在主观视觉和客观效果上对融合结果进行改进。考虑到不同类型的层包含不同的特征信息，在重构阶段，针对基础层和细节层分别进行。

基础层中残余有用的低频信息控制着融合图像的整体外观和对比度。上文引入 PCANet 设置融合规则获取基础层权值图 W_k^{b}，然后指导基础层 I_k^{b} 进行融合，得到融合后的基础层 F_k^{b}，能够很好地提取基础层中残余低频信息，使融合图像具有较好对比度。基础层重构可表示如下：

$$F_{\mathrm{b}} = \sum_{n=1}^{k} W_n^{\mathrm{b}} \times I_n^{\mathrm{b}}, k = 2 \tag{17}$$

其中：F_{b} 是融合基础层，W_n^{b} 表示 PCANet 提取基础层 I_n^{b} 得到的权值图，k 表示基础层的数量。

对于细节层重构，将显著图提取得到的最终权值图 $W_k^{d_i}$ 乘以相应的细节层 $I_k^{\mathrm{d},i}$。细节层融合可表示如下：

$$F_{\mathrm{d}} = \sum_{i=1}^{N-1} (I_1^{\mathrm{d},i} W_1^{d_i} + I_2^{\mathrm{d},i} W_2^{d_i}) \tag{18}$$

将融合的基础层和细节层叠加，得到最终的融合结果：

$$F = F_{\mathrm{b}} + F_{\mathrm{d}} \tag{19}$$

4　融合算法总体框架(The overall framework of the fusion algorithm)

融合算法框架图如图 5 所示。假设红外与可见光图像已经预配准，本文主要针对两张待融合源红外与可见光图像，并且当采用多张源图像时，融合算法的策略类似。本文所提融合算法的具体步骤如下。

(1)采用 RGF 将源红外图像 I_1 和源可见光图像 I_2 分解，得到分解之后源红外图像基础层 I_1^{b} 和多个细节层 $I_1^{\mathrm{d},i}$ $(i=1, 2, 3, 4)$，以及源可见光图像基础层 I_2^{b} 和多个细节层 $I_2^{\mathrm{d},i}$ $(i=1, 2, 3, 4)$。

(2)对于基础层融合，采用 PCANet 获取基础层权值图 W_1^{b} 和 W_2^{b}，然后通过基础层权值图指导基础层 I_1^{b} 和 I_2^{b} 进行融合得到融合基础层 F_{b}。

(3)对于细节层融合，首先使用拉普拉斯滤波和高斯滤波对源图像进行显著测量，得到源图像显著图 S_1 和 S_2。其次比较显著图，得到初始权值图 P_1 和 P_2。将源红外图像 I_1 和可见光图像 I_2 作为引导图像，对初始权值图 P_1 和 P_2 进行引导滤波处理，分别得到最终权值图 $W_1^{d_i}$ $(i=1, 2, 3, 4)$ 和 $W_2^{d_i}$ $(i=1, 2, 3, 4)$。最后通过最终权值图指导细节层 $I_1^{\mathrm{d},i}$ $(i=1, 2, 3, 4)$ 进行融合，得到融合细节层 F_{d}。

(4)将融合的细节层与基础层进行叠加，获取最终的融合图像为 F。

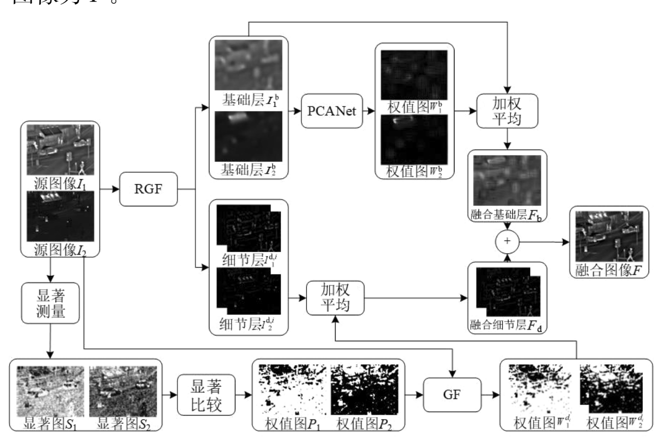


图 5　融合算法框架图

Fig. 5 The framework diagram of the proposed image fusion algorithm

5　实验结果(Experimental results)

本文采用的实验平台为 AMD(R) Ryzen(R) 5 3500X Geforce RTX 2070 SUPER 8 GB 卡、16 GB 3200 MHz 内存，在

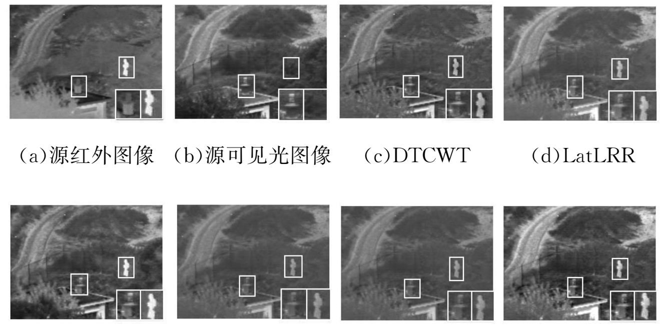
Windows 10 专业版的 64 位操作系统上用 MATLAB2020a 进行仿真实验。实验测试数据通过无人机调度平台采集数据集和在 TNO 数据集[17]上选择的三组典型红外与可见光图像，即 Nato_Camp、Bristol_Queen′s_Road 和 UN_Camp。

实验对比方法为 DTCWT[6]、LatLRR、NSCT_SR[8]、VGG-19[11]和 Resnet50[18]。DTCWT 和 LatLRR 属于多尺度分解方法；NSCT_SR 属于混合方法；VGG-19 和 Resnet50 属于深度学习网络的方法，主要利用预训练网络提取相应图像权值图，再将权值图和对应图像进行融合。以上所有对比算法采用的参数设置是基于参考文献的默认参数，本文所提算法对源图像采用 RGF 进行五级（$N=5$）分解，获得一个基础层和四个细节层，RGF 的平滑控制参数值为 $\sigma_s^i=\{648,108,18,3\}(i=1,2,3,4)$，所有的 σ_r 值设置为 0.2，GF 的参数值 $r_2=15$、$\varepsilon_2=0.1$、$r_3=5$、$\varepsilon_3=0.03$、$r_4=1$、$\varepsilon_4=0.01$。通过与目前红外与可见光图像融合的主流方法进行对比测试，能很好地验证本算法的有效性。

图 6 为第一组已经预配准 Nato_Camp 红外与可见光融合对比结果。图 6(a)和图 6(b)分别表示第一组测试源图像。图 6(c)至 6(h)分别表示不同红外与可见光融合的结果。可以看出，DTCWT 的融合方法整体对比度较差，目标物边缘比较模糊，融合质量很差；LatLRR 的融合方法中植物细节显示较差，目标物周围存在伪影；NSCT_SR 的融合方法可以很好地突出目标物，但图中植物的像素与源红外图像像素差异较大，表明该融合方法过多地将可见光信息引入融合图像，使得视觉效果变差；VGG-19、ResNet50 和本文所提方法的融合结果相对较好，并且采用本文提出的方法时图中的房屋和人物边缘信息保持较好、对比度较高且纹理细节较为清晰。

图 7 为第二组已经预配准 Bristol_Queen′s_Road 红外与可见光融合对比结果。图 7(a)和图 7(b)分别表示第二组测试源图像。图 7(c)至图 7(h)分别表示不同红外与可见光融合方法的结果。可以看出，上述融合方法基本能够保留图像中的目标人物和房屋边缘结构信息，但采用本文所提方法的图中框内的人物和房屋窗户细节信息上保留较完整，清晰度较高，具有很好的可视性。

图 8 为第三组已经预配准 UN_Camp 红外与可见光融合对比结果。图 8(a)和图 8(b)分别表示第三组测试源图像。图 8(c)至图 8(h)分别表示不同红外与可见光融合方法的结果。可以看出，DTCWT 的融合方法整体比较泛白，图中框内的天空处缺失红外图像的细节信息，屋顶边缘结构信息缺失；LatLRR 的融合方法图中框内的目标细节信息缺失，从而导致显示不够清晰，对比度较差；NSCT_SR 的融合方法图中框内的天空发生了变形，并且引入过多的伪影，清晰度非常差；VGG-19 和 ResNet50 的融合方法在结构信息上保持完整，噪声较少，但对比度较低；本文所提的融合方法可以很好地整合来自源图像的互补信息，有效保留大部分来自源图像的结构和细节信息，清晰度和对比度较高。

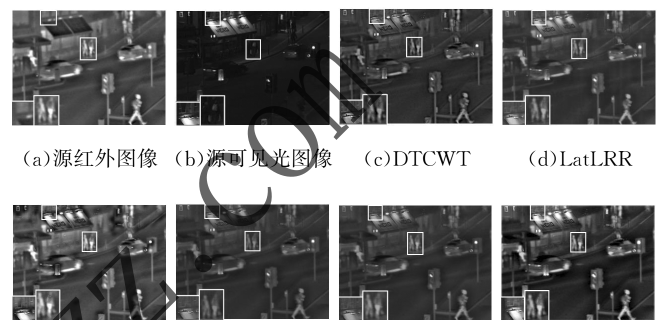


(a)源红外图像 (b)源可见光图像 (c)DTCWT (d)LatLRR

(e)NSCT_SR (f)VGG-19 (g)ResNet50 (h)本文方法

图 6 第一组图像不同融合算法主观对比结果

Fig. 6 Subjective comparison results of different fusion algorithms for the first group of images

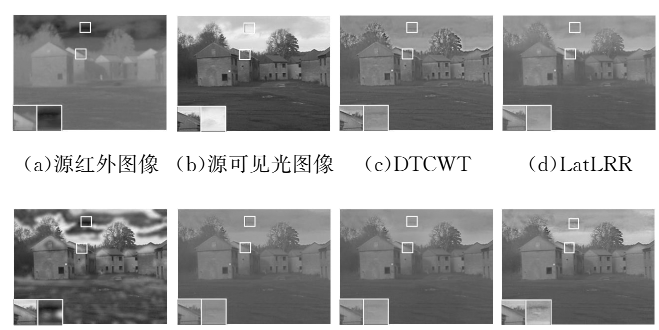


(a)源红外图像 (b)源可见光图像 (c)DTCWT (d)LatLRR

(e)NSCT_SR (f)VGG-19 (g)ResNet50 (h)本文方法

图 7 第二组图像不同融合算法主观对比结果

Fig. 7 Subjective comparison results of different fusion algorithms for the second group of images



(a)源红外图像 (b)源可见光图像 (c)DTCWT (d)LatLRR

(e)NSCT_SR (f)VGG-19 (g)ResNet50 (h)本文方法

图 8 第三组图像不同融合算法主观对比结果

Fig. 8 Subjective comparison results of different fusion algorithms for the third group of images

为了对实验结果进行定量评价，本文采用互信息（Mutual Information，MI）[19]、差异相关和（Sum of Correlations of Differences，SCD）[20]、信息熵（Entropy，EN）[21]、多尺度结构相似度测量（Multiscale Structural Similarity Measure，MS-SSIM）[22]和标准差（Standard Deviation，SD）[23]等经典评价指标对实验结果进行分析。其中，*MI* 主要通过比较融合图像与源图像之间的相关程度表示融合图像包含源图像的目标信息量，而且 *MI* 越大，表明融合结果包含目标信息量越多，其融合

效果就越好；EN 主要是基于信息论的角度评价融合图像所包含源图像的信息量，EN 越大，则表明信息量越多，融合效果越好，但该指标值会受到噪声的影响，融合图像中的噪声越多，该值也会越大；SCD 主要是通过差异相关性之和评价融合结果，同样 SCD 的值越大，表明融合的效果越好；$MS\text{-}SSIM$ 主要通过衡量结构信息评价融合结果，其值越大，表明融合图像效果越好，图像结构越接近源图像；SD 主要用来衡量融合结果对比度效果，SD 的值越大，表明融合结果对比度高和视觉更好。本文选取三组对比实验，客观评价结果如表 1 所示。

表1　不同对比融合算法的客观评价结果

Tab.1 Objective evaluation results of different comparison and fusion algorithms

图像	融合算法	*EN*	*MI*	*SD*	*MS-SSIM*	*SCD*
第一组	DTCWT	6.485	12.969	32.126	0.880	1.482
	LatLRR	6.554	13.107	36.332	0.903	1.733
	NSCT_SR	7.061	14.122	43.106	0.874	1.334
	VGG-19	6.243	12.485	27.876	0.872	1.491
	ResNet50	6.251	12.501	27.824	0.872	1.485
	本文方法	**7.200**	**14.400**	**54.415**	**0.956**	**1.992**
第二组	DTCWT	6.159	12.317	55.480	0.916	1.550
	LatLRR	6.030	12.060	60.347	0.904	1.739
	NSCT_SR	6.842	13.684	81.235	0.921	1.476
	VGG-19	5.925	11.850	46.639	0.890	1.480
	ResNet50	5.982	11.964	50.178	0.896	1.511
	本文方法	**6.847**	**13.694**	**88.221**	**0.970**	**1.958**
第三组	DTCWT	6.482	12.964	62.085	0.934	1.706
	LatLRR	6.375	12.750	62.417	0.872	1.749
	NSCT_SR	7.122	14.245	92.902	0.765	1.475
	VGG-19	6.219	12.438	56.927	0.890	1.715
	ResNet50	6.302	12.604	66.695	0.894	1.694
	本文方法	**7.165**	**14.330**	**111.261**	**0.969**	**1.999**

由表 1 可知，与其他五种方法相比，本文提出的方法在评价指标 EN、MI、$MS\text{-}SSIM$、SCD 和 SD 都表现出不同程度的领先优势，尤其评价指标 SD 和 SCD 领先较多，表明本文方法的融合结果具有高对比度，融合图像信息与源图像具有很强的相关性，证明本文所提算法具有有效性。

6　结论(Conclusion)

本文提出一种将深度特征提取与滚动引导滤波结合的多模态图像融合算法，可用于无人机自动巡检系统。通过开展实验，与目前主流的红外与可见光图像融合算法进行对比，本文所提方法在突出源图像目标物体和保留完整的边缘细节方面具有很好优势。实验结果表明，本文所提方法在信息熵、互信息、多尺度结构相似度测量、标准差和差异相关和等客观评价指标上具有明显的领先优势，主观评价具有较好的可视性，有利于后续开展目标识别和探测等任务。下一阶段的研究重点将放在分解层数的自适应选择上，为减少噪声、伪影和保留更多的有效细节信息，制定更加有效的融合规则。

参考文献(References)

[1] ZHANG X C, DEMIRIS Y. Visible and infrared image fusion using deep learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(8): 10535-10554.

[2] JIANG M X, DENG C, SHAN J S, et al. Hierarchical multi-modal fusion FCN with attention model for RGB-D tracking[J]. Information Fusion, 2019, 50: 1-8.

[3] 沈瑜，党建武，苟吉祥，等. 近红外与可见光双通道传感器信息融合的去雾技术[J]. 光谱学与光谱分析，2019，39(5)：1420-1427.

[4] LI G F, LIN Y J, QU X D. An infrared and visible image fusion method based on multi-scale transformation and norm optimization[J]. Information Fusion, 2021, 71: 109-129.

[5] NENCINI F, GARZELLI A, BARONTI S, et al. Remote sensing image fusion using the curvelet transform[J]. Information Fusion, 2007, 8(2): 143-156.

[6] LEWIS J J, O'CALLAGHAN R J, NIKOLOV S G, et al. Pixel- and region-based image fusion with complex wavelets[J]. Information Fusion, 2007, 8(2): 119-130.

[7] 冯鑫. Tetrolet 框架下红外与可见光图像融合[J]. 光子学报，2019，48(2)：76-84.

[8] UPLA K P, JOSHI M V, GAJJAR P P. An edge preserving multiresolution fusion: use of contourlet transform and MRF prior[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(6): 3210-3220.

[9] JIAN L H, YANG X M, ZHOU Z L, et al. Multi-scale image fusion through rolling guidance filter[J]. Future Generation Computer Systems, 2018, 83: 310-325.

[10] ZHANG Q, SHEN X Y, XU L, et al. Rolling guidance filter[C]//Springer. Computer Vision-ECCV 2014. Cham: Springer, 2014: 815-830.

[11] ZHOU J W, REN K, WAN M J, et al. An infrared and visible image fusion method based on VGG-19 network[J]. Optik, 2021, 248: 168084.

[12] MA J Y, YU W, LIANG P W, et al. FusionGAN: a generative adversarial network for infrared and visible image fusion[J]. Information Fusion, 2019, 48: 11-26.

[13] CHAN T H, JIA K, GAO S H, et al. PCANet: a simple deep learning baseline for image classification? [J]. IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society, 2015, 24(12): 5017-5032.

[14] LI M K, LI M, ZHANG P, et al. SAR image change detection using PCANet guided by saliency detection[J]. IEEE Geoscience and Remote Sensing Letters, 2019, 16(3): 402-406.

[15] LINDEBERG T. Scale-space theory: a basic tool for analyzing structures at different scales[J]. Journal of Applied Statistics, 1994, 21(1/2): 225-270.

[16] LI S T, KANG X D, HU J W. Image fusion with guided filtering[J]. IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society, 2013,

22(7):2864-2875.
[17] 冯鑫,方超,龚海峰,等. 二尺度分解和显著性提取的红外与可见光图像融合[J]. 光谱学与光谱分析,2023,43(2):590-596.
[18] LI H,WU X J,DURRANI T S. Infrared and visible image fusion with ResNet and zero-phase component analysis[J]. Infrared Physics & Technology,2019,102:103039.
[19] HOSSNY M,NAHAVANDI S,CREIGHTON D. Comments on “Information measure for performance of image fusion”[J]. Electronics letters,2008,44(18):1066-1067.
[20] ASLANTAS V,BENDES E. A new image quality metric for image fusion: the sum of the correlations of differences[J]. AEU - International Journal of Electronics and Communications,2015,69(12):1890-1896.
[21] ROBERTS J W,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.
[22] MA K D,ZENG K,WANG Z. Perceptual quality assessment for multi-exposure image fusion[J]. IEEE Transactions on Image Processing:a Publication of the IEEE Signal Processing Society,2015,24(11):3345-3356.
[23] RAO Y J. In-fibre Bragg grating sensors[J]. Measurement Science and Technology,1997,8(4):355-375.

作者简介:

张　薇(1989-),女,硕士,工程师。研究领域:深度学习,无人机目标识别。
汪雷丹(1992-),男,本科,助理工程师。研究领域:深度学习,智能巡检相关软件研发。
刘晓亮(1997-),男,硕士,工程师。研究领域:图像处理。
李坤煌(1999-),男,本科,工程师。研究领域:深度学习,目标识别。

(上接第24页)

4 结论(Conclusion)

本文在对VGG-net网络进行优化和改进的基础上构建s-VGG网络,采用桥小脑角区听神经瘤和脑膜瘤的T1WI增强图像和T2WI图像,分别训练s-VGG-T1和s-VGG-T2模型。测试集中s-VGG-T1与s-VGG-T2的P、R、S指标分别达到0.937、0.839、0.829与0.925、0.810、0.802,证明模型具有良好的分类性能。在分类模型结果的基础上,结合放射科与放疗科的临床诊断结果,构建深度学习辅助临床诊断模型DL-MDT,该诊断模型在10例肿瘤中的诊断准确率为100%,表明了本文构建的DL-MDT模型具有良好的诊断性能,可以在临床工作中提升诊断的准确性和临床工作的效率。

参考文献(References)

[1] 于洪泉,姚长义,邓玉生. 桥小脑角区脑膜瘤的诊断及治疗[J]. 临床医药文献电子杂志,2015,2(27):5650-5651.
[2] RADES D,FEHLAUER F,LAMSZUS K,et al. Well-differentiated neurocytoma: what is the best available treatment? [J]. Neuro-Oncology,2005,7(1):77-83.
[3] LITJENS G,KOOI T,BEJNORDI B E,et al. A survey on deep learning in medical image analysis[J]. Medical Image Analysis,2017,42:60-88.
[4] MOESKOPS P,VIERGEVER M A,MENDRIK A M,et al. Automatic segmentation of MR brain images with a convolutional neural network[J]. IEEE Transactions on Medical Imaging,2016,35(5):1252-1261.
[5] 莫修源,吴丽丽,陆志翔. 改进的V-C-Net卷积神经网络脑肿瘤图像多层次分割实验[J]. 软件工程,2022,25(12):37-43.
[6] 洪奇峰,施伟斌,吴迪,等. 深度卷积神经网络模型发展综述[J]. 软件导刊,2020,19(4):84-88.
[7] ARTHURS B J,LAMOREAUX W T,MACKAY A R,et al. Gamma knife radiosurgery for vestibular schwannomas: tumor control and functional preservation in 70 patients[J]. American Journal of Clinical Oncology: Cancer Clinical Trials,2011,34(3):265-269.
[8] FLEISSIG A,JENKINS V,CATT S,et al. Multidisciplinary teams in cancer care: are they effective in the UK? [J]. The Lancet Oncology,2006,7(11):935-943.
[9] HERENT P,SCHMAUCH B,JEHANNO P,et al. Detection and characterization of MRI breast lesions using deep learning[J]. Diagnostic and Interventional Imaging,2019,100(4):219-225.
[10] 赵尚义,王远军. 基于3D CNN的脑胶质瘤分类算法[J]. 光学技术,2019,45(6):749-755.
[11] 缪飞,沈天真,陈星荣,等. 桥小脑角区MRI解剖研究[J]. 中华放射学杂志,1997,31(8):520-523.
[12] ENGELHARD H H. Progress in the diagnosis and treatment of patients with meningiomas. Part I:diagnostic imaging,preoperative embolization[J]. Surgical Neurology,2001,55(2):89-101.

作者简介:

胡小洋(1989-),男,硕士生,主管技师。研究领域:医学图像处理,医学物理。
刘　颖(1980-),女,博士,副教授。研究领域:磁共振成像技术。本文通信作者。
倪春霞(1982-),女,硕士,副主任医师。研究领域:神经系统肿瘤放射治疗。
陈　淑(1978-),女,本科,主治医师。研究领域:神经系统肿瘤放射治疗。
董彬彬(1994-),男,本科,技师。研究领域:磁共振成像技术。