

基于滚动导向滤波和混合多尺度分解的红外与可见光图像融合方法

赵程¹, 黄永东^{1,2*}

¹北方民族大学图像处理与理解研究所, 宁夏 银川 750021;

²大连民族大学数学与信息科学研究中心, 辽宁 大连 116600

摘要 系统分析了滚动导向滤波的参数, 根据参数变化的规律提出一种基于滚动导向滤波的混合多尺度分解方法, 并基于此方法将红外与可见光图像分解到基本层、小尺度层和大尺度层。在基本层上使用像素能量与梯度能量相结合的融合规则, 在大、小尺度层上采用双通道自适应脉冲耦合神经网络(DAPCNN)的融合规则。通过逆混合多尺度分解方法得到融合图像。与已有的图像分解方法相比, 所提图像分解方法不仅可以很好地提取图像的纹理细节, 保留图像的边缘特征, 还可以避免在边缘位置上出现光晕现象。实验结果表明, 所提方法可以很好地提取图像的红外目标信息, 并将其融合到可见光图像中, 与已有的融合方法相比, 所提方法不仅在人眼视觉感知的主观评价方面具有明显优势, 而且在互信息、信息熵、标准差、非线性相关信息熵和 Chen-Varshney 指标等客观评价标准上也具有很大的优势。

关键词 图像处理; 图像融合; 滚动导向滤波; 混合多尺度分解; 双通道自适应脉冲耦合神经网络

中图分类号 TP751

文献标识码 A

doi: 10.3788/LOP56.141007

Infrared and Visible Image Fusion via Rolling Guidance Filtering and Hybrid Multi-Scale Decomposition

Zhao Cheng¹, Huang Yongdong^{1,2*}

¹*Institute of Image Processing and Understanding, North Minzu University, Yinchuan, Ningxia 750021, China;*

²*Center of Mathematics and Information Science, Dalian Minzu University, Dalian, Liaoning 116600, China*

Abstract We systematically analyze the parameters of rolling guidance filtering and propose a hybrid multi-scale decomposition method based on rolling guidance filtering, utilizing the law of parameter variation. First, the infrared and visible images are decomposed into base, small-scale, and large-scale layers using this method. Second, a fusion rule that combines the pixel and gradient energies is applied to the base layer, and another fusion rule based on an adaptive dual-channel pulse-coupled neural network (DAPCNN) is used to combine the large- and small-scale layers. Finally, the fused image is obtained via inverse hybrid multi-scale decomposition. Compared with other common image decomposition methods, the proposed method can not only extract the image's texture details and preserve its edge features but also prevent the halo phenomenon at the edges. The experimental results show that the proposed method can extract the target information from the infrared image and fuse it into the visible image. Compared with existing fusion methods, the proposed method shows clear advantages not only in a subjective human visual evaluation, but also in several objective evaluation metrics, namely the mutual information, information entropy, standard deviation, non-linear correlation information entropy, and Chen-Varshney indexes.

Key words image processing; image fusion; rolling guidance filtering; hybrid multi-scale decomposition; adaptive dual-channel pulse-coupled neural network

OCIS codes 100.2000; 110.3080; 110.2960; 110.3000; 110.6820

收稿日期: 2019-01-17; 修回日期: 2019-01-25; 录用日期: 2019-02-20

基金项目: 国家自然科学基金(11761001, 61462002)、宁夏科技创新领军人才项目(KJT2016002)、北方民族大学研究生创新项目(YCX18086)

* E-mail: Huang_yongdong@163.com

1 引言

通过不同的传感器可以获取不同类型的图像,但是在所获取的单幅图像中并不能提取到所需要的全部有用信息。如红外图像在低光或恶劣的天气条件下可以捕捉到场景目标的热源信息,而这在可见光图像中是不可能得到的;可见光图像则可以提供场景的更多细节信息,得到与人类视觉相一致的图像亮度和对比度信息。由此可见,将同一场景的红外图像与可见光图像整合,可以得到一幅场景细节信息更多、更加完整的图像,故红外与可见光图像融合方法应运而生。红外与可见光图像融合方法已广泛应用于军事、目标检测和情报收集等领域。

近年来,学者们先后提出了多种红外与可见光图像融合算法^[1-12]。其中,基于像素层红外与可见光图像的融合方法大体上可以分为基于空间域融合方法和基于变换域融合方法。基于空间域融合方法是在空间域对配准的高、低空间分辨率图像进行处理,包含直接融合法、加权融合法等。该类方法在融合处理过程中不对图像进行分解,可能会导致融合图像出现模糊,其相应的对比度较低。目前,使用最多的图像融合方法是基于多尺度变换的融合方法,其是针对变换域而言的,主要包含基于塔形分解的变换方法^[1]、基于小波变换的方法^[2-3]、基于 Curvelet 变换的方法^[4]、基于 Contourlet 变换的方法^[5]、基于双树复小波变换的方法(DTCWT)^[6-7]等。基于塔形分解的变换方法包括基于拉普拉斯金字塔变换^[8]、基于梯度金字塔变换^[9],以及基于低通比率金字塔(ROLP)变换^[10]等。然而上述方法均不具有平移不变性,在融合结果中表现出较明显的伪 Gibbs 现象。为了解决此类问题,学者们提出了 NSCT(non-subsampled contourlet transform)法与 NSST(non-subsampled contourlet transform)法^[11-12],这些方法具有较好的方向选择性和平移不变性。但是,其在融合的过程中没有充分考虑空间一致性,导致图像可能会出现亮度和颜色失真的现象。

为了克服上述缺陷,将基于边缘保留滤波的方法引入到图像融合中。这主要是因为边缘保留滤波具有平移不变性和好的边缘保留性能,且计算效率相对较高。基于边缘保留滤波的方法在应用中直接将原始图像分解到细节层和基本层,从而达到平滑纹理、保留边缘信息的目的。大部分边缘保留滤波是根据对比度平滑细节信息,而双边滤波^[13]和导向滤波^[14]等少数滤波方法则根据空间尺度对图像进行分解。因此,为了更好地根据图像的尺度信息保

留边缘,平滑图像的结构,Zhang 等^[15]提出了基于滚动导向滤波(RGF)的方法,其不仅可以更好地平滑图像纹理信息,还可以在实现过程中通过迭代来保留图像的边缘信息。因此可以很好地应用到图像处理中。常规的多尺度分解方法仅从单一通道对图像的信息进行分解,不能对图像信息进行精确分类,而混合多尺度分解方法可以弥补此缺陷,得到不同类型的图像信息。如 Zhou 等^[13]提出将高斯滤波与双边滤波相结合的混合多尺度分解方法,其主要优点有:1)在融合过程中尽可能多地把红外光谱信息融入到可见光图像中;2)保留了可见光图像的场景和细节信息。因此,结合滚动导向滤波和混合多尺度分解方法,可以充分地提取图像的纹理、细节信息,同时保留图像的边缘信息。

此外,在图像融合过程中,选择合适的融合规则同样至关重要。脉冲耦合神经网络(PCNN)因其具有全局耦合性和脉冲同步性,目前已广泛应用于图像处理中。然而,常规的 PCNN 模型不能同时处理两幅图像,且经常选用单一像素值作为外部刺激,选用常量作为链接强度,降低了对图像细节信息处理的准确性。为了解决上述问题,Cheng 等^[16]提出了一种改进的双通道自适应脉冲耦合神经网络模型(DAPCNN),并将其用于处理高频子带图像信息,利用局部图像信息充分提取图像的细节信息,使 DAPCNN 更加适用于图像融合,且产生较好的融合效果。

基于上述分析,本文综合滚动导向滤波和混合多尺度分解方法的优势,提出了一种新的红外与可见光图像融合方法。混合多尺度分解可以保留原始图像的细节信息,滚动导向滤波具有空间一致性,可以更好地平滑图像纹理信息,保留图像边缘信息。因此,通过结合两者的优势,所提分解方法可以充分提取图像同一层包含的不同类型信息,进而根据不同的信息类型选择不同的融合规则,避免处理过程中信息的丢失。本文还选用改进的 DAPCNN 模型对图像的大尺度层信息进行处理,进一步提高对图像细节信息的处理能力。同时,为了充分保留输入图像的能量信息,采用区域能量加权方法处理基本层信息。实验结果表明,相比于其他方法,本文所提方法无论是在主观评价还是客观评价都表现出很明显的优势。

2 RGF 和 DAPCNN 理论

2.1 滚动导向滤波

滚动导向滤波是 Zhang 等^[15]提出的一种新的边缘保留滤波。与其他边缘保留滤波的不同之处在

于,滚动导向滤波将迭代运用到实现过程中,可以很好地实现快速收敛。此滤波的实现过程简单、快速,在空间重叠特征中实现了尺度的分离,所得结果没有出现人工效应。

2.1.1 滚动导向滤波实现过程

滚动导向滤波包括小结构消除和边缘恢复两个过程,如图 1 所示。

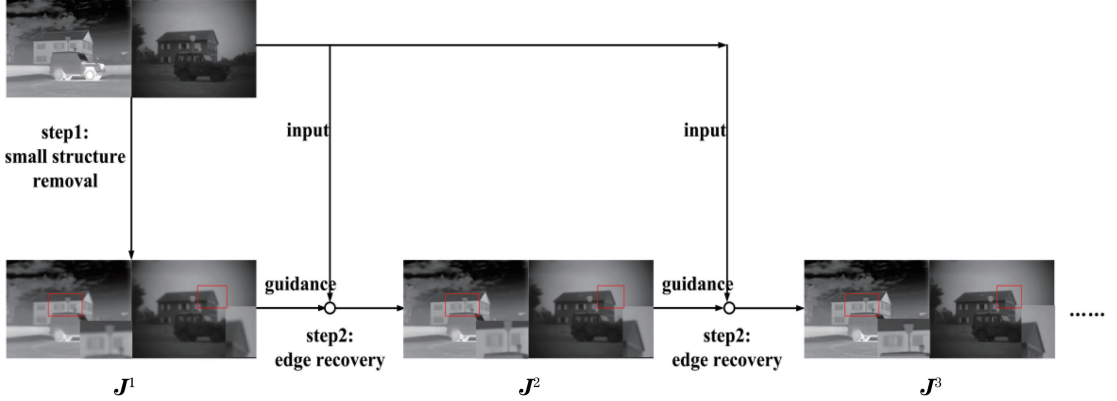


图 1 RGF 的基本框架

Fig. 1 Basic framework of RGF

式中: $K_p = \sum_{q \in N(p)} \exp\left(-\frac{\|p-q\|^2}{2\sigma_s^2}\right)$, 其中 $N(p)$ 为以像素 p 为中心的邻域像素集。值得注意的是,在尺度空间中,当尺度远远小于 σ_s 时,此滤波可以完全消除小结构。

2) 边缘恢复过程。此过程的实现方式包括两步:迭代以及联合滤波。选取导向滤波作为联合滤波,这是因为导向滤波不仅有高的计算效率,还有好的边缘保留性能。将 J^1 记为经高斯滤波处理后的输出图像,即 $J^1 = G$ 。 J^{t+1} 为第 t 次迭代的结果,可通过输入图像 I 和引导图像 J^t 经过导向滤波得到。具体实现方式可表示为

$$J^{t+1} = f_{GF}(J^t, I, \sigma_s, \sigma_r^t), t = 1, 2, \dots, T, \quad (2)$$

式中: f_{GF} 为导向滤波函数; σ_r 控制范围权重,设 $\sigma_r = 0.1$ 。将(2)式视为一个滤波过程,即通过 J^t 对输入图像 I 进行平滑处理后得到滤波图像 J^{t+1} 。

2.1.2 参数分析

为了更好地设计基于 RGF 的图像分解方法,需要对 RGF 的参数进行分析,并探讨 RGF 参数的选取对融合效果造成的影响。经查阅,对 RGF 参数系统分析的文献甚少,故需在此对其进行较为详细的分析。由 RGF 的构造可知,在小结构消除过程中,参数 σ_s 主要用于平滑图像的纹理细节信息,参数 σ_r 在这一步为无效参数;在边缘恢复过程中,使用引导

1) 小结构消除过程。通过使用高斯滤波实现这个过程。小结构信息通常为细节信息,具体包括纹理、噪声以及小目标等信息。将 I 表示输入图像, G 表示输出图像, p 和 q 表示图像的像素坐标矢量, σ_s 为标准差,则滤波的表达式可以写为

$$G(p) = \frac{1}{K_p} \sum_{q \in N(p)} \exp\left(-\frac{\|p-q\|^2}{2\sigma_s^2}\right) I(q), \quad (1)$$

滤波且通过 J^t 平滑输入图像 I 得到 J^{t+1} , 此时参数 σ_r 起着至关重要的作用,其可以用于判断边缘恢复的程度。通过对参数进行分析,可以确定合理的参数,以获得最优的平滑图像纹理,并保留更多的图像边缘信息。

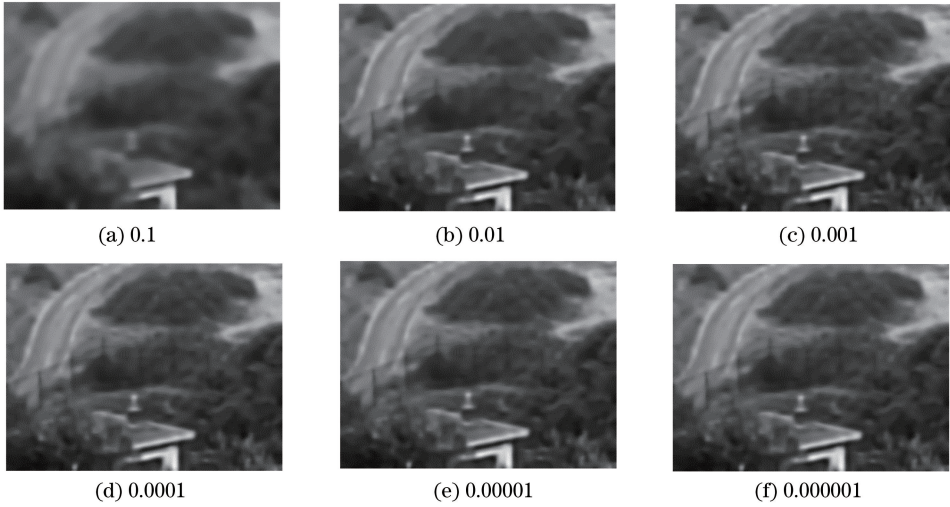
选择可见光图像(图 2)进行分析,针对不同的参数 σ_r 观察边缘恢复的程度。图 2 中给出了 σ_r 从 0.1 依次变化到 10^{-6} 所对应的滤波后的可见光图像,可以发现,当 σ_r 从 0.1 变化到 0.01 时,房子的边缘以及栅栏、烟囱等信息会逐渐变得清晰,当 σ_r 从 0.01 变化到 10^{-6} 时,除了房子、栅栏、烟囱的变化之外,处于中间位置的山体部分也逐渐变得清晰,且 σ_r 在 0.0001 到 10^{-6} 之间所对应图像的主观视觉一样。根据对图 2 的主观分析可知,随着 σ_r 的不断减小,滚动导向滤波的边缘恢复能力逐步增强,但是当 σ_r 减小到某一固定值后,滚动导向滤波的边缘恢复能力固定不变。

客观分析过程如下:

1) 固定 $\sigma_s = 2$, 并将 σ_r^0 的初值依次取为 0.1、0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9;

2) 根据 $\sigma_r^{c+1} = \sigma_r^c \times 10^{-1}$, 其中 $c = 0, 1, \dots, 25$, 获得在某一确定初始值的情况下所对应的所有 σ_r 值;

3) 选用确定的参数 σ_r, σ_s 作为滚动导向滤波的参数,对某一幅图像进行滤波得到滤波后的图像;

图2 不同 σ_r 值所对应的RGF平滑后的结果Fig. 2 Smooth results of RGF corresponding to different σ_r values

4) 选用标准差(STD)、空间频率(SF)和边缘强度(EI),对所得的每一幅滤波图像进行客观评价,计算同一幅图像在不同的 σ_r 情况下所对应的3种客观指标的值。

经过实验分析发现,当 $\sigma_r^0 = 0.1, 0.2, 0.3$ 时,上述3种指标均在 $\sigma_r = \sigma_r^0 \times 10^{-4}$ 时可以同时达到最优值,且在此之后所得指标值不再发生变化。当 $\sigma_r^0 = 0.4, 0.5, \dots, 0.9$ 时,上述3种指标均在 $\sigma_r = \sigma_r^0 \times 10^{-5}$ 时可以同时达到最优值并在此之后不再发生变化。此外,当 σ_r^0 越小时,所得的测试指标值越大。当 σ_r 的值小于 $\sigma_r^0 \times 10^{-5}$ 时, σ_r^0 取0.1~0.9之间的任意值,由此对应的 σ_r 产生的滤波图像所对应的3种客观指标的测试值均不再发生变化,且数值相等。因此,设定的最优值范围为 $\sigma_r < \sigma_r^0 \times 10^{-5}$ 。

为了更直观地证明上述结果,计算图2中对应图像在 σ_r 为0.1、0.01、0.001、0.0001和0.00001以及0.2、0.02、0.002、0.0002和0.00002时所对应的滤波图像的STD、SF和EI,并在表1给出相应的数据。由表1可知,对于同一幅可见光图像,当 σ_r 选择不同的参数值时,随着参数值的不断减小,图像对应的指标值不断增大,且当 $\sigma_r \leq 10^{-4}$ 时, σ_r 取不同的值对应的滤波图像的测试值不再发生改变。而这与上述分析结果一致,由此也验证了上述分析的有效性。

参数 σ_s 用于控制空间权重,具体用于小结构消除过程。选用与分析 σ_r 相同的方法进一步分析 σ_s 对图像的影响。选用同一幅可见光图进行测试, σ_s 的增长步长为5,实验发现随着 σ_s 的不断增大,滚动导向滤波对图像纹理信息的平滑效果更好,且当

$\sigma_s \geq 30$ 时,图像的平滑效果不再发生变化。

表1 不同 σ_r 值对应的可见光图像的客观评价指标
Table 1 Objective evaluation indicators of visible image corresponding to different σ_r values

σ_r	STD	SF	EI
0.1	0.1470	0.0305	0.0936
0.01	0.1514	0.0423	0.1290
0.001	0.1520	0.0443	0.1477
0.0001	0.1521	0.0443	0.1481
0.00001	0.1521	0.0443	0.1481
0.2	0.1457	0.0264	0.0909
0.02	0.1505	0.0399	0.1147
0.002	0.1520	0.0442	0.1466
0.0002	0.1521	0.0443	0.1481
0.00002	0.1521	0.0443	0.1481

2.2 双通道自适应脉冲耦合神经网络

DAPCNN是以PCNN为基础的一种新的脉冲耦合神经网络。PCNN是一种反馈型网络,模拟了哺乳动物大脑视觉皮层中神经元的同步行为现象,是一种二维人工神经网络模型,具有全局耦合性和脉冲同步性。但是,PCNN有两个缺陷:1)用单一像素值作为外部输入;2)采用固定值作为链接强度。因此,PCNN不具有自适应性。而DAPCNN改变了外部输入与链接强度的设置,同时考虑了两幅图像之间的联系,用于避免信息的遗失,弥补了PCNN不具有自适应性的不足,可以更有效地处理图像信息。DAPCNN的实现过程包括接收域、调制域和产生域3个部分。

1) 接收域的实现过程可表示为

$$F_{ij}^1(n) = S_{ij}^1(n), \quad (3)$$

$$F_{ij}^2(n) = S_{ij}^2(n), \quad (4)$$

$$L_{ij}(n) = \exp(-\alpha_L) \cdot L_{ij}(n-1) + V_L \sum_{ab} W_{ij,ab} Y_{ij}(n-1), \quad (5)$$

式中: n 为当前迭代次数; F_{ij}^k 为在 (i, j) 位置神经元的双通道反馈输入; L_{ij} 为链接输入; S_{ij}^k 为外部激励输入; $W_{ij,ab}$ 为联合权重强度, 组成矩阵 \mathbf{W} ; Y_{ij} 为在 (i, j) 位置的神经元输出; α_L 为衰减时间常数; V_L 为标准化常数; k 为 DAPCNN 的通道, 其中 $k = 1, 2$ 。

2) 调制域的实现过程可表示为

$$U_{ij}(n) = \max\{F_{ij}^1(n)[1 + \beta_1 L_{ij}(n)], F_{ij}^2(n)[1 + \beta_2 L_{ij}(n)]\}, \quad (6)$$

式中: U_{ij} 为总体内部活性度; $\beta_k \in [0, 1]$ 为权重系数, 用于判断两个通道所占的权重。

3) 产生域的实现过程可表示为

$$\theta_{ij}(n) = \exp(-\alpha_\theta) \theta_{ij}(n-1) + V_\theta Y_{ij}(n-1), \quad (7)$$

$$Y_{ij}(n) = \begin{cases} 1, & U_{ij}(n) > \theta_{ij}(n) \\ 0, & \text{others} \end{cases}, \quad (8)$$

式中: θ_{ij} 为神经元的动态阈值; Y_{ij} 为输出脉冲; α_θ 为衰减时间常数; V_θ 为正则化常量。若 $U_{ij} \geq \theta_{ij}$, 则神经元产生一个脉冲。

由于梯度可以反映图像细节信息, 需要引入 Sobel 算子来自适应选取链接强度。对图像的任意一点使用 Sobel 算子, 都会得到该点的梯度矢量或法矢量。图像 \mathbf{A} 在点 (i, j) 的边缘梯度为

$$G(i, j) = \sqrt{G_x^2(i, j) + G_y^2(i, j)}, \quad (9)$$

式中: $G_x(i, j) = \mathbf{E}_1 \cdot \mathbf{A}$; $G_y(i, j) = \mathbf{E}_2 \cdot \mathbf{A}$; \mathbf{E}_1 和 \mathbf{E}_2 分别为横向和纵向滤波算子, 其表达式可分别表示为

$$\mathbf{E}_1 = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}, \mathbf{E}_2 = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}。从$$

而链接强度 β_k 的计算公式可表示为

$$\beta_k(i, j) = \frac{1}{1 + \exp[-G(i, j)]}, k = 1, 2. \quad (10)$$

3 图像融合的基本过程

3.1 基于滚动导向滤波的混合多尺度分解框架的构造

滚动导向滤波可以通过调控所含参数达到保留图像边缘、平滑纹理的目的。因此, 通过设置不同的参数值对图像的边缘、纹理信息进行处理, 以期得到含有不同信息的滤波图像。此外, 常规的多尺度分

解方法是通过直接处理相邻层的滤波图像得到图像的细节层信息, 但不能更深入地提取图像的不同类型信息。而混合多尺度分解方法, 相当于对不同通道的滤波图像进行处理, 根据不同滤波图像包含信息量的多少设计具体的分解框架, 以此更充分地提取图像所包含的不同类型的细节信息, 从而可以更精确地为不同类型的信息选择不同的融合方法, 加强了后期图像的融合效果。结合上述两种方法的优势, 提出了一种基于滚动导向滤波的混合多尺度分解方法, 该方法不仅可以利用滚动导向滤波的优势提取出图像的细节信息, 还可以利用混合多尺度分解方法对细节信息进行分类, 得到细节信息对应的纹理信息和边缘信息, 具体框架如图 3 所示。其中用 \mathbf{I}_r 表示红外图像, 可见光图像的分解过程与红外图像分解过程一致。具体过程如下:

1) 使用一系列的滚动导向滤波去平滑输入图像 \mathbf{I} , 即 $\hat{\mathbf{I}}^{\omega+1} = f_{\text{RGF}, \hat{\sigma}_{\text{sw}}, \hat{\sigma}_{\text{rw}}}(\hat{\mathbf{I}}^\omega)$, 在这里 $\omega = 1, 2, \dots, n$ 且 $\hat{\mathbf{I}}^1 = \mathbf{I}$ 。

需要说明的是, 这一步主要考虑的是平滑处理图像的纹理信息, 因此设置 $\hat{\sigma}_r$ 为相对大的值以确保实现滚动导向滤波的平滑性能。此外, 为了使每一层的滤波图像包含相对多的信息, 此处设置 $\hat{\sigma}_s$ 为相对小的值。因此, 取 $\hat{\sigma}_s = \hat{\sigma}_{\text{sw}} = 5$ 和 $\hat{\sigma}_r = \hat{\sigma}_{\text{rw}} = 10^2$ 。

2) 使用一系列与 1) 不同参数的滚动导向滤波来平滑纹理细节信息, 保留图像的边缘信息, 即 $\mathbf{I}^{\omega+1} = f_{\text{RGF}, \sigma_{\text{sw}}, \sigma_{\text{rw}}}(\mathbf{I}^\omega)$ 。

此时, σ_s 主要用于平滑纹理并保留必要的边缘信息, σ_r 主要用于边缘恢复与保留。需要指出的是, 由于混合多尺度分解方法需要处理不同参数的滤波图像来得到精细的细节信息, 因此本研究没有直接使用上文分析的最优值处理图像, 而是根据参数的变化趋势设置具体的参数值。经过实验, 本研究的参数设置如下: $\sigma_{r1} = 0.5, \sigma_{s1} = 2, \sigma_{rw+1} = \sigma_{rw} / 2,$

$$\sigma_{sw+1} = 2\sigma_{sw}, f_{\text{RGF}, \hat{\sigma}_{s0}, \hat{\sigma}_{r0}}(\hat{\mathbf{I}}^0) = \mathbf{I}。$$

3) 将文献[13]中所提到的混合多尺度分解方法与滚动导向滤波相结合, 用于获得图像的纹理信息和边缘细节信息。经上述分析可知, 所取 $\hat{\sigma}_{\text{sw}}$ 与 $\hat{\sigma}_{\text{rw}}$ 为固定值, 而随着 σ_s 不断增大, 图像的纹理信息的平滑效果越好; 随着 σ_r 不断减小, 所保留的图像的边缘细节信息越多。因此, 当处于同一分解层时, 选取 σ_s 与 σ_r 为参数值的滚动导向滤波可以保留更多的边缘细节信息, 将其与同一分解层上选取

$\hat{\sigma}_{sw}$ 与 $\hat{\sigma}_{rw}$ 为参数值所得到的滤波图作差值,即可得到对应分解层的边缘细节信息。将前一分解层(参数为 $\hat{\sigma}_{sw}$ 与 $\hat{\sigma}_{rw}$)所对应的滤波图与当前分解层(参数为 σ_s 与 σ_r)的滤波图作差值即可得到纹理信息。因此,当进行第一层分解时,相当于用原始图与参数为 σ_s 与 σ_r 所得的滤波图(当前层)作差值,此时执行 2) 步时当前层是已被平滑的图像,其所含的纹理信息要少于原始图像,两者的差值即为初次分解的纹理信息。在此基础上对原始图像进行参数为 $\hat{\sigma}_{sw}$ 与 $\hat{\sigma}_{rw}$ 的滤波处理,同时对 2) 步滤波图再次进行参数为 σ_s 与 σ_r 的滤波处理,因为后者要比前者多执行一次滤波处理,且随着 σ_s 的不断

更新,其平滑纹理信息的能力会逐渐增强,而前者的参数值保持不变,两者相减即可得到后续分解层的纹理信息。综上所述,所提分解方法的具体分解框架的数学表达式为

$$\mathbf{D}^{(\omega,0)} = f_{\text{RGF},\hat{\sigma}_{sw-1},\hat{\sigma}_{rw-1}}(\hat{\mathbf{I}}^{\omega-1}) - f_{\text{RGF},\sigma_{sw},\sigma_{rw}}(\mathbf{I}^{\omega}), \quad (11)$$

$$\mathbf{D}^{(\omega,1)} = f_{\text{RGF},\sigma_{sw},\sigma_{rw}}(\mathbf{I}^{\omega}) - f_{\text{RGF},\hat{\sigma}_{sw},\hat{\sigma}_{rw}}(\hat{\mathbf{I}}^{\omega}), \quad (12)$$

式中: $\omega = 1, 2, \dots, n$ 为分解的层数; $\mathbf{D}^{(\omega,0)}$ 为第 ω 层的纹理细节信息; $\mathbf{D}^{(\omega,1)}$ 为第 ω 层的边缘细节信息。需要注意的是,在整个构造过程中多尺度分解的基本层为 $\mathbf{B} = f_{\text{RGF},\hat{\sigma}_{sh},\hat{\sigma}_{rh}}(\hat{\mathbf{I}}^n)$ 。

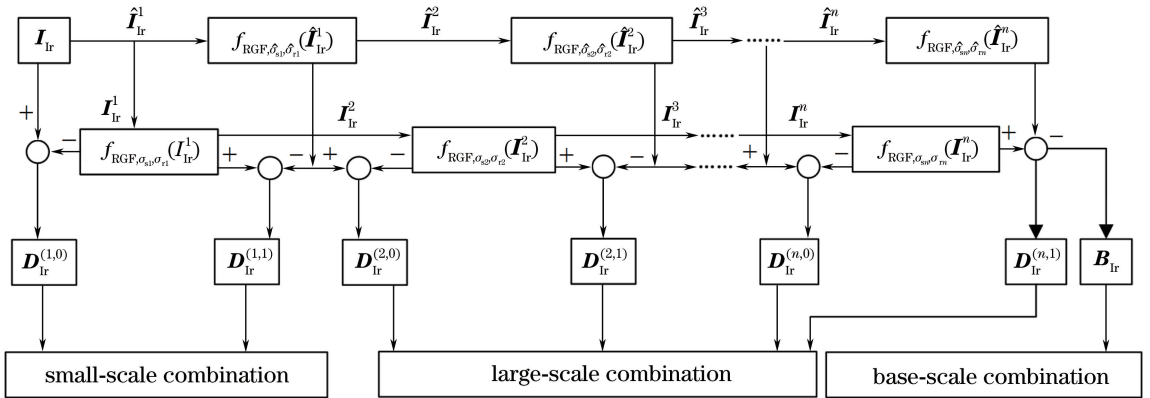


图 3 基于混合多尺度分解框架的构造

Fig. 3 Construction based on hybrid multi-scale decomposition

3.2 融合规则的选择

通过混合多尺度分解,将红外与可见光图像分解到小尺度层、大尺度层和基本层 3 个不同的尺度层上,并在不同尺度层上选择不同的融合规则来融合各层信息。

3.2.1 小尺度层的融合规则

小尺度层是图像分解的第一层,即 $\omega = 1$ 时对应的纹理信息和细节信息。由于该层主要包含图像的边缘细节信息和纹理信息,经常会使用 PCNN 对该尺度层进行处理。尽管 PCNN 模型具有视觉恢复机制,可以对高频子图进行仿生处理并在融合的过程中恢复图像的细节信息,然而其在实验过程中将单幅图像信息作为输入,并没有考虑到另一幅图像的信息,且其参数的选择依赖于实验调制,这就降低了对细节信息处理的准确性。为了避免上述问题,并在此基础上更好地处理图像较暗区域信息,提高整体算法的计算效率,本研究选择 DAPCNN 进行融合。由于边缘能量(EOE)^[17]与 Sobel 算子均是以图像的细节信息为基础,因此选用边缘能量作

为外部输入,选用 Sobel 算子作为链接强度,具体步骤如下:

- 1) 对实验中所需要的参数进行初始化,选用 Sobel 算子作为像素点的链接强度,即利用(10)式计算链接强度 $\beta_k(i, j)$;
- 2) 计算小尺度所对应系数的边缘能量,并将其作为 DAPCNN 的外部输入。

$$\begin{cases} F_{1,ij}^{(1,i_1)} = E_{\text{end},R}^{(1,i_1)}(i, j) \\ F_{2,ij}^{(1,i_1)} = E_{\text{end},V}^{(1,i_1)}(i, j) \end{cases}, \quad (13)$$

式中: $E_{\text{end},m}^{(1,i_1)}(i, j)$ 表示红外与可见光图像在尺度 i_1 位置 (i, j) 处所对应的边缘能量; $i_1 = 0, 1; m$ 为 R, V。边缘能量的定义为

$$E^{(1,i_1)}(i, j) = \sum_{i,j \in \Omega} \omega(i, j) E_L^{(1,i_1)}(i, j), \quad (14)$$

式中: Ω 为以像素位置 (i, j) 为中心的邻域窗口; $\omega(i, j)$ 为权重矩阵元,组成矩阵 ω , $\omega = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$; $E_L^{(1,i_1)}(i, j) = (\mathbf{L}_1 \cdot \mathbf{I}^{(1,i_1)})^2 +$

$$(\mathbf{L}_2 \cdot \mathbf{I}^{(1,i_1)})^2 + (\mathbf{L}_3 \cdot \mathbf{I}^{(1,i_1)})^2。其中 \mathbf{L}_1 = \begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix}, \mathbf{L}_2 = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}, \mathbf{L}_3 = \begin{bmatrix} -1 & 0 & -1 \\ 0 & 4 & 0 \\ -1 & 0 & -1 \end{bmatrix}。$$

3) 计算 L_{ij} 、 U_{ij} 、 θ_{ij} 和 Y_{ij} ，从而激活所有的神经元。选择相对应的融合系数为

$$D_F^{(1,i_1)}(i,j) =$$

$$\begin{cases} D_R^{(1,i_1)}(i,j)U^{(1,i_1)}(i,j) = U_R^{(1,i_1)}(i,j) \\ D_V^{(1,i_1)}(i,j)U^{(1,i_1)}(i,j) = U_V^{(1,i_1)}(i,j) \end{cases},$$

$$\begin{cases} U_R^{(1,i_1)}(i,j) = \\ F_1^{(1,i_1)}(i,j) [1 + \beta_1^{(1,i_1)}(i,j)L^{(1,i_1)}(i,j)] \\ U_V^{(1,i_1)}(i,j) = \\ F_2^{(1,i_1)}(i,j) [1 + \beta_2^{(1,i_1)}(i,j)L^{(1,i_1)}(i,j)] \end{cases},$$

(15)

式中： $D_R^{(1,i_1)}(i,j)$ 、 $D_V^{(1,i_1)}(i,j)$ 、 $D_F^{(1,i_1)}(i,j)$ 分别为 $\mathbf{D}_R^{(1,i_1)}$ 、 $\mathbf{D}_V^{(1,i_1)}$ 和 $\mathbf{D}_F^{(1,i_1)}$ 的矩阵元，代表小尺度层的红外系数、可见光系数和融合后的小尺度层系数。其他参数含义与 2.2 节中所代表的含义相同，在此不加赘述。

3.2.2 大尺度层的融合规则

大尺度层包括从 $w=2$ 到 $w=n$ 时对应的图像的纹理信息和细节信息。大尺度层中包含了原始图像的很多细节信息，如纹理、背景等。为了更好地检测和提取这些显著信息，使用 DAPCNN 模型对其进行处理。在融合过程中，红外图像的光谱特征经常会作为大尺度层的显著边缘，实际上并非所有与可见光图像融合的红外信息都是有用的。因此，为了避免在图像融合过程中合并过多不相关的红外细节和噪声，在上述基础上对所得到的融合结果使用加权最小二乘优化的方法 (WLS)^[18] 来加强融合效果。

1) 根据小尺度层的方法计算第 w 层的融合尺度层，记为 $\mathbf{H}^{(w,i_1)}$ ， $w=2, \dots, n, i_1=0, 1$ 。

2) 通过使用最小化成本函数得到第 w 层的融合系数尺度层 $\mathbf{D}^{(w,i_1)}$ ，这里最小化成本函数为

$$g_{cf} = \sum_p \{ (\mathbf{D}_p^{(w,i_1)} - \mathbf{H}_p^{(w,i_1)})^2 + \lambda a_p^w [\mathbf{D}_p^{(w,i_1)} - (\mathbf{D}_V^{(w,i_1)})_p]^2 \},$$

$$a_p^w = \left[\left| \sum_{p \in \Omega_p} D_R^{(w,i_1)}(i,j) \right| + \epsilon \right]^{-1}, \quad (16)$$

式中： p 为像素的空间位置矢量； ϵ 为 0.0001； Ω_p 为

以 p 为中心的窗口； $\mathbf{D}_V^{(w,i_1)}$ 为可见光图像的大尺度层； $\mathbf{D}_R^{(w,i_1)}$ 为红外图像的大尺度层； a_p^w 为空间变换权重； λ 为全局控制参数，用于平衡(16)式等号右边的两项，取 $\lambda=0.01$ 。将(16)式用矩阵的形式可表示为

$$\mathbf{g} = (\mathbf{D}^{(w,i_1)} - \mathbf{H}^{(w,i_1)})^T (\mathbf{D}^{(w,i_1)} - \mathbf{H}^{(w,i_1)}) + \lambda (\mathbf{D}^{(w,i_1)} - \mathbf{D}_V^{(w,i_1)})^T \mathbf{A}^{(w,i_1)} (\mathbf{D}^{(w,i_1)} - \mathbf{D}_V^{(w,i_1)}), \quad (17)$$

式中： $\mathbf{A}^{(w,i_1)}$ 为包含所有像素权重值的对角矩阵。利用优化理论与方法， $\mathbf{D}^{(w,i_1)}$ 的最优值可表示为

$$\{ 2\mathbf{I} + \lambda [\mathbf{A}^{(w,i_1)} + (\mathbf{A}^{(w,i_1)})^T] \} \mathbf{D}^{(w,i_1)} = 2\mathbf{H}^{(w,i_1)} + \lambda [\mathbf{A}^{(w,i_1)} + (\mathbf{A}^{(w,i_1)})^T] \mathbf{D}_V^{(w,i_1)}. \quad (18)$$

因为 $\mathbf{A}^{(w,i_1)}$ 是对角矩阵，所以有 $\mathbf{A}^{(w,i_1)} = (\mathbf{A}^{(w,i_1)})^T$ 。因此，(18)式可以写为

$$(\mathbf{I} + \lambda \mathbf{A}^{(w,i_1)}) \mathbf{D}^{(w,i_1)} = \mathbf{H}^{(w,i_1)} + \lambda \mathbf{A}^{(w,i_1)} \mathbf{D}_V^{(w,i_1)}, \quad (19)$$

式中： $\mathbf{D}^{(w,i_1)}$ 为大尺度层的最终融合系数矩阵。

3.2.3 基本层的融合规则

基本层 $\mathbf{B} = f_{\text{RGF}, \hat{i}_{\text{sp}}, \hat{i}_{\text{w}}}(\hat{\mathbf{I}}^n)$ 是平滑纹理的最后一层，包含了主要的能量信息和一些显著的细节信息，如边缘、轮廓等。为了在融合图像中显示更多信息，基本层的融合规则选用了像素能量和梯度能量的加权和，并对最后的结果进行一致性检验，从而避免信息的丢失与错位。具体过程如下：

1) 计算像素能量 E_P 与梯度能量 E_G ，计算公式表示为

$$E_P(i,j) = \sum_{a=-P}^P \sum_{b=-Q}^Q \omega'(i,j) \cdot [B(i+a, j+b)]^2, \quad (20)$$

$$E_G(i,j) = \sum_{a=-P}^P \sum_{b=-Q}^Q \omega'(i,j) \cdot \{ [\nabla B_x(i+a, j+b)]^2 + [\nabla B_y(i+a, j+b)]^2 \}, \quad (21)$$

$$\text{式中：} \omega' = \begin{bmatrix} 1 & 2 & 3 & 2 & 1 \\ 2 & 4 & 6 & 4 & 2 \\ 3 & 6 & 9 & 6 & 3 \\ 2 & 4 & 6 & 4 & 2 \\ 1 & 2 & 3 & 2 & 1 \end{bmatrix}, \text{其中 } \omega'(i,j) \text{ 为矩阵}$$

元； P, Q 为窗口大小； $\nabla B_x, \nabla B_y$ 分别为图像在 x 方向和 y 方向的梯度值。

2) 计算总体能量 E_{tol} 的公式可表示为

$$E_{\text{tol}}(i,j) = t \times E_G(i,j) + (1-t) \times E_P(i,j), \quad (22)$$

$$\text{式中: } t = \frac{E_G(i, j)}{E_G(i, j) + E_P(i, j)}.$$

通过总体能量去选择最终的融合系数,可表示为

$$B_F(i, j) = \begin{cases} B_{Ir}(i, j), E_{\text{tol}, Ir}(i, j) \geq E_{\text{tol}, Vis}(i, j) \\ B_{Vis}(i, j), \text{others} \end{cases}, \quad (23)$$

式中: $E_{\text{tol}, Ir}(i, j)$ 和 $E_{\text{tol}, Vis}(i, j)$ 分别为红外与可见光的总体能量; $B_{Ir}(i, j)$ 和 $B_{Vis}(i, j)$ 分别为红外与可见光的基本层系数。

3) 为避免某些像素点的最终融合系数与其邻域内大部分相邻像素的融合系数来源于不同的图像,利用一致性检验对所得的融合系数进行处理,并利用基本图与另外两个尺度图相加得到最后的融合图像,即

$$\mathbf{F} = \mathbf{B}_F + \sum_{w=1}^n (\mathbf{D}^{(w,0)} + \mathbf{D}^{(w,1)}). \quad (24)$$

3.3 融合算法

经过上述分析,所提的融合算法步骤如下:

1) 利用 3.1 节分别对红外与可见光图像进行混合多尺度分解,得到小尺度层 $\mathbf{D}_C^{(1,i_1)}$ 、大尺度层 $\mathbf{D}_C^{(w,i_1)}$ 和基本层 \mathbf{B}_C ,其中 C 为 R, V, $w=2, \dots, n, i_1=0, 1$ 。

2) 利用 3.2.1 节的融合规则融合小尺度层 $\mathbf{D}_V^{(1,i_1)}$ 和 $\mathbf{D}_R^{(1,i_1)}$,从而得到 $\mathbf{D}_F^{(1,i_1)}$ 。

3) 利用 3.2.2 节的融合规则融合大尺度层

$\mathbf{D}_V^{(w,i_1)}$ 和 $\mathbf{D}_R^{(w,i_1)}$,从而得到 $\mathbf{D}^{(w,i_1)}$ 。

4) 利用 3.2.3 节的融合规则融合基本层 \mathbf{B}_{Vis} 和 \mathbf{B}_{Ir} ,从而得到 \mathbf{B}_F 。

5) 利用(24)式对 $\mathbf{D}^{(w,i_1)}$ 、 $\mathbf{B}_F(i, j)$ 进行逆混合多尺度分解,得到最终的融合图像 \mathbf{F} 。

4 实验结果与比较

为了说明所提方法的有效性,本文使用主观评价与客观评价相结合的方法,对不同融合方法所产生的融合结果进行比较分析。与所设计算法作比较的融合方法分别为 DWT^[19]、CVT^[20]、DTCWT^[21]、NSCT-DAPCNN^[22]、NSDTST^[23]、LATLRR^[24]、GTF^[25] 和 MSVD^[26]。其中 NSCT-DAPCNN 方法将传统的 NSCT 模型与 DAPCNN 相结合。混合多尺度分解的分解层为 4,即 $N=4$ 。DAPCNN 的参数设置为 $\alpha_L=0.2, \alpha_\theta=0.2, V_L=1.5, V_\theta=25, \mathbf{W}=[0.707, 1, 0.707; 1, 0, 1; 0.707, 1, 0.707]$ 。

4.1 融合结果的视觉评价

为了评估所提融合方法的有效性,在实验中选取了 5 对红外与可见光图像,见图 4。在图 4 中第一行为 5 组红外图像,第二行为红外图像所对应的可见光图像,每一列为一对已配准的图像,依次命名为“Un camp”“Octec”“Steamboat”“Kayak”“Infrared4-Visible4”。在图 5~9 中分别展示了 5 组图像在不同方法下所对应的融合结果。

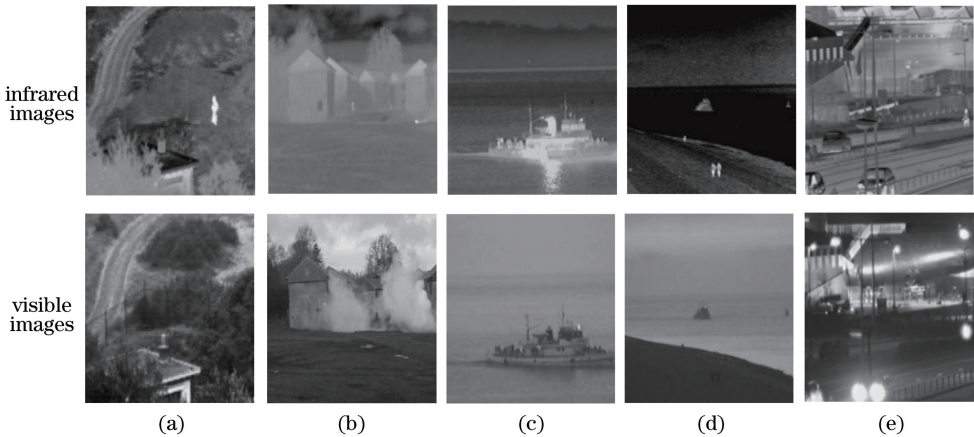


图 4 测试中所用到的 5 组红外与可见光图像。(a) Un camp; (b) Octec; (c) Steamboat; (d) Kayak; (e) Infrared4-Visible4
Fig. 4 Five sets of infrared and visible images used in test. (a) Un camp; (b) Octec; (c) Steamboat; (d) Kayak; (e) Infrared4-Visible4

图 5 是以“Un camp”图像作为输入图像,采用 9 种不同融合方法得到的融合图像。从图 5 可以看到,图 5(a)~(e)中树的辨识度相对较低。图 5(a)~(c)的房顶区域出现原始图像中不存在的斑点,以及人像区域出现了明显的光晕现象。

图 5(a)中有很明显的伪 Gibbs 效应。尽管图 5(d)中整体效果相对于其他图像效果较好,但此图像的对比度相对较低。图 5(e)中出现了很多明显的噪声。图 5(f)中没有显示左侧树枝的纹理信息。图 5(g)中整幅图像纹理信息模糊,尤其是在山体

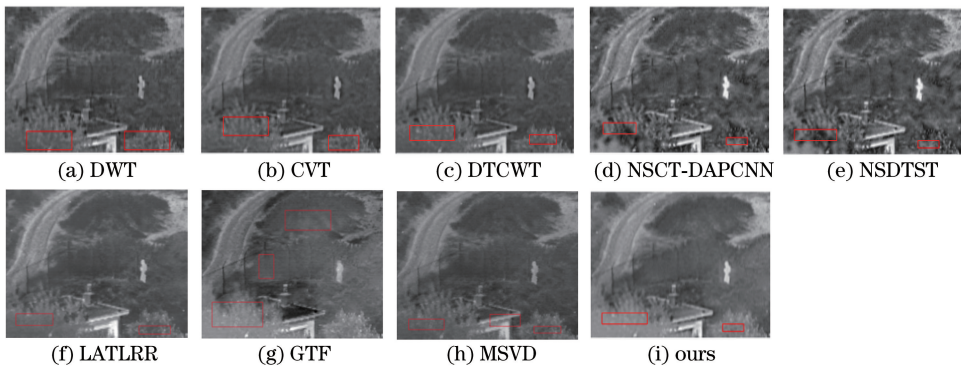


图5 “Un camp”图像在不同融合方法下获得的融合结果

Fig. 5 Fusion results obtained by different fusion methods applied in “Un camp” image

和树枝部分。图5(h)中整幅图像亮度较低。相比之下,图5(i)在“Un camp”图像的目标位置处(人像)没有出现光晕现象,且整幅图像没有出现人工效应,可以很好地将红外图像的光谱特征注入到融合图像中,同时也保留了可见光图像的场景和细节信息。

图6是以“Octec”作为输入图像,采用9种不同融合方法得到的融合图像。在图6(a)~(e)中,地面的道路痕迹(即每幅图第3个方框)均没有所提方法表现明显,且在图6(a)中出现了伪Gibbs效应。在中间方框处(人物与后边墙体部分),虽然人物表示出来了,但在原始图像中位于中间两座相邻的房子在图6中基本看不出来。在图6(b)~(e)中树枝

与天空的接壤部分,红外光谱信息的不适当注入导致此部分视觉效果非常不自然;同时还可以看出这4幅图像的对比度均不高;可见光图像中天空与云彩之间的对比度较低,致使图像不符合人类视觉感知。图6(e)中出现了很多噪声。图6(f)中可见光图像没有合理融入,导致天空部分的亮度较暗,与原图不符。图6(g)中红外背景信息的过多融入,以及红外目标信息的不合理提取,导致天空部分较暗,且人物目标不突出,融合效果极差。图6(h)所示的融合图像的整体亮度偏暗且模糊,人物目标不突出。相比之下,图6(i)所示的融合图像很清晰地显示了人物、房子、路痕以及树枝与天空接壤的区域,且此图像与人眼视觉相符,呈现的效果更自然。

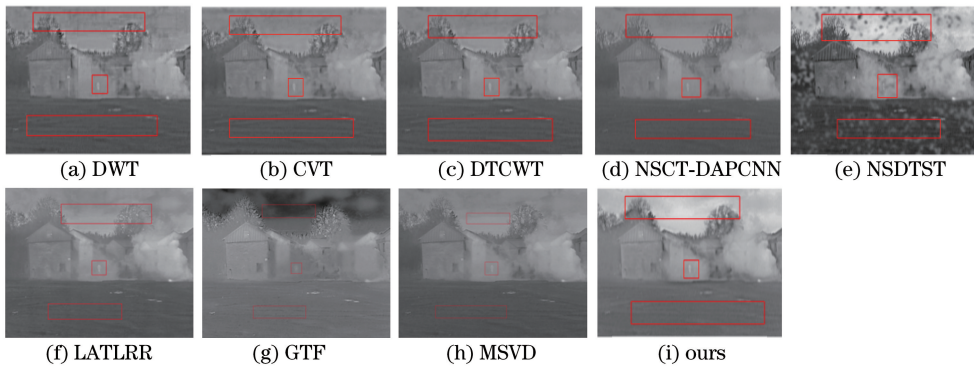


图6 “Octec”图像在不同融合方法下获得的融合结果

Fig. 6 Fusion results obtained by different fusion methods applied in “Octec” image

图7是以“Steamboat”作为输入图像,采用9种不同融合方法得到的融合图像。图7(a)~(e)所示的融合图像均不能很清晰地显示方框处栏杆所对应的窗口(中间方框)以及船身所对应的窗口(两边方框),且由于图像的亮度失真,图片中不能很好地表示出原始红外图像在海面上所反射的灯光,尤其是图7(d)和图7(e)。图7(a)中出现了伪Gibbs效应,在图7(e)中出现了噪声。图7(f)中船身的亮度较

暗。图7(g)中可见光图像信息的丢失,导致天空部分的融合极不自然,且船体的目标信息较为模糊。图7(h)中同样出现了亮度信息较低,整体图像信息模糊的现象。相比之下,图7(i)不仅可以清晰地看到方框中的目标,且保留了原始红外图像中的亮度信息,整体效果更为自然。因此所提方法可以将更多的红外图像信息融入到可见光图像中,得到符合人眼视觉感知的图像。

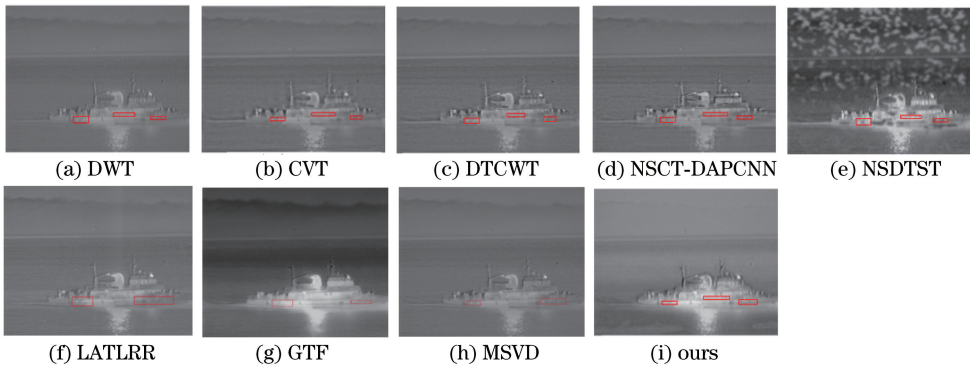


图 7 “Steamboat”图像在不同融合方法下获得的融合结果

Fig. 7 Fusion results obtained by different fusion methods applied in “Steamboat” image

图 8 是以“Kayak”作为输入图像,采用 9 种不同融合方法得到的融合图像。图 8(a)~(e)中船体四周出现了明显的人工效应;所注入的红外图像信息过多,导致在图 8(a)~(e)中海面的亮度以及天空的亮度过低,使得海面与沙滩的亮度过于接近,致使图像显示不符合人类的视觉感知。除此之外,图 8(e)中出现了明显的噪声。图 8(f)中

整体亮度较暗。图 8(g)中可见光信息丢失严重,整幅图像的识别度不高。图 8(h)中目标人物出现伪 Gibbs 效应。相比于其他几幅图像,图 8(i)的视觉效果较好,不仅把可见光图像中天空与海面的亮度展现出来,而且在船体周围也没有出现人工效应,使得最终的融合图像感观自然,与人类的视觉感知相吻合。

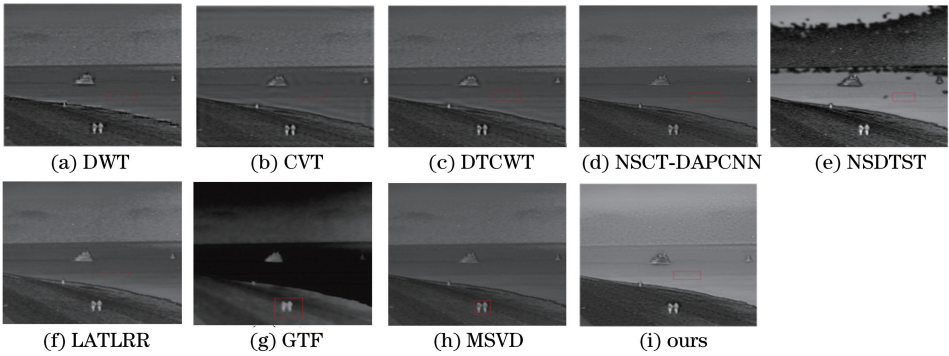


图 8 “Kayak”图像对在不同融合方法下获得的融合结果

Fig. 8 Fusion results obtained by different fusion methods applied in “Kayak” image

图 9 是以“Infrared4-Visible4”作为输入图像,采用 9 种不同融合方法得到的融合图像。在图 9(a)~(e)中,路灯(左上方的方框)和院子(右下

方的方框)处的整体亮度不高。这是因为原始可见光图像信息没有被充分地融入到融合图像中。除此之外,在图 9(a)中也出现了伪 Gibbs 效应,

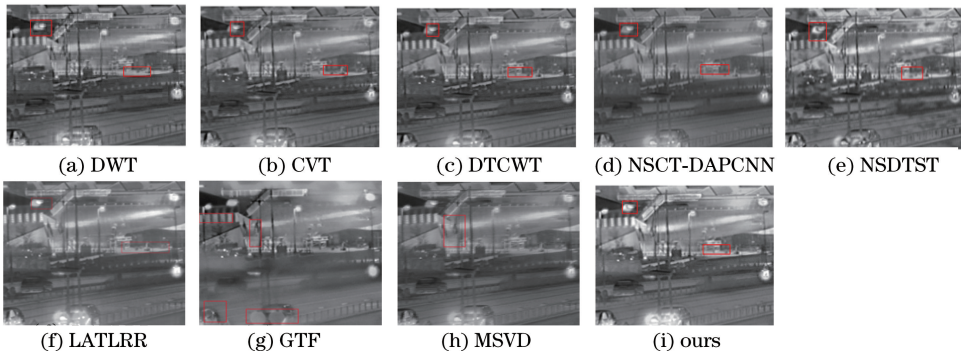


图 9 “Infrared4-Visible4”图像对在不同融合方法下获得的融合结果

Fig. 9 Fusion results obtained by different fusion methods applied in “Infrared4-Visible4” image

图 9(e)中出现了噪声。图 9(f)中整体信息较为模糊。图 9(g)中整体图像的红外信息丢失严重且整体模糊度高,而在路灯及人物部分,由于可见光信息的过多注入,目标轮廓不清晰。图 9(h)中整体亮度偏低,图像严重失真。而本文方法得到的图 9(i)更能充分地反映上述信息,且将原始图像中所包含的信息都表现出来,视觉效果更好。

4.2 融合图像的客观评价

本文选用互信息(MI)^[27]、信息熵(IE)^[28]、标准差^[29]、非线性相关信息熵(Q_{NCE})^[30]和 Chen-

Varshney 指标(Q_{CV})^[31]为客观评价指标。互信息是用于测量从原始图像转移到融合图像的信息数量。信息熵是用于测量融合图像中所包含的信息数量。标准差用于反映图像灰度相对于灰度平均值的离散情况。非线性相关信息熵则用于测量输入图像与融合图像的非线性关联度。Chen-Varshney 指标是基于人类视觉系统,用输入图与融合图之间的不同视觉来评估融合结果的好坏。前 4 种评价指标的所测值越大,说明融合结果越好;最后一个指标所测值越小,说明融合结果越好。

表 2 不同融合方法的客观评价指标

Table 2 Objective evaluation indicators of different fusion methods

Image	Index	DWT	CVT	DTCWT	NSCT-DAPCNN	NSDTST	LATLRR	GTF	MSVD	ours
Un camp	MI	0.1916	0.1893	0.2011	0.2072	0.1802	0.2410	0.2793	0.2165	0.2841
	STD	0.1029	0.0988	0.0974	0.0718	0.1332	0.1070	0.1051	0.0892	0.1152 (2)
	Q_{NCE}	0.8029	0.8029	0.8030	0.8030	0.8029	0.8034	0.8041	0.8031	0.8039 (2)
	Q_{CV}	425.1803	461.3446	426.053	423.8996	540.9333	332.3001	661.3364	424.1480	293.6759
	IE	0.4709	0.4685	0.4560	0.4580	0.6132	0.6268	0.6026	0.4500	0.8334
Octec	MI	0.3876	0.3758	0.4127	0.4283	0.3838	0.4472	0.5176	0.4628	0.5200
	STD	0.1167	0.1175	0.1136	0.1117	0.2132	0.1175	0.0959	0.1080	0.1573 (2)
	Q_{NCE}	0.8066	0.8064	0.8070	0.8071	0.8073	0.8073	0.8089	0.8075	0.8094
	Q_{CV}	648.2625	408.1675	629.2341	626.6379	391.1703	274.6485	2264.0	1902.5	192.0751
	IE	0.9613	0.9645	0.9672	0.9660	0.9635	0.9934	0.8337	0.8298	1.0000
Steamboat	MI	0.2576	0.2418	0.2695	0.2738	0.2911	0.2908	0.6284	0.2965	0.4514 (2)
	STD	0.0524	0.0482	0.0479	0.0449	0.124	0.0498	0.1331	0.0415	0.0769 (3)
	Q_{NCE}	0.8045	0.8043	0.8046	0.8046	0.8052	0.8047	0.8117	0.8047	0.8068 (2)
	Q_{CV}	163.4563	169.6701	167.6022	173.5677	298.9255	208.9697	331.4260	178.0732	138.259
	IE	0.3950	0.3742	0.3710	0.3567	0.6137	0.4527	0.5750	0.3516	0.7979
Kayak	MI	0.4471	0.3786	0.4492	0.5184	0.4923	0.4696	0.5469	0.6198	0.5195(3)
	STD	0.0731	0.0728	0.0696	0.0669	0.1707	0.0673	0.0864	0.0647	0.1444 (2)
	Q_{NCE}	0.8059	0.8049	0.8058	0.8070	0.8070	0.8061	0.8072	0.8087	0.8087
	Q_{CV}	876.5883	622.103	864.7246	791.0015	798.335	855.5583	3060.6	697.9614	135.6981
	IE	0.0279	0.0224	0.0207	0.0132	0.7651	0.0262	0.0191	0.0071	0.9980
Infrared4-Visible4	MI	0.2485	0.2461	0.2866	0.2754	0.2552	0.2943	0.2390	0.2852	0.3233
	STD	0.1465	0.1366	0.1364	0.1289	0.1697	0.1432	0.1330	0.1243	0.1707
	Q_{NCE}	0.8035	0.8034	0.8044	0.8038	0.8036	0.8040	0.8034	0.8039	0.8046
	Q_{CV}	1133.9	1056.2	1502.3	1076.9	1845.6	924.2383	1972.3	1045.83	718.6361
	IE	0.7643	0.7588	0.7511	0.7474	0.8973	0.8745	0.8450	0.7364	0.9504

表 2 是不同融合方法所得到的客观评价指标,最优值已用粗体表示,括号中数字表示所提方法排名情况。由表 2 可知,对于图像“Un camp”,本文方法在 MI、IE 和 Q_{CV} 3 个评价指标上是最优的,在 STD 和 Q_{NCE} 评价指标上排名第 2。对于图像“Octec”,其在 STD 评价指标上是次优,其他评价指标均为最优。对于图像“Steamboat”,其在 IE 和 Q_{CV} 评价指标上是最优,在 MI 和 Q_{NCE} 评价指标上则排名第 2。对于图像“Kayak”,其在 IE、 Q_{NCE} 和

Q_{CV} 评价指标上是最优,在 STD 评价指标上排名第 2,在 MI 评价指标上排名第 3。对于图像“Infrared4-Visible4”,其在所有评价指标上都是最优。与 DWT、CVT、DTCWT、NSCT-DAPCNN 和 LATLRR 5 种方法相比,本文方法在所有客观评价指标上均为最优的。在此,与 NSDTST、GTF 和 MSVD 3 种方法相比,本文方法在 EI 和 Q_{CV} 评价指标上是最优,偶尔在某一幅测试图中出现 MI 或 STD 或 Q_{NCE} 评价指标没有达到最优值的情况;然

而,这3种方法在主观评价中均有不同程度的噪声和模糊现象出现,所产生的视觉效果不符合人类视觉感知。整体上讲,与这8种方法相比,本文方法具有更好的融合效果。

为了更直观地分析本文方法在客观评价指标上的优越性,使用统计折线图(图10)进一步分析单个客观评价指标。图10(a)中本文方法在“Un camp”“Oetec”“Infrared4-Visible4”3幅图中对应的MI最大,说明本文方法可以从这3幅图像的原始图中提取更多信息到融合图中。图10(b)中本文方法在“Oetec”“Kayak”“Infrared4-Visible4”3幅图中对应的 Q_{NCIE} 最大,这说明相对于这3幅图像,本文方法的融合图与原始图之间的非线性关联度最大。图

10(c)中本文方法的IE均高于其他8种方法,这说明本文方法产生的融合图像包含了尽可能多的输入图像的原始信息。图10(d)中本文方法的STD在“Infrared4-Visible4”图像中达到最大值,在“Un camp”“Oetec”“Kayak”3幅图中,本文算法的STD值略低于NSDTST算法,在“Steamboat”图像中,本文算法的STD值略低于NSDTST和GTF算法,然而NSDTST和GTF算法的融合图均出现了噪声和模糊现象,因此,本文方法可以尽可能多地展示图像的灰度分散情况,且融合效果更好。从图10(e)可以看出,对于每一幅图,本文所提方法的 Q_{CV} 值最小,说明本文方法产生的融合图与原始图的差异最小,融合效果最好。

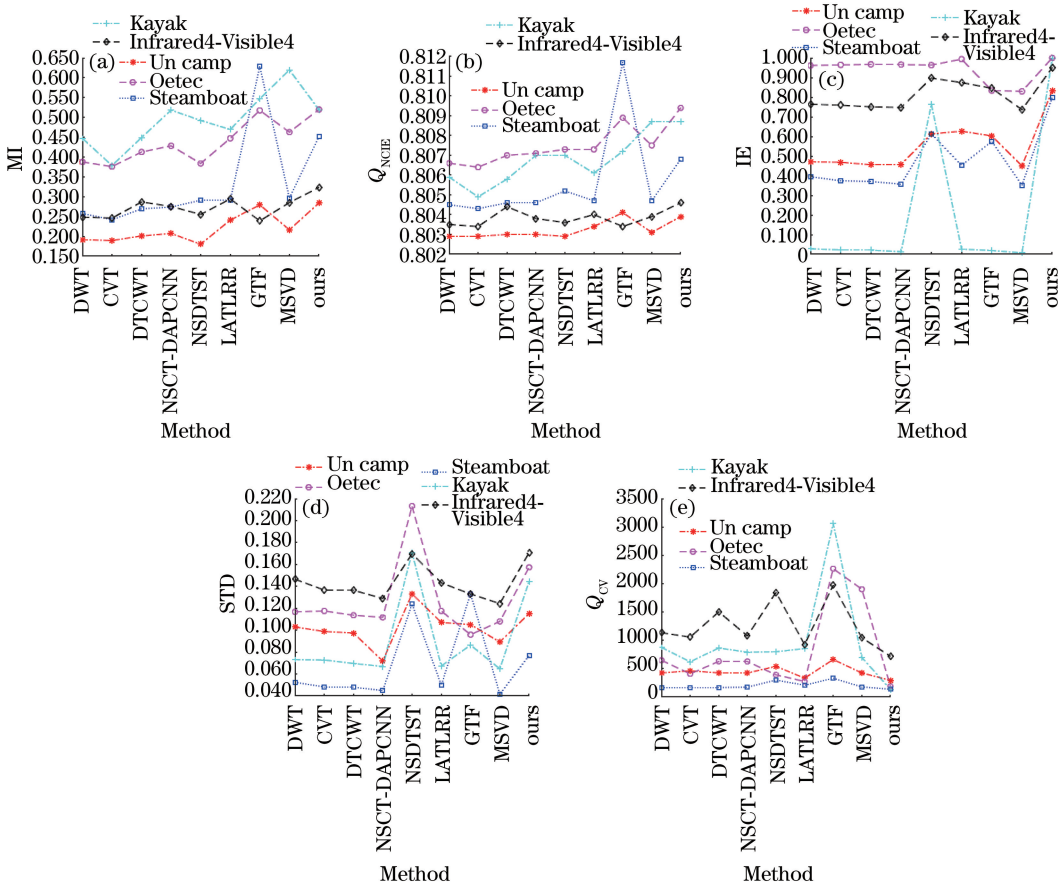


图10 不同客观评价指标的统计折线图

Fig. 10 Statistical polygonal images of different objective evaluation indicators

上述分析基于客观评价指标分析所提融合方法的优势。除此之外,本文选择2幅测试图像,并将其原始红外与可见光图像以及融合图像的三维数据图作比较,以此来进一步说明所提方法对图像信息融合的效果,如图11和图12所示。在这2幅图中,分别按红外图、可见光图和融合图的顺序进行排列。其中X和Y轴表示像素位置,Z轴表示像素灰度

值。在每幅图中可以观察到融合图像的数据图包含了红外与可见光图像的相关信息,例如在图11的框图处,以及图12的框图处,均可很明显地观察到红外图像目标信息与可见光图像背景信息融合到一起。因此,综合图10~12的研究结果,发现本文所提方法无论是在客观指标还是在整体性能上均具有较好的融合效果。

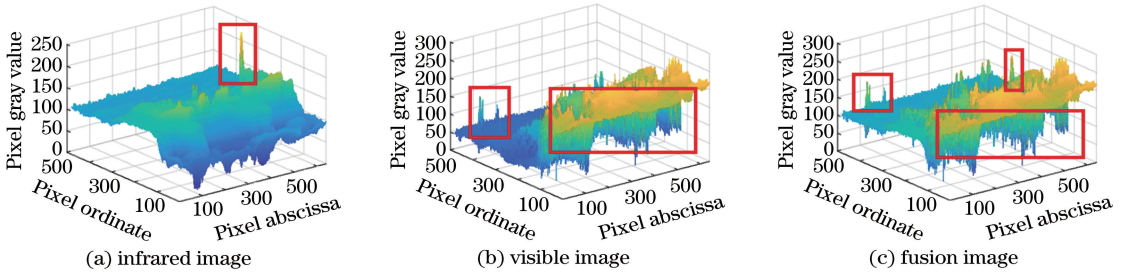


图 11 “Octec”的三维数据图

Fig. 11 Three-dimensional data images of “Octec”

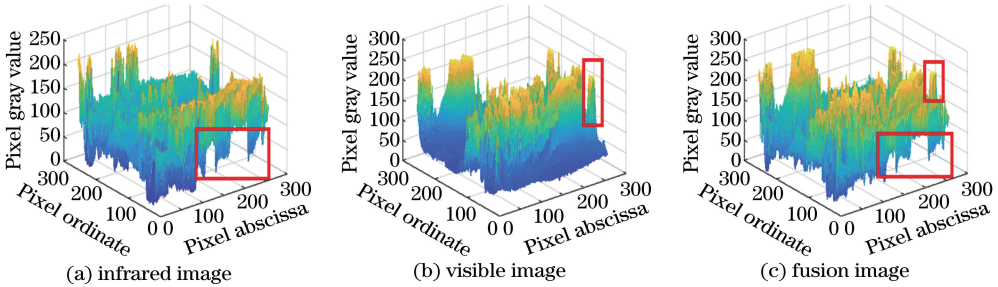


图 12 “Infrared4-Visible4”的三维数据图

Fig. 12 Three-dimensional data images of “Infrared4-Visible4”

4.3 运行时间比较

表 3 中给出了本文方法与其他 8 种方法的运行时间,由表 3 可知,本文方法的运行时间优于 NSCT-DAPCNN、NSDTST 和 LATLRR 这 3 种方法,排名第

6。由于混合多尺度分解和 DAPCNN 融合规则的实现相对耗时,本文方法的运行时间要多于其他 5 种方法。为了得到好的融合效果,并使其具有更高的质量和更广泛的应用价值,增加运行时间的代价是值得的。

表 3 所提方法与其他 8 中方法的运行时间

Table 3 Running-time of eight other methods and proposed method

Images	DWT	CVT	DTCWT	NSCT-DAPCNN	NSDTST	LATLRR	GTF	MSVD	ours
Un camp	0.6707	2.1643	0.5408	41.7589	101.5510	227.6768	1.5259	0.1645	19.0259 (6)
Octec	0.2924	1.8312	0.4142	139.7033	326.4207	218.3082	7.8089	0.5087	58.0978 (6)
Steamboat	0.2856	1.8511	0.4191	118.7886	267.4719	251.7716	7.1331	0.4355	49.8169 (6)
Kayak	0.2777	1.8518	0.3950	117.9741	257.4367	245.7336	6.2976	0.6104	49.4241 (6)
Infrare4-Visible4	0.3054	1.8840	0.4115	27.8120	67.3537	220.7137	1.1274	0.2653	11.6429 (6)

5 结 论

提出了一种基于滚动导向滤波的混合多尺度图像融合算法,使用文中所提出的基于滚动导向滤波的混合多尺度分解方法将红外图像与可见光图像分解到不同的尺度层,再选用不同的融合规则对不同尺度层进行融合,最后使用逆混合多尺度分解方法得到最终的融合图像。实验结果表明,无论是在主观分析还是在客观分析,所提方法相比于其他方法均能够得到更好的融合效果,且融合结果更符合人类的视觉感知。

参 考 文 献

[1] Xu H, Wang Y, Wu Y J, *et al.* Infrared and multi-

type images fusion algorithm based on contrast pyramid transform [J]. *Infrared Physics & Technology*, 2016, 78: 133-146.

[2] Daubechies I. The wavelet transform, time-frequency localization and signal analysis [J]. *IEEE Transactions on Information Theory*, 1990, 36(5): 961-1005.

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

[4] Wu Z G, Wang Y J. Image fusion algorithm using curvelet transform based on the edge detection [J]. *Optical Technique*, 2009, 35(5): 682-685, 690.

武治国, 王延杰. 一种基于边缘的曲波变换图像融合

- 方法[J]. 光学技术, 2009, 35(5): 682-685, 690.
- [5] Cai W, Li M, Li X Y. Infrared and visible image fusion scheme based on contourlet transform[C] // 2009 Fifth International Conference on Image and Graphics, September 20-23, 2009, Xi'an, Shaanxi, China. New York: IEEE, 2009: 516-520.
- [6] Aishwarya N, Bennila Thangammal C. Visible and infrared image fusion using DTCWT and adaptive combined clustered dictionary[J]. Infrared Physics & Technology, 2018, 93: 300-309.
- [7] 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.
- [8] Burt P J, Adelson E H. The Laplacian pyramid as a compact image code [J]. IEEE Transactions on Communications, 1983, 31(4): 532-540.
- [9] Petrovic V S, Xydeas C S. Gradient-based multiresolution image fusion[J]. IEEE Transactions on Image Processing, 2004, 13(2): 228-237.
- [10] Toet A. Image fusion by a ratio of low-pass pyramid [J]. Pattern Recognition Letters, 1989, 9(4): 245-253.
- [11] Adu J H, Gan J H, Wang Y, *et al.* Image fusion based on nonsubsampling contourlet transform for infrared and visible light image[J]. Infrared Physics & Technology, 2013, 61: 94-100.
- [12] Wu Y Q, Tao F X. Multispectral and panchromatic image fusion based on improved projected gradient NMF in NSST domain[J]. Acta Optica Sinica, 2015, 35(4): 0410005.
吴一全, 陶飞翔. 改进投影梯度 NMF 的 NSST 域多光谱与全色图像融合[J]. 光学学报, 2015, 35(4): 0410005.
- [13] Zhou Z Q, Wang B, Li S, *et al.* Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters [J]. Information Fusion, 2016, 30: 15-26.
- [14] Jiang Z T, Wu H, Zhou X L. Infrared and visible image fusion algorithm based on improved guided filtering and dual-channel spiking cortical model[J]. Acta Optica Sinica, 2018, 38(2): 0210002.
江泽涛, 吴辉, 周晓玲. 基于改进引导滤波和双通道脉冲发放皮层模型的红外与可见光图像融合算法[J]. 光学学报, 2018, 38(2): 0210002.
- [15] Zhang Q, Shen X Y, Xu L, *et al.* Rolling guidance filter[M] // Zhang Q, Shen X Y, Xu L, *et al.* Computer vision-ECCV 2014. Lecture notes in computer science. Cham: Springer, 2014, 8691: 815-830.
- [16] Cheng B Y, Jin L X, Li G N. Adaptive fusion framework of infrared and visual image using saliency detection and improved dual-channel PCNN in the LNSST domain[J]. Infrared Physics & Technology, 2018, 92: 30-43.
- [17] Wang J J, Li Q, Jia Z H, *et al.* A novel multi-focus image fusion method using PCNN in nonsubsampling contourlet transform domain[J]. Optik, 2015, 126(20): 2508-2511.
- [18] Ma J L, Zhou Z Q, Wang B, *et al.* Infrared and visible image fusion based on visual saliency map and weighted least square optimization [J]. Infrared Physics & Technology, 2017, 82: 8-17.
- [19] Li S T, Yang B, Hu J W. Performance comparison of different multi-resolution transforms for image fusion[J]. Information Fusion, 2011, 12(2): 74-84.
- [20] Nencini F, Garzelli A, Baronti S, *et al.* Remote sensing image fusion using the curvelet transform[J]. Information Fusion, 2007, 8(2): 143-156.
- [21] Sruthy S, Parameswaran L, Sasi A P. Image fusion technique using DT-CWT [C] // 2013 International Mutli-Conference on Automation, Computing, Communication, Control and Compressed Sensing (IMac4s), March 22-23, 2013, Kottayam, India. New York: IEEE, 2013: 160-164.
- [22] Zhang B H, Lu X Q, Jia W T. A multi-focus image fusion algorithm based on an improved dual-channel PCNN in NSCT domain[J]. Optik, 2013, 124(20): 4104-4109.
- [23] Yin M, Duan P H, Liu W, *et al.* A novel infrared and visible image fusion algorithm based on shift-invariant dual-tree complex shearlet transform and sparse representation [J]. Neurocomputing, 2017, 226: 182-191.
- [24] Li H, Wu X J. Infrared and visible image fusion using latent low-rank representation[EB/OL]. (2018-12-18) [2018-12-24]. <https://arxiv.org/abs/1804.08992?context=cs>.
- [25] Ma J Y, Chen C, Li C, *et al.* Infrared and visible image fusion via gradient transfer and total variation minimization [J]. Information Fusion, 2016, 31: 100-109.
- [26] Naidu V P S. Image fusion technique using multi-resolution singular value decomposition[J]. Defence Science Journal, 2011, 61(5): 479-484.
- [27] Qu G H, Zhang D L, Yan P F. Information measure

- for performance of image fusion [J]. *Electronics Letters*, 2002, 38(7): 313-315.
- [28] Feng X, Zeng Z M, Feng H, *et al.* A method for evaluating the disturbance in distributed vibration sensor based on wavelet information entropy [J]. *Acta Optica Sinica*, 2013, 33(11): 1106005.
冯欣, 曾周末, 封皓, 等. 基于小波信息熵的分布式振动传感系统的扰动评价方法[J]. *光学学报*, 2013, 33(11): 1106005.
- [29] Balasubramaniam P, Ananthi V P. Image fusion using intuitionistic fuzzy sets [J]. *Information Fusion*, 2014, 20: 21-30.
- [30] Wang H D, Yao X. Objective reduction based on nonlinear correlation information entropy [J]. *Soft Computing*, 2016, 20(6): 2393-2407.
- [31] Chen H, Varshney P K. A human perception inspired quality metric for image fusion based on regional information [J]. *Information Fusion*, 2007, 8(2): 193-207.