

引用格式: SUN Jiping, FAN Weiqiang. Mine Dual-band Image Fusion in MS-ADoG Domain Combined with ReNLU and VGG-16[J]. Acta Photonica Sinica, 2022, 51(3):0310002

孙继平, 范伟强. MS-ADoG 域结合 ReNLU 与 VGG-16 的矿井双波段图像融合算法[J]. 光子学报, 2022, 51(3):0310002

MS-ADoG 域结合 ReNLU 与 VGG-16 的矿井 双波段图像融合算法

孙继平, 范伟强

(中国矿业大学(北京)机电与信息工程学院, 北京 100083)

摘 要: 为满足矿井视频监控需求, 针对现有图像融合算法获取的融合图像存在伪目标、模糊目标、晕光遮挡目标等问题, 提出了一种基于多尺度和自适应高斯差分变换, 结合修正非线性单元和 VGG-16 的矿井双波段图像融合算法。设计了基于多尺度和自适应高斯差分变换的源图像分解模型, 将红外和可见光图像分解为基础图像和细节图像; 构造了一种修正非线性单元函数, 使红外基础图像的权值随可见光基础图像的灰度自动调节, 并采用“加权平均”的基础图像融合策略, 以获得消除光源干扰的融合基础图像; 利用预训练的 VGG-16 网络模型提取细节图像的 4 层网络深度特征, 获取 4 对具有不同网络深度特征的融合图像后, 通过“极大值选择”方法得到融合细节图像; 对融合基础图像和融合细节图像进行重构, 获取最终的融合图像。实验结果表明: 所提算法能够消除矿井下人工光源干扰, 得到场景清晰、特征显著的融合图像, 更符合人的视觉特性; 同时, 改善了异源图像的融合效率, 有利于图像的进一步分析处理。与其他典型算法相比, 该算法鲁棒性好, 克服了传统算法难以提取图像深度特征的缺点, 更易于消除光源干扰, 获得更加全面、可靠和丰富的场景信息。

关键词: 视频监控; 图像融合; 自适应高斯差分; 修正非线性单元; 神经网络; 深度特征; 池化反运算

中图分类号: TP391.41

文献标识码: A

doi: 10.3788/gzxb20225103.0310002

0 引言

煤炭是我国最主要的一次能源^[1]。煤炭行业是高危行业, 瓦斯、水、火、冲击地压、顶板、机电、运输等事故困扰着煤矿安全生产^[2-3]。减少煤矿井下作业人员是煤矿安全生产的有效措施, 是现代化煤矿的发展方向^[4]。视频监控是煤矿井下少人和无人作业的必然选择^[5-6]。矿用可见光(Visible, VIS)视频监控的图像具有分辨率高、色彩信息丰富的优点, 但易受到人工光源(如矿灯、巷道灯、车灯等)干扰^[3,7], 导致图像探测器出现“光致饱和”和“图像串扰”, 进而造成可见光图像中出现大面积亮斑和晕光等现象^[8-9]。矿用远红外(Far Infrared, FIR)视频监控的图像抗干扰能力强, 易于捕获暗环境或遮蔽后的热目标^[10], 但存在色彩单一、细节信息严重缺失、视觉效果不理想等问题^[3,7]。基于 VIS 和 FIR 图像优势互补的异源图像融合技术, 可有效提高低照度或复杂光照环境下目标与背景的信息^[11-12], 但无法消除矿井可见光图像的大面积亮斑和晕光。因此, 解决视频监控中的伪目标、模糊目标、光晕掩藏目标等问题, 提升视频图像质量, 是矿井图像智能分析和地面远程监视亟待解决的问题。

近年来, 国内外学者对不同谱段的图像融合技术进行了研究^[13-14], 并且在遥感监测、国防建设、工业视觉、医学诊断等领域得到了广泛应用^[15], 但其在矿井视频监控的图像应用方面却鲜有报道。目前, 图像融合方法主要包括基于空间域^[16]、变换域^[16]、稀疏表示^[17]、深度学习^[18]的图像融合。其中, 基于空间域的图像融合策略, 如: Independent Component Analysis(ICA)^[19]、Intensity Hue Saturation(IHS)^[20]、YUV^[21]、Brovey

基金项目: 国家重点研发计划项目(No. 2016YFC0801800)

第一作者: 孙继平(1958—), 男, 教授, 博士, 主要研究方向为矿井监控与监视、通信与定位。Email: sjp@cumtb.edu.cn

通讯作者: 范伟强(1992—), 男, 博士研究生, 主要研究方向为矿井监视与机器视觉。Email: fan_weiqiang@163.com

收稿日期: 2021-10-08; 录用日期: 2021-12-09

<http://www.photon.ac.cn>

Transform^[22]、Gram-Schmidt^[23]算法等,执行效率高,能够保留红外目标信息及可见光细节特征,但在融合过程中易引入红外热噪声并出现伪影现象。基于变换域的图像融合策略中,已广泛采用的DWT、Contourlet变换、Curvelet变换不具有平移不变性^[24],导致融合图像易出现伪吉布斯现象,影响图像的视觉效果。Non-Subsampled Contourlet Transform(NSCT)、Non-Subsampled Shearlet Transform(NSST)、非下采样双树复轮廓波变换等具有平移不变性,但存在图像融合过程耗时长、运算效率低等缺陷^[11,25-26]。基于稀疏表示的图像融合需要根据先验知识构造过完备字典和稀疏系数表示模型,计算过程复杂,模型具有较大局限性^[17-18]。基于深度学习的图像融合具有提取源图像深层特征的优点,但存在卷积网络层数多,深度特征提取效率低,以及特征融合过程复杂等缺点。

针对上述图像融合方法存在的问题,本文构建了基于多尺度和自适应高斯差分(Multi-scale and Adaptive Difference of Gaussian, MS-ADoG)的源图像分解模型,并通过该分解模型将源图像分解为基础图像和细节图像,设计了基于修正非线性单元(Rectified Non-Linear Units, ReNLU)的基础图像融合规则,使FIR基础图像的权值随VIS基础图像的灰度自动调节;采用预训练的VGG-16神经网络模型提取细节图像的多层深度特征,并根据不同层的深度特征构造对应的显著图,利用池化反运算和加权融合策略,获取具有不同深度特征的融合图像,并通过“极大值选择”策略得到融合细节图像;将融合的基础图像和细节图像进行重构,得到消除光源干扰和场景清晰、特征显著的融合图像。

1 图像融合算法基础

基于变换域的融合图像具有细节信息丰富、视觉效果好等优点,实际应用较多。该融合算法首先采用变换函数对FIR和VIS图像进行多尺度分解,接着设计不同的融合策略对各层次的子图像进行融合,最后利用逆变换函数获取融合图像。基于高斯拉普拉斯变换的FIR与VIS图像融合框架如图1所示。

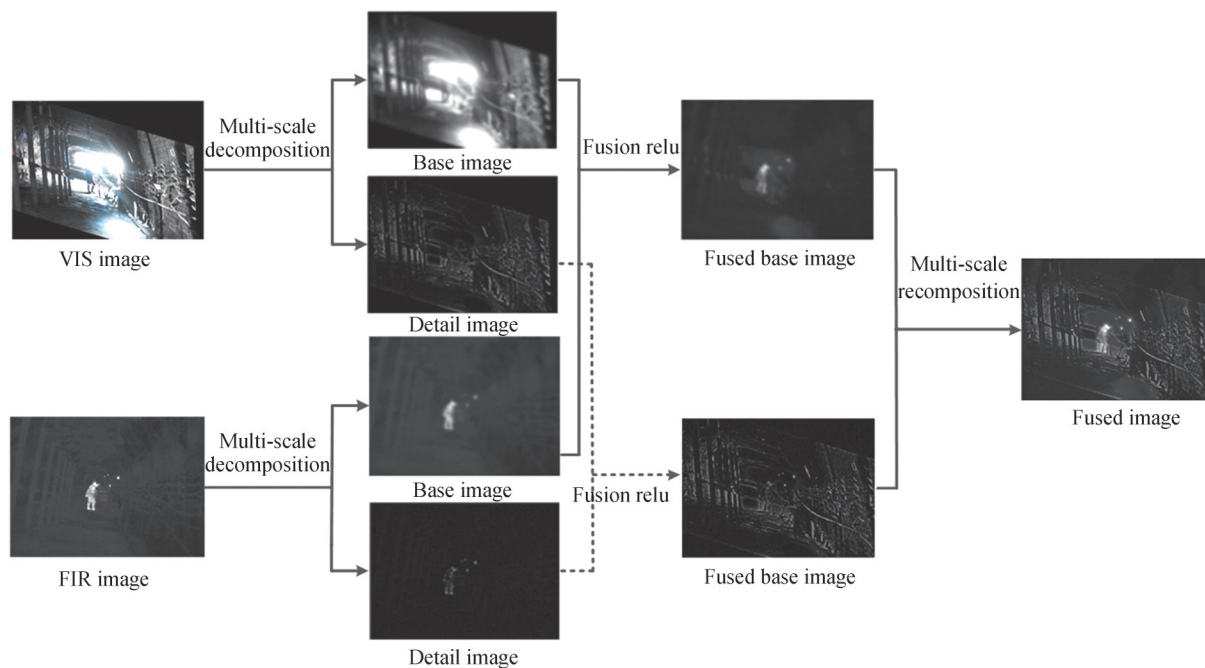


图1 基于高斯拉普拉斯变换的FIR与VIS图像融合框架

Fig. 1 Framework of FIR and VIS image fusion based on Laplace of Gaussian transform

基于深度学习的图像融合技术以FIR和VIS图像的深度特征作为一种显著特征,用于重构融合图像。首先,利用卷积神经网络获取FIR和VIS图像的深度特征;接着,采用特定的融合规则对同一深度的显著特征进行融合^[27];最后,对不同深度的融合特征图进行重构,获取融合图像。采用预训练的VGG-16神经网络进行图像融合,对应的图像融合框架如图2所示。

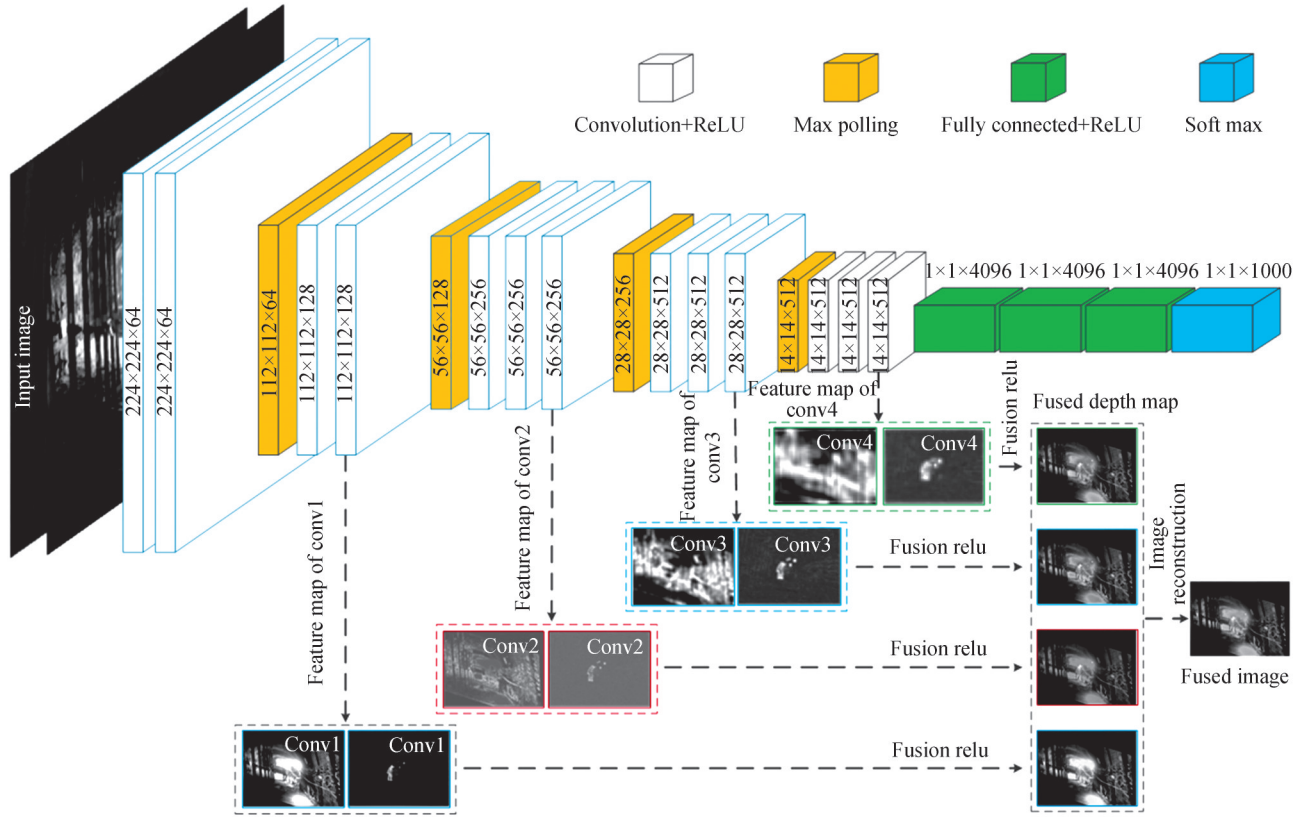


图2 基于VGG-16的FIR与VIS图像融合框架

Fig.2 Fusion framework of FIR and VIS image based on VGG-16

2 双波段图像融合算法

2.1 基于MS-ADoG的源图像分解模型

根据高斯拉普拉斯理论^[28]和矿井复杂光照环境下的成像特点^[29],通过高斯核函数对FIR和VIS源图像 I_k ($k=1$ 时,表示FIR图像; $k=2$ 时,表示VIS图像)进行多尺度模糊变换,并且构造MS-ADoG函数替代高斯拉普拉斯函数,使其获取最稳定的基础图像 I_k^b 和细节图像 I_k^d ,以及实现对VIS图像中光源干扰区域与细节特征分离的目的。2D源图像的高斯模糊变换定义为

$$I_k^b(x, y, \sigma) = G(x, y, \sigma) * I_k(x, y) \quad (1)$$

式中, $I_k^b(x, y, \sigma)$ 为源图像 I_k 经过高斯模糊变换后的基础图像, $G(x, y, \sigma)$ 为模糊度可调的高斯核函数, σ 为高斯核函数的标准差,其值越大,基础图像的模糊度越高, $I_k(x, y)$ 为源图像 I_k 在位置 (x, y) 处的像素值,*表示卷积。

假设2D高斯模板大小为 $(2r+1) \times (2r+1)$,则模板上像素 $I_k(x, y)$ 对应的高斯核函数为

$$G(x, y, \sigma) = \sum_{\beta=-r}^r \sum_{\alpha=-r}^r \exp \left[-\frac{[(x+\beta)^2 + (y+\alpha)^2]}{2\sigma^2} \right] \quad (2)$$

式中, r 为高斯核半径,用于控制高斯模板的大小。

由式(2)可知,高斯核函数采用无差别模式模糊图像的所有边缘。当 σ 较小, r 过大时,高斯模糊后图像平滑度低,且易出现高斯模板边缘区域的值为0,造成计算资源的过度耗费;当 σ 较大, r 过小时,高斯模糊后图像平滑度高,但高斯模板易产生截断,造成 I_k^b 中包含部分细节特征。因此,正确选择参数 σ 和 r 的取值,不仅能够提高图像融合效率,而且可以产生最稳定的 I_k^b 和 I_k^d 。

经过大量试验数据分析可知,高斯核函数的标准差 σ 与核半径 r 之间存在非线性关系,通过非线性拟合获取拟合函数,即

$$\sigma = 0.0039r^2 + 0.1044r + 0.0102 \quad (3)$$

由式(1)~(3),以及不同 σ 和 r 的高斯核函数与源图像卷积,构建自适应高斯差分函数,以获取源图像的细节图像,即

$$I_k^{\bar{d}}(x, y) = I_k^b(x, y, \sigma_2) - I_k^b(x, y, \sigma_1) = [G(x, y, \sigma_2) - G(x, y, \sigma_1)] * I_k(x, y) \quad (4)$$

式中, $I_k^{\bar{d}}(x, y)$ 为采用自适应高斯差分函数得到的细节图像, σ_1, σ_2 为高斯核函数的任意标准差,分别由高斯核半径 r_1 和 r_2 确定,其中 $\sigma_1 < \sigma_2$ 。

经过自适应高斯差分函数得到的 $I_k^{\bar{d}}$,其细节特征明显,但存在边缘伪影现象,且不能很好地实现图像空间重叠特征的有效分离。为此,采用单位阶跃函数消除 $I_k^{\bar{d}}$ 的边缘伪影,并采用MS-ADoG函数进行高斯模糊变换,获取图像空间重叠特征分离的 I_k^d 。则第 n 尺度下的细节图像 $I_{k,n}^d$ 为

$$I_{k,n}^d(x, y) = I_{k,n-1}^b(x, y, \sigma_{n-1}) - I_{k,n}^b(x, y, \sigma_n) \quad (5)$$

式中, σ_{n-1}, σ_n 分别为第 $n-1$ 尺度和第 n 尺度下的高斯标准差, $I_{k,n-1}^b, I_{k,n}^b$ 分别为第 $n-1$ 尺度和第 n 尺度下的基础图像, $I_{k,n}^d$ 为第 n 尺度下的细节图像。

由式(5)得到多尺度分解和消除边缘伪影的基础图像 I_k^b 和细节图像 I_k^d ,即

$$\begin{cases} I_k^b(x, y) = I_{k,n}^b(x, y, \sigma_n) \\ I_k^d(x, y) = \sum_{n=1}^S [f_{us}(I_{k,n}^d(x, y)) I_{k,n}^d(x, y)] \end{cases} \quad (6)$$

式中, f_{us} 为单位阶跃函数^[30], S 为高斯模糊变换的尺度数。

基于MS-ADoG的源图像分解模型将配准的源图像 I_k 分解为基础图像 I_k^b 和细节图像 I_k^d (如图3所示)后,需分别采用不同的融合策略对 I_k^b 和 I_k^d 进行图像融合。由于 I_k^b 表示图像的近似分量,反映视场的概貌特征; I_k^d 表示图像的细节分量,表达图像边缘、纹理等细节信息,也是人眼识别和机器视觉最敏感的部分^[31]。因此,选择 I_k^b 和 I_k^d 的融合规则对于图像融合质量至关重要。

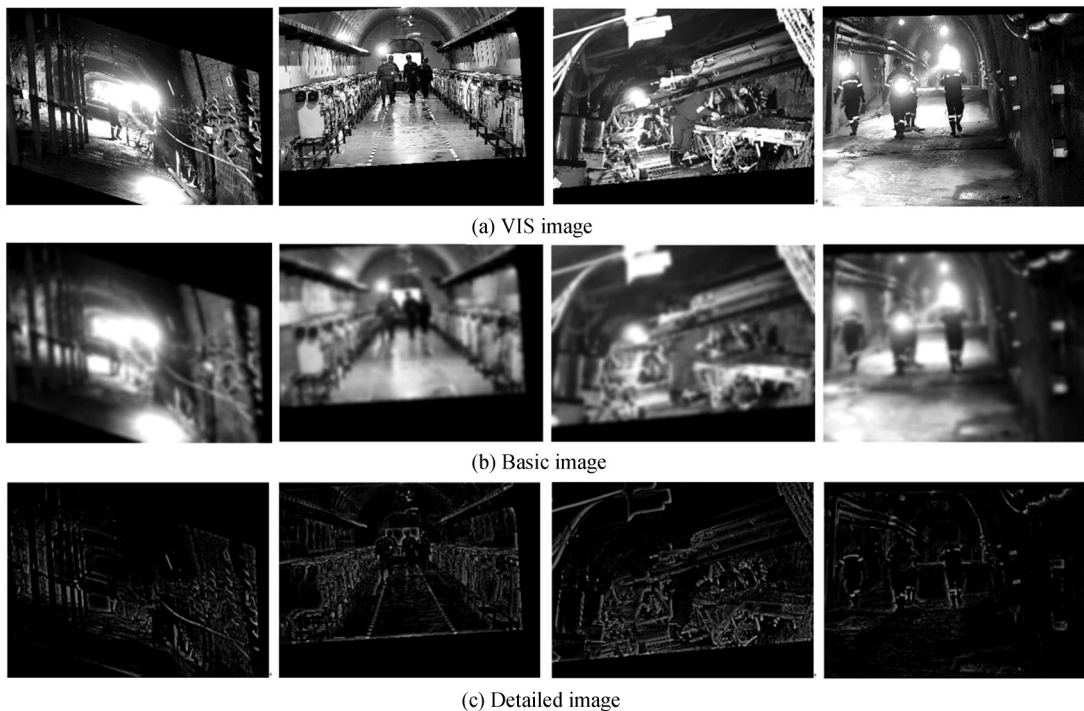


图3 VIS图像、基础图像和细节图像

Fig. 3 VIS image, basic image and detailed image

2.2 基于ReNLU的基础图像融合规则

由图3可知,矿井下VIS图像的光源干扰部分集中于可见光基础图像(VIS基础图像)中,并且在光源干

扰部分无任何有效的场景信息。为提高矿井下融合图像的整体对比度和信息丰富度,对 I_k^b 的融合策略要以消除VIS基础图像中光源干扰部分为目的。鉴于此,根据光源干扰部分的灰度明显大于非光源干扰部分,并且灰度值逐渐增大至趋近于恒常数的特点,提出了基于ReNLU函数的基础图像融合规则,使远红外基础图像(FIR基础图像)的权值随VIS基础图像的灰度自动调节。其主要融合思想为:在非光源干扰部分采用恒定权值方法实现VIS基础图像和FIR基础图像的信息融合;在光源干扰部分采用非恒定权值方法对两者进行融合,且随着干扰光源的灰度逐渐增大,FIR基础图像对应的信息所占权重逐渐增大,VIS基础图像对应的信息所占权重逐渐减小,当干扰光源中心区域对应的灰度趋于恒常数时,FIR基础图像对应的信息所占权重趋近于1,VIS基础图像对应的信息所占权重趋近于0,进而在融合后图像中消除光源干扰区域。构造的ReNLU函数为

$$W_{\text{IR}}^b(x, y) = \begin{cases} w & I_2^b(x, y) < T \\ aI_2^b(x, y)^3 + bI_2^b(x, y)^2 + cI_2^b(x, y) + d & I_2^b(x, y) \geq T \end{cases} \quad (7)$$

式中, $W_{\text{IR}}^b(x, y)$ 为FIR基础图像 I_1^b 在位置 (x, y) 处的权值, $I_2^b(x, y)$ 为VIS基础图像 I_2^b 在位置 (x, y) 处的灰度, w 为任意权值, T 为 I_2^b 光源干扰部分与非光源干扰部分的临界值, a, b, c, d 为任意常数。

由于矿井下环境光照条件复杂多变,临界值 T 为非恒常数。为此,通过OTSU算法^[32]自适应地获取临界值 T 。同时,由式(7)可知,临界值 T 将ReNLU函数划分为线性函数段和非线性函数段,为保证所构造函数的连续性和平滑性(函数可导性),ReNLU函数需满足条件

$$\begin{cases} aT^3 + bT^2 + cT + d = w \\ aI_{\text{max}}^3 + bI_{\text{max}}^2 + cI_{\text{max}} + d = 1 \\ 3aT^2 + 2bT + c = 0 \\ 3aI_{\text{max}}^2 + 2bI_{\text{max}} + c = 0 \end{cases} \quad (8)$$

式中, I_{max} 为VIS基础图像中像素的最大值。

当 w 和 I_{max} 已知时,通过式(8)求得ReNLU函数中 a, b, c, d 的数值。为了尽可能保留红外目标区域信息及可见光背景细节,并抑制可见光光源干扰区域, w 取0.5。则ReNLU函数关系如图4所示。

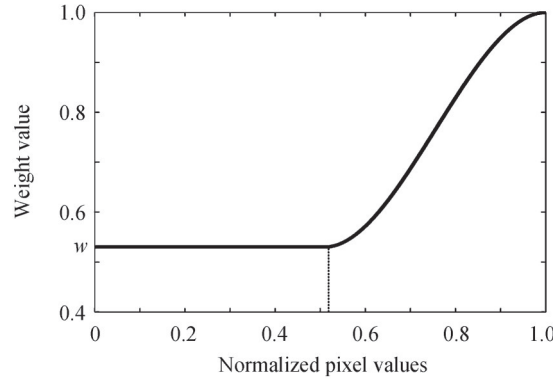


图4 权值 w 与临界值 T 的函数关系

Fig.4 The functional relationship between the weight w and the critical value T

由图4可知,当 I_2^b 小于临界值 T 时,即不存在干扰光源区域,FIR基础图像的 W_{IR}^b 等于 w ,即融合后图像能够最大程度地保留 I_1^b 和 I_2^b 的有用信息;当 I_2^b 大于临界值 T 时,即存在干扰光源区域,FIR基础图像的 W_{IR}^b 随着 I_2^b 的灰度逐渐增大直至趋近于1,即随着干扰光亮度的逐渐增大,融合后图像中 I_1^b 的信息占比越来越大, I_2^b 的信息占比越来越小,进而可有效去除光源干扰区域。进一步表明,提出的基础图像融合规则可实现FIR基础图像的权值随VIS基础图像的光源干扰强度自适应调节,进而提高融合算法的鲁棒性。

结合式(7)、(8)求得 I_2^b 的权值 $(1 - W_{\text{IR}}^b)$ 。根据 I_1^b 的权值 W_{IR}^b 与 I_2^b 的变化规律,对基础图像采用“加权平均”融合策略,得到消除 I_2^b 光源干扰部分后的融合基础图像 $F_b(x, y)$,即

$$F_b(x, y) = W_{\text{IR}}^b(x, y)I_1^b(x, y) + (1 - W_{\text{IR}}^b(x, y))I_2^b(x, y) \quad (9)$$

式中, $F_b(x, y)$ 为融合后的基础图像 F_b 在坐标 (x, y) 处的像素值。

由式(9)可知, 融合基础图像 F_b 在光源干扰部分主要取红外图像信息, 在非光源干扰部分对两者信息进行加权融合, 以最大程度地保留两波段图像的有用信息。

2.3 基于 VGG-16 的细节图像融合策略

由于 VGG 是在 AlexNet 网络的基础上开发的, 具有良好的泛化能力, 且采用的卷积网络能够很好地提取融合图像中包含不同信息的深度特征^[33]。为此, 基于 VGG-16 网络模型提出了一种新的细节图像融合策略, 如图 5 所示。首先, 采用预训练的 VGG-16 网络模型^[34]分别提取细节图像 I_k^d 的多层深度特征; 然后, 根据多层深度特征构造深度特征权重图; 最后, 通过不同权重图与细节图像的融合和“极大值选择”的重构得到细节图像 F_d 。

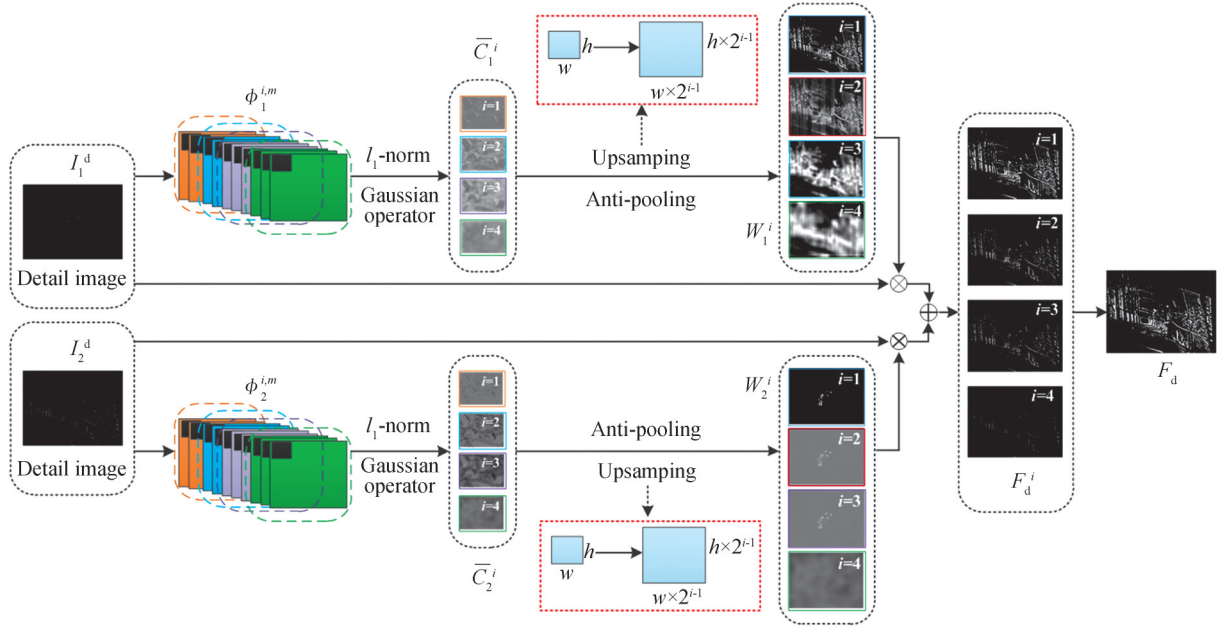


图 5 细节图像的融合策略框架

Fig. 5 The fusion strategy framework of detailed images

通过 VGG-16 网络模型获取细节图像 I_k^d 在第 i 层的深度特征图 $\phi_k^{i,m}$

$$\phi_k^{i,m} = f_i(I_k^d) \quad (10)$$

式中, $f_i(\cdot)$ 为 VGG-16 网络中第 i 个卷积层, $\phi_k^{i,m}$ 为 I_k^d 在第 i 个深度特征图, i 为神经网络的层数, $i \in \{1, 2, 3, 4\}$, m 为第 i 层的通道数目, $m \in \{1, 2, \dots, M\}$, $M = 64 \times 2^{i-1}$, $k \in \{1, 2\}$ 。

由于 l_1 -norm 具有正则化项的“稀疏解”特性, 更适用于特征筛选, 提取关键特征, 把不重要的特征置零。根据式(10)可知, $\phi_k^{i,m}$ 存在 M 维的特征层, 即 $\phi_k^{i,m}$ 在坐标 (x, y) 处的 M 维特征可采用特征向量 $\phi_k^{i,1:M}(x, y)$ 表示。

根据不同深度的特征向量 $\phi_k^{i,1:M}(x, y)$, 采用 l_1 -norm 计算 I_k^d 的粗略显著图 C_k^i 。

$$C_k^i(x, y) = \left\| \phi_k^{i,1:M}(x, y) \right\|_1 \quad (11)$$

式中, C_k^i 为细节图像 I_k^d 在第 i 层的粗略显著图, $C_k^i(x, y)$ 为 C_k^i 在坐标 (x, y) 处的显著程度。

为了提高图像融合算法的噪声抑制能力, 通过高斯算子计算 C_k^i 的精细显著图 \bar{C}_k^i 。

$$\bar{C}_k^i(x, y) = \frac{\sum_{\beta=-r_2}^{r_2} \sum_{\alpha=-r_2}^{r_2} G(x+\beta, y+\alpha) C_k^i(x+\beta, y+\alpha)}{(2r_2+1)^2} \quad (12)$$

式中, \bar{C}_k^i 为 C_k^i 对应的精细显著图, G 为高斯算子, r_2 为高斯算子的半径, r_2 越大, 融合算法的噪声抑制能力越强, 但细节信息丢失越多, 本文中取 $r_2=1$ 。

在 VGG-16 网络模型中,池化因子的步长设置为 2,即每一次池化运算将特征图的大小调整为源图像大小的 $1/2$ 倍。因此,第 i 个卷积层对应的深度特征图是 I_k^d 大小的 $1/2^{i-1}$ 倍(重叠池化运算)。 \bar{C}_k^i 在第 i 层的权重图 W_k^i 计算公式为

$$W_k^i(x, y) = \Phi \left(\frac{\bar{C}_k^i(x, y)}{\sum_n^K \bar{C}_n^i(x, y)} \right) \quad (13)$$

式中, W_k^i 为池化反运算后的权重图, $W_k^i(x, y)$ 为 W_k^i 在坐标 (x, y) 处的权重值, $\Phi(\cdot)$ 为基于上采样的池化反运算函数, K 为 \bar{C}_k^i 的总数,本文中取 $K=2$ 。

通过式(13)依次获取 4 对权重图 W_k^i 后,计算每一对 W_k^i 对应的初始融合细节图像 F_d^i ,即

$$F_d^i(x, y) = \sum_{k=1}^2 W_k^i(x, y) \cdot I_k^d(x, y) \quad (14)$$

选择 4 组 F_d^i 在同一坐标处的像素最大值,获取融合细节图像 F_d ,即

$$F_d(x, y) = \max [F_d^i(x, y) | i \in \{1, 2, 3, 4\}] \quad (15)$$

2.4 算法实现过程

MS-ADoG 域结合 ReNLU 与 VGG-16 的双波段图像融合方法实现过程包括:

- 1) 根据仿射变换原理^[35],以 FIR 图像作为参考图像,对同一监视场景同时刻采集的 FIR 和 VIS 图像进行配准;
- 2) 为避免 I_k^d 的部分细节特征丢失,并提高图像融合效率,采用基于 MS-ADoG 的源图像分解模型,对配准后的 FIR 和 VIS 图像分别进行 3 尺度模糊变换,获取基础图像 I_k^b 和细节图像 I_k^d ;
- 3) 通过设计的 ReNLU 函数,对基础图像进行“加权平均”融合,获取消除光源干扰的融合基础图像 F_b ;
- 4) 根据预训练的 VGG-16 网络模型,获取细节图像的 4 层深度特征,并采用多层融合策略,获取融合的细节图像 F_d ;
- 5) 对融合的基础图像 F_b 和细节图像 F_d 进行重构,获取融合图像 F ,即

$$F(x, y) = F_b(x, y) + F_d(x, y) \quad (16)$$

式中, $F(x, y)$ 为融合图像 F 在坐标 (x, y) 处的像素值。

融合算法框架如图 6 所示。

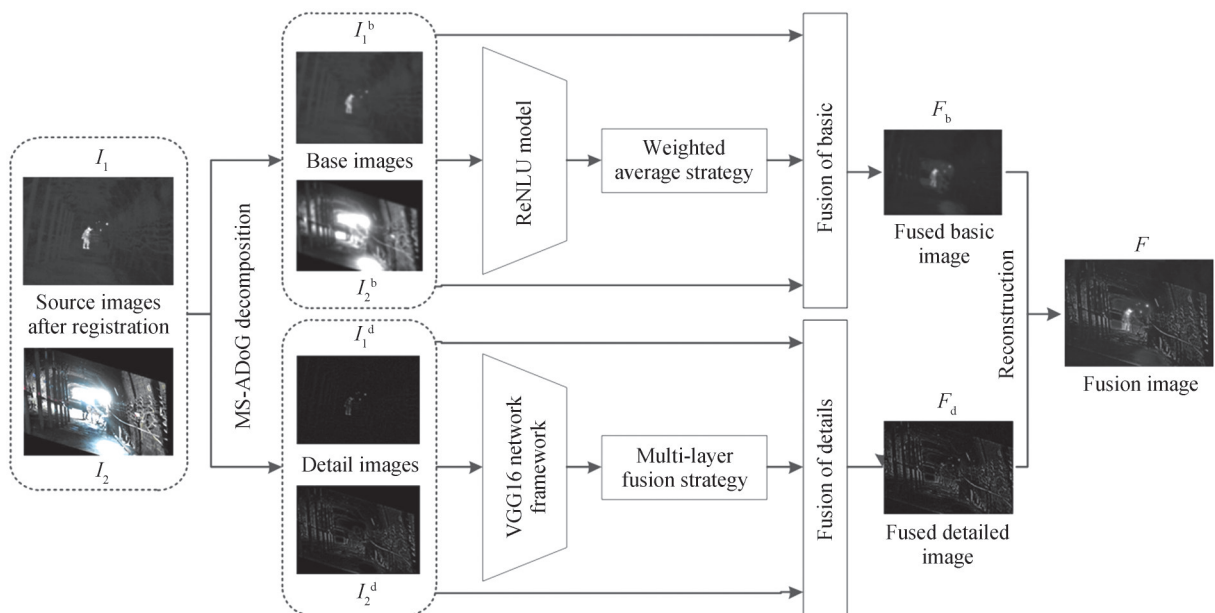


图 6 融合算法框架

Fig. 6 Framework diagram of fusion algorithm

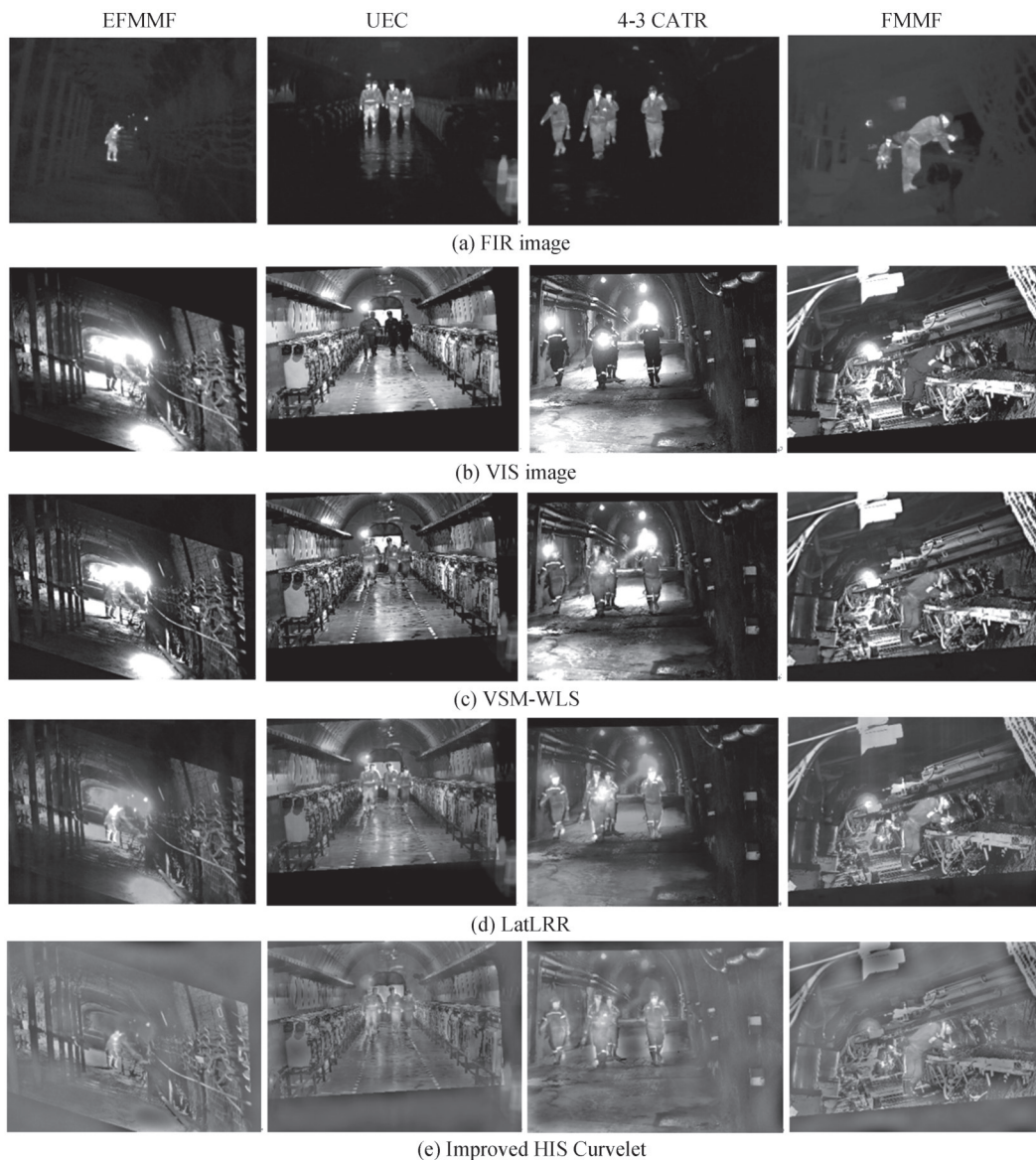
3 实验结果与分析

为了验证本文算法的有效性,于2021年6月在宁夏煤业双马煤矿井下摄取了巷道、回采工作面、机电硐室等可见光和远红外图像,分辨率为 384×288 像素。选取4组经过配准的源图像进行实验验证。融合实验均在 Windows 10 操作系统上进行测试,CPU 主频 3.20GHz,内存 16.00GB,编程工具为 Matlab R2020a。此外,将本文算法与其他 5 种典型融合算法(VSM-WLS^[36]、LatLRR^[37]、Improved HIS Curvelet^[24]、DLF-VGG16^[33]、ResNet50-ZCA^[38])的结果进行实验对比,从主观视觉效果和客观评价指标两方面对其进行分析讨论。实验中,为保证最佳的图像融合质量和融合效率,高斯模糊半径分别取 2、4、6,其他 5 种对比融合算法的参数设置均与原文献保持一致。

3.1 主观评价

从构建的测试数据集中选取 4 类典型的源图像进行融合实验,实验结果如图 7 所示。源图像分别采集于综采工作面端头(End of Fully Mechanized Mining face,EFMMF)、井下机电硐室(Underground Electromechanical Chamber,UEC)、4-3 煤辅助运输大巷(4-3 Coal Auxiliary Transportation Roadway,4-3 CATR)、综采工作面(Fully Mechanized Mining Face,FMMF)。

由各种融合算法的融合结果可知,图 7(c)中的光源干扰部分无法消除,无法提高 FIR 图像中热目标的显著性,图 7(d)、(f)、(g)中的场景信息丰富度得到了显著提高,但光源干扰部分的抑制程度未发生明显性



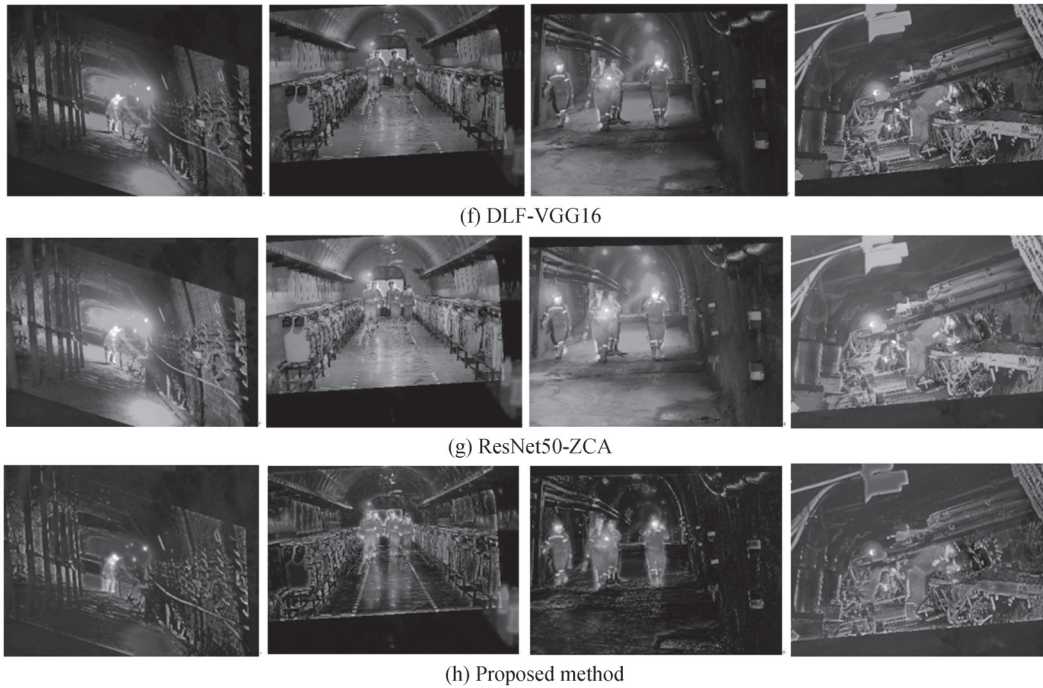


图7 FIR与VIS图像及其融合图像
Fig. 7 FIR and VIS images and fusion images

变化;图7(e)中的光源干扰部分消除不彻底,场景信息模糊,且整体清晰度较低;图7(h)中的光源干扰部分得到了可靠的消除,且场景信息的丰富度得到了显著提高,同时视场的概貌特征和目标的边缘、纹理等细节信息清晰度高,视觉感良好。

综上所述,5种典型融合算法的融合图像均具有不同的缺陷,其中VSM-WLS采用视觉显著图的低频融合策略提高了融合图像的亮度,但无法消除光源干扰对融合图像造成的影响;LatLRR、DLF-VGG16和ResNet50-ZCA采用的融合策略均侧重于对细节图像的融合,融合图像的细节信息得到了较好的融合,但未考虑低频部分中光源干扰对融合图像的影响;Improved HIS Curvelet采用低频系数权值自动调节的融合策略能够消除矿井下的光源干扰,但对高频系数直接采用“极大值选择”方法进行融合,导致融合后图像中引入大量噪声,造成目标轮廓模糊,整体清晰度不高;本文算法采用基于MS-ADoG的源图像分解模型,可相对彻底地分离出基础图像中的细节信息,采用基于ReNLU的基础图像融合策略可实现对光源干扰的消除,通过多层深度特征的细节图像融合策略较好地保留了背景和目标的细节特征,更符合人的视觉特性。

3.2 客观评价

为了验证算法在不同光照环境中的融合性能,选取5种典型指标对5种对比算法和本文算法的融合性能进行比较。同时,采用平均运行时间(Average Running Time, ART)作为一个重要指标,对不同融合算法的运行效率进行评价。

1)平均梯度(Average Gradient, AG)。AG用于描述图像的整体空间活跃度^[39],反映图像中的微小细节反差和纹理变化能力。AG越大,融合图像越清晰。AG的值 R_{AG} 表示为

$$R_{AG} = \frac{1}{2} \sum_{k=1}^2 \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N \sqrt{\nabla f_x^2(x, y) + \nabla f_y^2(x, y)} \quad (17)$$

式中, $M \times N$ 为图像尺寸, ∇f_x 、 ∇f_y 分别为图像在行、列方向上的一阶差分算子。

2)均方根误差(Root Mean Square Error, RMSE)。RMSE通过计算融合图像与参考图像之间像素误差的全局大小来衡量图像质量的好坏^[40]。RMSE的值越大,表明融合算法对抑制图像噪声和消除光源干扰部分的能力越强。RMSE的值 R_{RMSE} 表示为

$$R_{RMSE} = \frac{1}{2} \sum_{k=1}^2 \sqrt{\frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N [F(x, y) - I_k(x, y)]^2} \quad (18)$$

3) 相关系数(Correlation Coefficient, CC)。CC反映了融合图像与参考图像之间的相关程度^[39]。CC的值越大,表明两幅图像高度相关,说明对原始信息的保持越好。CC的值 R_{CC} 可表示为

$$R_{CC} = \frac{1}{2} \sum_{k=1}^2 \frac{\sum_{x=1}^M \sum_{y=1}^N (F(x, y) - \bar{F})(I_k(x, y) - \bar{I}_k)}{\sqrt{\sum_{x=1}^M \sum_{y=1}^N (F(x, y) - \bar{F})^2 (I_k(x, y) - \bar{I}_k)^2}} \quad (19)$$

4) 光晕消除度(Halation Elimination, HE)。HE用于评价融合图像对光源干扰部分的消除能力,能够更加客观评价图像的人眼主观效果^[41]。HE的值越大,融合图像对光源干扰部分的消除效果越好,表示为

$$R_{HE} = SSIM(F, I_1) - SSIM(F, I_2) + 0.5 \quad (20)$$

式中, R_{HE} 为光晕消除度指标值,SSIM(*)为结构相似性算子。

5) 偏差指数(Deviation Index, DI)。DI表示融合图像与参考图像的相对偏差。DI的值越大,两者在光谱信息上的相对差异越大^[40],即融合图像对参考图像的噪声抑制作用越大。DI的值 R_{DI} 可表示为

$$R_{DI} = \frac{1}{2} \sum_{k=1}^2 \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N \frac{|F(x, y) - I_k(x, y)|}{F(x, y)} \quad (21)$$

依据5种典型客观评价指标和ART,依次获取4类典型源图像所对应融合图像的客观评价指标值,如表1所示。

表1 不同算法在4类源图像下的性能指标
Table 1 Performance indicators of different algorithms under 4 types of source images

Images	Indicators	Algorithms					
		SVM-WLS	LatLRR	Improved IHS Curvetlet	DLF- VGG16	Resnet50- ZCA	Proposed Method
EFMMF	AG	0.023 0	0.008 3	0.005 9	0.006 6	0.005 5	0.008 0
	RMSE	0.175 4	0.161 8	0.319 3	0.160 9	0.160 2	0.222 2
	HE	0.490 7	0.497 4	0.502 7	0.499 8	0.499 6	0.504 5
	CC	9.735 9	9.655 6	8.787 3	10.167 0	10.303 0	8.467 1
	DI	1.424 0	0.451 4	1.502 3	0.552 4	0.560 0	1.148 3
	ART	0.910 1	25.795 1	0.859 8	2.855 6	1.797 6	1.692 4
UEC	AG	0.039 7	0.013 3	0.008 5	0.013 2	0.010 4	0.0108
	RMSE	0.205 3	0.195 0	0.278 1	0.194 5	0.192 5	0.249 2
	HE	0.490 4	0.496 1	0.503 7	0.499 6	0.499 6	0.507 0
	CC	11.247 7	11.337 8	10.576 0	12.534 3	12.290 2	9.943 5
	DI	1.339 4	0.590 7	1.562 8	0.705 6	0.725 4	1.720 5
	ART	0.460 7	23.764 7	0.228 3	2.339 7	1.352 4	1.671 8
4-3 CATR	AG	0.016 1	0.004 8	0.003 9	0.004 2	0.003 4	0.006 0
	RMSE	0.233 0	0.214 2	0.309 8	0.208 9	0.209 1	0.290 3
	HE	0.485 0	0.495 8	0.506 0	0.499 7	0.499 6	0.512 8
	CC	10.970 1	10.890 9	9.789 1	11.796 1	11.770 4	8.982 5
	DI	1.398 9	0.532 1	2.801 5	0.606 6	0.602 2	0.832 0
	ART	0.448 7	25.951 1	0.205 7	2.410 2	1.235 1	1.708 6
FMMF	AG	0.056 1	0.019 1	0.013 6	0.017 9	0.013 6	0.015 3
	RMSE	0.190 3	0.168 1	0.259 9	0.161 9	0.162 1	0.207 2
	HE	0.494 9	0.497 4	0.502 5	0.499 7	0.499 8	0.503 4
	CC	10.557 6	10.144 5	10.649 4	11.283 0	11.244 3	10.267 3
	DI	1.612 0	0.370 2	0.503 3	0.418 6	0.427 1	0.834 6
	ART	0.452 2	25.113 8	0.225 3	2.996 9	1.285 7	1.660 4

根据5种典型客观评价指标和ART求解公式,获取数据集中13组源图像在融合后的客观评价指标。结合表1数据,计算13组融合图像的客观评价指标平均值,如表2所示。

表2 不同融合方法的性能指标平均值
Table 2 Average performance indicators of different fusion methods

Fusion algorithms	Performance evaluation indicators					
	AG	RMSE	HE	CC	DI	ART
SVM-WLS	0.032 7	0.198 7	0.493 0	10.690 9	1.610 0	0.507 0
LatLRR	0.011 5	0.185 3	0.497 3	10.686 7	0.520 0	24.502 4
Improved IHS Curvetlet	0.007 4	0.271 1	0.503 5	10.290 9	1.585 7	0.288 8
DLF-VGG16	0.010 4	0.182 4	0.499 7	11.654 1	0.644 2	2.463 5
Resnet50-ZCA	0.008 2	0.182 2	0.499 7	11.572 6	0.619 3	1.334 5
Proposed method	0.010 3	0.242 3	0.506 1	9.753 2	1.698 8	1.689 0

根据表2数据绘制折线图,如图8所示。

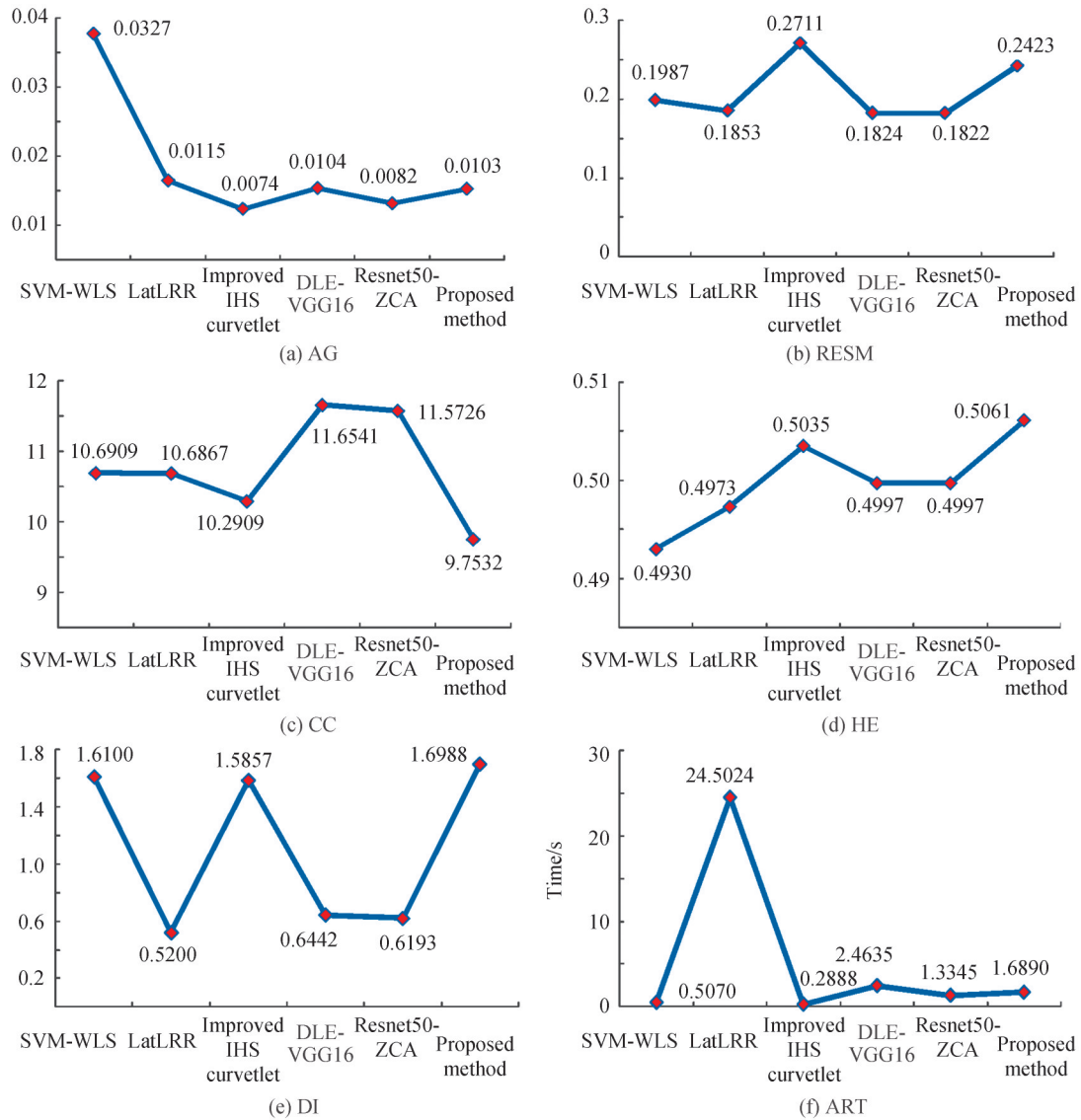


图8 评价指标折线图
Fig. 8 Line chart of evaluation indicators

根据表1~2和图8中的5种典型评价指标和ART可知,本文算法的AG、RMSE、HE、DI值较高,表明该算法比其他对比算法获得的融合图像质量更高。本文算法在抑制图像噪声和消除图像中的光源干扰时,在一定程度上降低了原图像和融合图像的相关程度CC值,但也反映了本文算法在增强图像清晰度、抑制图像噪声和消除图像中光源干扰方面具有明显的优势,这与主观视觉分析一致。SVM-WLS算法的AG、DI值很大,但HE值很小,即该类方法能够增强图像清晰度,但无法去除图像中的光源干扰;LatLRR、DLF-VGG16、ResNet50-ZCA算法的RMSE、HE、DI值较小,即该类方法对抑制图像噪声和消除光源干扰部分的能力较弱;Improved IHS Curvetlet算法的RMSE、HE、DI值较大,但AG较小,即该类方法能够消除图像中的光源干扰,但易导致图像概貌特征丢失,无法改善图像的清晰度。Improved IHS Curvetlet算法的CC值较小,这是由于在消除光晕现象时,一定程度上扭曲了源图像的结构信息,导致相关系数偏低。从ART的数据值和折线图来看,LatLRR算法的ART值较大,即该类图像融合方法易降低融合效率;本文算法与SVM-WLS、Improved IHS Curvetlet、DLF-VGG16、Resnet50-ZCA算法相比ART值较小,即该算法能够适用于实时性要求较高的场所。

通过主观视觉和客观评价可知,本文算法不仅具有较高的融合质量和融合效率,而且能够很好地保留背景和目标的细节特征,更符合人的视觉特性。此外,本文算法在抑制图像噪声和消除井下人工光源干扰方面优于对比算法,更能够满足矿井多源图像智能分析和地面远程监视的实际需求。

4 结论

本文提出了MS-ADoG域结合ReNLU与VGG-16的矿井双波段图像融合算法。利用设计的MS-ADoG分解模型,实现了将源图像分解为包含概貌特征的基础图像和纹理信息的细节图像;采用构造的ReNLU函数,实现了FIR基础图像的权值随VIS基础图像的灰度值自动调节的目的,并实现了对FIR和VIS基础图像的快速融合和干扰光源的消除;通过基于预训练的VGG-16网络模型获得细节图像的不同网络深度特征,并使用 l_1 范数和高斯算子得到了不同深度特征的显著图;利用池化反运算和加权融合策略获取每一对显著图对应的初始融合的细节图像,并采用“极大值选择”方法获取了融合的细节图像。将融合的基础图像和细节图像进行重构,得到了消除光源干扰和场景清晰、特征显著的融合图像。

选取煤矿井下4种不同场景中的源图像,结合5种典型的图像融合算法进行主观分析,并采用5种融合图像质量评价指标和平均运行时间进行客观评价。实验结果表明,本文提出的矿井双波段图像融合算法具有较高的融合质量和融合效率,在抑制图像噪声和消除井下人工光源干扰方面优于现有算法,可用于矿井多源图像智能分析和地面远程监视,还可用于消除地下空间、地下工程或夜间道路视频监控图像中的人工光源干扰问题。

参考文献

- [1] SUN Jiping, JIN Chunhai, CAO Yuchao. Research on mine flood identification and trend prediction method based on video image[J]. Industrial and Mining Automation, 2019, 45(7): 1-4+16.
孙继平, 靳春海, 曹玉超. 基于视频图像的矿井水灾识别及趋势预测方法研究[J]. 工矿自动化, 2019, 45(7): 1-4+16.
- [2] SUN Jiping, LI Yue. Binocular vision-based perception and positioning method of mine external fire[J]. Industrial and Mining Automation, 2021, 47(6): 12-16+78.
孙继平, 李月. 基于双目视觉的矿井外因火灾感知与定位方法[J]. 工矿自动化, 2021, 47(6): 12-16+78.
- [3] SUN Jiping, FAN Weiqiang. Research on explosion perception method of gas and coal dust using video images[J]. Industrial and Mining Automation, 2020, 46(7): 1-4+48.
孙继平, 范伟强. 基于视频图像的瓦斯和煤尘爆炸感知报警及爆源判定方法[J]. 工矿自动化, 2020, 46(7): 1-4+48.
- [4] SUN Jiping. Characteristics of coal mine accidents and new technologies of coal mine communication, personnel and monitoring[J]. Industrial and Mining Automation, 2015, 41(2): 1-5.
孙继平. 煤矿事故特点与煤矿通信、人员定位及监视新技术[J]. 工矿自动化, 2015, 41(2): 1-5.
- [5] SUN Jiping, TIAN Zijian. Image monitoring system and key technology in underground mine[J]. Coal Science and Technology, 2014, 42(1): 65-68.
孙继平, 田子建. 矿井图像监视系统与关键技术[J]. 煤炭科学技术, 2014, 42(1): 65-68.
- [6] SUN Jiping. Study on identified method of coal and rock interface based on identification[J]. Coal Science and Technology, 2011, 39(2): 77-79.
孙继平. 基于图像识别的煤岩界面识别方法研究[J]. 煤炭科学技术, 2011, 39(2): 77-79.

- [7] SUN Jiping, SUN Yanyu, FAN Weiqiang. Mine exogenous fire identification method based on visible light and infrared image[J]. *Industrial and Mining Automation*, 2019, 45(5): 1-5+21.
孙继平, 孙雁宇, 范伟强. 基于可见光和红外图像的矿井外因火灾识别方法[J]. *工矿自动化*, 2019, 45(5): 1-5+21.
- [8] ZHANG Zhen, CHENG Xiangai, JIANG Zongfu. Excessive saturation effect of visible light CCD[J]. *High Power Laser and Particle Beams*, 2008, 20(6): 917-920.
张震, 程湘爱, 姜宗福. 可见光 CCD 的光致过饱和现象[J]. *强激光与粒子束*, 2008, 20(6): 917-920.
- [9] GUO Bingtao, ZHANG Weiguo, WANG Jianjun, et al. Modeling and simulation of halo effect of night vision system[J]. *High Power Laser and Particle Beams*, 2016, 28(11): 170-175.
郭冰涛, 张卫国, 王健军, 等. 微光夜视系统光晕效应定量表征与建模方法[J]. *强激光与粒子束*, 2016, 28(11): 170-175.
- [10] TATTERSALL G J, DANNER R M, CHAVES J A, et al. Activity analysis of thermal imaging videos using a difference imaging approach[J]. *Journal of Thermal Biology*, 2020, 91(7): 102611.
- [11] ZHOU Xiaoling, JIANG Zetao. Infrared and visible image fusion based on pulse coupled neural network and guided filtering[J]. *Acta Optica Sinica*, 2019, 39(11): 1110003.
周晓玲, 江泽涛. 结合脉冲耦合神经网络与引导滤波的红外与可见光图像融合[J]. *光学学报*, 2019, 39(11): 1110003.
- [12] GUO Quanmin, DONG Liang, LI Daidi. Vehicles anti-halation system based on infrared and visible images fusion[J]. *Infrared and Laser Engineering*, 2017, 46(8): 171-176.
郭全民, 董亮, 李代娣. 红外与可见光图像融合的汽车抗晕光系统[J]. *红外与激光工程*, 2017, 46(8): 171-176.
- [13] GHASSEMIAN H. A review of remote sensing image fusion methods[J]. *Information Fusion*, 2016, 32(11): 75-89.
- [14] LIU Y, CHEN X, WANG Z F, et al. Deep learning for pixel-level image fusion: Recent advances and future prospects[J]. *Information Fusion*, 2018, 42(7): 158-173.
- [15] TANG Lili, LIU Gang, XIAO Gang. Infrared and visible image fusion method based on dual-path cascade adversarial mechanism[J]. *Acta Photonica Sinica*, 2021, 50(9): 0910004.
唐丽丽, 刘刚, 肖刚. 基于双路级联对抗机制的红外与可见光图像融合方法[J]. *光子学报*, 2021, 50(9): 0910004.
- [16] DOGRA A, GOAL B, AGRAWAL S. From multi-scale decomposition to non-multi-scale decomposition methods: a comprehensive survey of image fusion techniques and its applications[J]. *IEEE Access*, 2017, 5(8): 16040-16067.
- [17] SHEN Ying, HUANG Chunhong, HUANG Feng, et al. Infrared and visible image fusion: Review of key technologies[J]. *Infrared and Laser Engineering*, 2021, 50(9): 152-169.
沈英, 黄春红, 黄峰, 等. 红外与可见光图像融合技术的研究进展[J]. *红外与激光工程*, 2021, 50(9): 152-169.
- [18] TIAN Songwang, LIN Suzhen, LEI Haiwei, et al. Multi-band image synchronous super-resolution and fusion method based on improved WGAN-GP[J]. *Acta Optica Sinica*, 2020, 40(20): 2010001.
田嵩旺, 蔺素珍, 雷海卫, 等. 基于改进 WGAN-GP 的多波段图像同步超分与融合方法[J]. *光学学报*, 2020, 40(20): 2010001.
- [19] MITANOUDIS N, STATHAKI T. Optimal contrast correction for ICA-based fusion of multimodal images[J]. *IEEE Sensors Journal*, 2008, 6(1): 1-21.
- [20] GUO Quanmin, LI Xiaoling. Vehicles' anti-blooming method based on visible and infrared images fusion[J]. *Infrared Technology*, 2015, 37(6): 475-478.
郭全民, 李晓玲. 基于可见光和红外图像融合的汽车抗晕光方法[J]. *红外技术*, 2015, 37(6): 475-478.
- [21] LIU Jiani, JIN Weiqi, LI Li, et al. Infrared and visible image fusion algorithm based on self-adaptive reference image[J]. *Spectroscopy and Spectral Analysis*, 2016, 36(12): 3907-3914.
刘佳妮, 金伟其, 李力, 等. 自适应参考图像的可见光与热红外彩色图像融合算法[J]. *光谱学与光谱分析*, 2016, 36(12): 3907-3914.
- [22] CHEN Rong. The analysis of image fusion based on improved Brovery transform[C]. *The 2015 International Industrial Informatics and Computer Engineering Conference (IIICEC)*, 2015: 1131-113.
- [23] YILMAZ V, YILMAZ S C, GUNGOR O, et al. A genetic algorithm solution to the gram-schmidt image fusion[J]. *International Journal of Remote Sensing*, 2020, 41(4): 1458-1485.
- [24] GUO Quanmin, WANG Yan, LI Hanshan. Anti-halation method of visible and infrared image fusion based on improved IHS-Curvelet transform[J]. *Infrared and Laser Engineering*, 2018, 47(11): 440-448.
郭全民, 王言, 李翰山. 改进 IHS-Curvelet 变换融合可见光与红外图像抗晕光方法[J]. *红外与激光工程*, 2018, 47(11): 440-448.
- [25] DENG Linuan, YAO Xinfeng. Research on the fusion algorithm of infrared and visible images based on non-subsampled shearlet transform[J]. *Acta Electronica Sinica*, 2017, 45(12): 2965-2970.
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.
- [26] DENG Hui, WANG Changlong, HU Yongjiang, et al. Fusion of infrared and visible images based on non-subsampled dual-tree complex contourlet and adaption block[J]. *Acta Photonica Sinica*, 2019, 48(7): 0710006.
邓辉, 王长龙, 胡永江, 等. 基于非下采样双树复轮廓波与自适应分块的红外与可见光图像融合[J]. *光子学报*, 2019, 48

- (7):0710006.
- [27] LI Hui, WU Xiaojun. DenseFuse: a fusion approach to infrared and visible images[J]. IEEE Transactions on Image Processing, 2019, 28(5):2614-2632.
- [28] XU Hongming, LU Cheng, BERENDT R, et al. Automatic nuclei detection based on generalized Laplacian of Gaussian filters[J]. IEEE Journal of Biomedical and Health Informatics, 2017, 21(3):826-837.
- [29] FAN Weiqiang, LIU Yi. Fuzzy enhancement algorithm of coal mine degradation image based on adaptive wavelet transform [J]. Journal of China Coal Society, 2020, 45(12):4248-4260.
范伟强, 刘毅. 基于自适应小波变换的煤矿降质图像模糊增强算法[J]. 煤炭学报, 2020, 45(12):4248-4260.
- [30] VENETIS J. An analytic exact form of the unit step function[J]. Mathematics and Statistics, 2014, 2(7):235-237.
- [31] YAN Li, XIANG Tianzhu. Fusion of infrared and visible images based on edge feature and adaptive PCNN in NSCT domain[J]. Acta Electronica Sinica, 2016, 44(4):761-766.
闫利, 向天烛. NSCT 域内结合边缘特征和自适应 PCNN 的红外与可见光图像融合[J]. 电子学报, 2016, 44(4):761-766.
- [32] LIU Dongjun, YU Jian. Otsu method and K-means[C]. The 2009 Ninth International Conference on Hybrid Intelligent Systems (INSPEC), 2009:344-349.
- [33] LI H, WU X J, KITTLER J. Infrared and visible image fusion using a deep learning framework[C]. The 24th International Conference on Pattern Recognition (ICPR), 2018:2705-2710.
- [34] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]. International Conference on Learning Representations (ICLR), 2015, arXiv:1409.1556.
- [35] PFUGFELDER D, SCHARR H. Practically lossless affine image transformation [J]. IEEE Transactions on Image Processing, 2020, 29(3):5367-5373.
- [36] MA Jinlei, ZHOU Zhiqiang, WANG Bo, et al. Infrared and visible image fusion based on visual saliency map and weighted least square optimization[J]. Infrared Physics & Technology, 2017, 82(5):8-17.
- [37] LI H, WU X J. Infrared and visible image fusion using Latent low-rank representation[C]. Computer Vision and Pattern Recognition (CVPR), 2019, arXiv:1804.08992v4.
- [38] LI H, WU X J, DURRANIT S. Infrared and visible image fusion with ResNet and zero-phase component analysis[J]. Journal of Infrared Physics & Technology, 2019, 102(11):103039.
- [39] PANIGRAHY C, SEAL A, MAHATO N K. Fractal dimension based parameter adaptive dual channel PCNN for multi-focus image fusion[J]. Optics and Lasers in Engineering, 2020, 133(10):106141.
- [40] YU Xianchuan, PEI Wenjing. Performance evaluation of image fusion quality metrics for the quality of different fusion methods[J]. Infrared and Laser Engineering, 2012, 41(12):3416-3422.
余先川, 裴文静. 针对不同融合算法的质量评价指标性能评估[J]. 红外与激光工程, 2012, 41(12):3416-3422.
- [41] GUO Quanmin, CHAI Gaixia, LI Hanshan. Quality evaluation of night vision anti-halation fusion image based on adaption partition[J]. Journal of Electronics & Information Technology, 2020, 42(7):1750-1757.
郭全民, 柴改霞, 李翰山. 夜视抗晕光融合图像自适应分区质量评价[J]. 电子与信息学报, 2020, 42(7):1750-1757.

Mine Dual-band Image Fusion in MS-ADoG Domain Combined with ReLU and VGG-16

SUN Jiping, FAN Weiqiang

(School of Mechanical Electronic and Information Engineering, China University of Mining and Technology (Beijing), Beijing 100083, China)

Abstract: Mine video surveillance technology has significant advantages in promoting coal mine safety, high efficiency, and unmanned mining. However, the poor working conditions of cameras in coal mines have led to serious degradation of image quality. For this reason, multi-source fusion processing of mine monitoring video will solve the current problems and be helpful to promote the intelligent development of coal mines. In view of the low computational efficiency of existing image fusion algorithms and poor timeliness, the fusion images acquired by the existing image fusion algorithms have problems such as false targets, fuzzy targets, halo occluded targets, etc., which cannot meet the needs of mine video surveillance. This paper proposes a mine dual-band image fusion algorithm using multi-scale and adaptive Gaussian difference transform combined with rectified non-linear unit and VGG-16. A source image decomposition

model based on multi-scale and adaptive difference of Gaussian is designed. This model decomposes infrared and visible images into basic images and detailed images. Among them, the basic image represents the approximate components of the source image, reflecting the general features of the field of view. The detailed image represents the detailed components of the source image, including detailed information such as edges and textures, and is also the most sensitive part of human eye recognition and machine vision. To eliminate the interference part of the light source in the visible basic image and improve the overall contrast and information richness of the fusion image in the underground mine, a rectified non-linear unit function is constructed. The rectified non-linear unit function makes the weight of the infrared basic image automatically adjust with the gray level of the visible basic image, and the “weighted average” basic image fusion strategy is adopted to obtain the fused basic image that eliminates the interference of the light source and retains the general features of visible and infrared images. Meanwhile, the pre-trained VGG-16 network model is used to extract the 4-layer depth features of the detailed image, and the l_1 -norm and Gaussian operator are used to sequentially obtain the saliency maps corresponding to the 4-layer depth feature. After obtaining 4 pairs of fused images with different depth features through pooling inverse operation and weighted fusion, the fused detail image is obtained by the “maximum value selection” method. The fusion basic image and detail image are reconstructed to obtain the final fusion image. To verify the effectiveness of the proposed algorithm, the experiment selected the source images of the coal mine in four different scenes, combined five typical image fusion algorithms for subjective analysis, and used five quality evaluation indicators of fusion image and average running time for objective evaluation. The experimental results show that the proposed algorithm can eliminate the interference of artificial light sources and obtain fused images of underground mines with clear scenes and salient features, and the fused image is more in line with human visual characteristics. At the same time, it improves the fusion quality and fusion efficiency of heterogeneous images, which is conducive to the further analysis and processing of images. Compared with the other five typical algorithms, the proposed algorithm is more robust. It not only overcomes the shortcomings of traditional algorithms that cannot extract image depth features but also makes it easier to completely eliminate light source interference and obtain more comprehensive, reliable, and rich scene information. In addition, the proposed algorithm can be used for the intelligent analysis of multi-source images of mines and remote monitoring on the ground. It can also be used to eliminate the problem of artificial light source interference in underground space, underground engineering or night road video surveillance images.

Key words: Video surveillance; Image fusion; Adaptive difference of Gaussian; ReLU function; Neural network; Depth feature; Pooling inverse operation

OCIS Codes: 100.2960; 100.6820; 110.3000; 150.5495