

基于卷积自编码器和残差块的红外与可见光图像融合方法

江泽涛, 何玉婷*

桂林电子科技大学广西图像图形与处理智能处理重点实验室, 广西 桂林 541004

摘要 为了在红外与可见光图像融合中充分利用中间层提取的信息,防止信息过度丢失,提出一种新的基于卷积自编码器和残差块的图像融合方法。该方法采用由编码器、融合层和解码器三部分组成的网络结构。将残差网络引入编码器中,将红外与可见光图像分别送入编码器后,通过卷积层和残差块来获取图像的特征图;将得到的特征图采用改进的基于 L1-norm 的相似度融合策略进行融合,并将其整合为一个包含源图像显著特征的特征图;重新设计损失函数,利用解码器对融合后的图像进行重构。实验结果表明,与其他融合方法相比,该方法有效地提取并保留了源图像的深层信息,融合结果在主观和客观评价中都有着一定的优势。

关键词 机器视觉; 图像融合; 可见光图像; 红外图像; 残差块; 卷积自编码器

中图分类号 TP751

文献标识码 A

doi: 10.3788/AOS201939.1015001

Infrared and Visible Image Fusion Method Based on Convolutional Auto-Encoder and Residual Block

Jiang Zetao, He Yuting*

Guangxi Key Laboratory of Image and Graphic Intelligent Processing, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China

Abstract In order to make full use of the information extracted from the middle layer and prevent information from losing excessively, a new image fusion method based on a convolutional auto-encoder and a residual block is proposed, which is composed of an encoder, a fusion layer, and a decoder. First, the residual network is introduced into the encoder, the infrared and visible images are fed into the encoder, and the convolution layer and residual block are used to obtain the feature map of the image. Then, the obtained feature map is fused by using an improved fusion strategy based on L1-norm similarity, which is integrated into a feature map containing the salient features of the source image. Finally, the loss function is redesigned and the decoder is used to reconstruct the fused image. The experimental results show that compared with other fusion methods, the method effectively extracts and preserves the deep information of the source image, which makes the fusion result have certain advantages in both subjective and objective evaluation.

Key words machine vision; image fusion; visible image; infrared image; residual block; convolutional auto-encoder

OCIS codes 150.0155; 100.2980; 110.2960; 110.3000; 110.3080

1 引 言

红外与可见光图像的融合是图像处理领域中的重要应用之一,它尝试从源图像中提取出特征信息,然后通过适当的方法将这些特征信息融合到单个图像中^[1],使得融合图像兼具红外与可见

光图像的显著特征,从而可以提供更多的目标与背景信息,因此将红外与可见光图像进行融合的方法在目标识别、视频监控、军事应用等领域有着广阔的应用前景。

图像融合的关键问题是如何从源图像中提取出特征,并将其结合起来生成融合后的图像^[2]。

收稿日期: 2019-04-23; 修回日期: 2019-05-15; 录用日期: 2019-05-31

基金项目: 国家自然科学基金(61572147,61876049,61762066)、广西科技计划(AC16380108)、广西图像图形与智能处理重点实验室项目(GIIP201701,GIIP201801,GIIP201802,GIIP201803)、广西研究生教育创新计划(2019YCX043)

* E-mail: 839191881@qq.com

许多信号处理方法都可用于图像融合中,以提取图像的特征,如多尺度分解、稀疏表示等。基于多尺度分解的方法,首先采用图像分解的方法来提取图像特征信息,然后采用适当的融合策略来获得最终的融合图像^[3]。基于稀疏表示的方法通过构造冗余字典对图像进行稀疏编码并提取源图像特征,再根据一定的规则融合稀疏表示系数,然后通过字典和融合的稀疏表示系数重构融合图像。这些融合方法虽然能够取得良好的融合效果,但在提取特征时,仍会丢失有用的信息,不能较好地保留源图像的信息。

深度学习现已广泛应用于图像融合任务中,比经典方法实现了更好的融合性能和效率。基于深度学习的算法^[4]使用大量的图像对网络进行预训练,使得训练后的网络可以获得图像的显著特征。其中,卷积神经网络(CNN)可用于提取图像特征和重构融合后的图像。在基于 CNN 的融合方法中^[5],若只使用最后一层的结果作为图像的特征信息,将会丢失中间层获得的许多深层信息,从而影响最终融合结果。文献^[6]提出了一种基于 CNN 的多聚焦图像融合方法,输入图像是不同模糊位置的图像块,用于训练网络并获得决策图,然后使用决策图和源图像获得融合后的图像,但是该方法仅适用于多聚焦图像融合。文献^[7]提出一种基于 CNN 的多曝光图像融合方法,采用 Siamese 网络结构,其权重是相互关联的。首先将两个输入图像送入编码器中,得到两个特征映射序列,用加法策略进行融合,融合后的特征图由解码器的三层卷积层进行重构。该方法虽然取得了较好的性能,但仍存在两个主要缺点:1)网络架构过于简单,无法提取出深层特征;2)只使用了编码网络中最后一层的结果,中间层会丢失许多特征信息,网络越深,特征信息丢失的现象越严重。

基于上述问题,本文提出一种新的基于卷积自编码器和残差块的网络结构,用于红外图像与可见光图像融合。利用编码器来提取图像特征,利用解码器重构融合后的图像。编码器由卷积层和残差块构成,可以保存更多来自中间层的有用信息,并且易于训练。利用改进的基于 L1-norm 的相似度融合策略和包含 4 个卷积层的解码器对融合后的图像进行重构。将本文方法与其他融合方法的进行对比,验证本文算法在保留中间层信息方面具有一定的优势。

2 基于卷积自编码器和残差块的图像融合方法

2.1 残差块

在传统的 CNN 中,随着网络深度增加,网络的退化问题也随之暴露,即中间层提取的信息并没有得到充分的利用。为了解决退化问题,He 等^[8]提出深度残差学习框架,通过快捷连接和残差表示,直接将输入信息绕道传至输出,从而保护信息的完整性。该网络更容易优化,并通过增加深度提高了准确率。残差块的结构如图 1 所示。

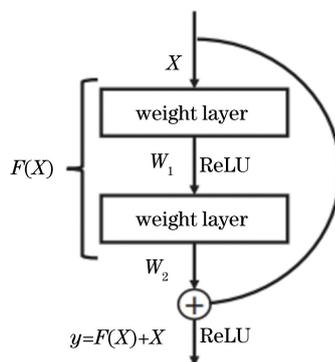


图 1 残差块的结构

Fig. 1 Architecture of residual block

残差块是在传统的线性网络结构基础上通过跳跃连接来实现的。一般地,通过直接求和操作可实现特征图的跳跃连接^[9],本文改用拼接操作。经过卷积操作后,网络会保留一些特征信息,但也会丢失一些信息,而丢失的这些可能是有用的信息,因此通过加入原始输入,使网络拥有更多的选择,从而使得丢失的信息得到了保留。为此,在编码器中加入残差块,使得网络可以保存更多来自中间层的特征信息,并且易于训练。

2.2 总体思路框架

将红外与可见光图像分别定义为 X_{IR} 和 X_{VIS} 。该网络结构由三部分组成:编码器、融合层和解码器。编码器包含三个部分——C1、ResidualBlock、C2,用于提取图像特征。其中,ResidualBlock 部分包含两个卷积层。编码器中每个卷积操作的滤波器大小和步长分别是 3×3 和 1,输入图像的大小不受限制。残差块可以尽可能地保留更多的特征,确保深层特征用于融合层。融合层的输出作为解码器的输入。解码器包含 4 个卷积层。所提出的网络框架如图 2 所示。

将 X_{IR} 和 X_{VIS} 分别送入编码器中。假设有 m 个

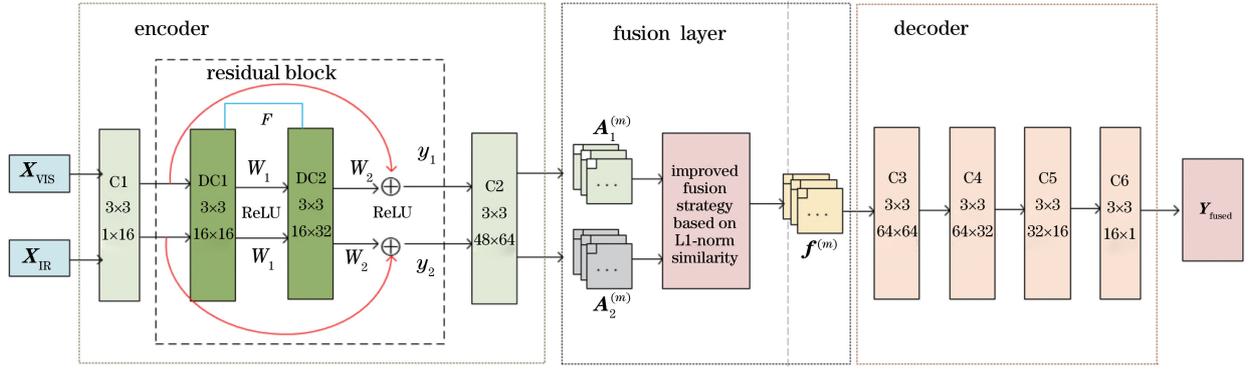


图 2 所提出的网络架构

Fig. 2 Network architecture of proposed method

卷积核,每个卷积核由卷积矩阵 $\mathbf{a}^{(m)}$ 和偏置 $\mathbf{b}^{(m)}$ 参数组成,激活函数为 δ ,将输入图像 \mathbf{X}_k 进行卷积后生成 m 个特征图 \mathbf{A} ,则有

$$\mathbf{A}_k^{(m)} = \delta[\mathbf{X}_k * \mathbf{a}^{(m)} + \mathbf{b}^{(m)}], \quad (1)$$

式中: m 表示上一层的输出通道数目, $m \in \{1, 2, \dots, M\}$, $M=64$ 代表特征图的数量; $\mathbf{A}_k^{(m)}$ 表示由编码器得到的特征图, $k \in \{1, 2\}$, 1、2 分别表示红外图像与可见光图像; $*$ 表示卷积。

将生成的特征图送进 ResidualBlock 中,经过第一层线性变化并激活后的输出 F 为

$$F = W_2 \delta[W_1 \mathbf{A}_k^{(m)}], \quad (2)$$

式中: W_1 为第一层的权重参数; W_2 为第二层的权重参数。

通过一个快捷连接和第二个非线性函数 ReLU,获得输出 y ,表达式为

$$y = F\{\mathbf{A}_k^{(m)}, \{W_t\}\} + \mathbf{A}_k^{(m)}, \quad (3)$$

式中: W_t 为权重参数, $t \in \{1, 2\}$

将由编码器得到的红外图像特征图 $\mathbf{A}_1^{(m)}$ 和可见光图像特征图 $\mathbf{A}_2^{(m)}$ 送入融合层,通过改进的基于 L1-norm 的相似度融合策略进行融合,得到融合特

征图 $\mathbf{f}^{(m)}$ 。

最后将融合特征图 $\mathbf{f}^{(m)}$ 送入解码器中,该部分由 4 个卷积层组成。将每张融合特征图 \mathbf{f} 与其对应的卷积核的转置 (\bar{w}) 进行卷积操作并对结果进行求和,再加上偏置 \mathbf{c} ,即得到特征重构之后的结果:

$$\mathbf{Y}_{\text{fused}} = \delta\left[\sum \mathbf{f}^{(m)} * \bar{w}^{(m)} + \mathbf{c}\right]. \quad (4)$$

要更新权值,需要确定损失函数,该操作是在训练框架部分完成的,即将输入的图像与最终利用特征重构出来的结果进行比较,通过反向传播算法进行优化,就可以得到一个完整的训练框架。

2.3 改进的基于 L1-norm 的相似度融合策略

在测试阶段,将红外图像与可见光图像分别送入编码器。当编码器和解码器的权值固定时,使用融合策略来融合编码器得到的特征。参照文献[10],提出改进的基于 L1-norm 的相似度融合策略来融合从编码器得到的特征图。融合策略如图 3 所示。

首先,由编码器得到的特征图 $\mathbf{A}_k^{(m)}$ 的 L1-norm 获得初始活动水平图 \bar{C}_k ,通过基于窗口的平均算子来计算得到最终的活动水平图 \bar{C}_k ;然后,对活动水

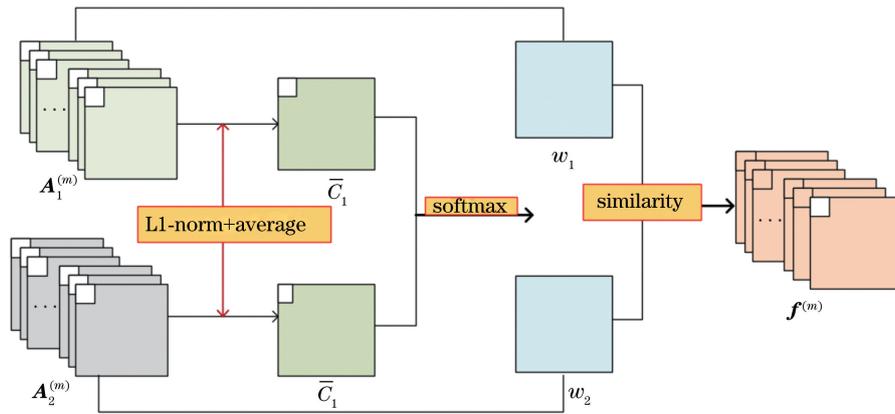


图 3 融合策略图

Fig. 3 Diagram of fusion strategy

平图 \bar{C}_k 进行 softmax 函数计算,得到初始权重图 w_k ,计算初始活动水平图 C_1 和 C_2 之间的相似度 S ;最终,根据相似度 S 选择融合规则,得到融合特征图 $f^{(m)}$ 。详细步骤如下。

1) 采用 $A_k^{1:M}(i,j)$ 的 L1-norm 作为源图像的活动水平测量,从而获得初始活动水平图 C_k :

$$C_k(i,j) = \|A_k^{1:M}(i,j)\|_1, \quad (5)$$

式中: $A_k^{1:M}(i,j)$ 是一个 M 维的向量,表示特征图中位置 (i,j) 处的 $A_k^{(m)}$; (i,j) 代表位置坐标。

2) 使用基于窗口的平均算子来计算最终的活动水平图 \bar{C}_k ,以使融合方法对错误配准具有稳健性。 \bar{C}_k 可表示为

$$\bar{C}_k(i,j) = \frac{\sum_{a=-r}^r \sum_{b=-r}^r C_k(i+a,j+b)}{(2r+1)^2}, \quad (6)$$

式中: a 为窗口横坐标取值; b 为窗口纵坐标取值; r 代表窗口大小。 r 决定块大小:当 r 值较大时,能考虑到图像的整体信息,但忽略了边缘细节信息,因此使用较大的 r 有着更强的稳健性,但会丢失一些细节,从而导致融合效果的细节信息比较模糊;当 r 值过小时,如选取 $r=0$,由于未考虑到图像的整体信息,忽略了周围特征信息的相关性,因此稳健性较差。在红外与可见光图像的融合中,由于小尺度细节的存在,采用较小的 r 效果更好,经过验证,本研究选取 $r=1$ 。

3) 根据得到的活动水平图 \bar{C}_k ,通过 soft-max 函数计算得到初始权重图 w_k :

$$w_k(i,j) = \frac{\bar{C}_k(i,j)}{\sum_{n=1}^k \bar{C}_n(i,j)}, \quad (7)$$

式中: n 为 k 的变量取值, k 的定义与(1)式相同。

4) 通过初始活动水平图 C_k 来计算初始活动水平图 C_1 和 C_2 之间的相似度 S :

$$S(i,j) = \frac{2\sqrt{C_1(i,j)C_2(i,j)}}{C_1(i,j)+C_2(i,j)}. \quad (8)$$

5) 根据得到的相似度 S 来选择融合规则,从而得到融合之后的特征图 $f^{(m)}$ 。

当 $S(i,j) \geq T$ (T 为阈值)时,有

$$f^m(i,j) = \omega_1(i,j)A_1^{(m)}(i,j) + \omega_2(i,j)A_2^{(m)}(i,j), \quad (9)$$

式中: $\omega_1(i,j) = \frac{\bar{C}_1(i,j)}{\bar{C}_1(i,j)+\bar{C}_2(i,j)}$; $\omega_2(i,j) =$

$$\frac{\bar{C}_2(i,j)}{\bar{C}_1(i,j)+\bar{C}_2(i,j)}.$$

当 $S(x,y) < T$ 时,有

$$f^{(m)}(i,j) = \begin{cases} A_1^{(m)}(i,j), & A_1^{(m)}(i,j) \geq A_2^{(m)}(i,j) \\ A_2^{(m)}(i,j), & A_1^{(m)}(i,j) < A_2^{(m)}(i,j) \end{cases}. \quad (10)$$

经过验证, $T=0.8$ 时融合效果最好。

2.4 损失函数设计

为了更精确地重构输入图像,选择最小化损失函数 L 来训练编码器和解码器。损失函数可用来估计网络的输入与输出之间的差异^[11]。常用的损失函数是均方误差,但其不能很好地学习到图像的目标细节信息,会导致输出图像产生模糊。在此基础上,引入结构损失和梯度损失,以更好地保留纹理信息和边缘信息。综合考虑图像的结构信息、边缘信息,以及内容信息之间的差异,最后的损失函数 L 为

$$L = L_p + L_{\text{grad}} + \lambda L_{\text{ssim}}, \quad (11)$$

式中: L_p 为均方误差损失; L_{grad} 为梯度损失; L_{ssim} 为结构相似性损失; λ 为比例系数。

因为均方误差损失和梯度损失与结构损失有3个数量级的差异,所以在训练阶段将 λ 分别设置为1,10,100,1000,发现 $\lambda=10$ 时效果最好。

均方误差损失 L_p 通过计算网络训练过程中输出图像与输入图像之间的差值,使得图像内容保持一致性^[12],表达式为

$$L_p = \frac{1}{N} \sum \|O - I\|^2, \quad (12)$$

式中: O, I 分别代表输出和输入图像; N 为训练样本的个数。

梯度损失 L_{grad} 是引入 Sobel 算子来计算输出图像和输入图像的梯度,通过二者之间的差值来优化图像的边缘^[13],表达式为

$$L_{\text{grad}} = \frac{1}{N} \sum |G_o - G_I|, \quad (13)$$

式中: G_o, G_I 分别为输出图像和输入图像的梯度,

$$G_o = \begin{bmatrix} 1 & 0 & -1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \cdot O + \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \cdot O;$$

$$G_I = \begin{bmatrix} 1 & 0 & -1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \cdot I + \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \cdot I.$$

G_o, G_I 分别用来检测水平和垂直方向的边缘。

结构损失是通过结构相似性损失 L_{ssim} 来评估的,它表示两幅图像之间的相似度。

$$L_{\text{ssim}} = 1 - f_{\text{SSIM}}(O, I), \quad (14)$$

$$f_{SSIM}(\mathbf{O}, \mathbf{I}) = \frac{2\mu_o\mu_I + D_1}{\mu_o^2 + \mu_I^2 + D_1} \cdot \frac{2\sigma_{oI} + D_2}{\sigma_o^2 + \sigma_I^2 + D_2}, \quad (15)$$

式中： $f_{SSIM}(\cdot)$ ^[14]表示两幅图像之间的相似度； μ_o 、 μ_I 分别是输出图像和输入图像的平均值； σ_o^2 、 σ_I^2 分别是输出图像和输入图像的方差； σ_{oI} 是输出图像和输入图像的协方差； D_1 、 D_2 是用来维持稳定的常数。

为了证明该损失函数的有效性，将该损失函数与传统的损失函数均方误差进行对比，给出训练过程中不同损失函数的下降曲线，如图4所示。

2.5 训练框架

在训练过程，仅考虑编码部分和解码部分，尝试通过训练编码器和解码器来重构输入图像。该训练策略的明显优势在于，可以为特定的任务设计合适的融合层，因此还具有进一步发展融合层的空间。

该网络的详细训练框架如图5所示。

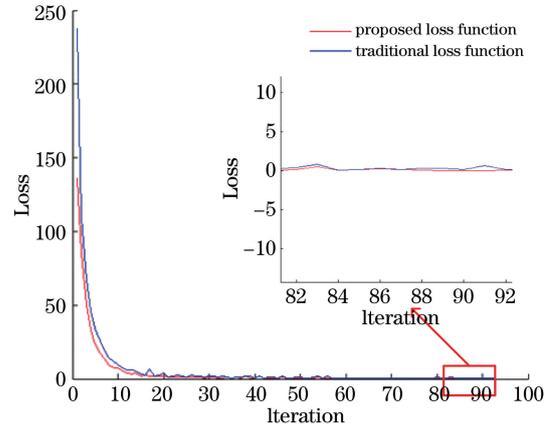


图4 训练过程中不同损失函数的下降曲线图
Fig. 4 Decline curves of different loss functions in training process

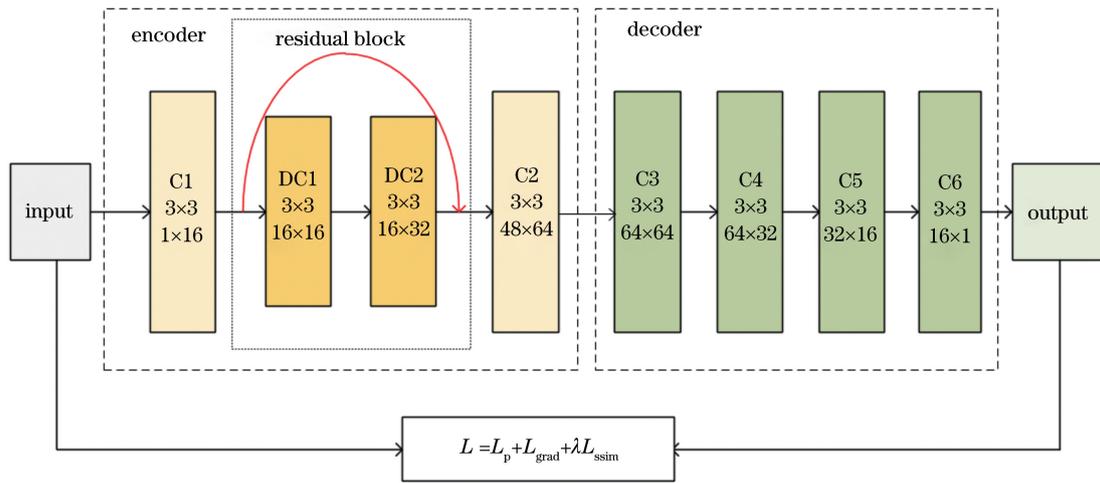


图5 训练流程框架
Fig. 5 Framework of training process

C1、C2是编码器中的卷积层，包含 3×3 的滤波器。DC1、DC2是残差块中的卷积层。解码器由C3、C4、C5和C6组成，用于重建输入图像。为了更精确地重建输入图像，选取最小化损失函数 L 来训练编码器和解码器。

3 实验结果与分析

3.1 实验设置

在训练阶段，使用MS-COCO数据集中的可见光图像来训练编码器和解码器的权重。随机选取 4×10^4 张图像作为输入图像，并将其转换成 $256 \text{ pixel} \times 256 \text{ pixel}$ 的灰度图像。在每一次迭代中用450张图像来验证网络的重构能力。

使用tensorflow搭建实验环境，并使用GTX1080Ti GPU进行训练。在网络训练过程中，

优化器为Adam，学习速率为0.0001，激活函数是ReLU，Batch Size为8，Epoch为20。

为了验证本文方法的性能，从20对测试图像中选取3组红外与可见光图像，与文献[15-18]的方法进行对比。

从主观视觉效果和客观评价标准两方面验证各方法的图像融合效果。

3.2 主观视觉评价

4组图像的实验结果如图6~8所示。图6~8中的(a)均为红外图像；(b)均为可见光图像；(c)均为基于CNN融合方法的融合结果；(d)均为基于视觉显著性图和加权最小二乘优化的融合方法的融合结果；(e)均为基于红外特征提取与视觉信息保存的融合方法的融合结果；(f)均为NSCT (Nonsubsampled Contourlet变换)和压缩感知的融

合方法的融合结果;(g)均为本文方法的融合结果。为了更直观地比较图像之间的差异,对图中的融合区域作标注。

对比图6中屏幕下的桌椅、地灯、行人以及电杆标志可以明显看出,图6(c)~(f)中这些地方比较模糊,存在噪声,有些甚至出现失真现象;而本文算法[图6(g)]具有较好的视觉效果,可以清晰地识别出屏幕下的桌椅、电杆标志等。

从图7可以看出,图7(c)中地面出现黑色阴影区域,未能保留出可见光图像的背景信息;图7(d)~(f)地面偏暗,两边草丛模糊;而本文算法[图7(g)]亮

度均匀,能清晰地看出地面和两边的草丛,较好地保留了背景信息。

从图8可以看出,图8(c)~(f)中车前方玻璃出现黑色区域,房子周围的树木模糊,部分区域出现失真现象;本文算法[图8(g)]可以清晰地看到玻璃和树木的纹理信息,充分融合了可见光图像的背景信息,相比其他方法具有明显优势。

受篇幅限制,为了进一步验证本文方法的优越性,图9仅展示部分融合结果。

3.3 客观评价

为了进一步对上述不同融合方法所得的实验结

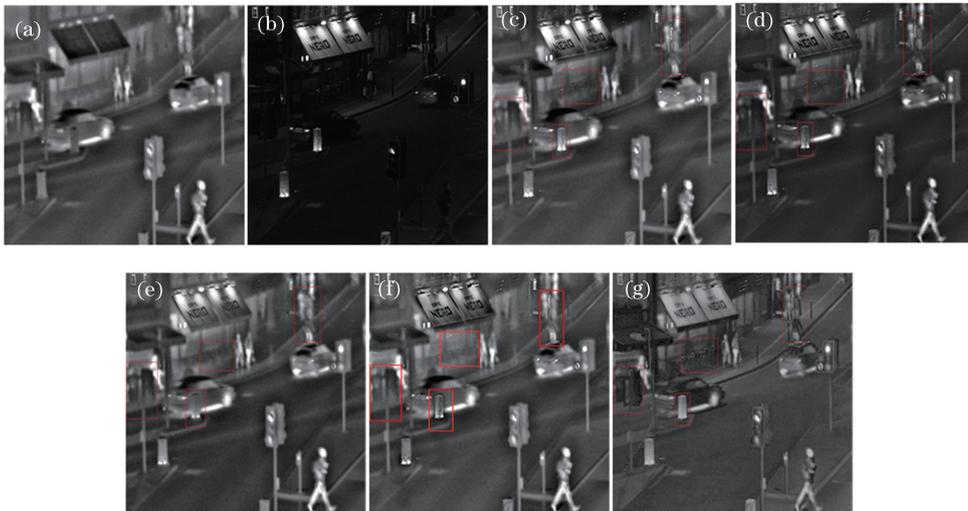


图6 第一组图像融合对比实验。(a)红外图像;(b)可见光图像;(c)文献[15]方法;(d)文献[16]方法;(e)文献[17]方法;(f)文献[18]方法;(g)本文方法

Fig. 6 The first group of image fusion contrast experiments. (a) Infrared image; (b) visible image; (c) method in Ref. [15]; (d) method in Ref. [16]; (e) method in Ref. [17]; (f) method in Ref. [18]; (g) proposed method

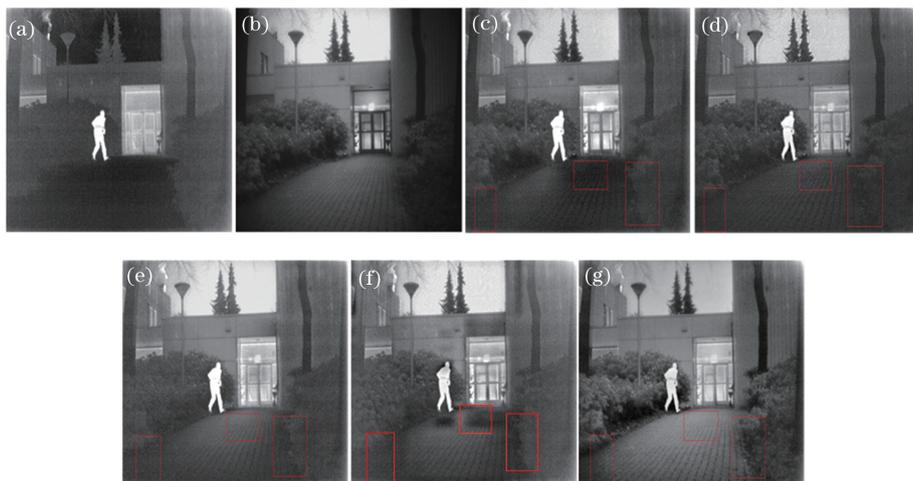


图7 第二组图像融合对比实验。(a)红外图像;(b)可见光图像;(c)文献[15]方法;(d)文献[16]方法;(e)文献[17]方法;(f)文献[18]方法;(g)本文方法

Fig. 7 The second group of image fusion contrast experiments. (a) Infrared image; (b) visible image; (c) method in Ref. [15]; (d) method in Ref. [16]; (e) method in Ref. [17]; (f) method in Ref. [18]; (g) proposed method

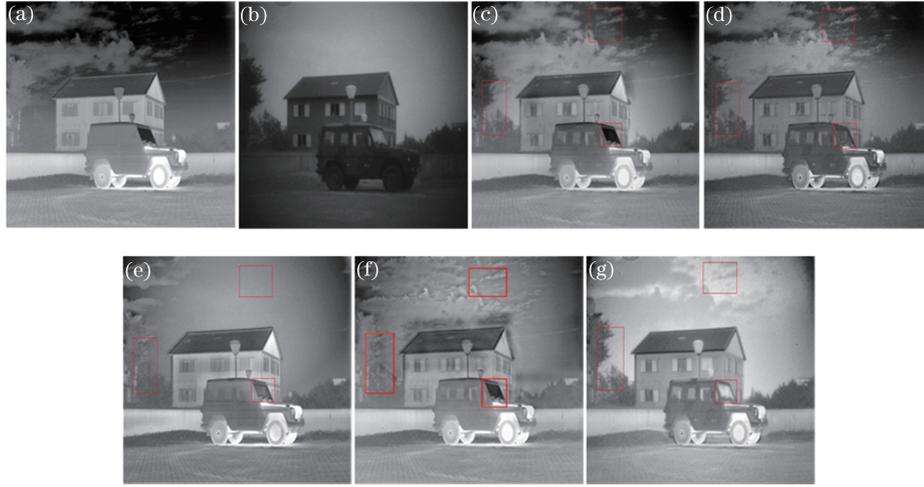


图 8 第三组图像融合对比实验。(a)红外图像;(b)可见光图像;(c)文献[15]方法;(d)文献[16]方法;
(e)文献[17]方法;(f)文献[18]方法;(g)本文方法

Fig. 8 The third group of image fusion contrast experiments. (a) Infrared image; (b) visible image; (c) method in Ref. [15]; (d) method in Ref. [16]; (e) method in Ref. [17]; (f) method in Ref. [18]; (g) proposed method

果进行客观评价,选取信息熵(IE)、标准差(SD)、平均梯度(AG)和互信息(MI)这 4 个指标进行客观评价,结果如表 1~3 所示。

从表 1~3 可以看出,本文方法在这 4 项评价指

标上明显优于其他融合算法。虽然本文方法在表 1 的 IE 指标小于文献[18]中的方法,在表 3 中的 IE 指标小于文献[15]的方法,但在其他指标上均有明显优势。

表 1 第一组对比实验的客观评价结果

Table 1 Objective evaluation indicators of the first group of comparative experimental results

Fusion method	IE	SD	AG	MI
Method in Ref. [15]	6.7072	37.049	3.5297	22.3850
Method in Ref. [16]	6.1381	32.436	3.3973	30.7770
Method in Ref. [17]	6.6984	33.590	3.3490	22.7620
Method in Ref. [18]	6.7510	35.192	3.4359	23.1432
Proposed method	6.6411	38.270	4.1455	31.8710

表 2 第二组对比实验的客观评价结果

Table 2 Objective evaluation indicators of the second group of comparative experimental results

Fusion method	IE	SD	AG	MI
Method in Ref. [15]	7.1000	53.673	3.8040	15.6260
Method in Ref. [16]	6.9540	49.513	4.2530	16.5480
Method in Ref. [17]	6.9710	52.781	3.7172	19.3090
Method in Ref. [18]	7.1947	51.664	4.4496	13.0329
Proposed method	7.4754	58.703	4.4842	24.5260

表 3 第三组对比实验的客观评价结果

Table 3 Objective evaluation indicators of the third group of comparative experimental results

Fusion method	IE	SD	AG	MI
Method in Ref. [15]	7.4114	39.817	3.0124	17.6580
Method in Ref. [16]	7.1459	35.916	3.6467	17.6800
Method in Ref. [17]	7.0764	34.879	3.1716	16.9430
Method in Ref. [18]	7.1703	37.727	3.3523	21.7977
Proposed method	7.3484	41.758	3.8451	27.3920

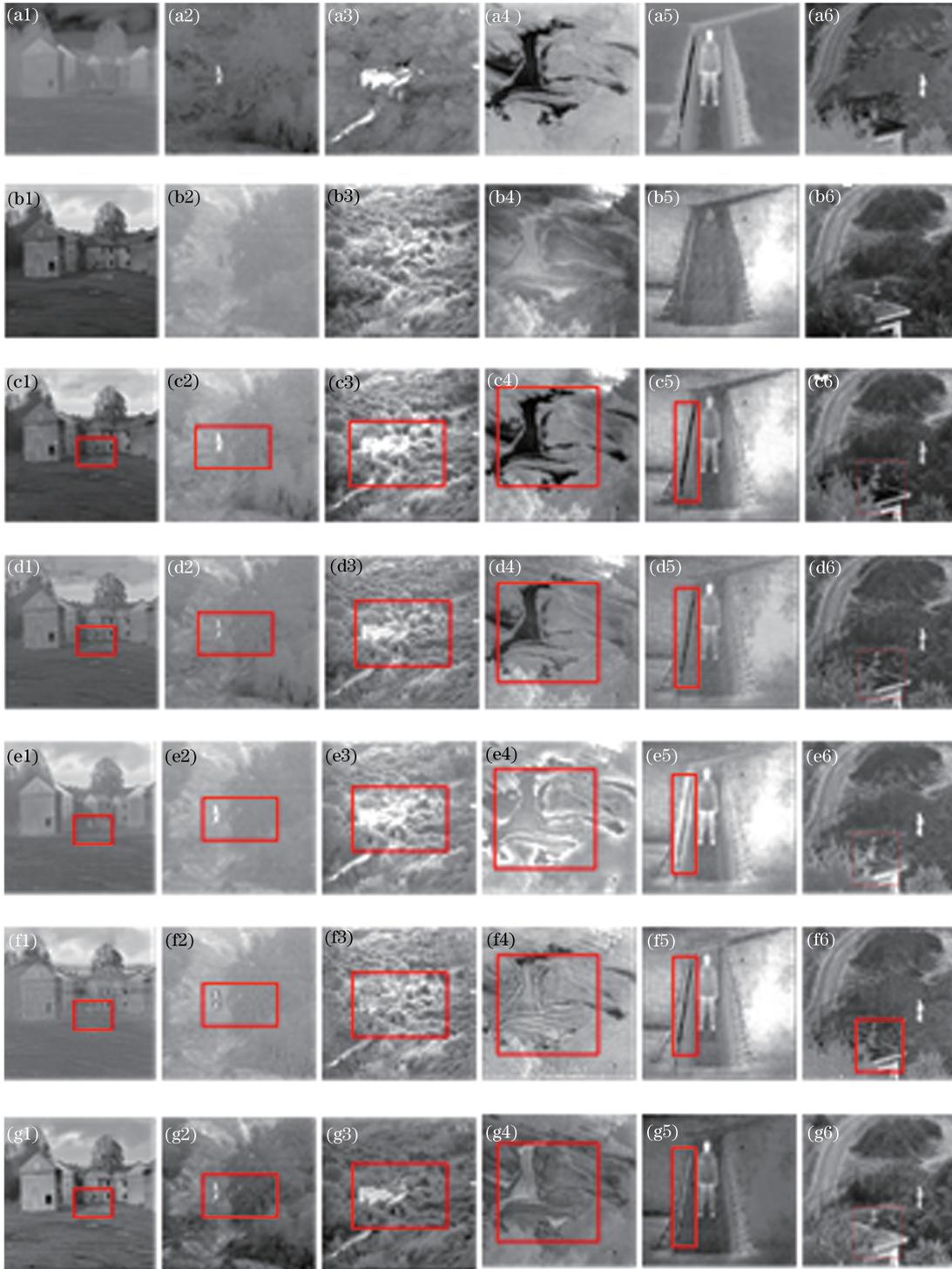


图 9 部分图像融合对比实验结果。(a1)~(a6)红外图像;(b1)~(b6)可见光图像;(c1)~(c6)文献[15]方法;
(d1)~(d6)文献[16]方法;(e1)~(e6)文献[17]方法;(f1)~(f6)文献[18]方法;(g1)~(g6)本文方法

Fig. 9 Partial results of image fusion contrast experiments. (a1)-(a6) Infrared images; (b1)-(b6) visible images; (c1)-(c6) method in Ref. [15]; (d1)-(d6) method in Ref. [16]; (e1)-(e6) method in Ref. [17]; (f1)-(f6) method in Ref. [18]; (g1)-(g6) proposed method

4 结 论

提出了一种基于卷积自编码器和残差块的红外

与可见光图像融合方法。将源图像用作编码器的输入,通过卷积层和残差块来获取图像的特征图;将得到的特征图采用改进的基于 L1-norm 的相似度融合

策略进行融合,将其整合为一个包含源图像显著特征的特征图;利用解码器对融合后的图像进行重构。实验结果表明,与其他融合方法相比,本文方法能够有效地保留图像的中间层信息,融合后的图像更加清晰自然,是一种有效的融合方法。接下来,将在保证效率的情况下进一步优化网络模型,使得网络的性能更好。同时,将继续研究融合层策略,使其能够更好地应用于其他融合任务,如多聚焦图像、遥感图像等。

参 考 文 献

- [1] 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.
- [2] Ma J Y, Ma Y, Li C. Infrared and visible image fusion methods and applications: a survey [J]. *Information Fusion*, 2019, 45: 153-178.
- [3] Liu X H, Chen Z B. Fusion of infrared and visible images based on multi-scale directional guided filter and convolutional sparse representation [J]. *Acta Optica Sinica*, 2017, 37(11): 1110004.
刘先红, 陈志斌. 基于多尺度方向引导滤波和卷积稀疏表示的红外与可见光图像融合 [J]. *光学学报*, 2017, 37(11): 1110004.
- [4] Lin S Z, Han Z. Images fusion based on deep stack convolutional neural network [J]. *Chinese Journal of Computers*, 2017, 40(11): 2506-2518.
蔺素珍, 韩泽. 基于深度堆叠卷积神经网络的图像融合 [J]. *计算机学报*, 2017, 40(11): 2506-2518.
- [5] Li H, Wu X J, Kittler J. Infrared and visible image fusion using a deep learning framework [C] // 2018 24th International Conference on Pattern Recognition (ICPR), August 20-24, 2018, Beijing, China. New York: IEEE, 2018: 2705-2710.
- [6] Liu Y, Chen X, Peng H, *et al.* Multi-focus image fusion with a deep convolutional neural network [J]. *Information Fusion*, 2017, 36: 191-207.
- [7] Prabhakar K R, Sai Srikr V, Babu R V. DeepFuse: a deep unsupervised approach for exposure fusion with extreme exposure image pairs [C] // 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 4724-4732.
- [8] 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 (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 770-778.
- [9] Lu Y S, Li Y X, Liu B, *et al.* Hyperspectral data haze monitoring based on deep residual network [J]. *Acta Optica Sinica*, 2017, 37(11): 1128001.
陆永帅, 李元祥, 刘波, 等. 基于深度残差网络的高光谱遥感数据霾监测 [J]. *光学学报*, 2017, 37(11): 1128001.
- [10] 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.
- [11] Zhao H, Gallo O, Frosio I, *et al.* Loss functions for image restoration with neural networks [J]. *IEEE Transactions on Computational Imaging*, 2017, 3(1): 47-57.
- [12] Dong C, Loy C C, He K M, *et al.* Image super-resolution using deep convolutional networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): 295-307.
- [13] Huang R, Zhang S, Li T Y, *et al.* Beyond face rotation: global and local perception GAN for photorealistic and identity preserving frontal view synthesis [C] // 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 2458-2467.
- [14] Tao L, Zhu C, Xiang G Q, *et al.* LLCNN: a convolutional neural network for low-light image enhancement [C] // 2017 IEEE Visual Communications and Image Processing (VCIP), December 10-13, 2017, St. Petersburg, FL, USA. New York: IEEE, 2017: 17614346.
- [15] 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.
- [16] 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.
- [17] 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: 227-237.
- [18] Chen M S. Image fusion of visual and infrared image based on NSCT and compressed sensing [J]. *Journal of Image and Graphics*, 2016, 21(1): 39-44.
陈木生. 结合 NSCT 和压缩感知的红外与可见光图像融合 [J]. *中国图象图形学报*, 2016, 21(1): 39-44.