

# 基于区域特征优化及边缘增强的多聚焦图像融合

王程<sup>1</sup>, 王巍<sup>2</sup>, 杨馨<sup>1</sup>, 刘晓文<sup>1</sup>

(1. 中国人民公安大学 信息网络安全学院, 北京 100038;  
2. 中国人民公安大学 公安遥感应用工程技术研究中心, 北京 100038)

**摘要:**针对多聚焦图像融合中存在的聚焦信息缺失、边缘特征表达不充分的问题,提出了一种基于聚焦区域特征优化及边缘增强的多聚焦图像融合方法。首先,提出了一种四流向的聚焦特征增强网络,在图像编码过程中实现聚焦特征的增强与补全;其次,提出了一种新的基于空间频率(NSF)的聚焦区域提取策略,实现对聚焦区域特征的高效提取;最后,提出了一种边缘特征增强模块,通过计算图像边缘梯度,实现对目标区域的边缘信息增强。在Lytro数据集与MFW数据集与8种融合方法进行定性与定量对比。实验结果表明,该融合方法较好地解决了多聚焦融合图像中聚焦特征表达不充分、纹理细节不完整和边缘模糊等问题,视觉效果显著。

**关键词:**多聚焦图像;图像融合;卷积神经网络;特征优化;边缘增强

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2024)04-0062-08

doi:10.20165/j.cnki.ISSN1673-629X.2024.0010

## Multi-focus Image Fusion Based on Feature Optimization of Region and Edge Enhancement

WANG Cheng<sup>1</sup>, WANG Wei<sup>2</sup>, YANG Xin<sup>1</sup>, LIU Xiao-wen<sup>1</sup>

(1. School of Information and Cyber Security, People's Public Security University of China, Beijing 100038, China;  
2. Remote Sensing Center of Public Security, People's Public Security University of China, Beijing 100038, China)

**Abstract:** In order to solve the problems of missing focus information and insufficient expression of edge details in multi-focus fusion images, a multi-focus image fusion method based on feature optimization of focus region and edge enhancement is proposed. Firstly, a four direction focusing feature enhancement network is proposed, which can enhance and complete the focusing feature during image coding. Secondly, a new focus region extraction strategy based on spatial frequency (NSF) is proposed, which can efficiently extract the features of focus region in the coding process. Finally, an edge feature enhancement module is proposed, which can effectively calculate the edge gradient of the image to enhance the edge information of the target area. Qualitative and quantitative comparisons were conducted with eight fusion methods in the Lytro dataset and the MFW dataset. The experimental results show the proposed fusion method can solve the problems of insufficient expression of focus features, incomplete texture details and fuzzy edges in multi-focus fusion images effectively, with remarkable visual effect.

**Key words:** multi-focus image; image fusion; convolutional neural network; feature optimization; edge enhancement

## 0 引言

多聚焦图像融合是图像处理中的重要课题。光学透镜的局限性表现为景深(DOF)范围内的物体在图像中具有聚焦和清晰的外观,而其他物体很可能被模糊<sup>[1]</sup>,因此,不同景深的目标很难在同一相机镜头中全部对焦<sup>[2]</sup>。研究人员提出了一种低成本的多聚焦图像融合方法,将同一场景但不同焦点的两组或多组源图像,通过一定算法融合成一幅全焦的图像<sup>[3]</sup>。多聚焦图像融合已经广泛应用于遥感、光学显微镜、视觉传感

器网络等众多领域。主要分为基于传统的方法和基于深度学习的方法。

基于传统的方法可以分为两类:空间域方法与变换域方法。空间域方法不需要将输入图像转换为其他表示形式,直接在空间域中对源图像像素值进行操作。但是对噪声较为敏感,由于子区域的错误决策,导致融合图像中的伪影和空间失真,具有一定局限性。变换域比空间域更能清晰准确地反映图像的显著特征,主要包括三个阶段,即图像变换、系数融合和逆变换。但

收稿日期:2023-07-18

修回日期:2023-11-21

基金项目:国家重点研发计划项目(2022YFC31400);高分辨率对地观测系统重大专项(01-Y30F05-9001-20/22,GFZX0404130307)

作者简介:王程(1999-),男,硕士研究生,研究方向为图像融合;通讯作者:王巍(1986-),男,博士,讲师,研究方向为公安遥感应用。

是需手工设计转换方法,表达源图像的能力有限,并且对输入的鲁棒性较差,计算复杂度高,需要较高的内存。上述传统方法均需要手工设计融合规则,且实际图像场景更为复杂,会极大限制融合方法的特征表达能力,影响融合规则的鲁棒性。

近些年,随着深度学习技术的发展,人们开始用深度学习解决多聚焦图像融合问题。Liu 等人<sup>[3]</sup>首次将卷积神经网络(Convolutional Neural Network, CNN)引入该领域,通过学习 CNN 模型,可以联合生成传统方法中需分别处理的活动水平测量和融合规则。深度学习可以通过深层特征学习焦点测量,在训练时自动学习融合规则,比传统方法更加合理,可分为两类:分类模型方法与回归模型方法。分类模型一般是将焦点检测作为分类问题,输入一个图像块,通过预先设计好的分类器判断图像块的聚散焦性质,因大多融合框架非端到端,后处理步骤需使用一些源于像素的空间域方法,以获得融合结果。回归模型旨在学习源图像到融合图像的端到端映射,通过单一网络模型与目标函数得到最终的融合结果,规避多模块时训练不一致、计算偏差等问题。目前大多数方法都是采用端到端的单流或者双流网络,因网络结构设计以及损失函数误差等因素,经常无法实现对聚焦信息的充分提取,出现信息缺失、图像模糊等问题。而且图像的边界形状不规则,区域聚焦特性十分复杂,许多方法都存在一定的边缘信息丢失,边界区域的焦点像素融合也是深度学习方法中的难题,对融合图像的视觉质量有着重要的影响。

为了解决传统方法的人工设计问题,与部分深度学习方法中特征提取不充分、图像模糊、边缘特征不明显等问题,该文提出了一种新的融合方法。首先,设计了一种端到端的基于聚焦信息增强的四流向特征提取网络,实现特征补充与增强;其次,设计了一种新的基于空间频率的聚焦区域提取策略,实现对聚焦区域信息的高效提取;此外,设计了一种边缘特征增强模块,作用于源图像实现对中间特征图的边缘信息增强。该方法可有效解决深度学习网络存在的特征提取不充分、图像模糊、边缘细节缺失等问题。

综上所述,该文的贡献有四个方面:

(1)为了解决深度学习网络在特征提取中存在的聚焦信息表达不充分、细节缺失等问题,设计了一种基于聚焦信息增强的四流向特征提取网络,实现对焦点信息的高效提取与补全。

(2)为了解决聚焦信息特征提取不充分的问题,提出了一种新的基于空间频率(New Spatial Frequency, NSF)的聚焦区域提取策略,实现对源图像聚焦信息的提取最大化,减少散焦区域冗余的影响。

(3)为了解决融合图像边缘信息表达不充分的问题,

设计了一种边缘特征提取模块,直接作用于源图像进行边缘提取,有效解决边界形状不规则、聚焦特性复杂的问题。

(4)在两个数据集中与传统方法、深度学习方法进行对比,视觉效果显著,各项指标良好,定性和定量实验结果充分表明该方法优于其他 8 种对比方法。

## 1 相关工作

### 1.1 基于传统的融合方法

传统方法分为空间域方法与变换域方法,目前发展较为成熟。

空间域方法直接对源图像像素值进行操作。变换域方法分类较多,主要有基于多尺度分解(MSD)的方法、基于稀疏表示(SR)的方法、基于梯度域(GD)的方法。Amin-Naji 等人<sup>[4]</sup>提出了基于 DCT 域奇异值分解的融合方法,适合于实时应用,同时在噪声条件下可得到良好的融合结果。随后, Amin-Naji 等人<sup>[5]</sup>又提出了基于拉普拉斯方差、能量和相关系数(DCT\_EOL)的融合方法,在 DCT 域中提出了两种强聚焦测量的计算方法与两种新的焦点测量方法。

### 1.2 基于深度学习的融合方法

深度学习的融合方法可分为基于分类模型的方法与基于回归模型的方法。

分类模型的方法通过模型鉴别像素块的聚焦特征。Amin-Naji 等人<sup>[6]</sup>提出了基于 CNN 的集成学习方法,通过追求模型和数据集之间的多样性,减少训练数据集中的过拟合问题。回归模型的方法旨在学习源图像到融合图像的端到端映射,通过网络训练得到最终的融合结果,目前有多种先进方法。Li 等人<sup>[7]</sup>提出了深度回归对学习模型,在传统的端到端 CNN 模型中,将输入图像分割成小块,从这些小块中获取信息,则将整个图像转换为二值掩膜后再进行融合。Zhang 等人<sup>[8]</sup>提出了梯度和强度比例保持的融合网络,将图像融合统一为源图像的纹理和强度比例保持的问题。将网络划分为梯度路径和强度路径进行信息提取,并引入路径传递块,进行不同路径间的信息交换。Ma 等人<sup>[9]</sup>提出了同时生成决策图和融合结果的端到端学习网络,引入梯度感知损失函数保留输出融合图像中的梯度信息;并设计了一种决策校准策略以减少融合中的时间消耗。

该文提出了一种新的融合方法。通过四流向特征提取网络,在编码过程中实现特征补充与增强;通过新空间频率(NSF)提取策略,实现对聚焦区域信息的高效抓取;通过边缘特征增强模块,实现对融合图像的边缘信息增强。该方法可在网络提取中尽量减少聚焦信息丢失,充分提取深层次特征,同时保留源图像聚焦区

域的边缘细节信息。

## 2 文中方法

文中方法的融合网络框架如图 1 所示,本节将对网络结构、模块结构、聚焦策略、损失函数和网络细节

进行详细阐述。

### 2.1 基于区域特征优化及边缘增强的网络结构

图 1 给出了基于区域特征优化及边缘增强的网络结构,具体步骤及细节如下:

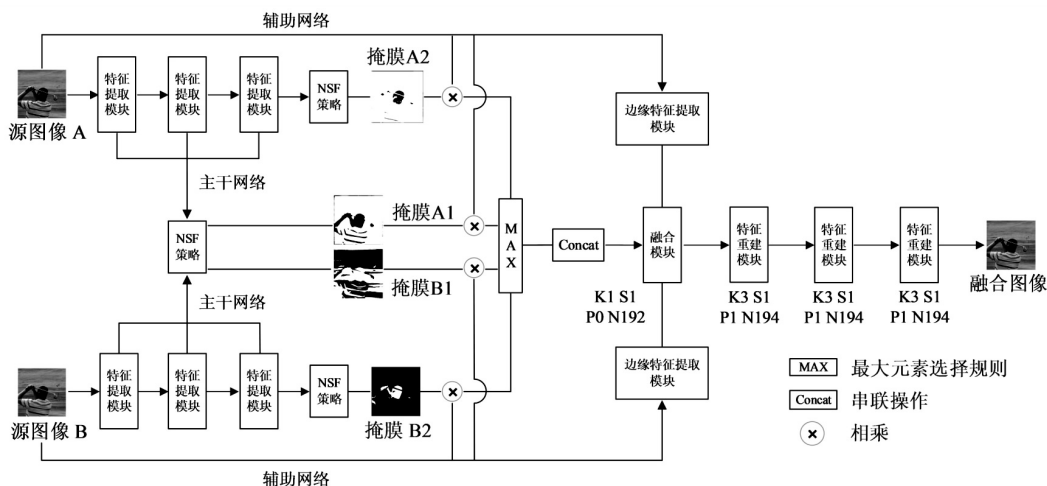


图 1 网络结构

(1)将源图像输入至网络,通过三组特征提取模块串联提取聚焦信息。卷积运算会损失部分信息,无法实现对源图像细节特征的充分提取,并且卷积块经过多次串联会造成信息的进一步丢失。该文在特征提取过程中将每组特征提取模块得到的中间结果进行相加运算,将各运算阶段、各通道的特征结果进行整合,实现特征增强以解决上述问题。

(2)在特征提取模块后,引入了一种新的基于空间频率(New Spatial Frequency, NSF)的聚焦区域提取策略,实现对聚焦区域特征的进一步优化。将 NSF 策略所得结果转化为二值掩膜 A1, B1, 通过掩膜操作放大聚焦信息,消除冗余信息干扰,实现对中间特征的细化<sup>[9]</sup>,通过与源图像相乘再次还原彩色特征,该文将此两条分流网络称之为主干网络。将多层串联的卷积提取网络作为辅助网络,即生成掩膜 A2, B2 的两条网络,实现对主干网络的信息补充与增强,减少卷积块多次串联运算造成的信息丢失。

(3)为了更好地保留场景的焦点信息,获得更高的视觉质量,采用元素最大选择规则对不同源图像提取的特征进行融合,元素最大选择规则此前已在卷积神经网络通用融合框架中被验证是一种有效的方法。

(4)为了解决边缘细节缺失、伪影等问题,引入了一种基于图像梯度的边缘增强模块,通过对源图像求取梯度图进行边缘提取。而后将边缘梯度图与经过元素最大选择规则、Concat 操作得到的最终特征图通过特征融合模块进行融合。

(5)将以上步骤所得的特征图进行特征重建得到最终的融合图像。

采用此端到端的网络结构,可规避多模块的固有缺陷,减少错误积累与网络的复杂程度;便于引入所设计的辅助网络以实现对主干网络的效果增强,以及 NSF 策略与边缘特征增强模块,保持网络结构的整体性与便捷性。网络中输入与输出通道数均为 16。特征提取模块中,Conv 大小为  $3 \times 3$ ,步长与填充设置为 1,并添加 LeakyReLU 激活层;Conv $1 \times 1$  步长设置为 1, Aconv 大小为  $3 \times 3$ ,步长为 1,填充为 1; Aatt 卷积设置为  $7 \times 7$ ,步长为 1,填充为 0。特征融合模块中卷积大小设置为 1,步长为 1,填充为 0,通道数位 192;特征重建模块卷积大小设置为  $3 \times 3$ ,步长为 1,填充为 1,通道数为 194,每个卷积后均使用 LeakyReLU 函数作为激活层。

### 2.2 特征提取模块

传统卷积根据卷积滤波器权重在局部接受域中利用聚合函数,权重在整个特征映射中进行共享,关注于局部特征。Transformer 模型中的自注意力机制采用基于输入特征上下文的加权平均操作,通过相关像素对之间的相似性函数动态计算注意力权重,使自注意力模块能够自适应地关注长距离相关区域,具有高度的灵活性。因此,卷积和自注意力机制通常被认为是彼此不同的对等方法,将二者的特点进行结合,可实现对聚焦信息的高效抓取。

Pan 等人<sup>[10]</sup>证明了二者之间存在强大的潜在联系,将自注意力与卷积的优势进行结合,同时与单纯卷积或自注意力相比又具有更小的计算成本。结合其思想,利用二者同时进行特征提取。以图 1 中第二、第三特征提取模块的部分输出结果为例,如图 2 所示。在



阶段一,首先通过一个  $3 \times 3$  卷积将结果分别输入至三个  $1 \times 1$  卷积进行映射,重塑为  $N$  个片段,获得了  $3 \times N$  组特征映射的中间特征集,所含信息丰富。在第二阶段,对于自注意力路径,收集到  $N$  组中间特征,其中每组包含三个特征, key、query 和 value 为对应的特征映

射,遵循传统的多头注意力模块;对于核大小为  $k$  的卷积路径,采用全连接层,生成  $k^2$  个特征映射,通过移动和聚合生成的特征,以卷积方式处理输入特征,从局部接受域收集信息。最后,将两条路径的输出进行相加,  $\alpha, \beta$  为可学习的标量。

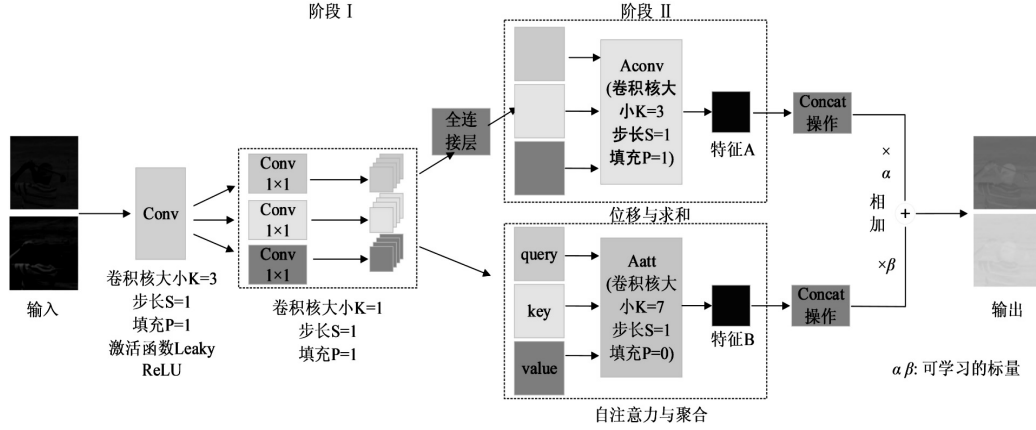


图2 特征提取模块

### 2.3 NSF 聚焦区域提取策略

原始的空间频率(SF)在单通道灰度图像进行计算,因此,对于卷积提取到的深度特征需修改SF的计算方法<sup>[9]</sup>。该文提出了一种新的空间频率(NSF)聚焦区域提取策略,继续提取中间特征图的深层次聚焦特征。将特征提取模块提取的特征表示为  $F$ ,  $F_{(x,y)}$  表示一个特征向量,  $(x,y)$  表示向量在图像中的坐标。从三个方面计算图像的频率,RF、CF 分别代表行向量频率、列向量频率,DF 为斜向量频率,结合 RF、CF 的特点,同时进行行与列的向量计算,完善提取策略,充分捕捉聚焦信息,具体公式如下:

$$RF_{(x,y)} = \sqrt{\sum_{a=-r}^r \sum_{b=-r}^r [F_{(x+a,y+b)} - F_{(x+a,y+b-1)}]^2} \quad (1)$$

$$CF_{(x,y)} = \sqrt{\sum_{a=-r}^r \sum_{b=-r}^r [F_{(x+a,y+b)} - F_{(x+a-1,y+b)}]^2} \quad (2)$$

$$DF_{(x,y)} = \sqrt{\sum_{a=-r}^r \sum_{b=-r}^r [F_{(x+a,y+b)} - F_{(x+a-1,y+b-1)}]^2} \quad (3)$$

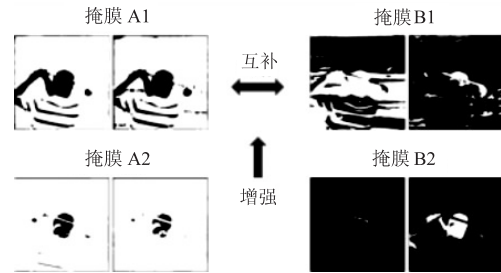
$$NSF_{(x,y)} = \sqrt{\frac{(CF_{(x,y)})^2 + (RF_{(x,y)})^2 + (DF_{(x,y)})^2}{(2r+1)^2}} \quad (4)$$

其中,  $r$  是核半径,文中  $r=5$ 。原始的空间频率是基于块的,文中方法是基于像素的,在特征图的边界采用“相同”填充策略。

### 2.4 掩膜生成

经过 NSF 策略,得到信息丰富的深层次聚焦特征图,通过 Sigmoid 函数将特征图转化为二值掩膜,细化特征,阈值设置为所有通道掩膜结果的均值。图3给出了部分特征掩膜的结果图。如图3所示,辅助网络

所得的二值掩膜  $A_2, B_2$ , 存在某些  $A_1, B_1$  中未提取到的细节信息,并且也存在  $A_1, B_1$  中已充分提取到的聚焦信息,  $A_2, B_2$  可通过叠加操作实现对主干网络  $A_1, B_1$  的特征增强,对于未曾提取到的细节信息,  $A_2, B_2$  可以实现对图中聚焦信息的有效补充增强。掩膜  $A_1$  与  $B_1$ 、掩膜  $A_2$  与  $B_2$  彼此互为近焦与远焦结果,可起到良好的特征互补效果。



### 2.5 边缘特征增强模块

该文设计了一种边缘特征增强模块,具体结构如图4所示。将源图像分别输入至 Laplacian 算子与 Sobel 算子求取梯度图,再将二者结果相加得到最终的边缘特征梯度图。Laplacian 算子通过灰度差分计算邻域内的像素,判断中心像素灰度值与它周围其他像素的灰度值,若中心像素灰度更高,则提升中心像素的灰度;反之降低其灰度,以实现图像锐化。该算子是各向同性的,可对任意方向的界线与线条进行锐化,突出强度快速变化的区域。但是对噪声敏感,对孤立像素的响应比对边缘或线的响应更强烈。Sobel 算子常用于边缘检测,结合高斯平滑和微分求导,可计算明暗程度近似值,根据边缘附近的明暗程度把该区域内超过

一定数量的特定像素点记为边缘。该算子的边缘定位更准确,常用于噪声较多、灰度渐变的图像,因为引入了类似局部平均的运算,因此对噪声具有平滑作用。根据上述两种算子的特点,将二者的优势结合,实现彼

此的特征互补,计算图像的近似梯度,梯度越大成为边缘的可能越高,最终将二者的计算结果相加,系数根据前后调整均设置为 1。

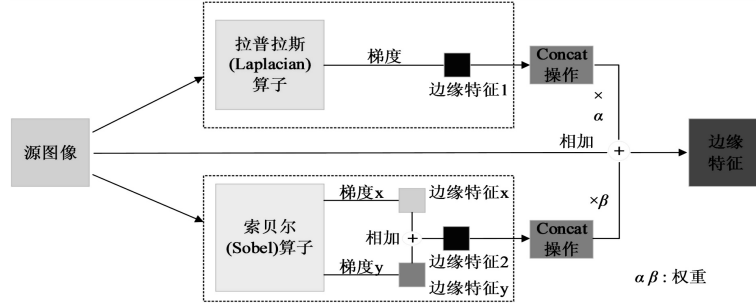


图 4 边缘特征提取模块

## 2.6 图像融合与特征重建

经过 NSF 提取策略与阈值分割后,可得到信息丰富的二值图,将其与源图像逐像素相乘后的结果送入最大元素选择规则进行筛选与特征拼接。元素融合规则对多个输入的卷积特征进行融合,在实际应用中,应根据图像数据集的特点选择融合规则,经常使用元素最大融合规则来融合多焦点图像、红外图像以及医学图像<sup>[10]</sup>。随后,将拼接结果送入至特征融合模块,实现信息融合。该模块可自适应地对四条分支网络掩膜图像中的冲突区域进行加权融合,实现自适应学习,将有意义的信息进行增强处理,无关部分进行削弱,保证信息最大化与完整性。特征融合后,将结果输入至特征重建模块进行特征解码重建,最终通过 Sigmoid 函数将特征图还原,得到最终的彩色 RGB 融合图像。

## 2.7 损失函数

为了精确地重建融合图像,保留源图像的结构信息,减小两者的差异,必须选择适当的损失函数以使误差最小化。损失函数  $L$  定义如下:

$$L = \alpha L_{SSIM} + \beta L_1 + \gamma L_{grad} \quad (5)$$

其中,  $L_{SSIM}$  表示结构相似性损失函数,  $L_1$  表示 L1 范数损失函数,  $L_{grad}$  表示梯度损失函数,  $\alpha, \beta, \gamma$  代表权重,经大量实验验证,  $\alpha, \beta, \gamma$  分别初始化为 0.8, 0.2, 0.4 时融合效果较好。

SSIM 损失函数<sup>[11]</sup>从三个维度(亮度、对比度和结构)比较两图像间的信息相似性,将数据进行归一化后,分别计算图像的亮度(均值),对比度(方差)和归一化后的像素向量三者相似度,然后再将三者相乘,具体公式如下:

$$SSIM(a, b) = \frac{(2\mu_a\mu_b + C_1)(\sigma_{ab} + C_2)}{(\mu_a^2 + \mu_b^2 + C_1)(\mu_a^2 + \mu_b^2 + C_2)} \quad (6)$$

$$L_{SSIM} = 1 - SSIM(a, b) \quad (7)$$

其中,  $a, b$  分别代表所预测的图像与真实值,  $SSIM(a, b)$  则代表二者的结构相似性,  $\mu$  表示二者的平均值,  $\sigma$

为协方差,  $C_1$  与  $C_2$  是常数,为了避免分母为 0。

L1 范数损失函数<sup>[12]</sup>用于计算融合图像与源图像之间的差异,将目标值与估计值的绝对差值总和最小化,即最小化误差,具体公式如下:

$$L_1 = |a - b| \quad (8)$$

其中,  $a, b$  分别代表所预测的图像与真实值图像。

为了使融合后的图像具有更清晰的外观,增加了约束梯度差的梯度损失函数<sup>[13]</sup>。梯度损失函数由所预测图像  $a$  和真实图像  $b$  的梯度变化间的平方 Frobenius 范数表示,具体公式如下:

$$L_{grad} = \frac{\|\nabla a - \nabla b\|_F^2}{HW} \quad (9)$$

其中,  $\nabla$  代表梯度变化,  $H, W$  代表图像的高度与宽度。

## 3 实验结果与分析

### 3.1 数据集与实验设置

基于监督学习的方法大都采用合成训练集, Zang 等人<sup>[14]</sup>提出了基于 DUTS 数据集生成多焦点数据集的新方法,由四组不同模糊程度的图像组成,比之前的工作更多样化,更接近现实,根据训练生成足够数量的多焦点图像对,每对生成的多焦点图像对都拥有 ground-truth 图像。该文使用 Zang 等人<sup>[14]</sup>公开的 216 幅高质量图像作为训练集,每次随机选取 4 幅图像,周期数为 600,每 10 个周期存储一次结果。使用 Pytorch1.8.1 框架实现提出的融合模型。实验设备 CPU 型号为 Intel(R) Xeon(R) Platinum 8255C CPU @ 2.50 GHz,算力型号 RTX 3080,显存 10 GB。在 Lytro 数据集<sup>[15]</sup>与 MFFW 数据集<sup>[16]</sup>进行测试,两组数据集分别有 20 组和 13 组不同场景的近焦图像与远焦图像。

### 3.2 消融实验

#### 3.2.1 特征增强网络消融实验

辅助网络可起到信息补充与增强作用。在图 2

中,不引入辅助网络,将上述结果与原网络结果进行对比,如图5所示。第3列中“Man”存在明显黑色噪声区域,“Dog”较为模糊且存在聚焦信息损失,其他结果均存在同样问题;第4列中“Man”未发现明显噪声块,“Dog”较为清晰,所含信息丰富,其他结果表现良好。结果充分证实了特征增强网络的有效性。

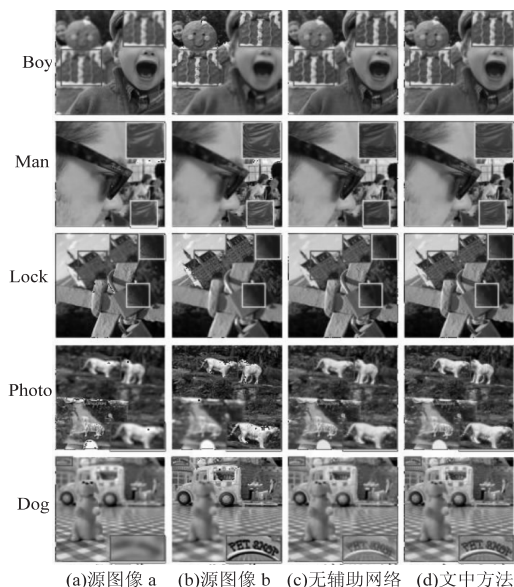


图5 特征增强网络消融结果

### 3.2.2 NSF 策略消融实验

为了验证 NSF 策略的有效性,在图1中首先不引入该策略,其它结构保持不变,将上述结果输出与原网络结果进行对比,如图6所示。第3列中,“Wine”与“Opera house”边缘特征模糊,聚焦信息提取不充分,其他图像所选区域饱和度过高;第4列中“Wine”与“Opera house”边缘轮廓完整清晰,聚焦信息丰富,贴近于源图像,其他图像色彩表现良好。验证了 NSF 策略的优越性。

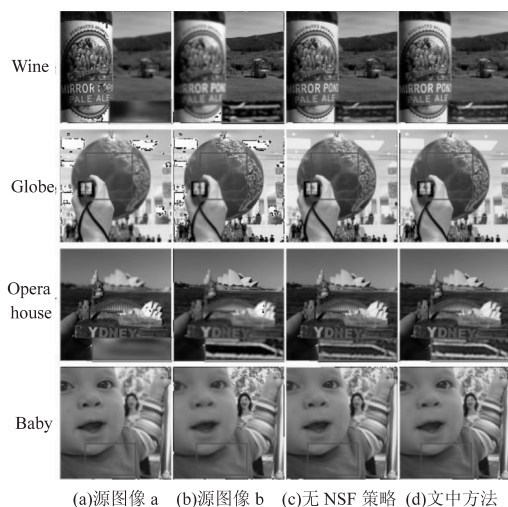


图6 NSF 策略消融结果

### 3.2.3 边缘特征增强模块消融实验

为验证该文设计的边缘增强模块的有效性,在图

1网络结构中首先不引入该模块,其它结构保持不变,将上述结果与原网络结果进行对比,如图7所示。第3列中,“Opera house”边缘不够清晰,且存在色差,“Baby”边缘存在伪影,轮廓不够平滑;第4列中“Opera house”较为清晰,轮廓细节及颜色特征更贴近源图像,“Baby”边缘特征表现良好。说明了边缘特征提取模块的可行性。

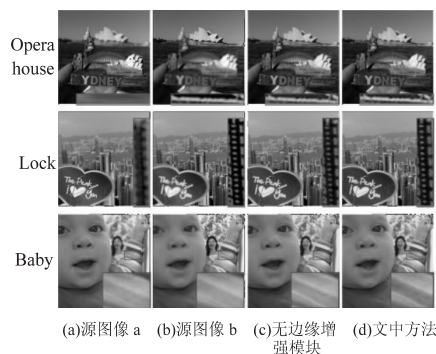


图7 边缘增强模块消融结果

### 3.3 对比实验

为评价文中方法在不同场景图像的聚焦融合性能,在 Lytro 数据集和 MFFW 数据集上进行对比实验,分别进行主观和客观评价。

#### 3.3.1 主观分析

选取了八种融合方法进行对比,具体为 NSCTRR、DCT\_EOL、SVDDCT、CNN、DRPL、ECNN、GCF 和 PMGI。GCF 是 Xu 等人<sup>[17]</sup>提出的梯度和连通区域的方法;NSCTRR、DCT\_EOL、SVDDCT 是基于传统的方法;CNN、DRPL、ECNN、GCF、PMGI 是基于分类与回归模型的深度学习方法。

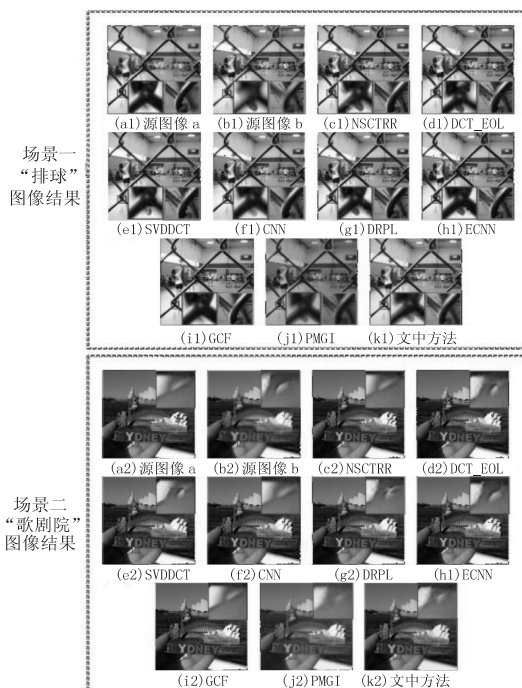


图8 Lytro 数据集图像结果对比



图 8 场景一中, (d)、(e)、(g)、(h)、(j) 存在噪声像素及伪边缘, 清晰度不足; (c)、(f)、(i) 未能成功提取铁网间的聚焦信息, 部分图像存在噪声块。场景二中, (d)、(e)、(g)、(h) 存在明显噪声块, 干扰聚焦像素, (c)、(f)、(i)、(j) 较为模糊, 未能成功识别。图 9 给出了文中方法与其他八种方法在 MFFW 数据集的结果, 该数据集包含大小不同的图像, 在此数据集进行测试也可验证文中方法的鲁棒性。图 9 场景一中, (f)、(j) 存在模糊、边缘区域伪影现象; 其他方法均存在明显的噪声。场景二中除 (j) 外, 在两区域均存在大范围噪声, 影响视觉效果; (j) 清晰度不足, 与源图像存在差距。经对比, 文中方法不存在上述问题, 适应性更强, 在聚焦区域及边界处表现优秀, 视觉效果显著。

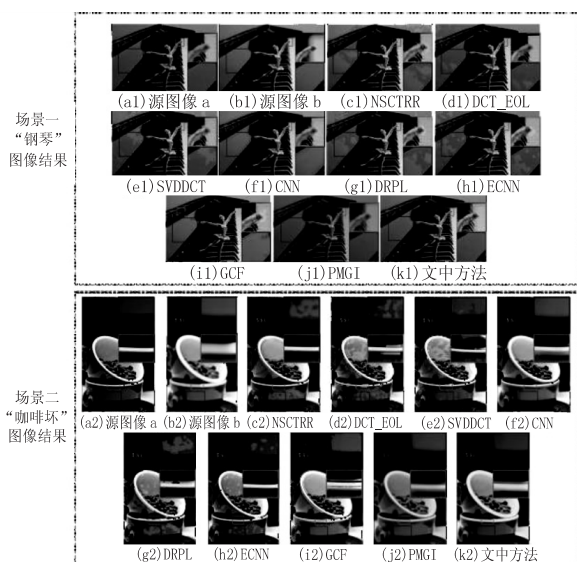


图 9 MFFW 数据集图像结果对比

### 3.3.2 客观分析

本节通过结构相似性(SSIM)、相关系数(CC)<sup>[18]</sup>、

峰值信噪比(PSNR)<sup>[19]</sup>和均方根误差(RMSE)<sup>[20]</sup>, 从图像色彩、结构、同源图像相关性、信息丰富度等方面进行评价。前三项指标数值越大融合效果越好; RMSE 数值越小融合效果越好。SSIM 常用于单通道灰度图像, 文中融合结果为 RGB 图像, 因此将分别计算 RGB 图像三通道的 SSIM 值, 并求得算术平均数得到每幅图像的 SSIM, 最后将融合图像与两幅源图像的计算结果相加得到最终所需要的 SSIM 值。CC 表达的是两个数值变量的线性相关性, 将向量标准化后求余弦夹角。PSNR 可衡量融合图像的失真效果。RMSE 反映变量间的差异程度, 可用于衡量融合图像与源图像的偏差, 反映图像的空间细节信息。表 1 给出了文中方法与八种对比方法在 Lytro、MFFW 数据集中所有融合图像的指标均值。

Lytro 中, 文中方法在 SSIM、PSNR、RMSE 上优于其他对比方法, CC 呈次优值, 此结果与主观视觉效果一致。SSIM 指标优于其他对比方法, 说明文中方法具有良好的对比度、亮度和结构, 具有丰富的色彩信息与边缘结构特征。CC 指标反映了文中方法与源图像的线性相关程度更为密切, 更加还原源图像信息。PSNR 指标说明文中方法受失真与噪声的影响较小, 聚焦区域特征清晰, 所选区域无明显噪声, 可充分印证。RMSE 指标证实了文中方法有着更加丰富的空间细节特征信息。以上指标均与图 8、图 9 主观视觉反应一致, 优于其他方法。MFFW 中, 文中方法在四项指标上均优于其他方法, 说明了所含聚焦特征的充分性和边缘信息的丰富性, 同时, 在不同尺寸的数据集测试其结果也说明了该方法的鲁棒性。通过在两组数据集的比较, 反映了客观指标与主观视觉的一致性, 充分验证了所提方法的优越性。

表 1 客观指标对比

	Lytro 数据集				MFFW 数据集			
	SSIM	CC	PSNR	RMSE	SSIM	CC	PSNR	RMSE
NSCTRR	1.750 1	0.976 1	28.524 1	10.596 8	1.601 6	0.953 8	25.265 3	15.384 9
DCT_EOL	1.749 2	0.975 9	28.498 2	10.618 0	1.623 9	0.951 1	25.151 6	15.591 9
SVDDCT	1.749 2	0.975 9	28.512 5	10.613 7	1.584 8	0.949 3	24.996 1	15.769 2
CNN	1.753 0	0.976 5	28.630 9	10.475 9	1.617 2	0.955 8	25.473 2	15.004 7
DRPL	1.750 7	0.976 0	28.500 0	10.631 5	1.623 2	0.955 1	25.363 5	15.309 6
ECNN	1.749 1	0.975 9	28.471 0	10.643 1	1.587 1	0.953 2	25.288 1	15.435 5
GCF	1.749 7	0.976 0	28.454 0	10.649 0	1.598 2	0.953 2	25.223 0	15.463 4
PMGI	1.692 9	0.980 1	23.519 5	18.468 8	1.479 6	0.963 9	20.758 7	18.468 8
Ours	1.780 5	0.979 7	28.979 7	9.440 4	1.648 1	0.964 8	26.680 4	12.855 5

## 4 结束语

针对多聚焦融合图像中存在的聚焦信息缺失、边

缘特征表达不充分等问题, 提出了一种端到端的基于聚焦区域特征优化及边缘增强的融合方法, 较好抑制了网络编码过程中的聚焦信息丢失、图像模糊, 实现了

对源图像深层次特征的充分提取,也实现了对目标图像的边缘特征强化,实现了图像高效融合。定性和定量实验结果表明,该方法可有效地保留聚焦区域信息及边缘细节信息,融合后的图像纹理细节丰富、特征明显,充分验证了算法的可行性、有效性。

#### 参考文献:

- [1] BHAT S, KOUNDAL D. Multi-focus image fusion techniques: a survey[J]. *Artificial Intelligence Review*, 2021, 54(8): 5735–5787.
- [2] 王 磊, 齐争争, 刘 羽. 深度学习多聚焦图像融合方法综述[J]. *中国图象图形学报*, 2023, 28(1): 80–101.
- [3] 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.
- [4] AMIN-NAJI M, RANJBAR-NOIEY P, AGHAGOLZADEH A. Multi-focus image fusion using singular value decomposition in DCT domain[C]//2017 10th Iranian conference on machine vision and image processing (MVIP). Isfahan: IEEE, 2017: 45–51.
- [5] AMIN-NAJI M, AGHAGOLZADEH A. Multi-focus image fusion in DCT domain using variance and energy of Laplacian and correlation coefficient for visual sensor networks[J]. *Journal of AI and Data Mining*, 2018, 6(2): 233–250.
- [6] AMIN-NAJI M, AGHAGOLZADEH A, EZOJI M. Ensemble of CNN for multi-focus image fusion[J]. *Information Fusion*, 2019, 51: 201–214.
- [7] LI J, GUO X, LU G, et al. DRPL: deep regression pair learning for multi-focus image fusion[J]. *IEEE Transactions on Image Processing*, 2020, 29: 4816–4831.
- [8] ZHANG H, XU H, XIAO Y, et al. Rethinking the image fusion: a fast unified image fusion network based on proportional maintenance of gradient and intensity[C]//Proceedings of the AAAI conference on artificial intelligence. New York: AAAI, 2020: 12797–12804.
- [9] MA B, YIN X, WU D, et al. End-to-end learning for simultaneously generating decision map and multi-focus image fusion result[J]. *Neurocomputing*, 2022, 470: 204–216.
- [10] PAN X, GE C, LU R, et al. On the integration of self-attention and convolution[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. New Orleans: IEEE, 2022: 815–825.
- [11] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600–612.
- [12] SCHMIDT M. Least squares optimization with L1-norm regularization[J]. *Optimization*, 2005, 98: 230–238.
- [13] XU H, MA J, LE Z, et al. FusionDn: a unified densely connected network for image fusion[C]//Proceedings of the AAAI conference on artificial intelligence. New York: AAAI, 2020: 12484–12491.
- [14] ZANG Y, ZHOU D, WANG C, et al. UFA-FUSE: a novel deep supervised and hybrid model for multifocus image fusion[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 1–17.
- [15] NEJATI M, SAMAVI S, SHIRANI S. Multi-focus image fusion using dictionary-based sparse representation[J]. *Information Fusion*, 2015, 25: 72–84.
- [16] XU S, WEI X, ZHANG C, et al. Mffw: a new dataset for multi-focus image fusion[J]. *arXiv:2002.04780*, 2020.
- [17] XU H, FAN F, ZHANG H, et al. A deep model for multi-focus image fusion based on gradients and connected regions[J]. *IEEE Access*, 2020, 8: 26316–26327.
- [18] TAYLOR R. Interpretation of the correlation coefficient: a basic review[J]. *Journal of Diagnostic Medical Sonography*, 1990, 6(1): 35–39.
- [19] JAGALINGAM P, HEGDE A V. A review of quality metrics for fused image[J]. *Aquatic Procedia*, 2015, 4: 133–142.
- [20] CHAI T, DRAXLER R R. Root mean square error (RMSE) or mean absolute error (MAE)? – arguments against avoiding RMSE in the literature[J]. *Geoscientific Model Development*, 2014, 7(3): 1247–1250.