

金字塔域双边加权引导滤波的多聚焦图像融合算法研究

尹建国¹, 郭立强^{1,2}

¹淮阴师范学院计算机科学与技术学院, 江苏 淮安

²淮阴师范学院江苏省智慧物联大数据工程研究中心, 江苏 淮安

收稿日期: 2025年3月3日; 录用日期: 2025年3月22日; 发布日期: 2025年4月3日

摘 要

为了克服传统多聚焦图像融合方法中的边缘失真和图像模糊问题, 提出了一种金字塔域双边加权引导滤波的融合方法。首先, 利用双边加权引导滤波对均值滤波和差分算子得到粗糙焦点图进行细化并作为金字塔融合过程中高频信息选择策略, 在金字塔下采样过程中, 采用双边加权引导滤波替代传统的高斯滤波, 更有效地保留图像的边缘信息。同时, 为了提升图像细节部分, 以原始待融合图像的均值作为参考, 进行差异运算, 构建了两个图像金字塔, 并对每个金字塔单独处理, 之后利用高频信息决策图进行叠加, 通过重建金字塔得到融合图像。实验结果表明, 无论是在主观视觉感知还是在客观评价指标方面, 本文所提出的方法在视觉保真度等多个指标上优于经典方法, 具有较好的图像融合效果。

关键词

双边加权引导滤波, 图像金字塔, 图像融合

Research on Multi-Focus Image Fusion Algorithm Based on Bilateral Weighted Guided Filtering in Pyramid Domain

Jianguo Yin¹, Liqiang Gou^{1,2}

¹School of Computer Science and Technology, Huaiyin Normal University, Huai'an Jiangsu

²Smart IoT Big Data Engineering Research Center of Jiangsu Province, Huaiyin Normal University, Huai'an Jiangsu

Received: Mar. 3rd, 2025; accepted: Mar. 22nd, 2025; published: Apr. 3rd, 2025

文章引用: 尹建国, 郭立强. 金字塔域双边加权引导滤波的多聚焦图像融合算法研究[J]. 图像与信号处理, 2025, 14(2): 232-244. DOI: 10.12677/jisp.2025.142022

Abstract

In order to overcome the problems of edge distortion and image blur in the traditional multi-focus image fusion method, a fusion method based on Bilateral Weighted guided filtering in the pyramid domain was proposed. Firstly, the bilateral weighted guided filter is used to refine the rough focus map obtained by the mean filter and the difference operator, and as a high-frequency information selection strategy in the pyramid fusion process, the double-sided weighted guided filter is used to replace the traditional Gaussian filter in the pyramid downsampling process, so as to retain the edge information of the image more effectively. At the same time, in order to improve the detail of the image, the average value of two original images with a different focus was used as a reference, and two image pyramids were constructed by different operations, and each pyramid was processed separately, and then the high-frequency information decision map was used to superimpose and obtain the fusion image by reconstructing the pyramid. Experimental results show that the proposed method is superior to the classical method in terms of visual fidelity and other indicators, and has a better image fusion effect.

Keywords

Bilateral Weighted Guided Filter, Image Pyramid, Image Fusion

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

图像融合是指从不同的源图像中提取然后组合最有意义的信息,旨在生成信息量更大且对后续应用更有益的单个图像[1],具体可以分为多聚焦图像融合、多曝光图像融合、红外与可见光图像融合、遥感图像融合以及医学影像融合等[2]。多聚焦图像融合(Multi-focus Image Fusion, MFIF)是图像融合领域中的重要研究方向之一[3],通过合成同一场景下不同聚焦层次的影像,生成一幅拥有更高分辨率和更广深度的图像,具体来说,由于相机的物理限制,一次拍摄只能使特定距离范围内的物体处于最佳聚焦状态,而其他部分则可能显得模糊,多聚焦图像融合技术旨在解决单次拍摄无法同时清晰呈现前景和背景的问题。因此这种技术不仅提高了图像的整体清晰度,还增强了图像的视觉效果,使其更加接近人眼直接观察的结果。多聚焦图像融合广泛应用于显微成像、显微镜检查、摄影艺术等领域,对于提高图像的质量和信量具有重要意义。

目前,多聚焦图像融合技术可分为传统融合模式和基于深度学习的融合模式两大类[4]。传统方法通常涉及在空间域或变换域中,通过一系列数学变换对活动水平进行手工分析,并设计与之相适应的融合策略[5]。例如,在空间域中,常采用简单的像素加权平均、最大值选择等方法;但在变换域中,通常依赖于图像的频域特征,如离散余弦变换(Discrete Cosine Transform, DCT)、主成分分析(Principal Component Analysis, PCA)等。这些方法通常依赖经验和规则,并根据图像的具体特性进行调整。与此相比,基于深度学习的方法则是通过数据驱动的方式,自动从大量图像数据中学习有效的特征表示和融合策略。在实际的图像融合过程中,可以根据信息流的形式将其分为三个主要层次:像素级融合、特征级融合和决策级融合[6]。在像素级融合中,直接对源图像的像素进行处理,通常采用加权平均等简单方法进行融合;

特征级融合则是在对图像进行特征提取之后进行融合；决策级融合是在图像处理的最后阶段，基于每个图像的处理结果做出最终决策。

对于基于空间域的融合方法，它通常直接对源图像进行融合操作，这类方法因为计算量相对较小、实时性较强，因此在许多实际应用中非常有用。例如，在实时视频处理或自动驾驶中，空间域的融合方法可以快速有效地实现目标。而在基于变换域的融合算法中，常使用多尺度分解的方法，如拉普拉斯金字塔[7]、比率低通金字塔[8]、和小波变换[9]等。这些方法通过将图像分解为不同尺度的分量，从而能够更精确地提取出图像的高频和低频信息，进而实现更加精细的融合。

基于深度学习的融合模式凭借其强大的特征学习能力，在众多图像处理和计算机视觉任务中取得了显著的成功。自 2017 年 Liu 等人[10]首次提出基于深度学习的多聚焦图像融合方法以来，该领域迅速成为研究热点，至今已有近百种相关方法相继问世。Shao 等人[11]提出了一种创新的深度学习网络架构，将 Transformer 与卷积神经网络(Convolutional Neural Networks, CNN)相结合；Ouyang 等人[12]首次将图卷积网络(Graph Convolutional Networks, GCN)应用于该领域，通过将块分割与像素优化相结合，使得决策图能够从粗略到精细过渡，从而降低了网络推理的难度；Huang 等人[13]提出了一种基于生成对抗网络(Generative Adversarial Network, GAN)的图像融合算法，并设计了具有自适应约束的生成对抗网络(Auxiliary Classifier GAN, ACGAN)，该方法在网络模型中通过引入自适应权重块，学习聚焦像素的分布。

尽管经典的图像融合算法能够生成较好的结果，但仍存在一些局限性。例如，在图像融合的过程中难以兼顾图像的细节和全局信息。为了解决这些问题，本文提出了一种联合双边加权引导滤波和金字塔结构的融合算法，该算法能够有效地利用图像中的多种信息来源，生成高质量的融合图像。

2. 背景知识

2.1. 拉普拉斯图像金字塔

图像拉普拉斯金字塔变换作为图像金字塔变换的一种，是在高斯金字塔的基础上发展而来的[14]，后者在处理时会丢失部分高频信息。为描述这些丢失部分的高频信息，拉普拉斯金字塔应运而生。设原始图像 $G_0(m, n)$ ($m \leq M, n \leq N$)， M 和 N 是图像的行列数，其基本步骤为：首先对原始图像 G_0 进行高斯低通滤波和下采样处理获得它的低分辨率图像 G_1 ；然后将其上采样至原始尺寸，接着通过计算原图 G_0 与上采样图像 E_1 图像之间高频信息的差值，得到第一层拉普拉斯金字塔的分解图像 L_0 。通过反复执行以上过程，最终迭代得到完整的拉普拉斯金字塔结构。如公式(1)、(2)、(3)所示：

$$G_{k+1} = \sum_{m=-c}^c \sum_{n=-c}^c \omega(m, n) G_k(2i+m, 2j+n), i \leq \frac{N}{2}, j \leq \frac{M}{2}, k \in N \quad (1)$$

$$G_{k+1}^* = 4 \sum_{m=-c}^c \sum_{n=-c}^c \omega(m, n) G_{k+1}\left(\frac{i+m}{2}, \frac{j+n}{2}\right), i \leq \frac{N}{2}, j \leq \frac{M}{2}, k \in N \quad (2)$$

$$L_k = G_{k+1} - G_{k+1}^*, k \in N \quad (3)$$

其中 $\omega(m, n)$ 表示大小为 $(2c+1) \times (2c+1)$ 窗口的卷积核。

2.2. 双边加权引导滤波

双边加权引导滤波[15]，是一种结合传统引导滤波技术和双边加权岭回归模型的图像处理方法。通过在引导滤波的目标函数中引入双边权重来增强对边缘的感知能力，从而改进传统方法在抑制如光晕等伪影方面表现不足的问题。这种改进使得引导图 I 与输出图像 q 之间的关系保持了原有的线性特征，并进一步强化了其在边缘细节处理上的性能。传统引导滤波技术可先假设引导图 I 与输出图像 q 之间保持线

性特征, 公式如下:

$$q_i = a_k I_i + b_k, \forall i \in \Omega_k \quad (4)$$

在该表达式中, q_i 表示输出图像的像素值, I_i 是引导图像的像素值, i 和 k 为像素的索引, Ω_k 为第 k 窗口, a_k 和 b_k 分别为窗口中心在第 k 时的线性函数系数。为了确保输出图像 q 能够准确地反映输入图像 p 的信息, 采用最小二乘法对像素间的线性关系进行拟合, 从而最小化输入图像与输出图像之间的差异。这种方法能够有效地确定目标函数, 从而提升图像融合效果并更好地保留细节。得到目标函数为

$$E(a_k, b_k) = \sum_{i \in \Omega_k} \left[(a_k I_i + b_k - p_i)^2 + \varepsilon a_k^2 \right] \quad (5)$$

再引入双边核权重 $\omega_{i,k}$ 来达成双边加权引导滤波的目标函数, 使得窗口内的不同像素点根据其重要性在线性拟合中贡献不同的程度。这种加权方式, 能够让新的目标函数更准确地反映图像的细节和结构。新的目标函数为

$$E(a_k, b_k) = \sum_{i \in \Omega_k} \left[\omega_{i,k} (a_k I_i + b_k - p_i)^2 + \frac{\varepsilon}{|\Omega|} a_k^2 \right] \quad (6)$$

其中 $\omega_{i,k}$ 表示如下:

$$\omega_{i,j} = \frac{1}{K} D_{i,j} R_{i,j} \quad (7)$$

$$D_{i,j} = \exp \left(-\frac{\|j-i\|^2}{2\sigma_d^2} \right) \quad (8)$$

$$R_{i,j} = \exp \left(-\frac{|I_j - I_i|^2}{2\sigma_r^2} \right) \quad (9)$$

在该表达式中, K 是量纲一化常数, σ_d 表示空间方差, σ_r 为范围方差。双边核函数结合了空间和范围两个核, 在图像像素变化不大的区域, 范围核的权重接近于 1, 空间核的作用更加显著, 使图像呈现出类似高斯模糊的效果; 而在像素值变化较大的边缘区域, 范围核的权重变得更为重要, 可以有效保留边缘细节。

2.3. 焦点聚集检测

2.3.1. 差分图像

图像在多聚焦图像中, 聚焦区域比散焦区域包含更多的显著特征信息[16]。本文通过结合双边加权引导滤波与图像差分原理, 利用图像的高频部分进行聚焦检测。为了便于演示, 本研究选取彩色图像的第一通道作为处理目标。如图 1 所示, (a)组图像展示的是待融合的两幅彩色源图像的第一通道。接下来, 对源图像 I1 和 I2 进行均值滤波, 得到(b)组结果, 滤波公式如下:

$$A_1(x, y) = I_1(x, y) * f_m \quad (10)$$

$$A_2(x, y) = I_2(x, y) * f_m \quad (11)$$

其中, f_m 为均值滤波器, $*$ 表示卷积操作, x, y 指的是像素点的索引。

图 1(c)组图像是将均值滤波后的图像 A1 和 A2 与源图像进行差异计算并取其绝对值, 以提取源图像中的高频信息。具体步骤如下:

$$D_1(x, y) = |I_1(x, y) - A_1(x, y)| \quad (12)$$

$$D_2(x, y) = |I_2(x, y) - A_2(x, y)| \quad (13)$$

本文选取双边加权引导滤波, 为了得到更多的高频部分, 对粗略焦点图 D_i 进行双边加权引导滤波, 获得细化的焦点图(d)组图像, 该过程如下:

$$M_{D_1}(x, y) = G_{r, \varepsilon, \sigma_d, \sigma_r}(I_1(x, y), D_1(x, y), r, \varepsilon, \sigma_d, \sigma_r) \quad (14)$$

$$M_{D_2}(x, y) = G_{r, \varepsilon, \sigma_d, \sigma_r}(I_2(x, y), D_2(x, y), r, \varepsilon, \sigma_d, \sigma_r) \quad (15)$$

这里 $G_{r, \varepsilon, \sigma_d, \sigma_r}$ 表示对图像 D_i 在差分滤波后进行双边加权引导滤波操作, 本文设置 $r=8$, $\varepsilon=0.1$, $\sigma_d=5$, $\sigma_r=0.06$, 在图像局部区域像素变化较大的地方, 范围核发挥主要作用, 保留图像的边缘细节, 同时避免过度强调边缘细节。

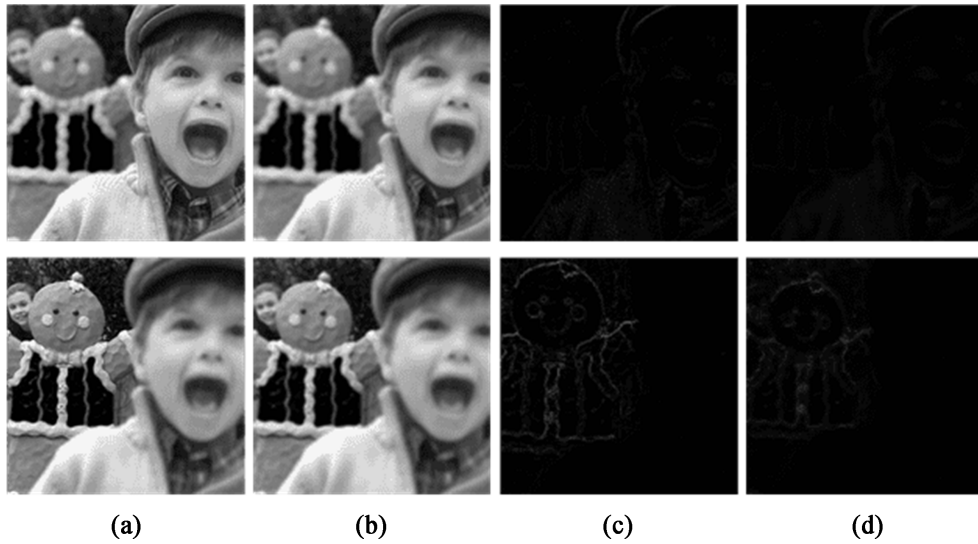


Figure 1. Differential image results of focus detection. (a) Original image; (b) Mean filtered image; (c) Difference image; (d) Double-weighted guided filtered image

图 1. 聚焦检测的差分图像结果。(a)原始图像; (b)均值滤波图像; (c)差值图像; (d)双边加权引导滤波图像

2.3.2. 高频信息决策图生成

采用上述聚焦区域检测方法, 并通过公式(15), 我们能够计算源图像 I_1 和 I_2 的聚焦度映射图 $M(x, y)$ 。

$$M(x, y) = \begin{cases} 1, & \text{if } (M_{D_1} > M_{D_2}) \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

尽管经过处理, 初始的决策图仍存在一些未能正确检测出的噪点, 如图 2(a)图像所示, 决策图中黑色背景中出现了白色像素, 而白色背景中也出现了黑色像素。之所以会出现这种现象, 是因为在缺乏明显的高频信息的区域, 聚焦和散焦往往出现很接近的状态, 难以进行区分。为了消除这些噪点, 我们采用小区域去除策略[17]对初始决策图进行优化。具体做法是通过局部区域的分析, 去除那些误判为聚焦或散焦的区域, 从而确保聚焦区域的准确性和决策图的清晰度。最后, 得到的高频信息决策图如图 2(b)。

3. 基于双边加权引导滤波的金字塔多聚焦图像融合

本文算法结合焦点聚集检测方法与图像金字塔原理, 对两幅不同聚焦的图像进行融合。首先, 通过

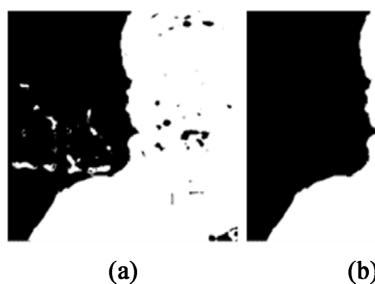


Figure 2. High-frequency information decision diagram. (a) Initial decision diagram; (b) Final Decision Diagram
图 2. 高频信息决策图。(a) 初始决策图; (b) 最终决策图

焦点区域检测方法, 利用双边加权引导滤波对源图像 $I1$ 和 $I2$ 的均值滤波和差分算子得到的粗糙焦点图进行细化, 生成图像金字塔第一层的高频信息决策图(High-frequency Information Decision Map, HFIDM)。接着, 将源图像 $I1$ 和 $I2$ 及其平均融合图像 F 作为底层图像, 利用双边加权引导滤波处理图像 $I1$ 、 $I2$ 与 F 。然后, 对经过滤波处理的图像 $I1$ 、 $I2$ 和 F 进行下采样, 得到图像 $G1$ 、 $G2$ 和 GF 。接着, 将 $G1$ 、 $G2$ 和 GF 作为输入, 反复进行双边加权引导滤波与下采样操作, 迭代形成金字塔结构。之后, 对平均融合金字塔的每一层图像进行上采样处理, 通过高斯卷积滤波得到预测图像, 并将每层预测图分别与相应的 G_i 图像做差处理, 得到 2 个拉普拉斯金字塔。最后, 将每层的拉普拉斯金字塔利用高频信息决策图(HFIDM)进行融合, 通过重建生成最终的融合金字塔。为了方便演示, 本研究继续选取彩色图像的第一通道作为处理目标, 具体过程如下。

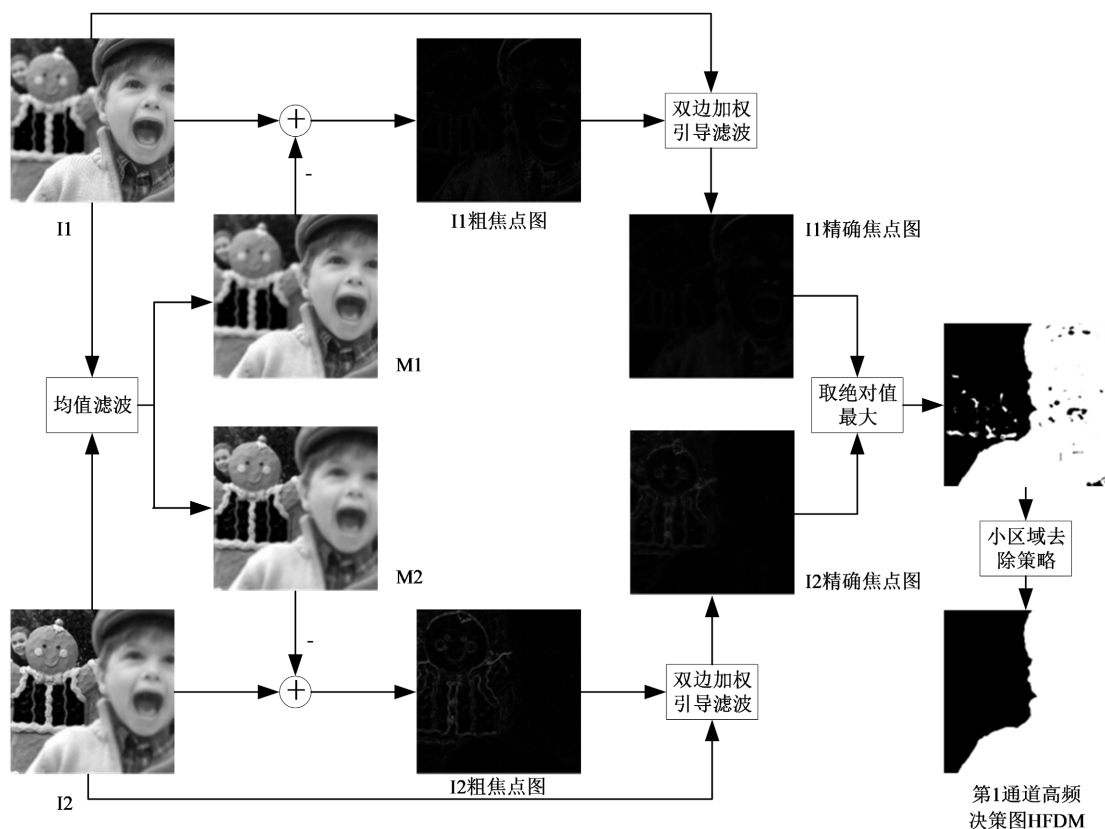


Figure 3. High-frequency information decision map generation process
图 3. 高频信息决策图生成过程

如图 3, 对原始图像 I_1 , I_2 使用焦点聚集检测方法与小区域去除策略可以得到第 1 通道图像高频信息决策图 HFDM, 我们选其作为金字塔域第 1 层的高频信息融合策略。接着, 获取 I_1 , I_2 的平均融合图 F 作为差异运算操作的参考。使用双边加权引导滤波处理 I_1 , I_2 和 F 图像, 并对 F 进行下采样, 将低分辨率的 F 图像上采样后的模糊图分别与 I_1 和 I_2 做差, 提取出两幅图像与融合图像 F 之间的高频信息差异, 通过高频信息决策图叠加图像的高频信息, 构建出第一层的高频融合图像 E_1 , 具体过程如图 4 所示。

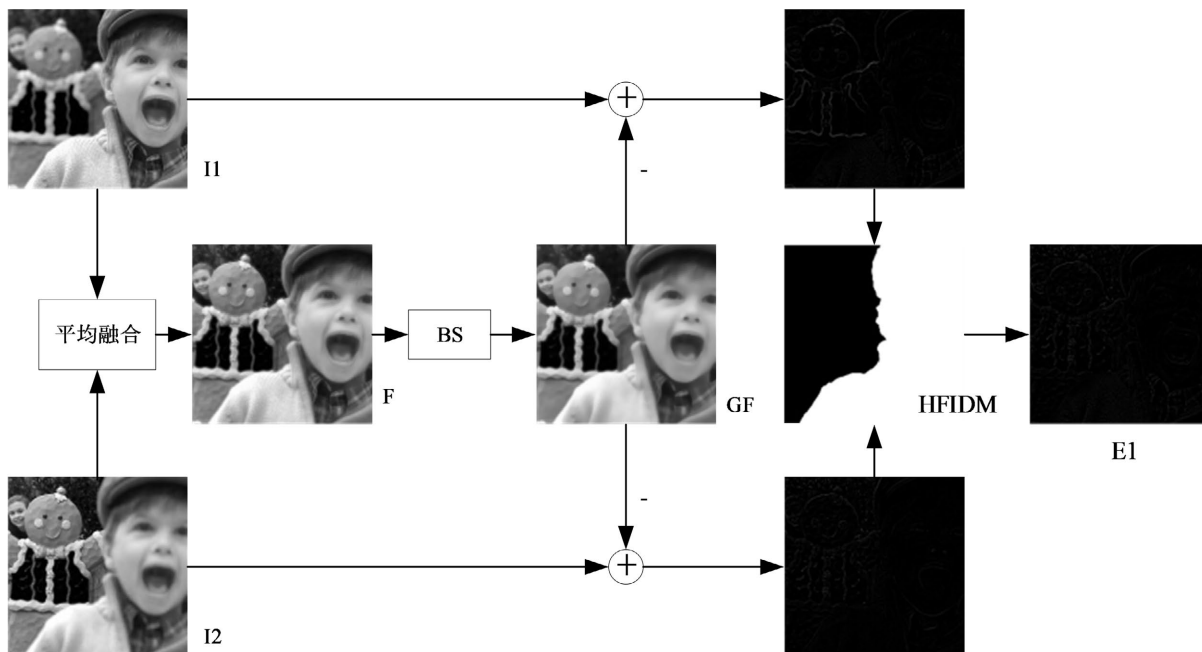


Figure 4. High-frequency fusion image generation process

图 4. 高频融合图像生成过程

图 4 中 BS 操作表示对 F 图像使用双边加权引导滤波处理后进行下采样再上采样后的操作, 最后再下采样 I_1 , I_2 , F 和 HFDM, 并将其作为输入, 重复图 4 过程三次, 得到三层高频融合图像 E_i , 融合图像的低频部分选择平均融合图 F 的顶层图像, 反向重建融合金字塔, 最终得到融合后的图像, 过程如图 5 所示。

4. 实验与分析

为了评估双边加权引导滤波的金字塔结构融合算法的效果和可行性, 本文使用了 Lytro 数据集[18]中的 20 组不同聚焦图像进行实验, 并与 7 种现有融合算法包括具有散焦扩散抑制机制的自适应区域划分图像融合算法(RDMF)[19]、基于均值滤波和引导滤波的图像融合算法(GFDF)[17]、基于引导滤波多尺度的图像分解算法(MGFF)[20]、各向异性扩散融合算法(ADF)[21]、基于相关系数的 DCT 域的多聚焦图像融合算法(DCT-Corr)[22]、基于赫西矩阵分解和显著差异聚焦检测的图像融合算法(HMD-SD)[23]和拉普拉斯金字塔图像融合算法(LP)进行了比较。我们使用了主观评价和 8 种客观指标来评估算法的性能, 这些指标包括结构相似性(Average Gradient, AG)[24]、相关一致性测度(Color Colorfulness Metric, CCM)[25]、色彩噪声指数(Color Naturalness Index, CNI)[26]、边缘强度(Edge intensity, EI)[27]、标准差(Standard Deviation, SD)[28]、结构保真度(Spatial Frequency, SF)[29]、视觉信息保真度(visual information fidelity, VIFF)[30]和信息融合质量测度(Image Contrast Metric, ICM)[25]。如表 1, 每一列数据代表同一客观评价指标下,

不同算法在 20 组融合结果中的平均值, 最后一行则是本文所提出的算法的评价结果。在对应的评价指标中, 表格中加粗的数值表示某一算法具有最佳的性能。通过这些数据, 本文对不同算法的融合效果进行了全面地对比分析。

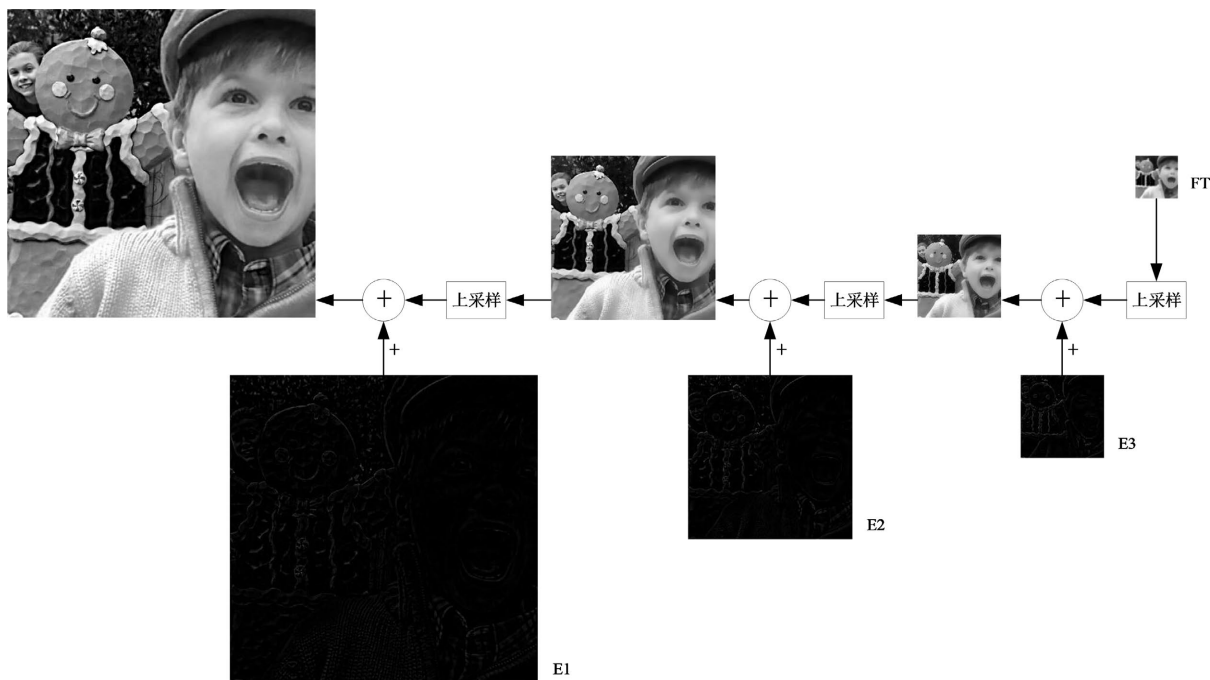


Figure 5. fusion image generation process

图 5. 融合图像生成过程

Table 1. The average of the 20 fusion results of the Lytro dataset

表 1. Lytro 数据集 20 组融合结果的平均值

融合方法	AG	CCM	CNI	EI	SD	SF	VIFF	ICM
RDMF	6.7981	24.8857	0.6906	70.5285	57.4950	19.3445	0.9425	0.3492
GFDF	6.7933	24.9585	0.6904	70.5195	57.5329	19.3279	0.9459	0.3476
MGFF	5.8586	26.3001	0.7087	61.6486	59.0839	16.9879	0.9842	0.3490
ADF	4.6976	23.1619	0.6750	49.4110	54.7480	12.6061	0.7776	0.3334
DCT-Corr	6.8282	25.0233	0.6912	70.8418	57.5240	19.4218	0.9428	0.3490
HMD-SD	6.8270	25.0013	0.6911	70.8247	57.5707	19.4351	0.9454	0.3489
LP	6.8274	24.9538	0.6923	70.8200	57.7735	19.3954	0.9547	0.3524
本文方法	8.0231	26.8502	0.7125	84.4405	60.9763	22.6068	1.0674	0.3512

从表 1 可以看出, 本章提出的算法在大多数评价指标上优于现有算法。“AG”、“EI”、“SD”、“SF”和“VIFF”指标均明显好于其他现有算法, “CCM”、“CNI”指标略有优势, “ICM”与最优值只有 0.3405% 的微幅差距, 因此本章算法的大多数客观评价指标都表现最好。总体而言, 本文提出的算法在融合结果上表现最佳。

接下来, 为了便于对不同融合算法进行主观评价和分析, 我们选取了 2 组融合结果进行详细的实验

对比。每组包括 2 副待融合图像和不同算法对应的融合结果，而最后一副图像展示了本文所提出的算法的融合结果。通过这种方式，我们能够直观地观察各算法在不同图像融合任务中的表现，并进一步分析它们的优势与不足。

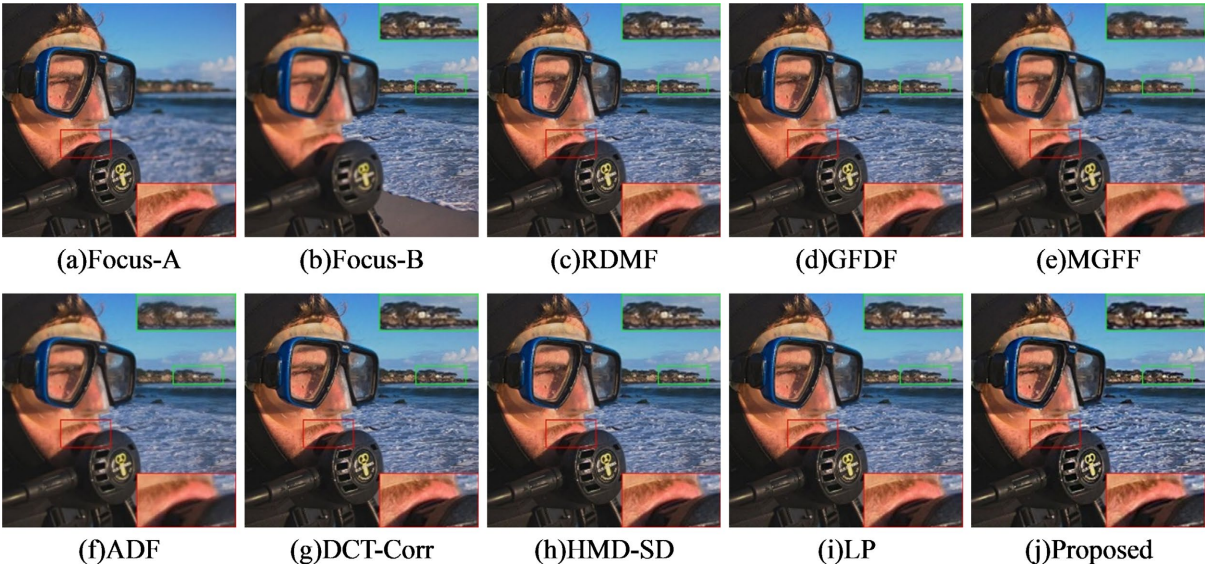


Figure 6. The first set of color image fusion results
图 6. 第一组彩色图像融合结果

从图 6 中可以观察到第一组图像的融合结果。原始待融合图像(图 6(a)和图 6(b))包含了近景的潜水员和远景的大海与礁石。图 6(c)至图 6(j)展示了多种算法的融合结果，其中包括本文提出的算法。为了方便地比较不同算法的融合效果，我们从待融合图像中提取了局部子图像，并进行了放大，得到不同聚焦物体的特写视图，并叠加生成融合图像。从图中可以观察到，MGFF 和 ADF 算法的融合效果较差，特别是在红色和绿色部分的融合中，图像质量较低，未能准确地融合边界信息。相比之下，RDMF、GFDF、DCT-Corr、HMD-SD 和 LP 算法的融合结果在整体上具有较高的质量。然而，本文提出的算法在红色和绿色特写视图中的融合效果尤为突出，特别是在礁石和胡子等细节信息的融合上表现更为充分，明显提升了视觉效果，同时也显著提升了融合质量。

表 2 中展示了第一组图像融合结果的客观评价指标。

Table 2. The evaluation index of the first set of color images under different methods
表 2. 第一组彩色图像在不同方法下的评价指标

融合方法	AG	CCM	CNI	EI	SD	SF	VIFF	ICM
RDMF	4.9924	18.1807	0.7964	52.4238	49.1728	14.1606	0.9742	0.2417
GFDF	4.9882	18.2120	0.7970	52.3892	49.1617	14.1237	0.9756	0.2416
MGFF	4.0912	18.0058	0.8017	43.5144	49.0885	11.4382	0.9355	0.2127
ADF	3.4084	17.2698	0.7909	36.7148	47.8684	8.9720	0.8217	0.1835
DCT-Corr	5.0007	18.2258	0.7961	52.4968	49.1697	14.1679	0.9750	0.2416
HMD-SD	5.0083	18.2229	0.7970	52.5704	49.1732	14.1948	0.9758	0.2416
LP	5.0173	18.2939	0.7993	52.6172	49.3474	14.1953	0.9818	0.2475
本文方法	6.0425	19.2163	0.8231	64.0161	51.2374	16.9692	1.1279	0.3095

从表 2 中可以看出, 在多项评价指标上, 本文提出的融合方法明显优于现有算法。具体来说, 本文方法在第一组图像的 8 个指标上均超越了其他算法, 特别是在 AG、CCM、EI、SD、SF 和 VIFF 指标上具有显著的优势。相比之下, RDMF、GFDF、MGFF、ADF、DCT-Corr、HMD-SD 和 LP 算法的表现相对较弱, 尤其是在融合结果的质量方面, 未能保持与源图像相似的整体结构和色彩。本文提出的融合方法在保留源图像丰富信息的同时, 有效融合了边缘细节, 保证图像的整体融合质量, 尤其在图像结构、色彩及细节保留方面表现突出。因此, 综合对比认为, 本文方法在第一组所有融合算法中效果最佳, 能够提供高质量的图像融合结果。

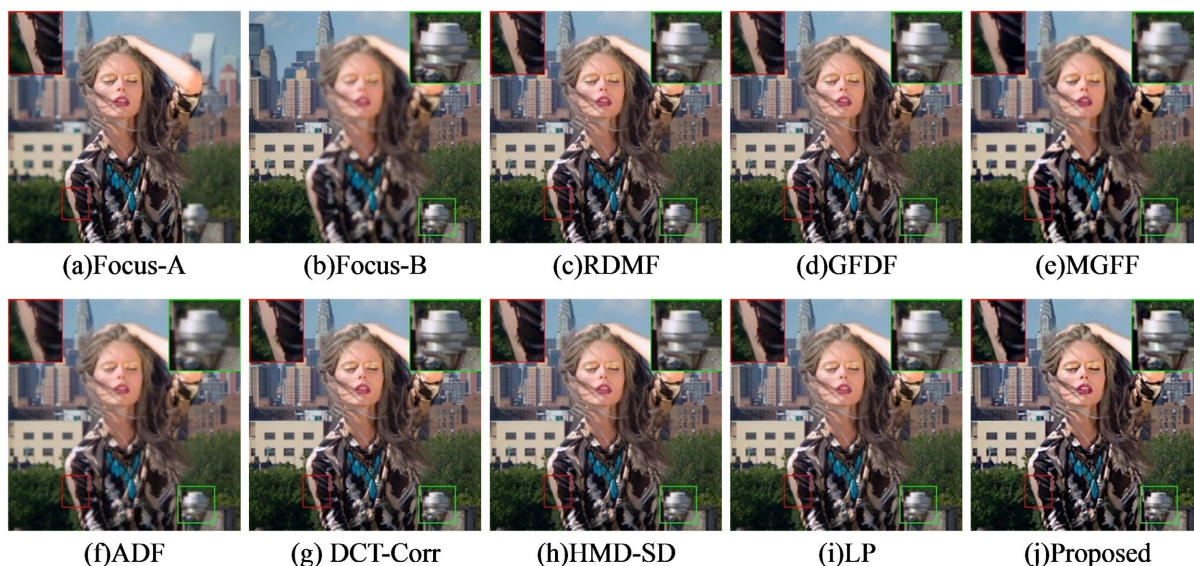


Figure 7. The second set of color image fusion results

图 7. 第二组彩色图像融合结果

从图 7 中可以观察到第二组图像的融合结果。原始待融合图像(图 7(a)和图 7(b))包含了近景的女模特和远景的背景区域。图 7(c)至图 7(j)展示了多种算法的融合结果, 其中包括本文提出的算法。从图中可以观察到, GFDF 算法、HMD-SD 算法和 LP 算法的整体融合结果质量高, 展现出较好的图像融合效果。RDMF 算法和 DCT 算法在绿色特写视图中未能正确融合边缘信息, 导致细节丢失。MGFF 算法和 ADF 算法的图像融合效果较差, 呈现出较为模糊的图像质量, 与之相比, 本文提出的算法在融合两幅多聚焦图像时, 不仅能够正确融合细节, 还在背景建筑物、窗户等物体的视觉感知上表现出更强的优势, 增强了图像的细节保留, 使融合结果看上去更清晰自然。

表 3 中展示了第一组图像融合结果的客观评价指标。

从表 3 中可以看出, 在多项评价指标上, 本文提出的融合方法明显优于现有算法。具体来说, 本文方法在“AG”、“CCM”、“CNI”、“EI”、“SD”、“SF”和“VIFF”等指标上均表现出优势, 尤其是在 AG、CCM、EI、SD、SF 和 VIFF 等指标上, 均优于对比算法。相比之下, RDMF、GFDF、MGFF、ADF、DCT-Corr、HMD-SD 和 LP 算法的表现相对较弱, 尽管它们在某些指标上与本文方法接近, 但在融合结果的质量和细节保留上未能达到相同水平, 尤其是在“EI”和“SD”等衡量图像质量的指标上。总体而言, 本文提出的融合方法不仅能够保留源图像的丰富信息, 还能有效融合边缘细节, 确保图像的整体质量。特别是在图像的结构、色彩以及细节保留方面, 本文方法的表现尤为突出。因此, 综合对比分析认为, 本文方法在所有融合算法中效果最佳, 能够提供更加精细和高质量的图像融合结果。

Table 3. The evaluation index of the second set of color images under different methods
表 3. 第二组彩色图像在不同方法下的评价指标

融合方法	AG	CCM	CNI	EI	SD	SF	VIFF	ICM
RDMF	6.0581	16.2649	0.6270	63.6822	58.5700	15.2612	0.9554	0.3395
GFDF	6.0389	16.2756	0.6290	63.5031	58.5530	15.2021	0.9564	0.3395
MGFF	5.1169	17.4752	0.7117	54.6654	60.2254	12.8778	0.9870	0.3260
ADF	4.0999	14.6979	0.6025	44.1572	56.6454	9.5662	0.8092	0.3171
DCT-Corr	6.0693	16.3276	0.6298	63.7843	58.5650	15.3001	0.9552	0.3394
HMD-SD	6.0849	16.3069	0.6292	63.9291	58.5980	15.3349	0.9570	0.3394
LP	6.1027	16.2642	0.6318	64.0782	58.6759	15.3326	0.9692	0.3536
本文方法	7.3503	18.7456	0.7128	77.9905	61.6969	18.5475	1.0900	0.3435

同时，为了评估金字塔域双边加权引导滤波融合算法在实际应用中的可行性，我们对其运行速度与其他算法进行了对比。实验在一台配备 i7-12700H CPU 和 16GB 内存的计算机上进行，使用 Lytro 数据集 20 组待融合图像，各个算法平均耗时如下表所示。

Table 4. Average time taken by the fusion algorithm for 20 groups in the Lytro dataset
表 4. Lytro 数据集 20 组融合算法的平均耗时

融合方法	RDMF	GFDF	MGFF	ADF	DCT-Corr	HMD-SD	LP	本文方法
运行时间	1.6772	0.1581	0.3802	0.3017	0.3316	0.8340	0.0587	0.6228

从表 4 可知，在对一组分辨率为 520×520 的彩色图像进行融合时，LP 算法的运行效率最高，耗时仅为 0.0587 秒，明显优于其他算法。由于本文方法采用了双边加权引导滤波，在计算双边核时不可避免地增加了计算成本，这是影响其运行效率的主要因素，但相较于 RDMF、HMD-SD 等算法仍具有一定优势。

5. 结论

本文提出了一种联合双边加权引导滤波和金字塔结构的多聚焦图像融合算法。在金字塔下采样的阶段中，采用双边加权引导滤波进行图像的处理。通过使用平均融合图作为基图，提取两幅图像与平均图像之间的高频差异，并结合焦点聚集检测生成高频信息的决策图实现图像的融合。实验结果表明，本文方法在主观视觉效果和客观评价指标上均优于现有经典算法。后续工作将进一步优化算法的计算复杂度的角度，以满足实时应用的需求。

基金项目

江苏省大学生实践创新训练计划项目(202410323027Z)。

参考文献

- [1] Zhang, H., Xu, H., Tian, X., Jiang, J. and Ma, J. (2021) Image Fusion Meets Deep Learning: A Survey and Perspective. *Information Fusion*, **76**, 323-336. <https://doi.org/10.1016/j.inffus.2021.06.008>
- [2] 苑津莎, 赵振兵, 高强, 孔英会. 红外与可见光图像配准研究现状与展望[J]. 激光与红外, 2009, 39(7): 693-699.
- [3] Zhang, X. (2022) Deep Learning-Based Multi-Focus Image Fusion: A Survey and a Comparative Study. *IEEE Transactions*

- on *Pattern Analysis and Machine Intelligence*, **44**, 4819-4838.
- [4] 唐霖峰, 张浩, 徐涵, 等. 基于深度学习的图像融合方法综述[J]. 中国图象图形学报, 2023, 28(1): 3-36.
 - [5] Ma, J., Ma, Y. and Li, C. (2019) Infrared and Visible Image Fusion Methods and Applications: A Survey. *Information Fusion*, **45**, 153-178. <https://doi.org/10.1016/j.inffus.2018.02.004>
 - [6] 王蓉, 柴玉华, 高立群, 等. 图像融合技术及融合效果评价的研究[J]. 农机化研究, 2006(1): 87-89.
 - [7] Burt, P. and Adelson, E. (1983) The Laplacian Pyramid as a Compact Image Code. *IEEE Transactions on Communications*, **31**, 532-540. <https://doi.org/10.1109/tcom.1983.1095851>
 - [8] Toet, A. (1989) Image Fusion by a Ratio of Low-Pass Pyramid. *Pattern Recognition Letters*, **9**, 245-253. [https://doi.org/10.1016/0167-8655\(89\)90003-2](https://doi.org/10.1016/0167-8655(89)90003-2)
 - [9] 张贵仓, 苏金凤, 拓明秀. DTCWT 域的红外与可见光图像融合算法[J]. 计算机工程与科学, 2020, 42(7): 8-17.
 - [10] Liu, Y., Chen, X., Peng, H. and Wang, Z. (2017) Multi-Focus Image Fusion with a Deep Convolutional Neural Network. *Information Fusion*, **36**, 191-207. <https://doi.org/10.1016/j.inffus.2016.12.001>
 - [11] Shao, X., Jin, X., Jiang, Q., Miao, S., Wang, P. and Chu, X. (2024) Multi-Focus Image Fusion Based on Transformer and Depth Information Learning. *Computers and Electrical Engineering*, **119**, Article ID: 109629. <https://doi.org/10.1016/j.compeleceng.2024.109629>
 - [12] Ouyang, Y., Zhai, H., Hu, H., Li, X. and Zeng, Z. (2025) FusionGCN: Multi-Focus Image Fusion Using Superpixel Features Generation GCN and Pixel-Level Feature Reconstruction CNN. *Expert Systems with Applications*, **262**, Article ID: 125665. <https://doi.org/10.1016/j.eswa.2024.125665>
 - [13] Huang, J., Le, Z., Ma, Y., Mei, X. and Fan, F. (2020) A Generative Adversarial Network with Adaptive Constraints for Multi-Focus Image Fusion. *Neural Computing and Applications*, **32**, 15119-15129. <https://doi.org/10.1007/s00521-020-04863-1>
 - [14] 费思成. 面向无线多媒体传感器网络低功耗图像融合算法研究[D]: [硕士学位论文]. 长沙: 湖南大学, 2016.
 - [15] 惠红军, 杨洋, 詹永照, 等. 双边加权引导图像滤波[J]. 江苏大学学报(自然科学版), 2024, 45(1): 60-66+76.
 - [16] 廖丽娜, 李伟彤, 项颖. 结合 SML 与差分图像的多聚焦图像融合算法[J]. 液晶与显示, 2023, 38(4): 524-533.
 - [17] Qiu, X., Li, M., Zhang, L. and Yuan, X. (2019) Guided Filter-Based Multi-Focus Image Fusion through Focus Region Detection. *Signal Processing: Image Communication*, **72**, 35-46. <https://doi.org/10.1016/j.image.2018.12.004>
 - [18] Nejati, M. (2016) Lytro Multi-Focus Image Dataset.
 - [19] Yu, W., Nie, Y. and Xiong, W. (2024) An Adaptive Region Division Multi-Focus Image Fusion Algorithm with Defocus Diffusion Mitigation Mechanism. *Proceedings of the 2024 2nd Asia Conference on Computer Vision, Image Processing and Pattern Recognition*, Xiamen, 26-28 April 2024, 1-6. <https://doi.org/10.1145/3663976.3664001>
 - [20] Bavirisetti, D.P., Xiao, G., Zhao, J., Dhuli, R. and Liu, G. (2019) Multi-Scale Guided Image and Video Fusion: A Fast and Efficient Approach. *Circuits, Systems, and Signal Processing*, **38**, 5576-5605. <https://doi.org/10.1007/s00034-019-01131-z>
 - [21] Bavirisetti, D.P. and Dhuli, R. (2016) Fusion of Infrared and Visible Sensor Images Based on Anisotropic Diffusion and Karhunen-Loeve Transform. *IEEE Sensors Journal*, **16**, 203-209. <https://doi.org/10.1109/jsen.2015.2478655>
 - [22] Naji, M.A. and Aghagolzadeh, A. (2015) Multi-Focus Image Fusion in DCT Domain Based on Correlation Coefficient. *2015 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI)*, Tehran, 5-6 November 2015, 631-638. <https://doi.org/10.1109/kbei.2015.7436118>
 - [23] Li, X., Wang, X., Cheng, X., Tan, H. and Li, X. (2022) Multi-Focus Image Fusion Based on Hessian Matrix Decomposition and Salient Difference Focus Detection. *Entropy*, **24**, Article No. 1527. <https://doi.org/10.3390/e24111527>
 - [24] Cui, G., Feng, H., Xu, Z., Li, Q. and Chen, Y. (2015) Detail Preserved Fusion of Visible and Infrared Images Using Regional Saliency Extraction and Multi-Scale Image Decomposition. *Optics Communications*, **341**, 199-209. <https://doi.org/10.1016/j.optcom.2014.12.032>
 - [25] Yuan, Y. (2011) Objective Quality Evaluation of Visible and Infrared Color Fusion Image. *Optical Engineering*, **50**, Article ID: 033202. <https://doi.org/10.1117/1.3549928>
 - [26] Huang, K., Wang, Q. and Wu, Z. (2006) Natural Color Image Enhancement and Evaluation Algorithm Based on Human Visual System. *Computer Vision and Image Understanding*, **103**, 52-63. <https://doi.org/10.1016/j.cviu.2006.02.007>
 - [27] Rajalingam B., and Priya R., (2018) Enhancement of Hybrid Multimodal Medical Image Fusion Techniques for Clinical Disease Analysis. *International Journal of Computer Vision and Image Processing*, **8**, 16-40. <https://doi.org/10.4018/ijcvip.2018070102>
 - [28] Mittal, A., Moorthy, A.K. and Bovik, A.C. (2012) No-Reference Image Quality Assessment in the Spatial Domain. *IEEE Transactions on Image Processing*, **21**, 4695-4708. <https://doi.org/10.1109/tip.2012.2214050>

- [29] Eskicioglu, A.M. and Fisher, P.S. (1995) Image Quality Measures and Their Performance. *IEEE Transactions on Communications*, **43**, 2959-2965. <https://doi.org/10.1109/26.477498>
- [30] Han, Y., Cai, Y., Cao, Y. and Xu, X. (2013) A New Image Fusion Performance Metric Based on Visual Information Fidelity. *Information Fusion*, **14**, 127-135. <https://doi.org/10.1016/j.inffus.2011.08.002>