ISSN 1004-9037,CODEN SCYCE4
Journal of Data Acquisition and Processing Vol. 41,No. 1,Jan. 2026,pp. 187－201
DOI:10. 16337/j. 1004-9037. 2026. 01. 013

http://sjcj. nuaa. edu. cn
E-mail:sjcj@nuaa. edu. cn
Tel/Fax:＋86-025-84892742

基于多重伪影抑制与多级融合的高动态范围成像

罗俊成，谢明鸿，张亚飞，李华锋

(昆明理工大学信息工程与自动化学院,昆明650500)

摘　要：由于现有成像设备的局限性,难以直接获取高动态范围(High dynamic range,HDR)图像。HDR成像技术旨在通过处理低动态范围(Low dynamic range,LDR)图像来生成HDR图像。现有的大多数方法通过融合多张不同曝光的图像来重建HDR图像。然而,由于前景和背景的相对运动,导致最终的重建结果中出现伪影。现有方法仅在融合多张不同曝光的图像前进行伪影消除。但这样会导致最终的HDR图像的质量严重依赖于融合前的伪影抑制结果。而伪影抑制的不理想导致融合过程中引入的伪影信息在后续重建过程中难以消除。基于此,提出了一种对重建特征进行伪影多重抑制和信息多级融合的网络框架,以高效重建HDR图像。通过多重伪影抑制块(Multiple artifact suppression block,MASB)来处理不同图像和特征之间的差异。与现有方法仅对融合前的图像或特征进行处理不同,在重建过程中对特征进行多重伪影抑制,从而进一步抑制重建特征中的伪影。同时,为了更好地利用非参考输入图像的特征,提出了多级融合块(Multilevel fusion block,MFB),在多级融合模块里进一步获取非参考图像中的互补信息。在多个数据集上的实验对比结果显示,所提方法在主观视觉效果和客观指标上均取得了更优异的表现。

关键词：高动态范围成像;深度学习;多重伪影抑制;多级融合

中图分类号：TP391.4　　**文献标志码:**A

引用格式：罗俊成,谢明鸿,张亚飞,等.基于多重伪影抑制与多级融合的高动态范围成像[J].数据采集与处理,2026,41(1):187-201. LUO Juncheng,XIE Minghong,ZHANG Yafei,et al. High dynamic range imaging with multiple artifact suppression and multilevel fusion[J].Journal of Data Acquisition and Processing,2026,41(1):187-201.

引　　言

图像的动态范围表示为图像中像素最大值与最小值的比。人类视觉系统的动态范围远高于成像设备的动态范围,而图像的动态范围过低就难以展现真实信息,如图像的过曝光区域和欠曝光区域都存在信息丢失的情况。因此,出现了高动态范围(High dynamic range,HDR)成像技术[1-7],该技术可以在低动态范围(Low dynamic range,LDR)图像的基础上生成HDR图像。

现有的HDR成像技术主要通过融合多张不同曝光水平的图像来重建HDR图像。该方法的优点是可以通过整合不同曝光程度的图像来达到扩大动态范围的目的。在静态场景下,不同曝光图像间没有前景运动和背景运动,通过对其进行融合可以产生质量很高的HDR图像[8-11]。但现实情况中,往往存在着前景运动(目标物的移动)和背景运动(相机的抖动),这就容易导致融合结果存在伪影。因此,如何抑制伪影就成了HDR成像技术的关键。

传统的HDR成像方法[12-19]主要依赖于手工设计的算法和规则来对图像进行处理,其中有基于图像

基金项目:国家自然科学基金(62161015, 62276120);云南省基础研究专项(202301AV070004)。
收稿日期:2024-11-26;**修订日期:**2025-02-25

配准的方法和基于运动检测的方法。

基于图像配准的方法先对不同的LDR图像进行配准，然后融合配准后的LDR图像得到HDR图像。Zimmer等[13]通过光流算法估计复杂的相机运动和场景移动得到光流场，然后根据光流场对非参考LDR图像进行配准。Sen等[14]提出了一种新颖的基于块的能量最小化公式，该公式将HDR成像过程中对齐和重建集成在联合优化中。Hu等[15]的方法与Sen等[14]类似，但在重建过程中传播强度和梯度信息，因此可以保留更多的细节。虽然这些方法可以取得一定的效果，但当LDR图像中存在大量过曝光和欠曝光情况时，这些方法就无法重建出理想的HDR图像。同时，这些方法的计算成本也普遍偏高。

基于运动检测的方法需要先对输入的图像进行区分，将像素分为动态像素和静态像素，然后在融合过程中舍弃掉动态像素，只将静态像素融入到HDR图像中。Raman等[16]提出了一种新颖的自下而上的分割算法，通过超像素分组检测场景运动区域，同时在梯度域中采用基于分段块的合成方法直接生成无伪影HDR图像。Yan等[17]通过稀疏表示的复合多曝光图像，直接生成曝光良好区域的HDR图像，同时将伪影区检测表述为对运动目标和噪声具有更强鲁棒性的稀疏表示问题。Ma等[18]将图像块分解为3个概念上独立的分量：信号强度、信号结构和均值强度，并根据强度、曝光度和结构一致性度量对每个分量进行处理。Li等[19]提出了一种多尺度递归降采样和处理平均强度图像的方法，以改善图像融合结果。这些方法舍弃了运动区域的信息，而这些信息中可能存在有用信息，因此这些方法重建出来的结果会丢失许多细节信息。

尽管传统方法在一些应用中效果较好，但存在鲁棒性差、计算复杂度高的缺点。随着深度学习在图像处理上的发展，基于深度学习的HDR成像方法也被相继提出。

深度学习模型[20-40]能够根据不同任务和数据自适应调整，使得其在不同类型的图像处理任务中都能取得良好的效果。其中，Kalantri等[27]通过经典的光流算法[28]对输入的不同曝光图像进行对齐。然后将对齐后的LDR图像输入到深度神经网络中进行重建，从而得到HDR图像。Peng等[29-30]使用光流网络（例如FlowNet[31]）来进行光流对齐，从而可以更加精确地对齐多样的LDR图像。Catley-Chandar等[32]在通过光流网络进行对齐的同时，增加了曝光不确定性建模，以此来对不确定性大的区域做进一步调整。

随着注意力[33]的提出，不少方法通过注意力机制来确定LDR图像之间的相关性，从而抑制伪影。Metwaly等[34]提出了注意力掩码来减少复杂区域的伪影。Yan等[35]提出了一个注意力模块，分别将非参考图像的特征和参考图像的特征输入到注意力模块中，得到非参考图像的注意力图。通过注意力的引导，可以有效避免引入非参考图像中会造成伪影的信息。Pu等[36]提出了一种合并了金字塔对齐和掩码融合的网络。基于注意力的多尺度金字塔模块可以充分运用不同尺度的信息来进行对齐，而自适应掩码策略可以对特征进行有效整合。Choi等[37]在使用金字塔注意力模块进行对齐的同时，在融合阶段设计了双激励机制，从而在空间和通道两个方面对特征做进一步的矫正。Yan等[38]采用了另一种方法，侧重于利用输入LDR图像中的非局部相关性来去除重影伪影。Chen等[39]为了减少潜在的伪影，仅对欠曝光和过曝光图像进行处理，通过单应性估计网络和注意力网络的结合来将过曝光图像与欠曝光图像对齐，从而解决伪影问题。ADNet[40]通过空间注意力处理LDR图像，同时使用PCD对准模块处理伽马矫正图像，最后通过可变形对齐模块来对动态帧进行对齐。

基于深度学习的网络多种多样，但各种网络在处理伪影上依旧存在问题。为了进一步抑制伪影、生成高质量HDR图像，本文设计了一个新的网络以实现HDR成像中的伪影高效抑制。首先，提出多重伪影抑制块（Multiple artifact suppression block，MASB）来抑制伪影，与以往的方法不同，以往的方法仅在图像融合之前对非参考图像进行处理，本文考虑到重建过程中的特征依然存在伪影，因此在重建

过程中继续对重建特征进行伪影抑制。同时,为了更有效地利用非参考图像,提出了多级融合模块(Multilevel fusion block,MFB),在对重建特征进行伪影抑制后,用该特征与非参考信息进一步进行融合。需要注意的是,在多级融合模块中,并不是直接对其进行融合,而是基于特征相关性来选取待融合信息,以此避免引入伪影信息。在3个数据集上验证模型的有效性,与多个先进方法进行定性定量对比,结果显示了所提模型的优越性。

本文的贡献主要包括:

(1)为解决动态场景中伪影难以抑制的问题,采用新的多重伪影抑制方法。多重伪影抑制除了在融合前对非参考特征进行伪影抑制,在重建过程中同样对重建特征进行伪影抑制。通过对特征进行多重伪影抑制操作,可以有效抑制伪影。

(2)为了更有效地利用非参考图像信息,引入多级融合块。在特征重建过程中,将非参考特征中与参考图像相关的信息再次融入到重建特征中,既避免了引入伪影信息,又能增强重建信息。

(3)在多个数据集上对所提方法进行了定量和定性评估,结果表明,所提方法在性能上优于现有的其他方法。

1 基于多重伪影抑制与多级融合的HDR成像

1.1 模型概述

所提出的网络可以分为以下几部分:特征提取、特征融合、特征重建、多重伪影抑制模块和多级融合模块。如图1所示,首先对输入的LDR图像提取特征,通过多重伪影抑制块对非参考特征做第一次处理,以初步抑制伪影。然后,将不同曝光的LDR信息融合。融合后的特征经过N次特征多层处理。如图1中紫色部分所示,每一次特征多层处理需要先对特征进行重建,然后对重建特征进行多重伪影抑制,最后对信息进行多级融合。为了更好地抑制伪影,通过多重伪影抑制块继续对重建过程中的特征进行伪影抑制。同时,为了更好地利用非参考图像的有效信息,通过多级融合模块对非参考信息做进一步融合。在多重伪影抑制块中,以参考图像即中间曝光图像的特征为参照,来处理非参考图像的特征或重建特征。而在多级融合模块中,采用先提取相关信息再融合的方式避免引入伪影信息。

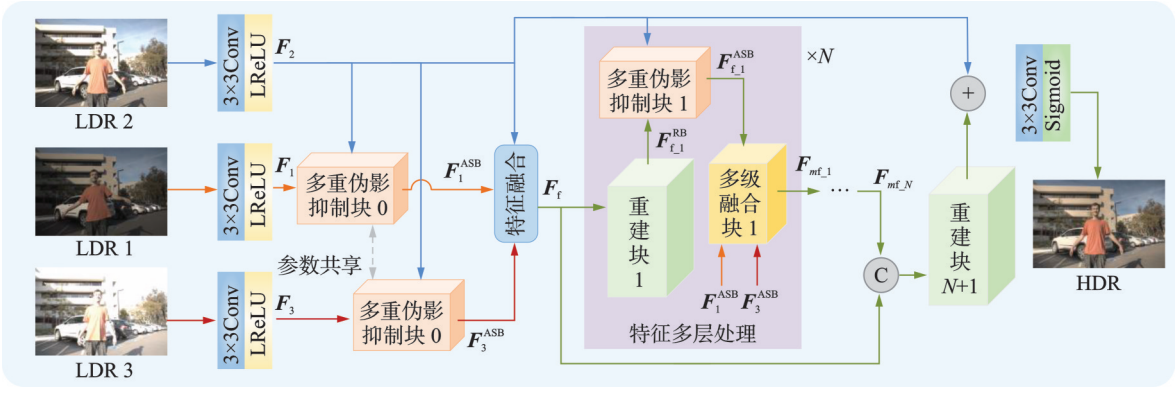


图1 本文方法的总体框架

Fig.1 Overall framework of the proposed method

1.2 数据预处理

给定的不同曝光图像序列$\{L_1,L_2,L_3\}$。因为不同曝光图像之间并不对齐,存在前景运动和背景运动,因此将中间曝光图像L_2作为参考图像,低曝光图像L_1和高曝光图像L_3作为非参考图像,最终通过

网络生成与参考图像严格对齐的HDR图像。为了得到更丰富的信息，先对LDR图像进行伽马矫正，将图像映射到HDR域，即

$$H_i=\frac{L_i^{\gamma}}{t_i} \qquad i=1,2,3 \tag{1}$$

式中：γ为矫正参数，根据经验将其设置为2.2；t_i为LDR图像L_i的曝光时间。最后将LDR与HDR域的图像拼接得到网络输入，即

$$X_i=[L_i,H_i] \qquad i=1,2,3 \tag{2}$$

1.3 网络结构

1.3.1 特征提取

如图1所示，对于每一个输入$X_i\in \mathbf{R}^{C\times H\times W}$($i=1,2,3$)，通过一个带LeakyReLU(LReLU)激活函数的3×3卷积层来进行编码。且对于每一个输入图像，使用的卷积层参数共享。其表示如下

$$F_i=\mathrm{Conv}L_{3\times 3}(X_i) \qquad i=1,2,3 \tag{3}$$

式中：$\mathrm{Conv}L_{3\times 3}(\cdot)$表示带LReLU激活函数的3×3卷积层。对于编码后得到的特征F_i($i=1,2,3$)，再对其进行伪影抑制、特征融合和特征重建。

1.3.2 多重伪影抑制

多重伪影抑制块(Multiple artifact suppression block, MASB)的结构以及在网络中的作用如图2所示，主要体现在不仅对非参考特征F_1、F_3进行伪影抑制，还对多个重建特征$F_{\mathrm{f}}^{\mathrm{RB}}$进行伪影抑制。因为前面伪影抑制不彻底，会导致重建特征中留有伪影信息。通过对重建特征进行多重伪影抑制，可以更好地重建出无伪影的高动态范围图像。

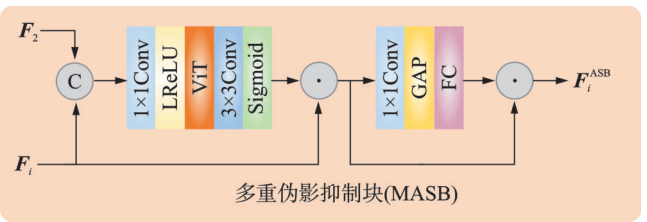


图2 多重伪影抑制块结构

Fig.2 Structure of multiple artifact suppression block

特征融合前的伪影抑制：对于融合前的特征F_i($i=1,2,3$)，将F_2作为参考特征，而另一个需要处理的特征(F_1或F_3)作为非参考特征。在F_2的引导下，对非参考特征进行伪影抑制，即

$$F_i^{\mathrm{ASB}}=\mathrm{MASB}(F_i,F_2) \qquad i=1,3 \tag{4}$$

特征多层处理中的伪影抑制：本文一共有N次特征多层处理，每一次处理过程中，都对重建特征进行伪影抑制。

尽管对F_1和F_3进行了伪影抑制，但F_i^{ASB}($i=1,3$)中伪影信息很难彻底消除。因此在第j($j=1,2,\cdots,N$)次特征多层处理中，将F_2作为参考特征，而将把重建特征$F_{\mathrm{f}_j}^{\mathrm{RB}}$($j=1,2,\cdots,N$)作为非参考特征，以消除重建特征中的伪影信息，即

$$F_{\mathrm{f}}^{\mathrm{ASB}}=\mathrm{MASB}(F_{\mathrm{f}_j}^{\mathrm{RB}},F_2) \tag{5}$$

通过多重伪影抑制操作，可以更好地抑制重建特征中的伪影信息。

伪影抑制：在每一次伪影抑制过程中，先将参考特征F_2和非参考特征F_i($F_i=F_1,F_3,F_{\mathrm{f}_j}^{\mathrm{RB}}$)拼接，然后通过卷积层和VIT对拼接后的特征进行学习，预测到第一次的偏移量off_1。然后，通过第一次预测的偏移量对非参考特征进行第一次调整，即

$$F_i^{\mathrm{mid_align}}=F_i\cdot \mathrm{off}_1 \tag{6}$$

在得到第一次调整的结果$F_i^{\mathrm{mid_align}}$后，对通道进行平均池化操作，来继续对伪影进行抑制，最后得到第二次估计的偏移量off_2。在与第一次的调整结果进行逐位素相乘后，得到最终伪影抑制后的特征

F_1^{ASB}、F_3^{ASB}和$F_{\mathrm{f}_j}^{\mathrm{ASB}}$,即

$$F_i^{\mathrm{ASB}} = F_i^{\mathrm{mid_align}} \cdot \mathrm{off}_2 \tag{7}$$

1.3.3 特征融合

在对非参考特征进行伪影抑制后,将特征融合在一起,以整合有用的信息。因为此时的非参考特征已经经过伪影抑制,所以在这里直接采用传统的融合方式来对特征进行融合。如图3所示,先采用简单的拼接方式来对特征初步整合。在拼接后使用一个带LReLU激活函数的3×3卷积来对特征进行更进一步的融合。图中的C表示拼接操作。该过程可表示为

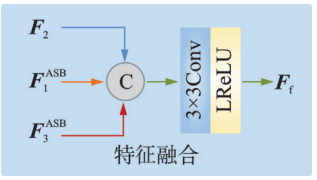


图3 特征融合过程

Fig.3 Feature fusion process

$$F_{\mathrm{f}} = \mathrm{Conv}L_{3\times3}([F_1^{\mathrm{ASB}}, F_2, F_3^{\mathrm{ASB}}]) \tag{8}$$

式中[·]表示拼接操作。在得到融合特征F_{f}后,通过对其进行重建操作,生成HDR图像。

1.3.4 多级融合模块

以往的方法对非参考图像特征只融合一次,这样虽然避免引入非参考特征中的伪影信息,但却无法很好地利用非参考信息。非参考特征中有不少互补的信息可以增强图像质量,为了进一步利用非参考特征的信息,对特征进行再次融合。

在N次特征多层处理中,每一次都对信息进行多级融合。以第j(j=1,2,…,N)次特征多层处理为例,首先将多重伪影抑制后的特征$F_{\mathrm{f}_j}^{\mathrm{ASB}}$和非参考特征$F_1^{\mathrm{ASB}}$、$F_3^{\mathrm{ASB}}$输入到多级融合模块中。多级融合模块的结构如图4所示。先使用卷积和LN(Layer normalization)操作对输入进行处理。将$F_{\mathrm{f}_j}^{\mathrm{ASB}}$映射成$\boldsymbol{Q}_{\mathrm{f}}$;将$F_1^{\mathrm{ASB}}$映射切分成$\boldsymbol{K}_1$、$\boldsymbol{V}_1$;将$F_3^{\mathrm{ASB}}$映射切分成$\boldsymbol{K}_3$、$\boldsymbol{V}_3$。其中Chunk表示对特征进行通道上的切分。

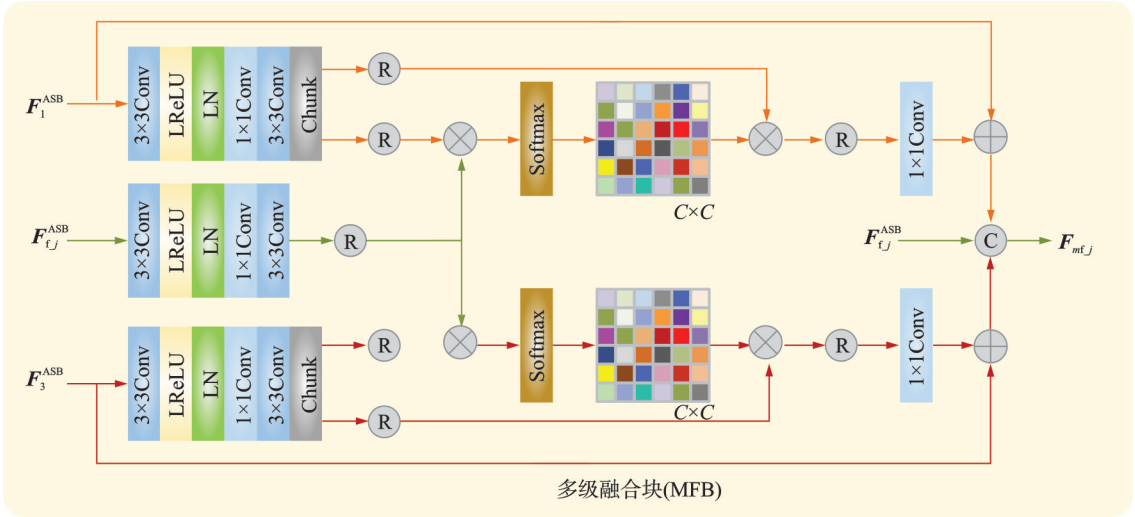


图4 多级融合块结构图

Fig.4 Structure of multilevel fusion block

将$\boldsymbol{Q}_{\mathrm{f}}$、$\boldsymbol{K}_1$、$\boldsymbol{V}_1$和$\boldsymbol{Q}_{\mathrm{f}}$、$\boldsymbol{K}_3$、$\boldsymbol{V}_3$分别进行计算,得到大小为$C\times C$的相关矩阵。通过相关矩阵,获取非参考特征中与$F_{\mathrm{f}_j}^{\mathrm{ASB}}$相关的信息$\boldsymbol{M}_1$和$\boldsymbol{M}_3$。其计算过程可表示如下

$$\begin{cases} \boldsymbol{A}_1 = \boldsymbol{V}_1 \times \mathrm{softmax}\left(\dfrac{\boldsymbol{K}_1 \times \boldsymbol{Q}_\mathrm{f}}{d}\right) \\ \boldsymbol{A}_3 = \boldsymbol{V}_3 \times \mathrm{softmax}\left(\dfrac{\boldsymbol{K}_3 \times \boldsymbol{Q}_\mathrm{f}}{d}\right) \end{cases} \tag{9}$$

$$\begin{cases} \boldsymbol{M}_1 = \boldsymbol{F}_1^{\mathrm{ASB}} + \mathrm{conv}_{1\times1}(\mathrm{R}(\boldsymbol{A}_1)) \\ \boldsymbol{M}_3 = \boldsymbol{F}_3^{\mathrm{ASB}} + \mathrm{conv}_{1\times1}(\mathrm{R}(\boldsymbol{A}_3)) \end{cases} \tag{10}$$

式中：d为尺度因子，R(•)为reshape操作，“×”为矩阵乘积。最后，还需要将$\boldsymbol{M}_1$和$\boldsymbol{M}_3$与重建特征$\boldsymbol{F}_{\mathrm{f}_j}^{\mathrm{ASB}}$进行融合，从而得到多级融合模块的输出$\boldsymbol{F}_{mf_j}$

$$\boldsymbol{F}_{mf_j} = [\boldsymbol{M}_1, \boldsymbol{F}_{\mathrm{f}_j}^{\mathrm{ASB}}, \boldsymbol{M}_3] \tag{11}$$

在每一次特征多层处理中，都对非参考特征进行融合。通过多级融合操作，可以更好地利用非参考图像的信息。

1.3.5 特征重建

特征多层处理主要可以分为3个部分，除了重建块外，还有多重伪影抑制块和多级融合模块。

重建块结构如图5所示，主要由卷积层构成，其中增加了一个对空间进行的平均池化操作。在第j(j=1,2,…,N)次特征多层处理中，输入待重建特征后，重建块对其进行特征重建，即

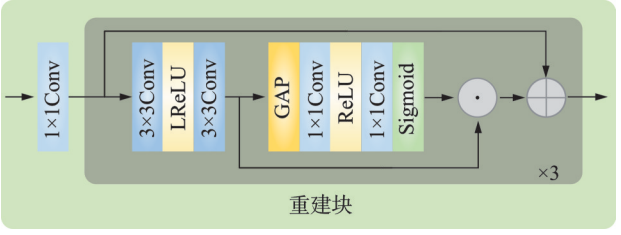


图5 重建块结构图

Fig.5 Structure of reconstruction block

$$\boldsymbol{F}_{\mathrm{f}_j}^{\mathrm{RB}} = \begin{cases} \mathrm{RB}(\boldsymbol{F}_\mathrm{f}) & j=1 \\ \mathrm{RB}(\boldsymbol{F}_{mf_j-1}) & j=2,3,\cdots,N \end{cases} \tag{12}$$

式中RB(•)表示重建块。得到$\boldsymbol{F}_{f_j}^{\mathrm{RB}}$后，先对其进行多重伪影抑制(如1.3.2节所述)，得到$\boldsymbol{F}_{\mathrm{f}_j}^{\mathrm{ASB}}$，然后对$\boldsymbol{F}_{\mathrm{f}_j}^{\mathrm{ASB}}$进行多级融合(如1.3.4节所述)，得到$\boldsymbol{F}_{mf_j}$。将上述3个操作记作一次完整的特征多层处理。

在经过N(N=2)次特征处理后，再使用一个重建块将之前的重建特征整合为$\boldsymbol{F}_{\mathrm{final}}^{\mathrm{RB}}$，然后通过一个带Sigmoid激活函数的3×3卷积层，将特征重建为最终的HDR图像$\boldsymbol{H}$。最终的图像重建过程可表示为

$$\boldsymbol{H} = \mathrm{Conv}S_{3\times3}(\boldsymbol{F}_{\mathrm{final}}^{\mathrm{RB}} + \boldsymbol{F}_2) \tag{13}$$

式中：$\mathrm{Conv}S_{3\times3}(\cdot)$表示带Sigmoid激活函数的3×3卷积层。

1.4 损失函数

对于生成的HDR图像，通过3个损失对齐进行约束。首先是重建损失，Kalantari等[27]研究表明，对HDR图像在色调映射图像上进行约束更为有效。因此本文使用μ-law色调映射来对$\boldsymbol{H}$进行处理，即

$$T(\boldsymbol{H}) = \frac{\log(1+\mu H)}{\log(1+\mu)} \tag{14}$$

式中：μ为压缩参数，$T(\boldsymbol{H})$表示色调映射后的图像。根据文献[27]，本文将μ设为5 000。

首先是重建损失，使用L1损失计算网络生成的结果$\boldsymbol{H}$和标签$\boldsymbol{H}_{\mathrm{GT}}$之间的差距，可表示为

$$\ell_{\mathrm{re}} = \| T(\boldsymbol{H}) - T(\boldsymbol{H}_{\mathrm{GT}}) \|_1 \tag{15}$$

同时,为了在结构上对$\boldsymbol{H}$进行约束,引入了结构相似度损失

$$\ell_{\text{ssim}}=1-\text{SSIM}(T(\boldsymbol{H}),T(\boldsymbol{H}_{\text{GT}})) \tag{16}$$

式中SSIM($\boldsymbol{H},\boldsymbol{H}_{\text{GT}}$)表示计算两个输入的结构相似度。为了对图像的边缘信息做约束,还使用了梯度损失

$$\ell_{\text{g}}=\|\nabla(\boldsymbol{H})-\nabla(\boldsymbol{H}_{\text{GT}})\|_1 \tag{17}$$

式中$\nabla(\cdot)$表示梯度算子。最后可以得到网络的总损失为

$$\ell=\ell_{\text{re}}+\lambda_1\ell_{\text{ssim}}+\lambda_2\ell_{\text{g}} \tag{18}$$

式中λ_1和λ_2为超参数。

2 实验结果与分析

2.1 实验设置

数据集:使用文献[27]中的训练集对网络进行训练,Kalantari数据集分为训练集和测试集。训练集共含74组图像,其中每一组包含3张不同曝光的LDR图像$\{\boldsymbol{L}_1,\boldsymbol{L}_2,\boldsymbol{L}_3\}$,以及与中间曝光图像完全对齐的HDR图像$\boldsymbol{H}_{\text{GT}}$。每一张LDR图像的曝光时间是已知的。测试集与训练集类似,共有15组图像。为了进一步测试所提方法的有效性,在无标签数据集Sen[14]和Tursun[41]上测试模型的性能。Sen和Tursun数据集仅有LDR图像,而没有对应的HDR标签。

评价指标:对于有标签数据集Kalantari的测试结果,直接使用峰值信噪比(Peak signal to noise ratio, PSNR)、结构相似度(Structural similarity, SSIM)、HDR-VDP-2[42]和MEF-SSIMd[43]作为评价指标。为了更全面地评价结果,对映射前和映射后的HDR图像都计算指标,最后的指标可表示为PSNR_L、SSIM_L、PSNR_μ、SSIM_μ、HDR-VDP-2和MEF-SSIMd。其中_L表示对线性域图像的评价结果,_μ表示对μ律映射图像的评价结果。对于无标签数据集Sen和Tursun的测试结果,采用UDQM[41]、BTMQI[44]和MEF-SSIMd作为评价指标。

实验细节:在训练过程中,将图像裁剪为256×256像素大小,并随机对图像进行翻转,从而实现数据增强。批次大小设置为3。在使用Adam优化器[45]来更新网络参数的同时,采用warm-up学习率调整策略。初始学习率设置为0.000 5,衰减率为0.2。在1 000个训练轮次后衰减为0.000 1,2 000个训练轮次后衰减为0.000 02。

将模型训练5 000轮,选出性能最好的那一轮参数作为最终模型。所提算法通过PyTorch框架实现,其中CUDA版本为11.4,Torch版本为1.13.1。在单张NVIDIA 3090 GPU上训练模型。

2.2 对比实验

对比方法选取了基于块的Sen[14],以及基于深度学习的Kalantari[27]、AHDRNet[35]、HDR-GAN[3]、HDRI[25]、HDR-Transformer[5]和SGARN[46]。所有代码均使用公开的代码进行测试。

2.2.1 在有标签数据集上的对比结果

如图6所示,Sen、Kalantari和HDR-Transformer所生成的HDR图像均有较大面积的伪影。在Sen和Kalantari的结果中,树枝旁边的伪影都很明显;而HDR-Transformer的墙体边缘,同样有着较大面积的伪影。AHDRNet和HDRI生成的结果在树枝旁同样存在着伪影,而在边缘细节处,伪影也同样明显;HDR-GAN和SGARN的结果相对较好,但边缘处的信息(图6绿框)依然无法很好重建。相较而言,所提方法所生成的HDR图像没有明显的伪影,同时能很好地呈现出边缘信息,在视觉上最接近标签。从表1的定量指标评价也可以看到,所提方法在各个性能上都达到了最优。

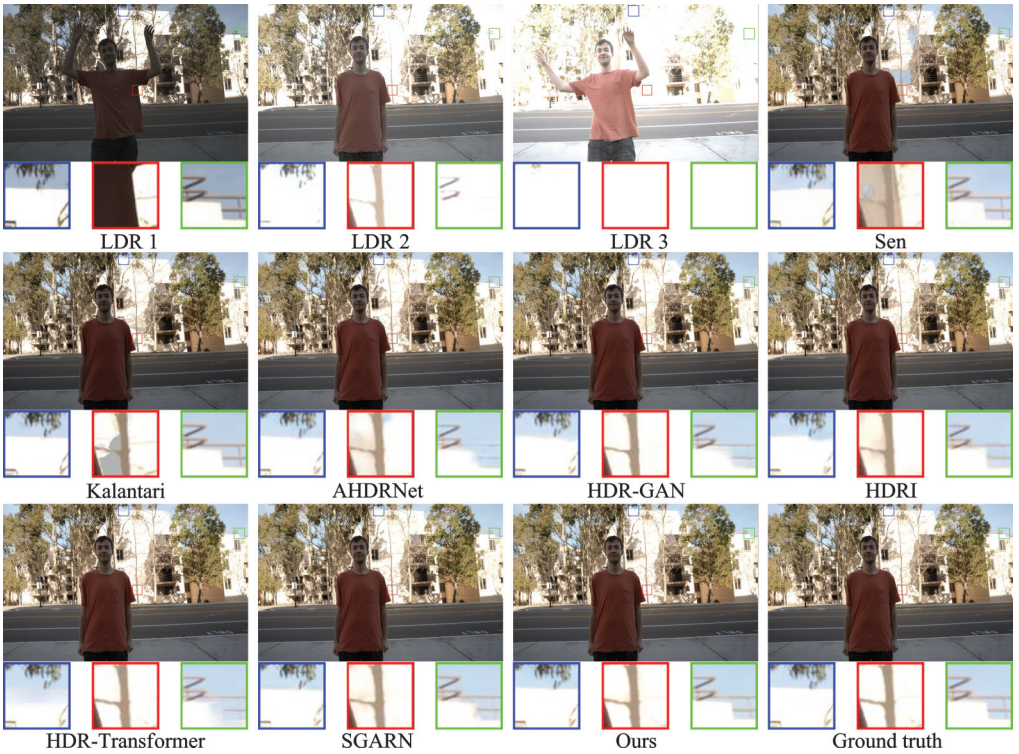


图 6 各方法在 Kalantari 测试集的视觉效果展示

Fig.6 Visual presentation of each method in Kalantari test set

表 1 Kalantari 测试集上不同方法的重建结果定量评估

Table 1 Quantitative evaluation of reconstruction results by different methods on the Kalantari test set

方法	PSNR_μ/dB	PSNR_L/dB	SSIM_μ	SSIM_L	HDR-VDP-2	MEF-SSIMd
Sen[14]	40.95	38.31	0.983 2	0.975 3	60.33	0.638 4
Kalantari[27]	42.74	40.72	0.987 7	0.982 4	62.87	0.640 7
AHDRNet[35]	43.61	41.04	0.990 0	0.970 2	63.51	0.651 1
HDR-GAN[3]	43.92	41.57	0.990 5	0.986 5	64.70	0.652 0
HDRI[25]	43.65	41.67	0.989 4	0.986 7	64.46	0.653 4
HDR-Transformer[5]	44.08	42.10	0.991 7	0.988 8	64.63	0.652 8
SGARN[46]	43.92	41.46	0.990 8	0.987 4	65.12	0.655 9
Ours	44.41	42.43	0.992 0	0.990 2	65.13	0.657 8

2.2.2 在无标签数据集上的对比结果

如图 7 和图 8 所示，Sen 对于过曝光区域无法很好重建；Kalantari 在图 7 中有大量信息丢失，而在图 8 红框中，伪影十分明显；AHDRNet 和 HDR-GAN 在图 7 中无法重建边缘信息，同时 AHDRNet 和 HDR-Transformer 在图 8 中的结果同样存在伪影，HDR-GAN 和 HDRI 则无法很好重建曝光区域信息；SGARN 没有明显伪影，但对于过曝光区域重建不理想，导致图像亮光区域不够自然。综上所述，所提方法所生成的 HDR 图像在视觉上比其他方法伪影更少，细节信息更丰富。从表 2 的定量指标评价中也可以看到，所提方法在各个性能上都达到了最优。

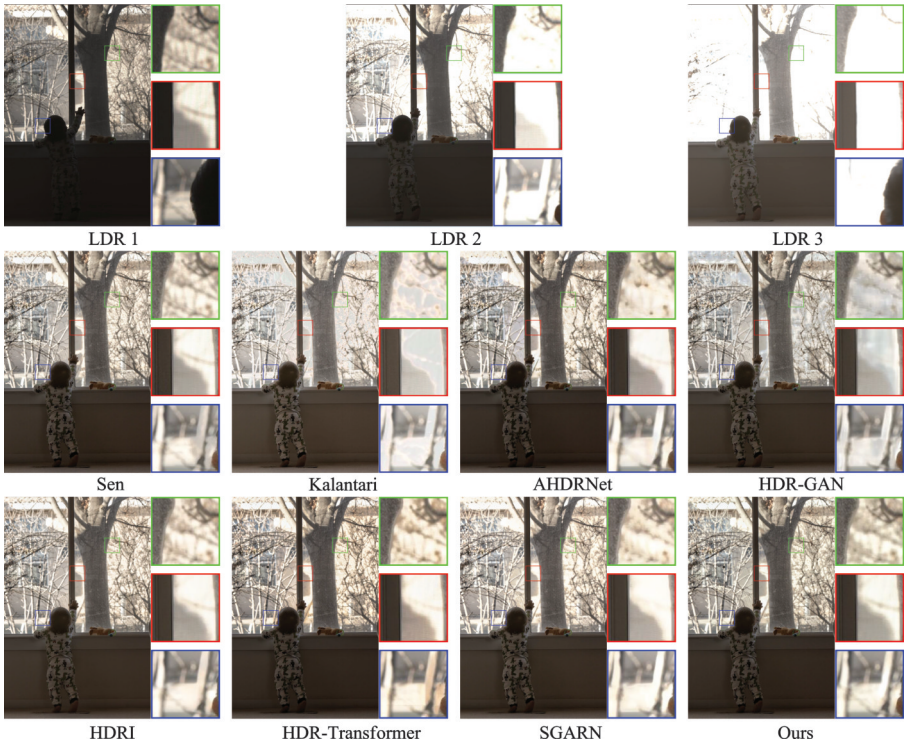


图7 各方法在Sen数据集中的视觉效果展示

Fig.7 Visual presentation of each method in Sen dataset

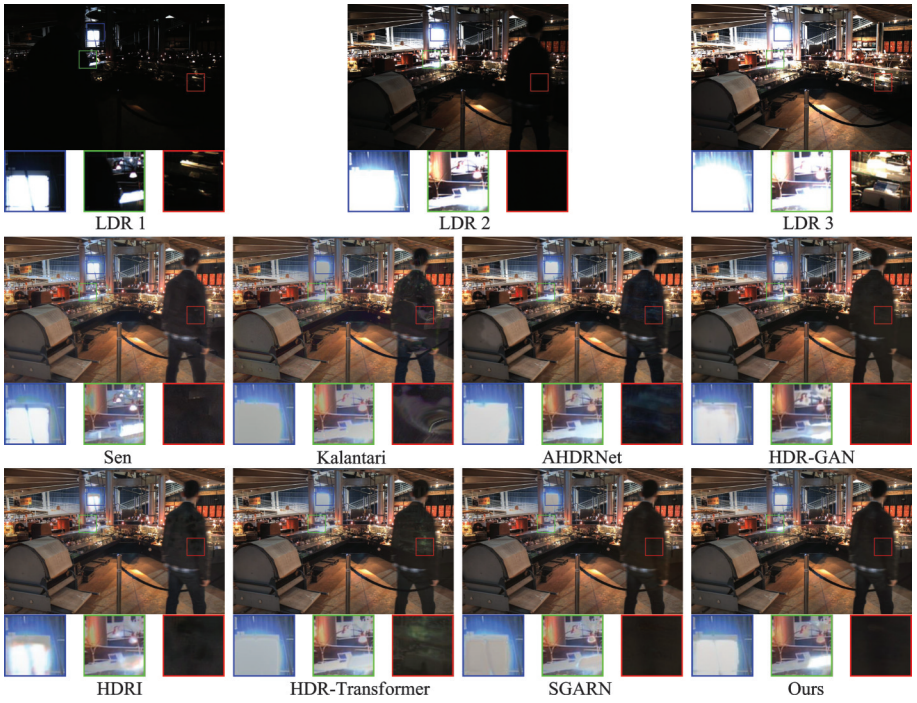


图8 各方法Tursun数据集中的视觉效果展示

Fig.8 Visual presentation of each method in Tursun datasets

表2 Sen和Tursun测试集上不同方法的重建结果定量评估

Table 2 Quantitative evaluation of reconstruction results by different methods on Sen and Tursun test sets

方法	Sen数据集		Tursun数据集		
	BTMQI↓	MEF-SSIMd↑	BTMQI↓	MEF-SSIMd↑	UDQN↑
Sen	3.724 2	0.789 6	4.277 1	0.655 7	0.393 8
Kalantari	3.767 9	0.763 6	3.940 9	0.653 2	0.398 4
AHDRNet	3.687 7	0.795 7	3.920 5	0.657 0	0.426 0
HDR-GAN	7.134 0	0.799 2	7.033 0	0.666 3	0.411 8
HDRI	5.503 6	0.802 4	5.042 3	0.628 3	0.409 8
HDR-Transformer	2.905 6	0.795 6	3.406 0	0.637 2	0.414 1
SGARN	3.647 3	0.800 3	3.781 5	0.661 2	0.431 1
Ours	2.801 9	0.803 6	3.300 2	0.670 3	0.434 4

2.3 消融实验

为了验证所提模块的有效性，对多重伪影抑制块(MASB)和多级融合块(MFB)进行了消融：w/o MASB & MFB表示同时没有多重伪影抑制操作和多级融合块；w/o MASB表示网络中没有多重伪影抑制操作，但存在着多级融合块。此时用$F_{\mathrm{f}}^{\mathrm{RB}}$代替$F_{\mathrm{f}}^{\mathrm{ASB}}$；w/o MFB表示模块中没有多级融合块。

多重伪影抑制的有效性：如图9所示，没有了多重伪影抑制操作，HDR图像中存在明显的伪影。表3的数据也表明，没有了多重伪影抑制操作，HDR图像的质量在下降。

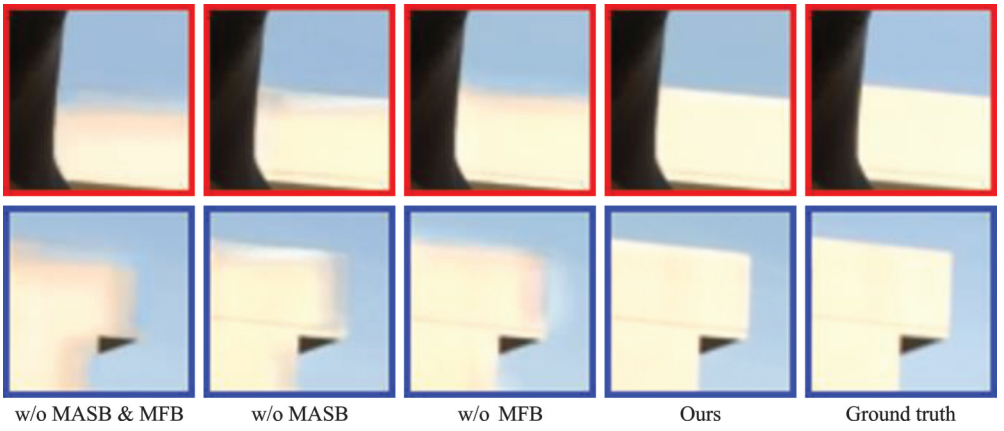


图9 验证本文各部分有效性的视觉效果

Fig.9 Visual results validating the effectiveness of different components of this paper

表3 各部分有效性的客观评价

Table 3 Objective evaluation of the effectiveness of each component

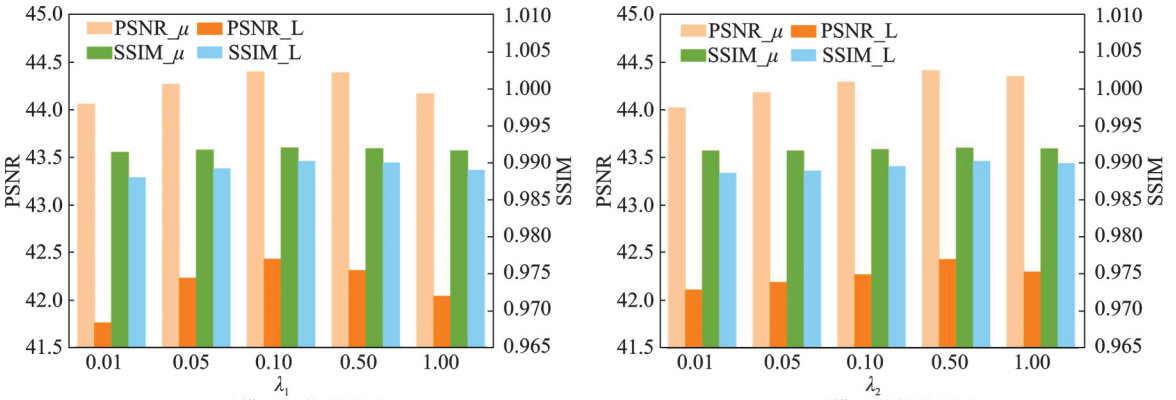
消融情况	PSNR_μ/dB	PSNR_L/dB	SSIM_μ	SSIM_L	HDR-VDP-2
w/o MASB&MFB	44.08	41.75	0.991 4	0.988 1	64.88
w/o MASB	44.12	42.13	0.991 7	0.988 9	64.93
w/o MFB	44.26	42.25	0.991 7	0.989 0	64.95
Ours	44.41	42.43	0.992 0	0.990 2	65.13

多级融合块的有效性:如图9所示,没有了多级融合块,最终结果在边缘处信息有缺失,这是由于没有利用好非参考特征的信息所导致的。表3的指标也说明,对非参考特征的再次融合,可以补充图像的信息。

两个步骤的联合有效性:由图9可以看出,当多重伪影抑制操作和多级融合块都没有的情况下,模型的质量是最差的,不仅有大量信息丢失,还有明显的伪影。而在各个模块齐备的情况下,重建的结果没有伪影,视觉效果也和标签很接近。从表3的定量指标也可以看出,在两个模块同时作用下,所提模型达到最佳的性能。

2.4 参数分析

两个超参数λ_1和λ_2用来平衡损失函数中ℓ_{ssim}和ℓ_g的贡献。对于超参数的选择,在Kalantari数据集上进行分析。图10显示了不同的取值对模型性能的影响。可以看出,当$\lambda_1=0.1$,$\lambda_2=0.5$时,模型的性能达到最优。因此本文将λ_1设置为0.1,将λ_2设置为0.5。



(a) λ_1对模型性能的影响
(a) Effect of λ_1 on model performance

(b) λ_2对模型性能的影响
(b) Effect of λ_2 on model performance

图10 在Kalantari测试集上超参数变化对模型性能的影响

Fig.10 Effect of hyperparameter variation on model performance on Kalantari test set

2.5 模型效率分析

为了对比不同方法的效率,从模型参数量和推理时间两个方面进行了比较,具体结果如表4所示。其中,使用Kalantari测试集中的15张图像计算了平均推理时间。对比结果显示,尽管所提算法模型参数量并非最小,但相较于HDR-GAN和HDRI,模型的大小是可接受的。推理时间方面,尽管尚未达到最优状态,但与Sen、Kalantari和HDRI相比推理时间也是适当的。总体而言,所提算法在效率上表现适中,结合其优越的性能,所提模型在实际应用中是可行的。

表4 不同方法在模型参数量和推理时间上的性能比较

Table 4 Performance comparison of different methods on the number of model parameters and inference time

方法	Sen	Kalantari	AHDRNet	HDR-GAN	HDRI	HDR-Transformer	SGARN	Ours
模型参数量/10^6	—	0.3	1.44	7.67	6.70	1.46	0.89	1.68
推理时间/s	61.81	29.14	0.30	0.69	20.61	4.92	0.53	0.66

2.6 局限性分析

基于多曝光融合的HDR成像方法,使用的数据集基本为Kalantari[27]、Sen[14]和Tursen[41]等。在这些数据集中,目标移动所导致的伪影都可以得到有效抑制。但所提算法的局限性主要体现于对缺失信息的恢复。当某一区域在3个输入图像中都无法获得有效信息时,网络难以恢复出该区域的信息。例如,图像中的某一区域在低曝光和中间曝光图像中因为曝光过低而导致信息缺失,而高曝光图像的该部分区域,因为目标的移动,导致信息被遮挡。这种情况下,该部分的信息很难恢复。这不仅和目标物的移动有关,还需要考虑其他曝光图像的情况。因此,在以后的工作中,将致力于解决这一问题。

3 结束语

考虑到动态场景下伪影难以抑制和融合不充分的问题,提出了一个全新的HDR图像重建网络。对于伪影难以抑制的问题,提出了多重伪影抑制策略,在参考图像的指导下对重建过程中的特征进行伪影再抑制,有效地抑制了重建特征中的伪影信息。为了进一步利用非参考图像的信息且不引入伪影信息,提出了多级融合模块。先提取和重建特征具有相关性的信息,然后将其融合到重建特征中,从而充分利用非参考图像的信息。所提方法在3个公共的数据集上进行了实验,实验结果表明所提方法相较于对比方法在主观和客观评估上均取得了更好的性能。其中在应用最广泛的Kalantari数据集上,所提方法的结果在PSNR_μ和PSNR_L上分别超过了次优算法(HDR-Transformer)0.33 dB。所提算法为高质量HDR图像的生成提供了新的有效途径。

参考文献:

[1] 付争方, 朱虹, 余顺园, 等. 基于灰度级映射函数建模的多曝光高动态图像重建[J]. 数据采集与处理, 2019, 34(3): 472-490.
FU Zhengfang, ZHU Hong, YU Shunyuan, et al. Multi exposure HDR image reconstruction based on gray scale mapping function modeling[J]. Journal of Data Acquisition and Processing, 2019, 34(3): 472-490.

[2] YAN Q S, WANG B, ZHANG L, et al. Towards accurate HDR imaging with learning generator constraints[J]. Neurocomputing, 2021, 428: 79-91.

[3] NIU Y, WU J B, LIU W X, et al. HDR-GAN: HDR image reconstruction from multi-exposed LDR images with large motions [J]. IEEE Transactions on Image Processing, 2021, 30: 3885-3896.

[4] LI R, WANG C, WANG J, et al. UPHDR-GAN: Generative adversarial network for high dynamic range imaging with unpaired data[J]. IEEE Transactions on Circuits and Systems for Video Technology , 2022, 32(11): 7532-7546.

[5] LIU Z, WANG Y L, ZENG B, et al. Ghost-free high dynamic range imaging with context-aware transformer[C]//Proceedings of European Conference on Computer Vision. [S.l.]: Springer, 2022: 344-360.

[6] SONG J W, PARK Y, KONG K, et al. Selective TransHDR: Transformer-based selective HDR imaging using ghost region mask[C]//Proceedings of European Conference on Computer Vision. [S.l.]: Springer, 2022: 288-304.

[7] YE Q, SUGANUMA M, XIAO J, et al. Learning regularized multi-scale feature flow for high dynamic range imaging[EB/OL]. (2022-07-06)[2024-01-14]. https://doi.org/10.48550/arXiv.2207.02539.

[8] MA K, ZENG K, WANG Z. Perceptual quality assessment for multi-exposure image fusion[J]. IEEE Transactions on Image Processing, 2015, 24(11): 3345-3356.

[9] PRABHAKAR K R, SRIKAR V S, BABU R V. DeepFuse: A deep unsupervised approach for exposure fusion with extreme exposure image pairs[C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2017: 4724-4732.

[10] DENG X, DRAGOTTI P L. Deep convolutional neural network for multi-modal image restoration and fusion[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 43(10): 3333-3348.

[11] DENG X, ZHANG Y T, XU M, et al. Deep coupled feedback network for joint exposure fusion and image super-resolution[J]. IEEE Transactions on Image Processing, 2021, 30: 3098-3112.

[12] WARD G. Fast, robust image registration for compositing high dynamic range photographs from hand-held exposures[J]. Journal of Graphics Tools, 2003, 8(2): 17-30.

[13] ZIMMER H, BRUHN A, WEICKERT J. Freehand HDR imaging of moving scenes with simultaneous resolution enhancement[C]//Proceedings of Computer Graphics Forum. Oxford: Blackwell Publishing Ltd., 2011, 30(2): 405-414.

[14] SEN P, KALANTARI N K, YAESOUBI M, et al. Robust patch-based HDR reconstruction of dynamic scenes[J]. ACM Transactions on Graphics, 2012, 31(6): 203.

[15] HU J, GALLO O, PULLI K, et al. HDR deghosting: How to deal with saturation?[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2013: 1163-1170.

[16] RAMAN S, CHAUDHURI S. Reconstruction of high contrast images for dynamic scenes[J]. The Visual Computer, 2011, 27: 1099-1114.

[17] YAN Q S, SUN J Q, LI H S, et al. High dynamic range imaging by sparse representation[J]. Neurocomputing, 2017, 269: 160-169.

[18] MA K, LI H, YONG H, et al. Robust multi-exposure image fusion: A structural patch decomposition approach[J]. IEEE Transactions on Image Processing, 2017, 26(5): 2519-2532.

[19] LI H, MA K, YONG H W, et al. Fast multi-scale structural patch decomposition for multi-exposure image fusion[J]. IEEE Transactions on Image Processing, 2020, 29: 5805-5816.

[20] K S G R, BISWAS A, PATEL M S, et al. Deep multi-stage learning for HDR with large object motions[C]//Proceedings of the IEEE International Conference on Image Processing. [S.l.]: IEEE, 2019: 4714-4718.

[21] LEE S H, CHUNG H, CHO N I. Exposure-structure blending network for high dynamic range imaging of dynamic scenes[J]. IEEE Access, 2020, 8: 117428-117438.

[22] DENG Y P, LIU Q, IKENAGA T. Attention-guided network with inverse tone-mapping guided up-sampling for HDR imaging of dynamic scenes[J]. Multimedia Tools and Applications, 2022, 81: 12925-12944.

[23] HUANG Y M, CHIANG J C, CHEN S G. HDR-AGAN: Ghost-free high dynamic range imaging with attention guided adversarial network[C]//Proceedings of the IEEE International Conference on Image Processing. [S.l.]: IEEE, 2022: 3316-3320.

[24] TAN X, CHEN H, XU K. High dynamic range imaging for dynamic scenes with large-scale motions and severe saturation[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-15.

[25] CHUNG H, CHO N I. High dynamic range imaging of dynamic scenes with saturation compensation but without explicit motion compensation[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. [S.l.]: IEEE, 2022: 61-71.

[26] CHEN J, YANG Z F, CHAN T N, et al. Attention-guided progressive neural texture fusion for high dynamic range image restoration[J]. IEEE Transactions on Image Processing, 2022, 31: 2661-2672.

[27] KALANTARI N K, RAMAMOORTHI R. Deep high dynamic range imaging of dynamic scenes[J]. ACM Transactions on Graphics, 2017, 36(4): 144-1.

[28] LIU C. Beyond pixels: Exploring new representations and applications for motion analysis[D]. [S.l.]: MIT, 2009.

[29] PENG F Y, ZAHNG M J, LAI S M, et al. Deep HDR reconstruction of dynamic scenes[C]//Proceedings of IEEE 3rd International Conference on Image, Vision and Computing. [S.l.]: IEEE, 2018: 347-351.

[30] PRABHAKAR K R, ARORA R, SWAMINATHAN A, et al. A fast, scalable, and reliable deghosting method for extreme exposure fusion[C]//Proceedings of the IEEE International Conference on Computational Photography. [S.l.]: IEEE, 2019: 1-8.

[31] DOSOVITSKIY A, FISCHER P, ILG E, et al. FlowNet: Learning optical flow with convolutional networks[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S.l.]: IEEE, 2015: 2758-2766.

[32] CATLEY-CHANDAR S, TANAY T, VANDROUX L, et al. FlexHDR: Modelling alignment and exposure uncertainties for flexible HDR imaging[J]. IEEE Transactions on Image Processing, 2022, 31: 5923-5935.

[33] ASWANI A V, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of NIPS. [S.l.]: MIT Press, 2017: 6000-6010.

[34] METWALY K, MONGA V. Attention-mask dense merger (attendense) deep HDR for ghost removal[C]//Proceedings of

ICASSP. [S.l.]: IEEE, 2020: 2623-2627.

[35] YAN Q S, GONG D, SHI Q F, et al. Attention-guided network for ghost-free high dynamic range imaging[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2019: 1751-1760.

[36] PU Z Y, GUO P Y, ASIF M S, et al. Robust high dynamic range (HDR) imaging with complex motion and parallax[C]//Proceedings of the Asian Conference on Computer Vision. [S.l.]: Springer, 2020: 134-149.

[37] CHOI S, CHO J, SONG W, et al. Pyramid inter-attention for high dynamic range imaging[J]. Sensors, 2020, 20(18): 5102.

[38] YAN Q S, ZHANG L, LIU Y, et al. Deep HDR imaging via a non-local network[J]. IEEE Transactions on Image Processing, 2020, 29: 4308-4322.

[39] CHEN S Y, CHUANG Y Y. Deep exposure fusion with deghosting via homography estimation and attention learning[C]//Proceedings of ICASSP. [S.l.]: IEEE, 2020: 1464-1468.

[40] LIU Z, LIN W, LI X, et al. ADNet: Attention-guided deformable convolutional network for high dynamic range imaging[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2021: 463-470.

[41] TURSUN O T, AKYUZ A O, ERDEM A, et al. An objective deghosting quality metric for HDR images[C]//Proceedings of Computer Graphics Forum. 2016, 35(2): 139-152.

[42] MANTIUK R, KIM K J, REMPEL A G, et al. HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions[J]. ACM Transactions on Graphics, 2011, 30(4): 1-14.

[43] FANG Y, ZHU H, MA K, et al. Perceptual evaluation for multi-exposure image fusion of dynamic scenes[J]. IEEE Transactions on Image Processing, 2019, 29: 1127-1138.

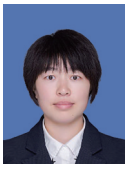
[44] GU K, WANG S, ZHAI G, et al. Blind quality assessment of tone-mapped images via analysis of information, naturalness, and structure[J]. IEEE Transactions on Multimedia, 2016, 18(3): 432-443.

[45] KINGMA D P, BA J. Adam: A method for stochastic optimization[EB/OL]. (2017-01-30). https://doi.org/10.48550/arXiv.1412.6980.

[46] TANG L F, HUANG H, ZHANG Y F, et al. Structure-embedded ghosting artifact suppression network for high dynamic range image reconstruction[J]. Knowledge-Based Systems, 2023, 263: 110278.

作者简介:

罗俊成(1999-),男,硕士研究生,研究方向:计算机视觉、图像处理,E-mail: 1509693938@163.com。



张亚飞(1981-),女,博士,副教授,硕士生导师,研究方向:图像处理、模式识别。

谢明鸿(1976-),**通信作者**,男,博士,副教授,硕士生导师,研究方向:计算机视觉、行人重识别、图像融合,E-mail: minghongxie@163.com。

李华锋(1983-),男,博士,教授,研究方向:图像处理、计算机视觉。

(编辑:夏道家)

High Dynamic Range Imaging with Multiple Artifact Suppression and Multilevel Fusion

LUO Juncheng, XIE Minghong*, ZHANG Yafei, LI Huafeng

(School of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: Due to the limitations of existing imaging equipment, it is difficult to obtain high dynamic range (HDR) images directly. High dynamic range imaging technology is designed to generate HDR images by processing low dynamic range (LDR) images. Most existing deep learning methods reconstruct HDR images by fusing multiple images with different exposures. However, due to the relative movement of foreground and background, artifacts appear in the final reconstruction result. Existing methods only perform artifact elimination before fusing multiple images with different exposures, which leads to a heavy dependence of the final HDR image quality on the artifact suppression results before fusion. Moreover, the artifact information introduced during the fusion process is difficult to eliminate in subsequent reconstruction due to unsatisfactory artifact suppression. To address this, we propose a network framework for multi-artifact suppression of reconstructed features and multilevel information fusion to efficiently reconstruct HDR images. First, we handle the differences between different images and features through multiple artifact suppression. Unlike existing methods that only process images or features before fusion, we perform multiple artifact suppression block (MASB) on the features during the reconstruction process to further suppress artifacts in the reconstructed features. Simultaneously, to better utilize the features of non-reference input images, we propose a multilevel fusion block (MFB), through which complementary information from non-reference images can be further extracted. Experimental comparisons on multiple datasets demonstrate that the proposed method achieves better performance in both subjective visual effects and objective metrics.

Highlights:

1. A novel network framework is proposed for HDR image reconstruction, integrating multi-artifact suppression and multi-level feature fusion.
2. Multiple artifact suppression block (MASB) is applied during reconstruction to further reduce artifacts in features.
3. Multilevel fusion block (MFB) is designed to better exploit complementary information from non-reference images, enhancing the utilization of multi-exposure inputs.

Key words: high dynamic range imaging; deep learning; multiple artifact suppression; multilevel fusion

Foundation items: National Natural Science Foundation of China (Nos.26161015, 62276120); Yunnan Province Basic Research Special Project (No.202301AV070004).
Received: 2024-11-26; **Revised:** 2025-02-25
***Corresponding author. E-mail:** minghongxie@163.com.