

6-16-2022

An Unsupervised Deep Neural Network for Image Fusion

Peipei Zhou

School of Electrical and Information Engineering, Changzhou Institute of Technology, Changzhou 213032, China; zhoupp@czu.cn

Xinglin Hou

School of Electrical and Information Engineering, Changzhou Institute of Technology, Changzhou 213032, China; houxl@czu.cn

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research, Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

An Unsupervised Deep Neural Network for Image Fusion

Abstract

Abstract: Due to the low dynamic range of camera, can not be expressed in the different region of the high dynamic scene a single-exposure image. *An unsupervised depth neural network is constructed to fuse the multi-exposure images into a high dynamic image.* Based on the VGG-Net, *encoding and decoding sub-networks are designed.* Guided by the structural similarity of the images before and after fusion, a loss function suitable for image fusion is designed by introducing the weight factors based on the local image information, and the valid information of the different input images is given consideration. Compared with the other methods, the subjective visual experience and objective quantitative indicators of the fused images are improved significantly.

Keywords

pattern recognition, high dynamic scene, image fusion, unsupervised deep network, loss function

Recommended Citation

Peipei Zhou, Xinglin Hou. An Unsupervised Deep Neural Network for Image Fusion[J]. Journal of System Simulation, 2022, 34(6): 1267-1274.

一种用于图像融合的非监督深度神经网络

周培培, 侯幸林*

(常州工学院 电气信息工程学院, 江苏 常州 213032)

摘要: 为解决相机动态范围较小, 单次曝光的图像往往无法表达高动态场景的不同区域的问题, 构建了一种非监督的深度神经网络, 把多次曝光的图像融合成一幅高动态图像。以VGG-Net (visual geometry group-Net)为基础网络, 设计编码器网络和解码器网络; 以融合前后图像的结构相似度为目标导向, 通过引入基于图像局部信息的权重因子, 定制适用于图像融合的损失函数, 融合图像可兼顾不同输入图像的有效信息。在基准数据集上与多种方法相比, 融合图像在主观视觉体验和客观量化指标上均取得了明显提升。

关键词: 模式识别; 高动态场景; 图像融合; 非监督深度网络; 损失函数

中图分类号: TP391.4 文献标志码: A 文章编号: 1004-731X(2022)06-1267-08

DOI: 10.16182/j.issn1004731x.joss.20-1062

An Unsupervised Deep Neural Network for Image Fusion

Zhou Peipei, Hou Xinglin*

(School of Electrical and Information Engineering, Changzhou Institute of Technology, Changzhou 213032, China)

Abstract: Due to the low dynamic range of camera, can not be expressed in the different region of the high dynamic scene a single-exposure image. An unsupervised depth neural network is constructed to fuse the multi-exposure images into a high dynamic image. Based on the VGG-Net, encoding and decoding sub-networks are designed. Guided by the structural similarity of the images before and after fusion, a loss function suitable for image fusion is designed by introducing the weight factors based on the local image information, and the valid information of the different input images is given consideration. Compared with the other methods, the subjective visual experience and objective quantitative indicators of the fused images are improved significantly.

Keywords: pattern recognition; high dynamic scene; image fusion; unsupervised deep network; loss function

引言

场景动态范围是场景中最大与最小光强度的比值, 自然场景的动态范围往往可达 10^7 。相机的动态范围在 $10^2 \sim 10^4$ 之间。用低动态范围(low dynamic range, LDR)的相机去拍摄高动态范围(high dynamic range, HDR)的场景, 往往使图像中

某一区域曝光过度或曝光不足, 造成信息丢失。多曝光图像融合(multi-exposure fusion, MEF)是解决该问题最有效的方法之一, 该方法将不同曝光时间下获取的多幅图像中的有效信息提取, 重建一幅表达高动态场景信息的图像。

在多曝光图像融合领域, 研究人员做了大量

收稿日期: 2020-12-31 修回日期: 2021-04-16

基金项目: 国防科技重点实验室基金(6142401200301); 江苏省高校自然科学面上项目(20KJB520033, 18KJB510002); 常州市应用基础研究计划(CJ20190052, CJ20200043)

第一作者: 周培培(1991-), 女, 博士, 讲师, 研究方向为数字图像处理与高动态成像。E-mail: zhoup@czu.cn

通讯作者: 侯幸林(1987-), 男, 博士, 讲师, 研究方向为模式识别与图像处理。E-mail: houxl@czu.cn

工作^[1-3]，其关键技术在于多曝光图像的权值分配方法。典型的图像融合方法大致分为2类：①是基于区域分割的多曝光融合方法^[4-5]。该类方法将图像划分为若干块，并提取相应块中信息量最大的进行拼接，获取结果图像，然而，该方法易受块效应的影响；②基于像素权值的多曝光融合方法。例如，基于局部能量和金字塔之间的相关性来计算得到图像每个像素的权重^[6]，使用边缘保持滤波器(如双边滤波器)进行场景合成的方法^[7]，基于概率模型的多曝光图像融合方法^[8]，基于引导滤波的图像融合技术^[9-10]，基于拉普拉斯金字塔的混合曝光权重测量方法^[11]。该类方法可以较好地保持边缘细节，但是容易遗留光晕伪影。综合来看，传统方法极大依赖于设计人员的经验且需要大量的计算资源，因此，该类方法面临着巨大的挑战。

卷积神经网络在计算机图像处理领域取得了长足的发展，不仅在目标跟踪等高级视觉处理领域表现出了极大优势^[12-14]，而且在图像滤波等低级图像处理领域也被广泛应用^[15-17]。Prabhakar等^[18]以无参考的图像质量指标作为损失函数，提出了一个无监督的多曝光融合深度学习框架；Li等^[19]针对多光源图像融合，提出了一种DenseFuse融合网络，其可用于多曝光融合；Ma等^[20]提出了一种可用于任意空间分辨率和曝光次数的静态图像序列融合的快速MEF-Net网络。然而，由于网络框架与损失函数的设计方法欠佳，已有的多曝光融合

网络性能仍有待提升。

本文提出了一种端到端的深度网络框架，以VGG-Net(visual geometry group-Net)为基础模型，扩展并设计了编码子和解码子网络。将输出归一化为权值图像，并以融合结果与原始图像的结构相似度为优化目标，设计损失函数，最终实现了端到端的输出。该方法无需对训练图像进行标注，以融合结果的质量为优化目标，实现了无监督的训练，极大缩短了该类方法的开发周期。

1 无监督深度神经网络

本文设计了一个端到端的图像融合框架，以不同曝光条件下获取的2幅图像为输入，经过编码子网络和解码子网络后，输出两幅图像的权值并对原始图像加权求和得到融合结果，在此基础上，设计针对融合结果图像质量的优化目标，最终构成整个网络。

1.1 网络结构设计

为了能无监督地学习图像融合的权值，本文设计了如图1所示的网络结构。首先将两幅彩色图像转换为灰度图像，以避免彩色图像融合时产生色偏，并将这组灰度图像作为网络的输入，而输出的权值是与输入图像等大的两幅权值图像。为此，网络结构分为编码子网络和解码子网络，编码子网络中交替使用卷积层和池化层，解码子网络中则交替使用反卷积层和上采样层。

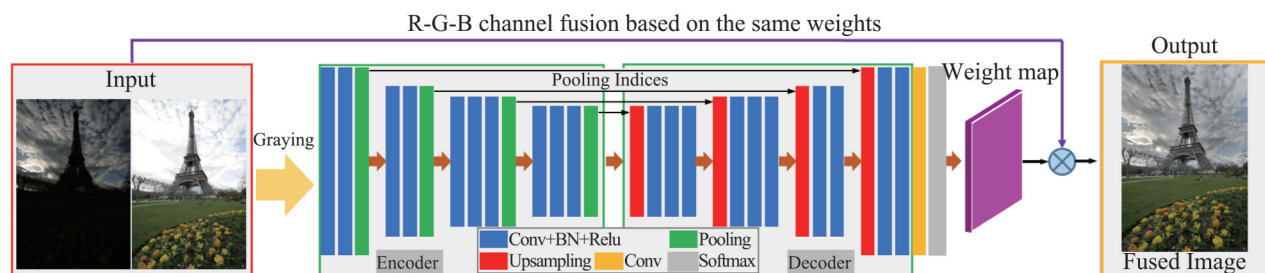


图1 本文网络的整体结构图

Fig. 1 General architecture of proposed network

编码器网络中卷积层(Conv)、归一化操作(batch normalization, BN)^[21]和修正线性单元(ReLU, $f(x)=\max(0, x)$)^[22]组成了一个子模块A, 如图1中蓝色区域所示。模块A后的最大池化操作可保证输入图像存在小空间位移时的平移不变性, 编码器网络中连续下采样使得其输出特征图中存在大量的空间信息, 在此过程中, 将保留最大池化的索引用于解码器网络中的上采样。编码器网络中包含10个卷积层, 分别对应于VGG-Net^[23]网络的前10个卷积层。

解码器网络与编码器网络一一对应, 将编码结果进行上采样并重复经过子模块A, 得到和输入图像大小($H \times W$)一致的64维的特征图像($H \times W \times 64$), 最终经过卷积层和逻辑回归操作得到权值图。

1.2 损失函数设计

本文方法需在无标注数据的前提下进行无监督学习, 因此, 如何令融合结果向最优的方向收敛是损失函数设计的根本任务。图像结构相似度(SSIM)往往被用来评价图像融合结果的图像质量。假设2幅欠曝光和过曝光的原始图像分别记为 f_u 和 f_o , 融合结果为 F , 为了能使融合图像中包含2幅图像的全部信息, 2幅图像中细节丰富区域的信息需在融合结果中保留, 因此, 求融合图像与2幅原始图像的局部结构相似度^[24], 并以原始图像中局部信息的丰富程度设计权值, 将2个局部相似度加权求和并训练, 使其最大化, 即可得到最优的融合结果, 损失函数的设计为

$$\text{loss} = 1 - \frac{1}{15} \sum_{c \in \{R, G, B\}} \sum_w (\alpha_w S(f_u(w), F(w), w) + (1 - \alpha_w) S(f_o(w), F(w), w)) \quad (1)$$

式中: α_w 为根据图像局部信息丰富程度设计的权值系数; $S(x, y, w)$ 为图像 x 和 y 在窗口 w 内的局部相似度, 其定义为

$$S(x, y, w) = \frac{(2\bar{w}_x \bar{w}_y + C_1)(2\sigma_{w_x w_y} + C_2)}{(\bar{w}_x^2 + \bar{w}_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (2)$$

式中: C_1 和 C_2 为经验值, 分别设置为 1×10^{-4} 和

9×10^{-4} ; w_x 为图像 x 中的窗口区域; \bar{w}_x 为该区域图像灰度的均值; σ_x^2 和 $\sigma_{w_x w_y}$ 分别为区域 w_x 内图像灰度的方差和 w_x 与 w_y 内图像灰度的协方差。结构相似度的取值范围为 $[0, 1]$, 且图像越相似则相似度取值越趋于1。

为了提高系统的稳定性, 使用平均梯度来表达图像局部区域的曝光质量。权值系数的设计将以不同曝光下2幅图像中对应区域的平均梯度来设计, 平均梯度越大, 所对应图像区域的权值就越大, 针对图像 f , α_w 的设计为

$$\alpha_w = \frac{g(G_{\text{ave}}(w_x))}{g(G_{\text{ave}}(w_x)) + g(G_{\text{ave}}(w_y))} \quad (3)$$

图像的梯度计算方法及图像计算窗口 w 内的平均梯度为

$$G_{\text{ave}}(w) = \text{mean}_w(G(i, j)) \quad (4)$$

$$G(i, j) = \sqrt{[f(i, j) \otimes \mathbf{h}_\perp]^2 + [f(i, j) \otimes \mathbf{h}_\parallel]^2} \quad (5)$$

式中: 2种梯度算子 $\mathbf{h}_\parallel = [1, 0, -1]$, $\mathbf{h}_\perp = [1, 0, -1]^T$ 。为了避免出现奇点, 设置 $g(x) = \max(x, 0.0001)$ 。

2 实验与分析

2.1 数据准备与网络训练

本文采用文献[5, 9–10, 18]中的56组, 其中45组用来训练网络参数, 其余11组作为测试基准数据集, 用来测试本文算法与已有算法的融合性能。本文搜集的图像数据集序列涉及不同的HDR场景, 如室内外、自然光和人工光照、线性和非线性曝光等, 图像的分辨率从 $231 \times 338 \sim 2128 \times 1416$ 不等。在训练阶段, 随机使用45个图像组来训练网络, 其余11组用于测试。训练图像被裁剪成大小为 60×60 的图像块, 生成超过30万个训练数据块。此外, 使用翻转(上下和左右)和旋转(0° 、 90° 、 180° 和 270°)进行数据扩充。本文网络的直接输入为2幅图像, 由不同曝光时长的2幅图像组成。

为了提高网络的稳定性, 本方法在训练过程

中使用不同尺度的窗口评价融合结果与原始图像的结构相似度。文中选择了5种大小为 3×3 、 5×5 、 7×7 、 9×9 、 11×11 的窗口，并对R、G、B 3个通道计算结构相似度，最终将结果取均值，为此，式(1)中的系数为 $1/15$ 。编码器网络中所有卷积核的大小均为 3×3 ，卷积核的数量分别为64、64、128、128、256、256、256、512、512和512，解码器网络中卷积核的数量与编码器网络中一一对应。训练过程使用了最小批处理大小为64的Adam优化器^[25]，学习率最初设定为0.000 1，并以0.99的速率呈指数衰减，训练50次。

所有对比方法的源代码均由原作者提供，本文模型及测试的源代码见文献[26]。

2.2 客观评价指标

为了验证本文算法的有效性，选取已有的6种多曝光融合方法进行比较，其中包括4种传统方法(Shen11^[8]、Li13^[9]、Ma15^[5]和Hou16^[10])和2种基于深度学习的方法(DF19^[19]和MEF-Net20^[20])。

单幅图像质量的客观评价指标较多，但是多曝光图像融合的评价指标需要考虑以下因素：

(1) 融合前后的图像信息对比：其难点在于该指标能否抓取多幅图像不同位置的有效信息并与融合后的信息做比较；

(2) 与视觉效果的一致性：量化指标应可以较好地匹配视觉效果；

(3) 指标的影响力：提出一个好的多曝光融合量化指标很困难，能在多篇有影响力的文中被采纳，可以间接说明一个指标的有效性。

本文选取多曝光融合结构相似度(MEF structural similarity, MEF-SSIM)得分^[27]作为评价融合图像质量的量化指标。MEF-SSIM值越高，结果越好。

2.3 两次曝光图像的融合

本文设计的深度网络的直接输入为欠曝光和过曝光的2幅图像，因此，针对两次曝光的图像

输入时，首先比较本文算法与已有算法得到的融合图像的主观视觉效果，需要说明的是，实验中采用的两次曝光图像是在公开数据集的该组高动态场景曝光图像中任意选取的2幅图像。

针对一个室外的HDR场景Balloons，单次曝光难以兼顾天空和地面2个亮度差异较大的区域，图2(a)为曝光不足和曝光过度的2幅图像，不同方法的融合结果如图2(b)~(g)所示。图2(b), (e), (f)所示的3幅图像整体色调偏暗，视觉效果较差。图2(g)中天空区域明显偏亮，云已几乎看不清楚，融合后的细节丢失较多。Li13的方法中细节得到了充分增强，但是该图像的天空和地面景色的交界处存在明显的白色边缘，即不同亮度的区域过渡区域的融合效果较差。此外，在Ma15和MEF-Net20方法的融合图像中，2个较大的热气球与天空的交界处有明显的白色光晕存在。基于DenseFuse^[19]，DF19-add和DF19-L1分别采用加法和L1-norm的融合策略，DF19-L1的融合图像在地面景物比DF19-add的效果更亮，但是天空细节丢失严重，二者都不能呈现出满意的全局景色。相比之下，本文方法(图2(i))能够捕捉到大范围的亮度变化，保持全局对比度，使天空与地面2个不同亮度区域的景色均得到了较好恢复，且在天地交接处没有出现伪影，产生更自然、更生动的视觉图像，较其他方法，本方法融合图像的视觉效果更好，说明本文设计的网络结构和损失函数是有效的。

此外，针对室内场景“Room”欠曝光与过曝光的2幅图像，图3(b)~(i)为不同融合方法的对比效果图，单次曝光难以很好地同时捕捉窗内外的细节，该场景的融合难点在于窗外景色的获取以及窗外到室内场景的自然过渡。从图3可见，图3(b), (d), (e)~(g) 5幅融合图像的窗外景色均看起来整体偏白，细节丢失较多，究其原因，主要是由于对2幅原始的曝光图像进行权值分配时，过曝光图像中该部分的权值过大造成的。图3(c)显著提升了细节方面的效果，但是图中明显存在光晕伪影，即该方法在亮区和暗区的过渡区域权值设计不合理。

MEF-Net20 是专为多曝光静态图像融合设计的深度神经网络, 可适用于任意分辨率和任意曝光次数的图像序列融合, 图 3(h)可见该方法的融合图像中细节部分处理较好, 且图像整体颜色鲜艳, 但是在中间的百叶窗左侧边缘部分有轻微的光晕存在, 比较而言, 图 3(i)所示的本文方法的融合效果无光晕存

在, 即较好地处理了亮区与暗区的过渡区域。此外, 本文方法的融合图像颜色保真度好, 整体可见颜色无违和感。从表 1 中“Room”的量化指标可见, 本文方法的融合图像得到了最高的 MEF-SSIM 分数, 说明融合过程中的细节保留更多, 证明我们设计的融合网络是有效的。

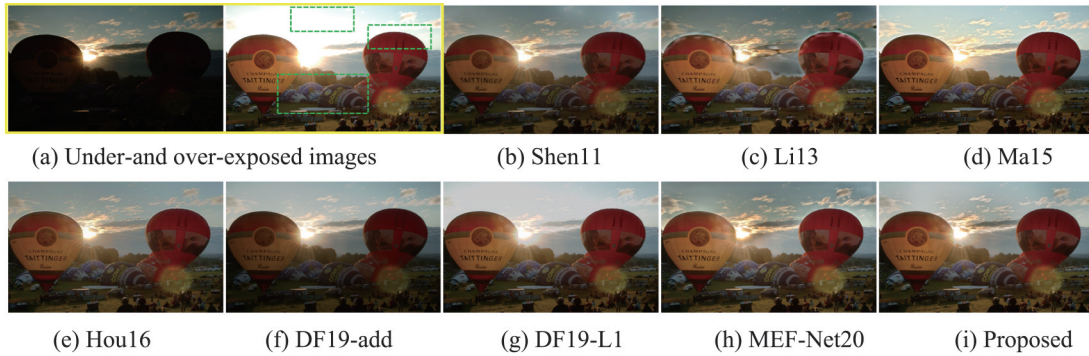


图 2 针对室外“Balloons”的两幅曝光图像, 本文方法与 6 种已有融合方法的融合效果比较

Fig. 2 Comparison of proposed method with 6 previous different fusion methods on outdoor "Balloons" image sequence



图 3 针对室内“Room”的两幅曝光图像, 本文方法与 6 种已有融合方法的融合效果比较

Fig. 3 Comparison of proposed method with 6 previous different fusion methods on indoor "Room" image sequence

表 1 在 11 个高动态场景下, 两次曝光输入下, 不同算法得到的融合图像的 MEF-SSIM 分值比较

Table 1 MEF-SSIM scores of fusion images with two-exposure images using different methods on 11 test sequences.

HDR 场景	Shen11	Li13	Ma15	Hou16	DF19-add	DF19-L1	MEF-Net20	本文方法
Balloons	0.932 3	0.908 2	0.987 9	0.932 7	0.947 8	0.963 2	0.978 3	0.989 4
Belgium house	0.926 6	0.884 9	0.978 6	0.918 5	0.937 9	0.958 4	0.974 3	0.989 0
Chinese garden	0.922 9	0.917 9	0.993 5	0.923 5	0.941 3	0.985 3	0.987 1	0.993 3
House	0.925 2	0.830 6	0.964 2	0.916 5	0.901 3	0.765 7	0.935 0	0.966 5
Landscape	0.899 8	0.905 2	0.994 8	0.921 9	0.984 2	0.990 8	0.981 6	0.989 4
Lighthouse	0.912 8	0.914 9	0.984 3	0.933 4	0.952 5	0.953 3	0.969 2	0.986 3
Madison capitol	0.893 5	0.800 5	0.981 2	0.871 2	0.914 0	0.906 6	0.958 1	0.984 1
Office	0.904 4	0.725 6	0.983 9	0.893 8	0.948 6	0.912 3	0.955 5	0.983 9
Room	0.920 6	0.853 3	0.978 2	0.915 1	0.938 4	0.939 6	0.967 8	0.987 1
Tower	0.912 8	0.858 6	0.986 0	0.910 1	0.915 1	0.921 9	0.959 1	0.988 5
Venice	0.903 9	0.834 9	0.987 9	0.904 6	0.935 9	0.961 4	0.965 7	0.982 7
Average	0.914 1	0.857 7	0.983 7	0.912 8	0.937 9	0.932 6	0.966 5	0.985 5

客观评价指标方面, 本文选用 MEF-SSIM 评分法^[27]对基于不同图像序列的 7 种无监督方法的性能进行了评价, 本文方法与 6 种已有方法之间的定量比较如表 1 所示。比较室外场景图 2 与室内场景图 3 中的不同融合图像的 MEF-SSIM 得分可以看出, MEF-SSIM 得分越高, 融合图像的视觉效果也越好。从表 1 可见, Li13 方法的 MEF-SSIM 得分明显低于其他方法, 通过图 2(c) 和图 3(c) 的视觉效果发现该方法的伪影比较明显, 导致伪影处细节丢失较多。作为一种传统的方法, Ma15 的 MEF-SSIM 得分较高。然而, 基于深度学习的 DF19 是端到端的图像融合网络, 该方法的得分值低得多, 主要原因是该方法为多曝光图像分配权值的策略较差, 无法兼顾不同亮度区域。表 1 中可以看出本文方法优于已有方法, 这是因为本文所设计的网络结构更适用于多曝光图像融合领域。与已有方法相比, 本文方法的平均值取得了最高分, 说明本文方法融合的信息量更大。综合主观视觉效果和客观评价指标, 本文方法大部分情况下的效果都优于已有算法, 究其原因, 本文所构造的损失函数更符合人类视觉的感知规律, 大幅提升了网络性能。

2.4 多次曝光图像的融合

虽然本网络的直接输入是 2 幅图像, 但通过逐一融合, 可以很容易扩展到任意数量的输入图像, 据此, 设计了以下实验进行验证。首先, 选取 11 组基准测试集中有挑战性较大的 1 组多输入图像序列, 如图 4(a) 所示 “House” 的 4 个不同曝光时间下的图像序列。在 MEF-SSIM 得分中, 图 4(f)~(g) 是得分最低的 2 幅图像, 均来自于 19DF 方法, 视觉上来看, 图像在窗外景色部分明显苍白, 细节丢失严重, 视觉效果较差。图 4(c) 光晕伪影严重, 图 4(e) 的全局对比度较差。通过对更多输入图像进行融合, 一个好的融合方法应提升 MEF-SSIM 的得分, 与表 1 中 “House” 的 MEF-SSIM 得分进行比较, 本文方法的得分得到提升, 说明本文方法适合多曝光融合。综合来看, 本文方法和 MEF-Net20 的融合图像在室内外均能呈现比较好的视觉效果, 与 MEF-SSIM 评分的客观指标相符。此外, 与专为多曝光融合设计的深度网络 MEF-Net20 相比, 本文方法融合图像得分稍高, 说明本网络可以很好地用于多曝光图像融合。

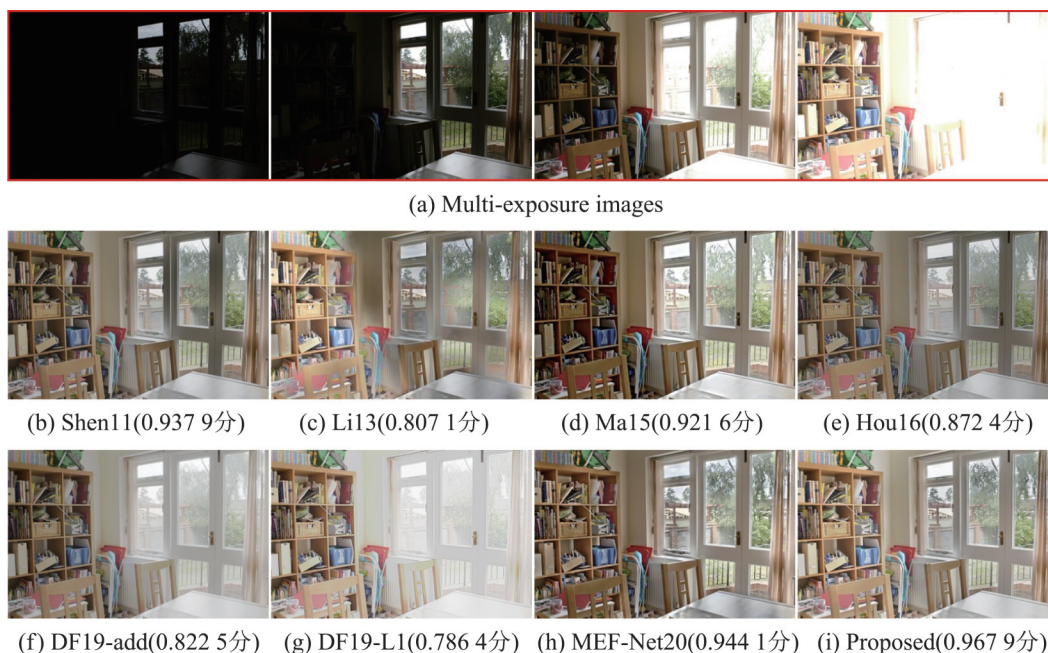


图 4 与 6 种不同融合方法的融合效果比较
Fig. 4 Comparison with 6 previous different fusion methods

2.5 参数分析

本文设计了一个多尺度加权损失函数。其中最重要的参数是 α_w , 本文的目标是为损失函数设计一个参数 α_w , 使其具有3个理想的性质: ①可度量性。 α_w 应反映局部图像的曝光度; ②梯度可微性。 α_w 应该是可微的, 以确保错误反向传播的过程; ③简单性。 α_w 应尽可能简单, 以减少训练阶段时间损耗。

本文根据以上规则, 利用平均梯度定义了 α_w , 也可以用其他指标来构造该参数, 例如, 曝光度、局部方差等。

3 结论

针对静态高动态场景的多曝光图像融合任务, 构建了一个端到端的无监督深度网络, 主要包括编码和解码2个子模块。此外, 针对多曝光图像融合的特点, 设计了特定的损失函数, 用以实现网络的无监督训练, 提升网络性能。在公开基准数据集上, 针对两输入和多输入的图像序列的多种实验对比发现, 本文融合方法在定性和定量2种评价方法上都优于现有的融合方法。可见, 本文提出的无监督深度网络适用于多曝光图像融合任务。然而, 在有动态目标的高动态场景中, 本文方法的融合效果不佳, 该问题将纳入未来的研究任务中。

参考文献:

- [1] Ma K, Li H, Yong H, et al. Robust Multi-Exposure Image Fusion: a Structural Patch Decomposition Approach[J]. IEEE Transactions on Image Processing(S1057-7149), 2017, 26(5): 2519-2532.
- [2] 魏利胜, 张平改. 基于分层模型的图像快速融合方法研究[J]. 系统仿真学报, 2016, 28(6): 1372-1379.
Wei Lisheng, Zhang Pingai. Fast Fusion Method for Multi-focus Image Based on Hierarchical Model[J]. Journal of System Simulation, 2016, 28(6): 1372-1379.
- [3] 马夏一, 范方晴, 卢陶然, 等. 基于图像块分解的多曝光图像融合去鬼影算法[J]. 光学学报, 2019, 39(9): 132-140.
Ma Xiayi, Fan Fangqing, Lu Taoran, et al. Multi-Exposure Image Fusion De-Ghosting Algorithm Based on Image Block Decomposition[J]. Acta Optica Sinica, 2019, 39(9): 132-140.
- [4] Goshtasby A A. Fusion of Multi-Exposure Images[J]. Image and Vision Computing(S0262-8856), 2005, 23(6): 611-618.
- [5] Ma Keda, Zhou Wang. Multi-Exposure Image Fusion: A Patch-Wise Approach[C]//2015 IEEE International Conference on Image Processing (ICIP). Quebec: IEEE, 2015: 1717-1721.
- [6] Burt P J, Kolczynski R J. Enhanced Image Capture Through Fusion[C]//1993 (4th) International Conference on Computer Vision. IEEE, 1993: 173-182.
- [7] Raman S, Chaudhuri S. Bilateral Filter Based Compositing for Variable Exposure Photography[C]//Eurographics. The Eurographics Association, 2009: 1-4.
- [8] Shen R, Cheng I, Shi J, et al. Generalized Random Walks for Fusion of Multi-Exposure Images[J]. IEEE Transactions on Image Processing(S1057-7149), 2011, 20(12): 3634-3646.
- [9] Li S, Kang X, Hu J. Image Fusion with Guided Filtering[J]. IEEE Transactions on Image processing(S1057-7149), 2013, 22(7): 2864-2875.
- [10] Hou X, Luo H, Qi F, et al. Guided Filter-Based Fusion Method for Multiexposure Images[J]. Optical Engineering (S0091-3286), 2016, 55(11): 113101.
- [11] Shen J, Zhao Y, Yan S, et al. Exposure Fusion Using Boosting Laplacian Pyramid[J]. IEEE Transactions on Cybernetics(S2168-2267), 2014, 44(9): 1579-1590.
- [12] Wang L, Xiong Y, Wang Z, et al. Temporal Segment Networks for Action Recognition in Videos[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence(S0182-8828), 2018, 41(11): 2740-2755.
- [13] Lin T Y, Goyal P, Girshick R, et al. Focal Loss for Dense Object Detection[C]//IEEE International Conference on Computer Vision. IEEE, ICCV, 2017: 2980-2988.
- [14] Fu J, Liu J, Tian H, et al. Dual Attention Network for Scene Segmentation[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2019: 3146-3154.
- [15] Divakar N, Venkatesh Babu R. Image Denoising Via CNNs: an Adversarial Approach[C]// IEEE Conference on Computer Vision and Pattern Recognition Workshops. IEEE, 2017: 80-87.
- [16] Prabhakar K R, Babu R V. Ghosting-Free Multi-Exposure Image Fusion in Gradient Domain[C]// 2016 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2016: 1766-1770.
- [17] Eilertsen G, Kronander J, Denes G, et al. HDR Image Reconstruction from a Single Exposure Using Deep

- CNNs[J]. *Acm Transactions on Graphics*(S0730-0301), 2017, 36(6): 1-15.
- [18] Prabhakar K R, Srikar V S, Babu R V. DeepFuse: A Deep Unsupervised Approach for Exposure Fusion with Extreme Exposure Image Pairs[C]// 2017 IEEE International Conference on Computer Vision. IEEE Computer Society, 2017: 4714-4722.
- [19] Li H, Wu X J. DenseFuse: A Fusion Approach to Infrared and Visible Images[J]. *IEEE Transactions on Image Processing*(S1057-7149), 2018, 28(5): 2614-2623.
- [20] Ma K, Duanmu Z, Zhu H, et al. Deep Guided Learning for Fast Multi-Exposure Image Fusion[J]. *IEEE Transactions on Image Processing*(S1057-7149), 2019, 29(1): 2808-2819.
- [21] Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift[C]//International Conference on Machine Learning. PMLR, 2015: 448-456.
- [22] Nair V, Hinton G E. Rectified Linear Units Improve Restricted Boltzmann Machines Vinod Nair[C]// International Conference on International Conference on Machine Learning. Omnipress, 2010: 807-814.
- [23] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J/OL]. [2020-12-14]. <https://arxiv.org/abs/1409.1556v6>.
- [24] 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*(S1057-7149), 2004, 13(4): 600-612.
- [25] Zhang J, Shao J, Chen J, et al. PFNet: an Unsupervised Deep Network for Polarization Image Fusion[J]. *Optics Letters*(S0146-9592), 2020, 45(6): 1507-1510.
- [26] 周培培. 无监督深度多曝光图像融合网络[EB/OL]. [2020-12-14]. <https://github.com/HouXinglin2018/Multi-Exposure-image-fusion-1>.
Zhou Peipei. Unsupervised Deep Multi-Exposure Image Fusion Network [EB/OL]. [2020-12-14]. <https://github.com/HouXinglin2018/Multi-Exposure-image-fusion-1>.
- [27] Ma K, Zeng K, Wang Z. Perceptual Quality Assessment for Multi-Exposure Image Fusion[J]. *IEEE Transactions on Image Processing*(S1057-7149), 2015, 24(11): 3345-3356.