

## 基于改进拉普拉斯金字塔的 HDR 图像色调映射算法

张博文, 夏振平\*, 张跃渊, 程成, 刘宇杰  
苏州科技大学电子与信息工程学院, 江苏 苏州 215009

**摘要** 为进一步提升高动态范围图像在普通显示器上的呈现效果, 提出了一种基于改进拉普拉斯金字塔的高动态范围图像色调映射算法。该算法将预处理后的图像分解为高频层和低频层, 分别输入 2 个特征提取子网络, 将 2 个包含不同特征的输出图像融合后再输入微调网络, 最终得到感知效果优越的低动态范围图像。此外, 该算法设计了自适应分组卷积模块以增强子网络提取局部和全局特征的能力。测试结果表明: 与现有的先进算法相比, 所提算法可以更好地压缩高动态范围图像的亮度, 保留更多图像细节, 拥有更加优越的客观质量指标和主观感知效果。

**关键词** 图像处理; 高动态范围图像; 色调映射; 拉普拉斯金字塔; 深度学习

中图分类号 TP391.9

文献标志码 A

DOI: 10.3788/LOP230581

## Tone Mapping Algorithm for High Dynamic Range Images Based on Improved Laplacian Pyramid

Zhang Bowen, Xia Zhenping\*, Zhang Yueyuan, Cheng Cheng, Liu Yujie

College of Electronics and Information Engineering, Suzhou University of Science and Technology,  
Suzhou 215009, Jiangsu, China

**Abstract** A tone mapping algorithm for high dynamic range (HDR) images based on the improved Laplacian pyramid is proposed to enhance the rendering effect of HDR images on ordinary displays. The algorithm decomposes the preprocessed image into high-frequency and low-frequency layers, which are then fed into two feature extraction sub-networks. The algorithm combines their output images having different features via a fine-tuning network and finally obtains a low dynamic range image with a superior perceptual effect. Furthermore, the algorithm designs an adaptive group convolution module to enhance the ability of the sub-network to extract local and global features. The test results show that, compared to the existing advanced algorithms, the proposed algorithm can compress the brightness of the HDR image better, retain more image details, and achieve superior objective quality and subjective perception.

**Key words** image processing; high dynamic range image; tone mapping; Laplacian pyramid; deep learning

## 1 引言

高动态范围(HDR)图像的亮度范围远超低动态范围(LDR)图像,在诸多领域<sup>[1]</sup>扮演着重要的角色。普通显示设备受硬件成本限制,无法直接显示HDR图像。为此,学者们提出一种用于压缩HDR图像亮度范围的算法,此类算法被统称为HDR图像色调映射算法<sup>[2]</sup>。

最初的色调映射算法采用全局映射法,即将同一个函数应用于图像中所有的像素。Lee等<sup>[3]</sup>引入了一种非对称曲线,改进了传统视网膜模型,并基于此提出了新型全局色调映射算法。Khan等<sup>[4]</sup>提出了一种结

合人类视觉系统灵敏度模型的全局色调映射算法。全局色调映射算法易于实现,不容易产生伪影,但由此生成的低动态范围图像通常对比度较低,容易丢失图像细节。与全局映射算法相比,局部映射算法加入了局部区域的像素统计信息,能够更好地保留图像的局部细节。Reinhard等<sup>[5]</sup>在映射时加入了邻域内的像素信息,提出了一种经典的局部色调映射算法。

随着深度学习技术在图像领域的发展,HDR图像色调映射算法愈发丰富。Montulet等<sup>[6]</sup>将U-Net和Patch-GAN分别作为生成器与鉴别器,提出深度卷积生成对抗网络(DCGANs),该网络取得较好的效果。Cao等<sup>[7]</sup>提出一种自适应色调映射算法,在条件生成对

收稿日期: 2023-02-06; 修回日期: 2023-03-04; 录用日期: 2023-03-10; 网络首发日期: 2023-03-20

通信作者: \*xzp@usts.edu.cn

抗网络(cGAN)的基础上进行相应优化。上述方法虽然取得较好的效果,但基于生成对抗网络(GAN)结构的网络计算开销较大,在处理高分辨率HDR图像时难以取得速度与效果间的平衡。Liang等<sup>[8]</sup>提出了一种将HDR图像进行多层分解后再重组的算法,避免了光晕伪影的问题,但该算法在HDR图像低亮度区域的图像细节保留度较低。同时,色调映射算法性能的最佳标准是人类的主观感知,上述算法在构建的过程中未充分考虑主观感知对算法性能的影响。

针对上述问题,本文提出了一种基于改进拉普拉斯金字塔<sup>[9]</sup>的色调映射算法。该算法将预处理后的HDR图像进行拉普拉斯分解,分解为高频层和低频层,分别输入全局提取网络与局部调整网络,然后将融合后的输出图像输入微调网络,得到主观感知效果优秀的LDR图像。为增强子网络的特征提取能力,设计了自适应分组卷积模块(AGCM),通过自适应选取卷积核尺寸来获取更丰富的图像结构信息。在数据集构建过程中引入主观感知实验<sup>[10]</sup>,搭建主客观相结合的质量评价模型来优化所构建的数据集,最终取得较好的效果。

## 2 色调映射算法

### 2.1 网络架构

所提的色调映射网络由3个子网络组成,分别为全局提取网络、局部调整网络以及微调网络,如图1所示。

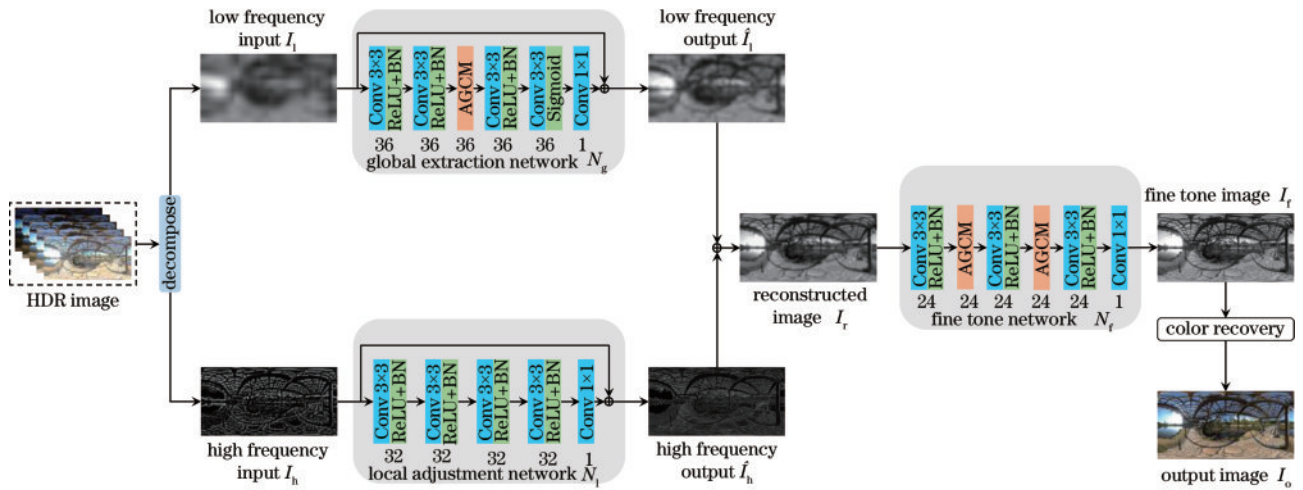


图1 色调映射算法网络架构

Fig. 1 Network architecture of tone mapping algorithm

为使网络能够快速高效地获取图像特征并对图像进行调整,针对3个子网络设计了不同的损失函数。 $L_1$ 损失和 $L_2$ 损失都是深度学习中最著名且常用的损失函数。 $L_1$ 损失用于度量输出图像和参考图像之间差值的绝对值之和, $L_1$ 可表示为

$$L_1 = \frac{1}{n} \sum_{i=1}^n |f(x_i; \theta) - y_i| \quad (1)$$

式中: $n$ 为输入图像的数量; $f(x_i; \theta)$ 为网络输出图像,

首先,网络将HDR图像进行预处理,交由改进的拉普拉斯金字塔进行分解,图像被分解为低频图像和高频图像(参考图像预处理后也采用相同的方法进行分解,分解后的图像分别作为全局网络和局部网络的参考图像,预处理后的图像作为微调网络的参考图像)。其中,低频图像主要包含原图中的亮度信息,而高频图像则包含结构信息。然后,将低频图像和高频图像分别输入全局提取网络和局部调整网络进行优化,将2个输出图像融合为重建图像并输入微调网络。微调网络完成优化后进行色彩还原,得到对应的LDR图像。

局部调整网络主要由5个卷积层组成,其目的是优化HDR图像中的高频信息部分,即分解后的高频层 $I_h$ 。除了最后一层卷积外均采用卷积核大小为3的卷积并使用ReLU激活函数和批标准化对卷积结果进行处理。最后一层采用卷积核大小为1的卷积将卷积结果融合,加入跳跃连接结构来保持图像的结构。全局提取网络的结构与局部调整网络类似,倒数第2个卷积层中将激活函数修改为Sigmoid函数以增强卷积结果的非线性。AGCM模块是为了增强网络对局部信息和全局信息的综合理解能力。微调网络由2个AGCM模块和4个卷积层构成,用于优化由高频层输出图像和低频层输出图像融合后的重建图像。图1中,3个网络中各个卷积层以及模块下方的数字代表输出通道数,每个子网络初始的输入通道数均为1。

$x_i$ 为输入图像, $\theta$ 为网络参数; $y_i$ 为参考图像。 $L_2$ 损失可以度量输出图像和参考图像之间差值平方的和, $L_2$ 可表示为

$$L_2 = \frac{1}{n} \sum_{i=1}^n [f(x_i; \theta) - y_i]^2 \quad (2)$$

特征损失<sup>[11]</sup>使用已经在ImageNet数据集上预先训练好的16层VGG网络来衡量两幅图像在语义信息上的差异。该损失函数能够衡量两幅图像之间不同层

次特征之间的相似性,相较于上述 2 种损失,其更注重图像的高维特征。特征损失  $L_{\text{feat}}(x_g, y_g)$  可表示为

$$L_{\text{feat}}(x_g, y_g) = \sum_{i=1}^5 \frac{1}{W_i H_i C_i} \|\phi_i(x_g) - \phi_i(y_g)\|_1, \quad (3)$$

式中:  $x_g$  为输出图像;  $y_g$  为参考图像;  $\phi_i(x_g)$ 、 $\phi_i(y_g)$  分别为输出和参考图像在 VGG16 网络中第  $i$  个激活层输出的特征图;  $W_i$ 、 $H_i$ 、 $C_i$  分别为特征图的宽度、高度、通道数。此外,还引入了  $L_2$  正则化损失  $R$ , 以防止网络参数过拟合,具体可表示为

$$R = \lambda \|\theta\|_2^2, \quad (4)$$

式中:  $\theta$  为网络的参数;  $\lambda$  为该函数的系数。局部调整网络的损失函数  $L_{\text{local}}$  可表示为

$$L_{\text{local}} = \alpha L_1 + \beta L_{\text{feat}} + \lambda R, \quad (5)$$

式中:  $\alpha=0.5$ ;  $\beta=0.5$ ;  $\lambda=0.2$ 。全局提取网络和局部调整网络的最终目标是一致的,故在 2 个网络中都采取相同的损失函数。重建图像与参考图像已经较为相似,但颜色偏移、区域模糊以及其他伪影问题也都有可能存在于重建图像中。为了帮助微调网络对重建图像进行优化,将  $L_i$  作为其损失函数,  $L_i$  可表示为

$$L_i = \gamma L_2 + \eta L_{\text{feat}}, \quad (6)$$

式中:  $\gamma=0.6$ ;  $\eta=0.4$ 。

### 2.2 改进的拉普拉斯金字塔

传统拉普拉斯金字塔可以将图像的全局亮度信息压缩到较低分辨率而不牺牲细节,另一方面,对低分辨率图像进行卷积运算可以有效降低计算复杂度,从而降低对计算设备的要求。利用拉普拉斯金字塔可以将 HDR 图像分割成  $n$  个不同的频带,最低频带包含原始图像的亮度信息,而其他频带则包含了局部细节和纹理信息。然而,采用传统拉普拉斯金字塔进行分解也具有劣势,低频层的信息需要被极大地压缩,高频层需要被局部调整,仅使用单个网络无法完成所有任务。如果使用  $n$  个不同的网络来处理不同频带的图像,会大大增加整体网络的复杂度和计算量。

为此,改进传统拉普拉斯金字塔,具体如图 2 所示。对预处理后的 HDR 图像采用 6 层拉普拉斯金字塔进行分解,将保留亮度信息的最低频带图像上采样至原图大小,再将次低频带图像上采样至上一级图像大小,然

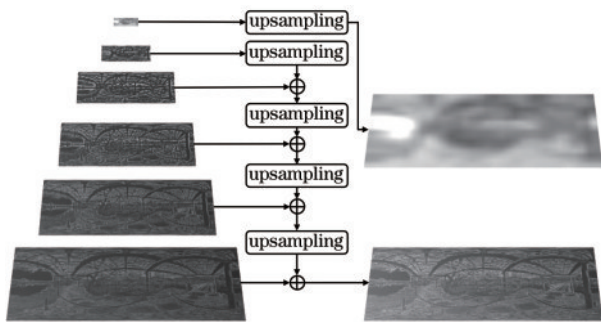


图 2 改进的拉普拉斯金字塔示意图

Fig. 2 Schematic diagram of improved Laplacian pyramid

后与上一级图像融合后继续上采样,直至最高频带。至此,传统的  $n$  层拉普拉斯金字塔减少到 2 层,分别包含了原 HDR 图像中的低频信息和高频信息。

### 2.3 自适应分组卷积模块

全局提取网络和微调网络需要兼具获取图像全局特征和局部特征的能力。Zhang 等<sup>[12]</sup>采用双分支来提取不同的图像信息并通过空洞卷积扩大感受野来提升网络获取全局信息的能力。ExpandNet<sup>[13]</sup>采用了 3 种卷积核大小不同的分支结构来获取不同的图像信息,并且在局部和细节分支中采用了编解码器结构来捕获全局特征。上述方法虽然都能获取全局特征信息,但都存在一些问题。空洞卷积会导致输出图像出现棋盘格伪影,而编解码器结构需要固定大小的输入,这使得网络需要调整输出图像的大小来保证与原图的一致性,可能破坏原有的全局信息。

受 Ioannou 等<sup>[14]</sup>提出的分组卷积启发,提出了一种增强网络获取全局和局部信息能力的自适应分组卷积模块,如图 3 所示。首先,自适应分组卷积模块将具有  $C$  个通道的图像平均分成 3 份,然后分别使用卷积核大小为  $1, 3, k$  的卷积对图像进行卷积运算,通道数以及顺序保持不变,随后按原顺序将 3 组图像连接在一起,再使用卷积核大小为 3 的卷积对其进行卷积运算,最后与输入的  $C$  个通道的图像逐像素相加。1 和 3 是卷积神经网络中常用的卷积核尺寸参数。参数  $k$  根据输入图像的通道数  $C$  计算得出,  $k$  可表示为

$$k = \begin{cases} \lfloor \lg C + b \rfloor, & \lfloor \lg C + b \rfloor \bmod 2 = 1 \\ \lfloor \lg C + b \rfloor + 1, & \lfloor \lg C + b \rfloor \bmod 2 = 0 \end{cases}, \quad (7)$$

式中:  $b=1$ 。式(7)参考了 ECA-Net<sup>[15]</sup>提出的通道注意力机制,有效增强模块的自适应性。由此,AGCM 通过分组卷积减少了参数量,同时借助能自适应调整尺寸的卷积核,保证了卷积运算时不同大小的感受野能够提取出不同的图像特征。

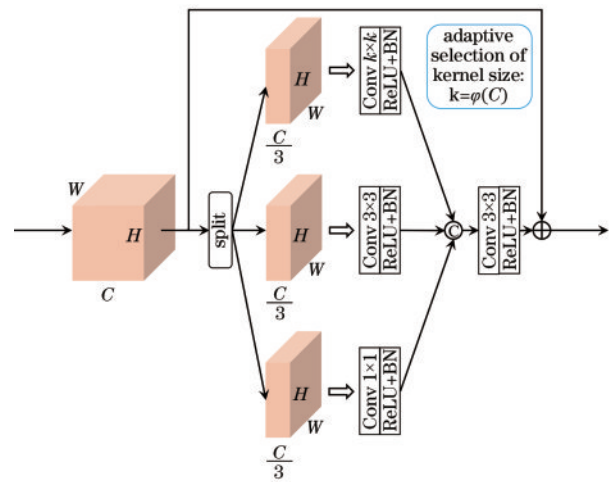


图 3 自适应分组卷积模块

Fig. 3 Adaptive group convolution module



## 2.4 色调映射数据集构建

参考 Hu 等<sup>[16]</sup>和 Yang 等<sup>[17]</sup>提出的方法,首先从网络上收集了大量 HDR 图像(其中大约 80% 的图像来自公开数据集,例如拉瓦尔室内 HDR 数据集<sup>[17]</sup>,其余图像来自免费的摄影图片网站),将其统一剪裁并调整大小,共收集 HDR 图像 1040 张。挑选了 4 种色调映射方法来生成对应的 LDR 图像,包含了 Durand TMO<sup>[18]</sup>、Reinhard TMO<sup>[5]</sup>、Liang 等<sup>[8]</sup>算法以及 Yang 等<sup>[17]</sup>算法,参考图像将从这些图像中产生。鉴于数据集规模较大,人工筛选成本高,构建一个适用于所提的色调映射评价模型是十分必要的。

首先,从 1040 张 HDR 图像中选择 60 张不同主题的高质量 HDR 图像,这些图像的主题分为 3 类,分别为室内场景、人像以及自然场景,每种主题包含 20 张 HDR 图像。然后,采用 4 种算法进行映射,共生成 240 张图像,使用同一张 HDR 图像映射得到的 4 张 LDR 图像为一组。每次实验前,随机调整图像组的顺序,每组图像内部也随机排序,每幅图像不设置观察时间限制,观察结束后受测者对图像进行评分。受测者的观测位置如图 4 所示,图像处于屏幕正中央,受测者正对显示屏中央位置观看图像,与显示屏的距离为 3 倍图像高度。实验采用 5 分制评分<sup>[10]</sup>,5 分为极佳,1 分为糟糕,表 1 为具体的评分标准。受测者在正式实验前需进行小规模(5 组图像)实验,确保其熟悉实验

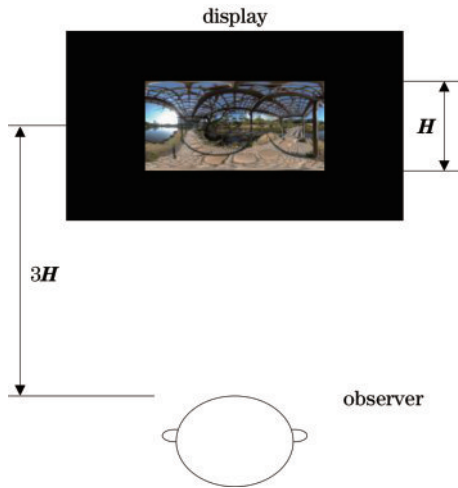


图 4 视觉感知实验设置

Fig. 4 Visual perception experiment setting

表 1 实验评分标准

Table 1 Grading standard of experiment

Grade	Image quality
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

步骤与方法。实验总共有 29<sup>[19]</sup>人参加,其中包括了 21 名男性和 8 名女性,年龄在 22~37 岁之间,平均年龄为 24.7 岁,受测者双眼视力均正常或矫正至正常水平。实验完成后每张图片对应 29 个主观评分,计算平均得分作为图像的平均主观得分(MOS)。

鉴于所提的色调映射评价模型与无参考图像质量评价任务<sup>[20]</sup>非常相似,决定通过建立多种图像特征与 MOS 之间的映射关系来完成评价模型。色调映射算法的目的是在降低亮度范围的同时尽可能保留原有的图像信息,亮度信息、边缘信息、图像细节与图像失真都是需要关注的特征。采用平均亮度值  $L_{\text{mean}}$  代表图像的亮度信息, $L_{\text{mean}}$  可表示为

$$L_{\text{mean}} = \frac{1}{n} \sum_{i=1}^n C(p_i), \quad (8)$$

式中: $C(p_i)$ 为图像中像素点  $p_i$  的亮度值。Tenengrad 函数是一种常用的基于梯度的图像清晰度评价函数<sup>[21]</sup>。其采用 Sobel 算子提取水平和垂直方向上的梯度值,水平和垂直方向上的算子分别为  $\mathbf{G}_x$  和  $\mathbf{G}_y$ 。图像在点  $(x, y)$  处的梯度即为  $S(x, y)$ 。

$$\mathbf{G}_x = \frac{1}{4} \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad \mathbf{G}_y = \frac{1}{4} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}, \quad (9)$$

$$S(x, y) = \sqrt{\mathbf{G}_x \times I(x, y) + \mathbf{G}_y \times I(x, y)}, \quad (10)$$

式中: $I(x, y)$ 为图像像素值。

在图像处理领域,清晰的图像一般被认为具有更明晰的边缘,也就具有更大的梯度函数值,因此将其作为表示边缘信息的特征。

$$I_{\text{tenen}} = \frac{1}{n} \sum_x \sum_y S(x, y)^2. \quad (11)$$

色调映射算法是对 HDR 图像进行亮度压缩的算法,为衡量 LDR 图像所保留的图像细节,采用 Yeganeh 等<sup>[22]</sup>提出的统计自然度模型。

$$I_{\text{nature}} = \frac{1}{K} P_m P_d, \quad (12)$$

式中: $K = \max\{P_m, P_d\}$ ,而  $P_m, P_d$  可表示为

$$P_m = \frac{1}{\sqrt{2\pi} \sigma_m} \exp\left(-\frac{m - \mu_m}{2\sigma_m^2}\right), \quad (13)$$

$$P_d = \frac{(1-d)^{\beta_d-1} d^{\alpha_d-1}}{B(\alpha_d, \beta_d)}, \quad (14)$$

式中: $\mu_m = 115.94$ ;  $\sigma_m = 27.99$ ;  $\alpha_d = 4.4$ ;  $\beta_d = 10.1$ ;  $B(\cdot, \cdot)$ 为 Beta 函数; $m, d$  分别为图像平均亮度和图像亮度的标准差。图像失真一直是图像质量评价领域关注的重点,所提的评价任务中不包含参考图像,选取 BRISQUE<sup>[23]</sup>和 NIQE<sup>[24]</sup>2 种经典的无参考图像质量评价方法作为失真评价指标。

使用多元非线性回归模型来学习从上述多种特征到 MOS 的映射,模型采用 sklearn 包来实现。训练集和测试集按照 8:2 的比例随机抽取,通过计算预测值

和真实值的斯皮尔曼线性相关系数(SROCC)和皮尔逊线性相关系数(PLCC)来判断模型性能,重复实验 100 次取平均值,PLCC 为 0.9177, SROCC 为 0.9168。将训练好的模型用于对剩余所有图像的评分预测,同一组中预测得分最高的图像作为 HDR 图像的 LDR 参考图像,至此完成了数据集的构建。

### 3 算法性能分析

#### 3.1 图像预处理及实现细节

由于构建数据集的图像来源较为复杂,为便于统一处理,将分辨率较高(图像分辨率大于 1024 pixel×512 pixel)的图像裁剪至 1024 pixel×512 pixel,调整分辨率略低于该尺寸的图像至 1024 pixel×512 pixel,删除分辨率过低的图像。数据集中随机抽取 90% 的图像对作为训练集,剩余 10% 的图作为测试集。为进一步增强网络针对不同尺寸图像的泛化能力,将训练样本进行随机下采样,下采样范围为原有尺寸的 50%~100%,每张图像会生成 9 张子图像,与原图一起输入网络,共 9360 个训练图像样本。子网络的初始输入图像为单通道灰度图像,以此降低网络运算量,提升训练速度。网络最后的输出图像亦为灰度图像,经过颜色恢复<sup>[25]</sup>得到最终图像,具体可表示为

$$I_{\text{color}} = \left( \frac{I_f}{h} \right)^s l, \quad (15)$$

式中: $I_f$ 为微调网络的输出图像; $h$ 为预处理后的灰度图像; $l$ 为 HDR 图像; $s=0.6$ 。表 2 为构建算法时计算机的系统配置。训练框架为 TensorFlow 1.11.0,训练优化器为 Adam,学习率为 0.001。训练过程中先训练全局提取网络和局部调整网络,然后再结合微调网络进行端到端的训练。

表 2 系统配置

Table 2 System configuration

Software and hardware configuration	Parameter
Operating system	Windows 10 Professional
CPU	AMD Ryzen 7 4800H
GPU	NVIDIA GeForce GTX 2060
RAM	16 GB
Python version	3.6

#### 3.2 客观性能分析

为验证色调映射算法的客观性能,采用 3 个评价指标,分别为 TMQI<sup>[22]</sup>、FSITM<sup>[26]</sup>、HDR-VDP2<sup>[27]</sup>。TMQI 是一个著名的色调映射算法评价指标,其综合了结构相似度和统计自然度来对映射后的 LDR 图像进行评价,其值越接近 1,说明 LDR 图像质量越好。FSITM 通过计算不同通道间 HDR 图像与对应 LDR 图像的相似度来评价色调映射算法的结果,其值越接近 1 则说明对应图像间的特征越相似。HDR-VDP2

是一种模拟 HVS 解剖结构的视觉指标,也可以用于度量色调映射算法的性能,其值越大,则说明算法性能越好。

选取 4 种 HDR 图像色调映射领域先进的方法来进行对比。Reinhard TMO 通过考察每个像素不同尺寸的邻域,调整亮度不适宜的区域,以此来完成色调映射。Durand TMO 是一个经典的色调映射算法,其将 HDR 图像分解为基础层和细节层,通过双边滤波器将基础层进行对比度压缩,再和细节层进行融合。该算法通过单个参数控制映射效果,本文中该参数采用文献[18]中的默认值。Liang 等提出了一种更加先进的多尺度色调映射方案,所得到的图像具有良好的视觉感知效果且产生的伪影较少。Yang 等提出了一种通过频带高低划分 HDR 图像层次并以此进行分解,采用多个卷积神经网络来完成色调映射任务的算法,能够保留更多的图像细节。

表 3 为上述及本文方法在测试集上的表现,包括 Durand TMO、Reinhard TMO、Liang 的方法、Yang 的方法和本文方法的平均 TMQI 值、平均 FSITM 值以及平均 HDR-VDP2 值。结果表明,与对比算法相比,测试集中的图像经本文算法处理后,在平均 TMQI 值和平均 HDR-VDP2 值 2 个评价指标上表现最佳。FSITM 分别计算各个通道间(RGB 三通道)LDR 图像与对应 HDR 图像的相似性,以此来判断算法的性能。本文方法为降低算法的计算开销,在预处理时将 HDR 图像进行了灰度化,输出时采用颜色恢复的方法得到彩色图像,此过程中可能产生了轻微的颜色偏移,导致本文方法的 FSITM 指标低于部分对比方法。

表 3 不同方法的客观指标对比

Table 3 Comparison of objective indexes of different methods

Method	TMQI	FSITM	HDR-VDP2
Durand <sup>[18]</sup>	0.864	0.784	36.57
Reinhard <sup>[5]</sup>	0.791	<b>0.848</b>	41.89
Liang <sup>[8]</sup>	0.886	0.792	43.67
Yang <sup>[17]</sup>	0.872	0.824	38.16
Proposed	<b>0.905</b>	0.814	<b>43.96</b>

图 5 为本文算法与其他算法的可视化对比结果。Durand TMO 在映射时能够保留丰富的图像细节,但其得到的 LDR 图像色彩过于鲜艳且伪影较多,导致图像整体质量不佳。Reinhard TMO 得到的 LDR 图像整体亮度偏低且图像细节被阴影区域遮蔽较多,虽然存在的伪影较少,但这与本文对于色调映射算法的预期差距较大。Liang 等的方法在处理 HDR 图像中的高亮区域和低亮区域时存在缺陷,难以在上述两种区域中保留图像细节。例如,在第 2 组图像中,Liang 等的方法所得图像整体亮度较低,玻璃杯底部区域接近黑色,导致图像细节很难被人眼获取。该方法所得人像整体亮度较高,虽然从全局看图像较为完整,但是面部的皱





图 5 算法效果对比图

Fig. 5 Comparison of the algorithms effect

纹、胡须等细节较难辨认。Yang 等的方法得到的 LDR 图像色彩鲜艳且较为自然,局部细节也较为清晰,映射效果较好,但在高亮区域的表现一般,无法清楚展示图像细节。第 1 组和第 4 组的局部放大图较好地证明了这一点,过高的亮度导致图像整体偏向于白色,在显示时会使人产生眼部不适。本文方法得到的 LDR 图像色彩鲜艳自然,局部细节清晰,与对比方法相比,在高亮区域和低亮区域所获取到的信息最为丰富,是一种更为优越的 HDR 图像色调映射算法。

表 4 为本文方法与对比方法的平均运行时间对比

表 4 算法平均运行时间对比

Table 4 Comparison of average running time of algorithms

Method	Time / s
Durand <sup>[18]</sup>	0.362
Reinhard <sup>[5]</sup>	0.387
Liang <sup>[8]</sup>	1.691
Yang <sup>[17]</sup>	0.082
Proposed	0.080

比。将测试集中的 HDR 图像进行映射,并计算出每种方法处理单张图像平均所需的时间。其中, Yang 等的方法和本文方法是面向图形处理器设计的卷积神经网络,所以表 4 中展示的是这两种方法在图形处理器上的运行时间,其余方法均是在中央处理器上的运行时间。本文方法和 Yang 等的方法处理单张图像的时间相当,但本文方法的参数量为 67.1 KB, Yang 等的方法参数量为 132.7 KB, 本文方法在参数量上具有一定的优势。

### 3.3 主观性能分析

主观感知实验是比较不同 HDR 图像色调映射算法结果的有效方法。从测试集中随机抽取 30 个编号并选取编号对应的 LDR 图像(由 4 种对比方法和本文方法生成), 总共 150 幅图像。实验过程中, 每组图像出现的顺序随机, 组内图像出现的顺序亦随机。实验设置和评分标准参考图 4、表 1。该实验总共有 8 名女性和 16 名男性参与, 受测者年龄在 22~37 岁之间, 均值为 24.8 岁。所有受测者视力水

平均满足实验要求。练习实验参照 2.4 节中的小规模实验。

实验结束后采用 SPSS 软件对实验结果进行分析,软件版本号为 20.0。分析结果如表 5 所示。表 5 中:df 为自由度;F 为组间均方差除以组内均方差的比值;Sig. 为显著性。分析结果表明,色调映射算法、图像内容及二者的交叉作用均是 LDR 图像质量的重要影响因素( $p < 0.001$ )。HDR 图像经本文算法处理后,所得 LDR 图像质量最佳,如图 6 所示,其 MOS 为 3.65,而次优方法(Yang)的 MOS 为 3.30,归一化后本文方法相较于次优方法的 MOS 上升了 15%。实验结果表明,本文方法能较好地将 HDR 图像映射为 LDR 图像,且有明显优势。

表 5 各变量因素对图像质量评分的 ANOVA 分析结果  
Table 5 Results of ANOVA analysis for different factors on image quality score

Factor	Image quality score		
	df	F	Sig.
Method	4	416.890	0
Image content	29	23.183	0
Cross interaction	116	5.706	0

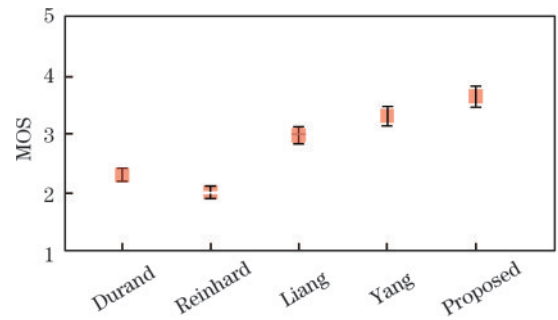


图 6 各组 LDR 图像的质量评分平均值

Fig. 6 Mean scores of image quality of LDR images in each group

### 3.4 消融实验

为了验证提出的自适应分组卷积模块在两个子网络中的有效性,设计了相应的消融实验。消融实验采用的子网络结构如图 7 所示,除图 7 中的模块外,其余部分均保持一致。采用客观性能分析中使用的 3 个指标来对照分析不同网络结构下整体网络的性能。图 7 中的网络结构以及本文结构在测试集上的表现如表 6 所示,结构 A、B、C 分别对应图 7(a)、图 7(b)、图 7(c)。实验结果表明,所提的自适应分组卷积模块能够提升网络获取局部和全局特征的能力,提高特征利用率,以此提高整体网络的性能。

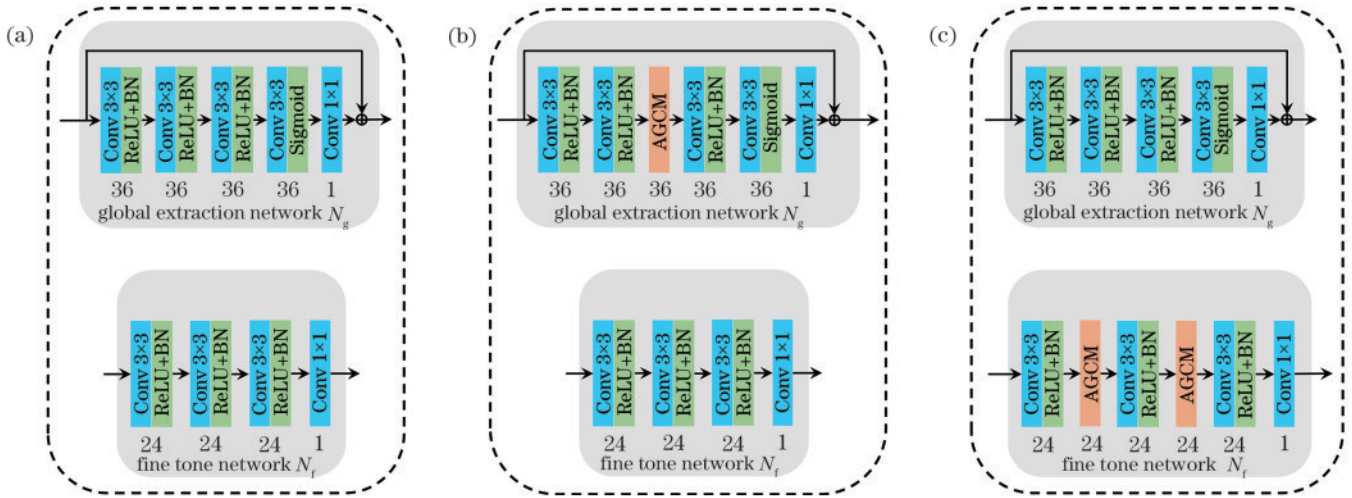


图 7 消融实验的网络结构。(a)无 AGCM 模块;(b)仅全局提取网络包含 AGCM 模块;(c)仅微调网络包含 AGCM 模块

Fig. 7 Network structure of ablation experiment. (a) Without AGCM module; (b) only global extraction network contains AGCM module; (c) only fine tone network contains AGCM module

表 6 消融实验结果对比

Table 6 Comparison of the results of ablation experiment

Structure	TMQI	FSITM	HDR-VDP2
Structure A	0.876	0.790	43.01
Structure B	0.893	0.809	43.36
Structure C	0.890	<b>0.823</b>	43.46
Proposed	<b>0.905</b>	0.814	<b>43.96</b>

为了更好地探究网络中各个子网络的作用,对各个子网络也进行消融实验,实验结果如图 8 所示。为使得该示意图能够更好地展示消融实验的结果,对所有的输出图像都进行颜色恢复。图 8(a)、图 8(b)、图 8(c)分别为仅包含单个子网络的实验结果,该结果表明:单个子网络的实际功能与本文对网络的预期保持一致,能够较好地完成各自所负责的信息提取和图



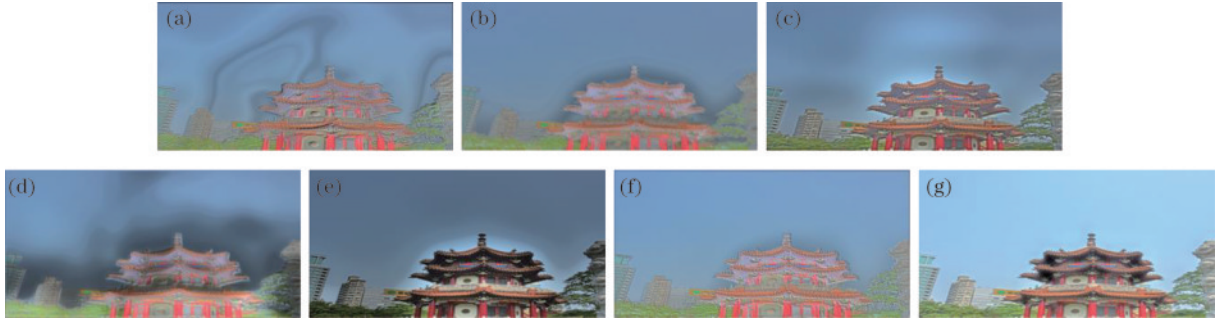


图 8 子网络消融实验。(a)仅包含全局提取网络;(b)仅包含局部调整网络;(c)仅包含微调网络;(d)不包含全局提取网络;(e)不包含局部调整网络;(f)不包含微调网络;(g)完整的网络

Fig. 8 Ablation experiment of subnets. (a) Only contain global extraction network; (b) only contain local adjustment network; (c) only contain fine tone network; (d) without global extraction network; (e) without local adjustment network; (f) without fine tone network; (g) whole network

像优化任务。图 8(d)、图 8(e)、图 8(f)分别为依次去除单个子网络后的输出图像。从主观来看,图 8(e)虽然缺失了局部调整网络,但所得图像最为接近完整网络的输出图像,说明全局网络和微调网络在整体中所占据的影响权重较大。图 8(f)为颜色恢复后的重建图像,该图像具备基本的亮度信息和纹理信息,但与最优结果的差距较大,表明所提的微调网络是色调映射算法中至关重要的一环。

为探究损失函数中各个子函数的系数对算法性能的影响,采用不同的系数组合进行实验。初始时将

式(5)中的 $\alpha$ 和 $\beta$ 分别设置为0.1和0.9,并将该参数组合标注为组合1,其余参数均设置为1且在 $\alpha$ 和 $\beta$ 的探究实验中保持不变。此后每次实验, $\alpha$ 递增0.1, $\beta$ 递减0.1,组合编号递增1,具体实验结果的数值如图9所示。由图9可知,当 $\alpha$ 和 $\beta$ 均取0.5时,网络表现最佳。然后,设置 $\lambda$ 为0.1,每次实验依次递增0.1,直至递增至0.9,其余参数均保持不变。经实验发现,当 $\lambda$ 为0.2时,算法性能最佳。式(6)中的2个参数按照 $\alpha$ 和 $\beta$ 的设置方式进行实验,当 $\gamma=0.6,\eta=0.4$ 时,算法表现最佳。

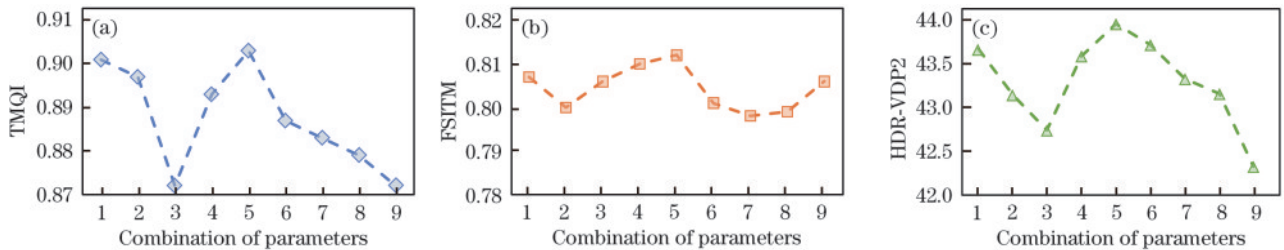


图 9  $\alpha$ 和 $\beta$ 的参数实验结果。(a)参数组合对应的 TMQI 数值;(b)参数组合对应的 FSITM 数值;(c)参数组合对应的 HDR-VDP2 数值

Fig. 9 Parametric experimental results of  $\alpha$  and  $\beta$ . (a) Value of TMQI corresponding to the parameter combination; (b) value of FSITM corresponding to the parameter combination; (c) value of HDR-VDP2 corresponding to the parameter combination

## 4 结 论

提出了一种基于改进拉普拉斯金字塔的 HDR 图像色调映射算法。算法借助拉普拉斯金字塔将 HDR 图像分解为高频层和低频层,分别使用局部调整网络和全局提取网络进行处理,融合后得到感知效果优越的 LDR 图像。该算法针对单个子网络难以同时获取图像局部和全局特征的问题,设计了自适应分组卷积模块来提升网络的特征提取能力。此外,在数据集构建过程中引入主观感知实验,优化了构建的数据集。经过论述,所提算法得到的 LDR 图像色彩鲜明且自然,保留的图像细节也较为丰富,从客观上看,本文算法的 TMQI、HDR-VDP2 均有优势,从主观上看,本文

方法具有更加优越的主观感知效果。未来将进一步优化网络结构,使其在参数量更小的情况下具备更好的性能。

## 参 考 文 献

[1] 汤璐, 卢荣胜, 史艳琼, 等. 基于 YCbCr 空间融合的高动态范围成像方法[J]. 激光与光电子学进展, 2022, 59(14): 1415029.  
Tang L, Lu R S, Shi Y Q, et al. High dynamic range imaging method based on YCbCr color space fusion[J]. Laser & Optoelectronics Progress, 2022, 59(14): 1415029.

[2] 赵鹏, 余新, 杨亚涛. 基于光转向的高动态范围激光显示[J]. 激光与光电子学进展, 2022, 59(5): 0512002.  
Zhao P, Yu X, Yang Y T. High dynamic range laser



- display based on light steering[J]. *Laser & Optoelectronics Progress*, 2022, 59(5): 0512002.
- [3] Lee D H, Fan M, Kim S W, et al. High dynamic range image tone mapping based on asymmetric model of retinal adaptation[J]. *Signal Processing: Image Communication*, 2018, 68: 120-128.
- [4] Khan I R, Rahardja S, Khan M M, et al. A tone-mapping technique based on histogram using a sensitivity model of the human visual system[J]. *IEEE Transactions on Industrial Electronics*, 2018, 65(4): 3469-3479.
- [5] Reinhard E, Stark M, Shirley P, et al. Photographic tone reproduction for digital images[C]//*Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, July 23-26, 2002, San Antonio, Texas. New York: ACM Press, 2002: 267-276.
- [6] Montulet R, Briassouli A. Deep learning for robust end-to-end tone mapping[EB/OL]. [2022-11-05]. [https://cris.maastrichtuniversity.nl/ws/portalfiles/portal/79731329/Briassouli\\_2019\\_Deep\\_learning\\_for\\_robust\\_end\\_to\\_end.pdf](https://cris.maastrichtuniversity.nl/ws/portalfiles/portal/79731329/Briassouli_2019_Deep_learning_for_robust_end_to_end.pdf).
- [7] Cao X, Lai K, Yanushkevich S N, et al. Adversarial and adaptive tone mapping operator for high dynamic range images[C]//*2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, December 1-4, 2020, Canberra, ACT, Australia. New York: IEEE Press, 2021: 1814-1821.
- [8] Liang Z T, Xu J, Zhang D, et al. A hybrid 11-10 layer decomposition model for tone mapping[C]//*2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 4758-4766.
- [9] Jin C, Deng L J, Huang T Z, et al. Laplacian pyramid networks: a new approach for multispectral pansharpening [J]. *Information Fusion*, 2022, 78: 158-170.
- [10] Di Gregorio F, Trajkovic J, Roperti C, et al. Tuning alpha rhythms to shape conscious visual perception[J]. *Current Biology*, 2022, 32(5): 988-998.e6.
- [11] Wang Y, Karimi H A. Perceptual loss function for generating high-resolution climate data[J]. *Applied Computing and Intelligence*, 2022, 2(2): 152-172.
- [12] Zhang Z M, Han C, He S F, et al. Deep binocular tone mapping[J]. *The Visual Computer*, 2019, 35(6): 997-1011.
- [13] Marnierides D, Bashford-Rogers T, Hatchett J, et al. ExpandNet: a deep convolutional neural network for high dynamic range expansion from low dynamic range content [J]. *Computer Graphics Forum*, 2018, 37(2): 37-49.
- [14] Ioannou Y, Robertson D, Cipolla R, et al. Deep roots: improving CNN efficiency with hierarchical filter groups [C]//*2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 5977-5986.
- [15] Wang Q L, Wu B G, Zhu P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks [C]//*2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11531-11539.
- [16] Hu L T, Chen H J, Allebach J P. Joint multi-scale tone mapping and denoising for HDR image enhancement[C]//*2022 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW)*, January 4-8, 2022, Waikoloa, HI, USA. New York: IEEE Press, 2022: 729-738.
- [17] Yang J, Liu Z, Lin M, et al. Deep reformulated laplacian tone mapping[EB/OL]. (2021-01-31)[2022-11-09]. <https://arxiv.org/abs/2102.00348v1>.
- [18] Durand F, Dorsey J. Fast bilateral filtering for the display of high-dynamic-range images[J]. *ACM Transactions on Graphics*, 2002, 21(3): 257-266.
- [19] Yang Z Y, Dai Q H, Zhang J S. Visual perception driven collage synthesis[J]. *Computational Visual Media*, 2022, 8(1): 79-91.
- [20] Chen B, Zhu L, Kong C, et al. No-reference image quality assessment by hallucinating pristine features[J]. *IEEE Transactions on Image Processing*, 2022(31): 6139-6151.
- [21] Lü M N, Gan H, Liu X, et al. Selection of optimal focusing evaluation function based on sparse microscopic images[J]. *Journal of Physics: Conference Series*, 2022, 33: 6139-6151.
- [22] Yeganeh H, Wang Z. Objective quality assessment of tone-mapped images[J]. *IEEE Transactions on Image Processing*, 2013, 22(2): 657-667.
- [23] Mittal A, Moorthy A K, Bovik A C. No-reference image quality assessment in the spatial domain[J]. *IEEE Transactions on Image Processing*, 2012, 21(12): 4695-4708.
- [24] Mittal A, Soundararajan R, Bovik A C. Making a “completely blind” image quality analyzer[J]. *IEEE Signal Processing Letters*, 2013, 20(3): 209-212.
- [25] Fattal R, Lischinski D, Werman M. Gradient domain high dynamic range compression[C]//*Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, July 23-26, 2002, San Antonio, Texas. New York: ACM Press, 2002: 249-256.
- [26] Ziaei Nafchi H, Shahkolaei A, Farrahi Moghaddam R, et al. FSITM: a feature similarity index for tone-mapped images[J]. *IEEE Signal Processing Letters*, 2015, 22(8): 1026-1029.
- [27] 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.