

渐进式多尺度特征级联融合颜色恒常性算法

杨泽鹏, 解凯*, 李桐

北京印刷学院信息工程学院, 北京 102600

摘要 颜色恒常性是实现目标检测、三维物体重建、自动驾驶等计算机视觉任务的重要前提。为充分利用图像中不同尺度的特征信息估计光源,提出渐进式多尺度特征级联融合颜色恒常性算法,通过三个卷积网络分支从不同尺度提取图像中的特征信息,通过特征融合得到更加丰富的特征信息,通过级联方式将图像中的浅层边缘信息和深层细粒度特征信息进行融合,提高了颜色恒常性算法的精确性。渐进式网络结构基于加权累计角度误差损失函数提高了算法在面对极端场景光照下光源估计的鲁棒性。在重处理的 ColorChecker 和 NUS-8 数据集上的实验结果表明,本文算法在各项评价指标上均优于目前的颜色恒常性算法,可应用于需要进行颜色恒常性预处理的其他计算机视觉任务。

关键词 视觉光学; 颜色恒常性; 光源估计; 多尺度; 特征融合; 加权累加损失

中图分类号 O439

文献标志码 A

doi: 10.3788/AOS202242.0533002

Progressive Multi-Scale Feature Cascade Fusion Color Constancy Algorithm

Yang Zepeng, Xie Kai*, Li Tong

School of Information Engineering, Beijing Institute of Graphic Communication, Beijing 102600, China

Abstract Color constancy is an important prerequisite for computer vision tasks such as object detection, three-dimensional object reconstruction, and automatic driving. In order to make full use of the feature information of different scales in the image to estimate the light source, a progressive multi-scale feature cascade fusion color constancy algorithm is proposed. The feature information in the image is extracted from different scales by three convolution network branches to fuse and get more abundant feature information. By cascading the shallow edge information and the deep fine-grained feature information in the image, the accuracy of the color constancy algorithm is improved. The progressive network structure improves the robustness of the algorithm for the light source estimation in extreme scenes by weighted cumulative angle error loss function. Experimental results on the reprocessed ColorChecker and NUS-8 datasets show that the proposed algorithm outperforms the current color constancy algorithm in terms of various evaluation indexes, and can be applied to other computer vision tasks requiring color constancy preprocessing.

Key words visual optics; color constancy; illumination estimation; multi-scale; feature fusion; weighted cumulative loss

1 引言

图像或视频在成像过程中受到场景光照、物体表面反射率,以及成像传感器的响应函数等多方面的影响。同一物体在不同的场景光源照射下会呈现出不同的颜色,颜色恒常性算法的目的是消除场景

中光照对物体颜色的影响,使物体无论在何种光源照射下都可以呈现出物体本身的颜色,即物体在标准白光下的颜色,这为安防、无人驾驶以及目标追踪等视觉任务^[1-3]提供稳定的特征信息,对于计算机视觉或者模式识别具有重要的理论意义和实际意义。

颜色恒常性是该领域研究人员的一个长期的热

收稿日期: 2021-08-17; 修回日期: 2021-09-13; 录用日期: 2021-09-23

基金项目: 北京市教委科研项目(KM201810015011)

通信作者: *2596898130@qq.com

点问题。目前颜色恒常性算法可以分为基于统计学的方法和基于学习的方法。基于统计学的方法利用图像的低阶特征信息,根据图像的统计特征及物理特性对光源进行估计,之后移除场景光源,使图像恢复到标准白光下的颜色。基于统计学的算法主要包括灰色世界法(grey-world)、灰色阴影法(shades-of-grey)、灰色边缘法(grey-edge)^[4]等,这类算法的运算速度快,但对于不同的场景,这类算法具有局限性。

基于学习的颜色恒常性算法可以分为基于手工提取特征的传统学习方法和基于深度学习的方法。基于手工提取特征的颜色恒常性算法通过对大量已知场景光照的图像样本进行统计和学习,根据先验知识估计场景光源,这类方法需要有目的地提取目标信息,非常依赖特定场景光源的特异性以及设计特征的专家的经验。基于传统学习的颜色恒常性算法主要包括色域映射法^[5]和贝叶斯推理算法^[6]。

深度学习可以从大数据或高质量的小样本数据中先学习简单的特征,之后逐渐学习更为复杂、抽象的特征,不依赖于人工提取特征工程。与基于传统学习的颜色恒常性算法相比,基于深度学习的方法进一步提高了颜色恒常性算法的精确性与稳定性。Bianco 等^[7]将图像分割成大小相等的块,通过 3 层卷积网络结构估计出每个图像块的光源,经过池化后得到全局光源。Qiu 等^[8]将深度学习充分提取图像特征的能力与基于假设的模型可解释性有效结合,通过特征图重载单元为图像中的图像块施加权重,以准确估计出场景光源。Afifi^[9]将图像中的语义信息和颜色空间特征信息相结合,通过卷积神经网络估计出场景光源,进一步提高了颜色恒常性算法的精确性。Afifi 等^[10]的研究表明采用颜色恒常性算法的预处理可以提高图像语义分割的精确性,证明了颜色恒常性预处理对于其他计算机视觉任务的重要性。Laakom 等^[11]证明了预训练网络对颜色恒常性算法泛化能力的重要性,提出无监督以及半监督的自编码网络,用于解决颜色恒常性算法中数据量不足的问题。对抗生成网络通过博弈的方式生成图像,是近年来在无监督领域具有广阔前景的方法之一。Das 等^[12-13]使用对抗网络直接生成标准白光下的图像,通过比较各种对抗生成网络模型在颜色恒常性算法上的性能,分析了将对抗生成网络应用于颜色恒常性算法的优势以及局限性。Hu 等^[14]提出置信度加权网络用于估计场景光源,基于置信度加权并充分利用图像中对光源估计有用的特征估

计出光源,这加快了颜色恒常性算法的运算速度,提高了颜色恒常性算法的精确性。

目前基于深度学习的颜色恒常性算法大多是由分类或者检测的卷积网络结构改进而来。这些基于分类或者检测的卷积网络结构^[15-20]由卷积层逐层串联,是特征递进式的学习算法。神经网络中浅层的卷积层神经元直接从输入数据中学习一些低层次的简单特征,而深层神经元则基于已学到的浅层特征学习更高级的特征,从计算机的角度学习深层的语义信息。但由于颜色的特殊性,将串联式的卷积网络结构应用于颜色恒常性存在着一定的局限性。其主要原因有:

1) 由于颜色的特殊性,图像中的浅层边缘特征以及更深层次的细粒度特征都可以为光源估计提供更多的信息,在对光源进行估计时,需要同时利用图像中物体的浅层边缘纹理信息以及更高级的细粒度特征信息提高场景光源估计的准确性。串联式的卷积网络结构的网络层数过少时,只能提取到图像中的边缘特征信息,忽略了更深层的细粒度特征,网络层数过多时,网络更多关注图像中的细粒度特征信息,忽略浅层的边缘特征信息,二者无法同时兼顾。

2) 图像中的颜色特征受到场景光照以及设备成像传感器的影响,同时自然界场景光源复杂丰富,应用于颜色恒常性算法的串联式卷积网络结构在面对同一成像设备下的单一场景光照时,可以准确估计出场景光源,但在面对极端场景光源或者复杂场景光源时无法准确估计出场景光源,复杂的场景光源使网络模型的鲁棒性较差。

为了解决这些问题,本文提出基于渐进式多尺度特征级联融合颜色恒常性算法。本文通过三个卷积网络分支从不同尺度提取图像中的特征信息,通过特征融合得到更加丰富的特征信息,之后通过级联方式将浅层的边缘特征信息与深层的细粒度特征信息相结合估计光源,以进一步提高颜色恒常性算法的精确性。本文还提出渐进式网络结构,通过加权累计角度误差损失函数实现了由粗到细的网络优化,解决了颜色恒常性算法在面对复杂场景光源或者极端场景光源时无法准确估计出光源的问题,提高了颜色恒常性算法在面对复杂场景光照环境以及极端光照环境时算法的鲁棒性以及稳定性。与目前现有的颜色恒常性算法相比,渐进式网络结构在面对复杂场景光源以及极端场景光照时依然可以准确估计出场景光源。

2 渐进式多尺度特征级联融合颜色恒常性算法

为了进一步提高颜色恒常性算法的精确性、解决颜色恒常性算法在面对复杂场景光源以及极端光照环境时无法准确估计出场景光源的问题,本文提出渐进式多尺度特征级联融合颜色恒常性算法。该算法的网络结构如图 1 所示。本文算法通过三个卷积网络分支从不同尺度提取图像中的特征信息,通过特征融合得到更加丰富的特征信息,之后使用级联方式将浅层的边缘特征信息和深层细粒度特征信息进行融合并估计光源,这提高了颜色恒常性算法的精确性。渐进式网络结构基于加权累计角度误差损失函数,提高了颜色恒常性算法在

面对复杂场景光源以及极端光源环境时的鲁棒性和稳定性,使颜色恒常性算法在面对极端光照环境时也可以准确估计出场景光源。算法的整体流程为:1)对数据集进行预处理;2)将预处理后的图像输入到渐进式多尺度特征级联融合网络结构中,通过三个卷积网络结构分支提取出不同尺度的特征信息,通过特征融合得到更加丰富的特征信息;3)通过级联方式将浅层的边缘特征信息与深层的细粒度特征信息进行融合;4)对融合后的特征进行置信度加权,充分利用可以为光源估计提供更多信息的特征以准确地估计出光源;5)使用加权累计角度误差损失函数对网络进行训练;6)对图像进行校正,将估计出的场景光源进行移除,得到标准白光下的图像。

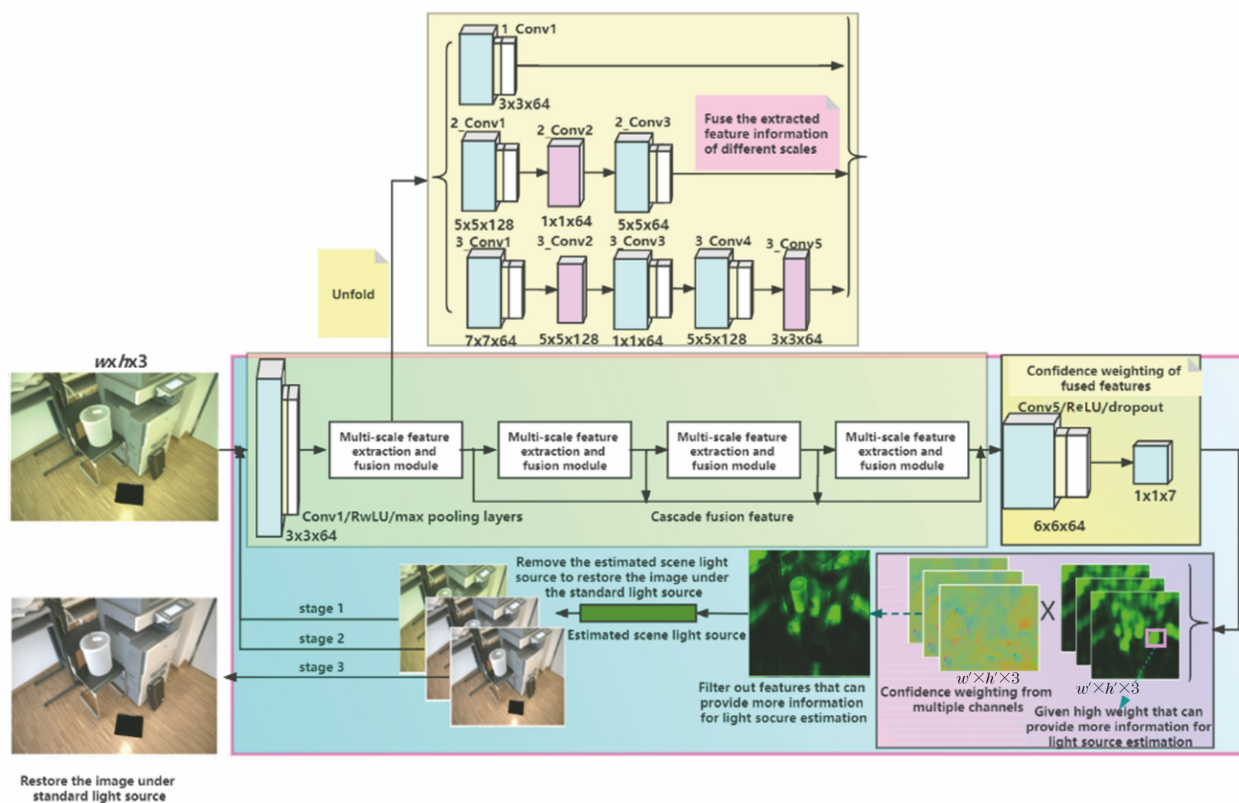


图 1 网络结构图

Fig. 1 Network structure

2.1 多尺度特征级联融合

利用神经网络中不同大小的卷积核可以提取到不同的特征信息,通过特征融合可以得到更加丰富的特征信息。同时,利用神经网络中浅层的神经元可以提取出图像中的边缘特征信息,深层的神经元基于已学到的浅层边缘特征继续学习更高级的细粒度特征信息。利用边缘特征可以学习到最显著的外观表现,但是一些细节特征容易被

忽略,而细粒度特征则含图像中更多的细节特征信息。目前应用于颜色恒常性算法的网络结构大多是由经典网络结构改进而来,例如, Bianco 等^[7]使用三层的卷积神经网络估计光源;Hu 等^[14]使用 AlexNet 和 Squeeze 网络结构估计光源。不同于以往使用串联式的卷积网络结构从图像中提取特征对场景光源进行估计,本文算法通过三个卷积网络分支充分提取图像中不同尺度的特征信息,通

过特征融合得到更加丰富的特征信息,通过级联方式将图像中浅层的边缘特征信息与深层的高级细粒度特征信息相结合以估计光源,从而进一步提高颜色恒常性算法的精确性。

渐进式多尺度特征级联融合网络结构包含 4 个多尺度特征提取融合模块,每个多尺度特征提取融合模块包含三个卷积分支,各个卷积分支所包含的网络层数和卷积核的个数如表 1 所示。使用最近邻插值方法对卷积分支 2 与卷积分支 3 提取到的特征进行上采样,使特征大小与卷积分支 1 所提取的特征大小相同,之后使用 concat 的方法对三个卷积分支的通道进行拼接,对三个卷积分支提取的特征进行融合。三个卷积网络分支分别从不同尺度提取图像中的特征信息并将其进行融合,得到更加丰富的特征信息。为了同时利用图像中浅层边缘特征信息与深层细粒度特征信息估计光源,对前三个多尺度特征提取融合模块提取到的特征进行缩放,使其与第 4 个多尺度特征提取融合模块提取到的特征具有相同的尺寸,将 4 个多尺度特征提取融合模块提取到的特征从通道维度进行级联融合。在两个公开数据集上的实验结果表明,渐进式多尺度特征级联融合颜色恒常性算法通过三个卷积网络分支将不同尺度图像中的特征信息进行融合,得到更加丰富的特征信息,通过级联方式将图像中的浅层边缘特征信息和深层细粒度特征信息进行融合并估计光源,这提高了颜色恒常性算法的精确性。

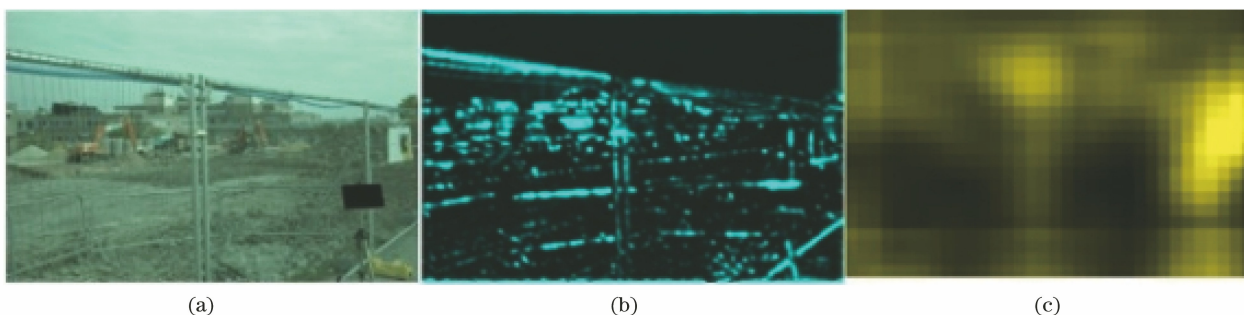


图 2 不同网络结构提取到的特征。(a)输入图像;(b)本文算法提取的特征;(c)使用串联式卷积网络提取到的特征

Fig. 2 Features extracted from different network structures. (a) Input image; (b) features extracted by algorithm in this paper; (c) features extracted by series convolution network

表 2 卷积支路网络层数

Table 2 Layer numbers of convolution branch network

Branch	Conv1	Conv2	Conv3	Conv4	Conv5	Conv6	Conv7
1	$3 \times 3 \times 64$	—	—	—	—	—	—
2	$5 \times 5 \times 128$	$1 \times 1 \times 64$	$5 \times 5 \times 64$	—	—	—	—
3	$7 \times 7 \times 64$	$5 \times 5 \times 128$	$1 \times 1 \times 64$	$5 \times 5 \times 128$	$3 \times 3 \times 64$	—	—
4	$7 \times 7 \times 64$	$5 \times 5 \times 128$	$3 \times 3 \times 64$	$3 \times 3 \times 128$	$1 \times 1 \times 64$	$3 \times 3 \times 128$	$1 \times 1 \times 64$

表 1 各个卷积分支网络结构

Table 1 Structures of convolution branch networks

Branch	Conv1	Conv2	Conv3	Conv4	Conv5
1	$3 \times 3 \times 64$	—	—	—	—
2	$5 \times 5 \times 128$	$1 \times 1 \times 64$	$5 \times 5 \times 64$	—	—
3	$7 \times 7 \times 64$	$5 \times 5 \times 128$	$1 \times 1 \times 64$	$5 \times 5 \times 128$	$3 \times 3 \times 64$

为了证明多尺度特征级联融合网络结构相比于以往串联递进式网络结构的有效性,在实验过程中将多尺度特征级联融合网络中的三个卷积层分支进行串联,提取图像中的特征,将其与多尺度级联特征融合网络提取到的特征进行对比,并对提取到的特征进行可视化,如图 2 所示。图 2(a)表示输入的图像,图 2(b)表示使用多路径特征融合网络提取到的特征信息,图 2(c)表示使用串联式卷积网络提取到的特征信息。从图 2 中可以看出,与单一串联式的网络结构相比,多尺度级联特征融合网络结构提取到的特征更加丰富,可以在提取到深层细粒度特征的同时提取图像中浅层边缘特征信息。实验结果表明,多尺度级联特征融合网络结构可以通过将不同尺度的特征信息进行融合得到更加丰富的特征信息,并通过级联方式将图像中浅层边缘纹理特征信息与深层细粒度特征信息相结合以对光源进行估计,这进一步提高颜色恒常性算法的精确性。

在实验过程中分别对 2 个卷积支路、3 个卷积支路、4 个卷积支路进行了训练及测试,各个卷积支路中卷积层数量以及卷积核的数量如表 2 所示。

不同卷积支路在重处理的 ColorChecker 数据集上的角度误差均值 (mean)、中位数 (med)、三均值 (tri)、最优 25% 均值 (best25%)、最差 25% 均值 (worst25%) 和 95th 百分位数 (95th) 6 个评价指标的实验结果如图 3 所示。实验结果表明,与 2 个卷积支路相比,3 个卷积支路在 6 个评价指标上分别提升了 19.9%, 22.1%, 32.0%, 30.5%, 28.7%, 19.5%。3 个卷积支路提取的特征更加充分,更加有利于准确地估计出场景光源。4 个卷积支路在角度误差的 6 个评价指标上取得了与 3 个卷积支路的网络结构相近的结果,分析其原因是:3 个卷积支路网络结构已经可以充分提取图像中的特征信息并进行光源估计。因此,本文算法采用三个卷积支路的网络结构估计光源。

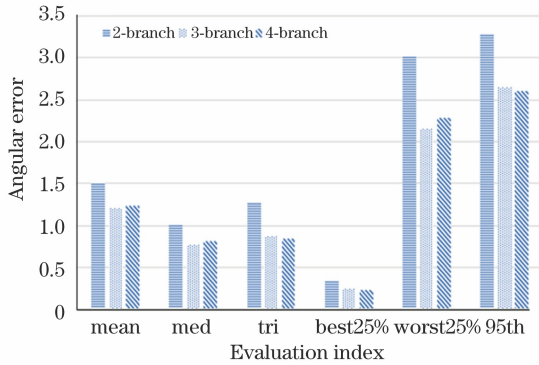


图 3 不同卷积分支数对应的角度误差值

Fig. 3 Angular error of different convolutional branch number

2.2 渐进式网络结构

光源估计任务不同于计算机视觉中的分类或者分割任务,由于光源颜色的特殊性,目前的颜色恒常性算法在面对单一场景光照时可以达到较高的准确性,但在面对极端光照(强光照或者多种颜色光源混合下的光照)时无法准确估计出场景光源,场景光源的复杂性会影响颜色恒常性算法的鲁棒性以及稳定性,本文统计了目前颜色恒常性算法中角度误差的 best25% 以及 worst25% 评价指标值,如图 4 所示,其中横坐标表示颜色恒常性算法,纵坐标表示角度误差值。best25% 代表算法在面对单一场景光源时算法的精确性, worst25% 代表颜色恒常性算法在面对复杂场景光源和极端光照环境时算法的精确性和泛化性,其值越大说明颜色恒常性算法的泛化能力越差。图 4 中同一算法的 worst25% 角度误差均值和 best25% 角度误差均值存在较大差异,表明目前的颜色恒常性算法在面对单一简单场景光源时可以准确估计出场景光源,可以达到较高的准确性,但在

面对复杂场景光源和极端光照环境时颜色恒常性算法估计出的场景光源与真实场景光源存在较大偏差,因此可以得出颜色恒常性算法没有很好的泛化性的结论。

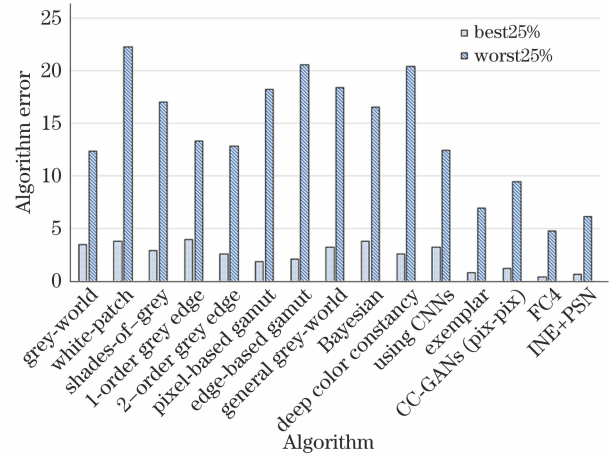


图 4 worst25% 和 best25% 的角度误差值

Fig. 4 Angle error of worst25% and best25%

为了解决这个问题,本文提出渐进式训练网络结构,通过加权累计角度误差损失函数提高颜色恒常性算法在面对复杂场景光源以及极端光照环境下光源估计的准确性,以进一步增强颜色恒常性算法的鲁棒性,使颜色恒常性算法在面对复杂场景光照或者极端光源时也可以准确估计出场景光源。加权角度误差损失函数为

$$L = \sum_i w_i A[f(x_i), y], \quad (1)$$

式中: x_i 表示输入图像; $f(x_i)$ 表示渐进式多尺度特征级联融合颜色恒常性算法估计出的场景光源; y 表示真实的场景光源标签; $A[f(x_i), y]$ 表示角度误差,可以表示为

$$A[f(x_i), y] = \arccos \left[\frac{f(x_i) \cdot y}{|f(x_i)| \cdot |y|} \right], \quad (2)$$

式中: $f(x_i) \cdot y$ 表示内积操作; $|\cdot|$ 表示向量之间的欧几里得距离。 w_i 表示每个阶段的权重,可以表示为

$$w_i = \frac{f(x_i) \cdot y}{\sum_i f(x_i) \cdot y}. \quad (3)$$

渐进式多尺度特征级联融合颜色恒常性算法分为三个阶段。第一阶段,该算法将数据集中的图像进行数据增强,之后将其输入到网络结构中,通过多尺度特征融合模块提取图像中的特征信息,估计出场景光源,根据(2)式得到第一阶段的角度误差损失值,并通过估计出的场景光源对图像进行校正。第二阶段,该算法将第一阶段校正后的图像进行数据

增强,将其再次输入到网络结构中用于估计光源,基于(2)式得到第二阶段的角度误差损失值,通过估计出的光源对图像进行校正。第三阶段,该算法将第二阶段校正后的图像输入到网络结构中并进行光源估计,得到第三阶段的角度误差损失值,基于(3)式计算出每个阶段损失值的权重,基于(1)式进行累计加权,得到网络模型最后的损失值。通过反向传播对网络模型进行训练。这样做的优势在于:对于有些复杂场景光源(如强光照或者多种混合光源)下的图像,第一阶段估计出的场景光源可能与真实的场景光源存在较大偏差(角度误差值较大),基于加权累计误差损失函数,第二阶段估计出的光源是对第一阶段估计出的光源的修正,第三阶段估计出的光源是对第二阶段估计出的光源的修正,三个阶段可以学习到图像中潜在的光源特征联系,实现网络从

粗到细的优化,使颜色恒常性算法在面对复杂场景光照时可以准确地估计出场景光源,提高了颜色恒常性算法的鲁棒性与稳定性。

从重处理的 ColorChecker 的数据集中选择复杂场景光源或极端光照下的图像,使用渐进式多尺度特征级联融合颜色恒常性算法对其进行校正,并将其与其他颜色恒常性算法进行比较。极端场景光源指强光照下的图像,复杂场景光源指多种混合光源下的图像。实验结果如图 5 所示。图 5(a)表示复杂场景光源以及极端场景光源下拍摄的图像,图 5(b)表示使用 grey-world 算法^[21]校正后的图像,图 5(c)表示使用 grey-edge 算法^[4]校正后的图像,图 5(d)表示使用 FC4^[14]算法校正后的图像,图 5(e)表示使用目前效果最好的 C4 算法^[22]校正后的图像,图 5(f)表示使用本文算法校正后的图像。



图 5 复杂场景光照下颜色恒常性算法对比。(a)输入图像;(b) grey-world 算法;(c) grey-edge 算法;(d) FC4 算法;(e) C4 算法;(f)本文算法

Fig. 5 Comparison of color constancy algorithms under complex scene illumination. (a) Input images; (b) grey-world algorithm; (c) grey-edge algorithm; (d) FC4 algorithm; (e) C4 algorithm; (f) proposed algorithm

图 5(e)、(f) 右上角的数字表示角度误差值。从视觉效果以及角度误差值可以看出,目前效果最好的 C4 算法^[22]在面对复杂场景光源或者极端光照时估计出的场景光源与真实的场景光源存在一定偏差。渐进式特征级联融合网络结构基于加权累计角度误差损失函数,在面对复杂场景光照以及极端光源时估计出的场景光源与真实场景光源更为接近,进一步增强了颜色恒常性算法在面对复杂场景光源以及极端光照时进行光源估计的准确性。

2.3 网络结构

渐进式多尺度特征级联融合颜色恒常性算法的网络结构由输入层、4 个多尺度特征提取模块、2 个全卷积网络层组成。网络结构如图 1 所示。输入层接收大小为 512×512 的三通道彩色图像。多尺度特征提取模块中 3 个卷积网络分支层分别从不同尺度提取图像中的特征信息,用于光源估计。各个卷积分支中所包含的卷积核的大小和数量如表 1 所示,其中 1×1 的卷积核可以实现信息的跨通道交互和整合,对卷积核通道数进行降维和升维,以减少参数量。3 个卷积网络分支分别从不同的尺度提取图像中的特征信息,通过特征融合得到更加丰富的特征信息,通过级联方式将多尺度特征模块提取到的浅层边缘特征与深层细粒度特征进行融合,通过连接卷积层 conv_4、conv_5 对提取到的特征从三个通道进行置信度池化加权,对提取到的特征中可以为光源估计提供更多信息的特征赋予更高的权重,充分利用图像各个通道中可以为光源估计提供更多信息的特征以准确地估计出光源。Conv_4 由 64 个大小为 6×6 的卷积核组成。Conv_5 由 6 个大小为 1×1 的卷积核组成,最后输出估计出的场景光源。渐进式网络结构通过加权累计角度误差损失函数进行训练,使网络结构在面对复杂场景光源或者极端光源时也可以准确估计出场景光源。

3 实验与结果分析

为验证本文算法的有效性,在两组单光源数据集上进行实验,并将本文方法与其他颜色恒常性算法的误差进行对比,分析本文算法的有效性。

实验环境为 3.00 GHz 的英特尔 i7CPU,16 GB 内存,GPU 为 NVIDIA GeForce GTX 2080,12 GB 显存,操作系统为 Ubuntu 20.02,文本提出的渐进式多尺度特征级联融合颜色恒常性算法是基于 pytorch 框架实现的。

3.1 数据集

1) 重处理的 ColorChecker 数据集^[23]。本文选用重处理的 ColorChecker 数据集是 Shi 等在原始 ColorChecker 数据集 RAW 图像的基础上经过线性化处理得到的。数据集中包括佳能 1D 和佳能 5D 拍摄的 568 张 RGB 图像、每张图片所对应的真实场景光源,以及图像中麦克白色卡的坐标信息。麦克白色卡可以为图片提供真实的场景光源信息。在将图像输入到网络模型进行训练之前,需要将每张图像中的麦克白色卡进行遮挡,以防对网络训练造成干扰。将数据集中的 568 张图像分为三部分用于三重交叉验证。第一部分包括 189 张图像,这些图像经数据增强后用于网络训练;第二部分包括 191 张图像,用于网络模型测试;第三部分包括 188 张图像,用于验证网络模型。三重交叉验证可以更加充分地测试网络模型对场景光源估计的准确性以及稳定性。

2) NUS-8 数据集^[24]。NUS-8 数据集中包括了用 8 种相机拍摄的 1736 张图像。数据集中的每张图像都含有麦克白色卡,用于测量每张图片的真实场景光源。与重处理的 ColorChecker 数据集相比,NUS-8 数据集包含了使用多种相机在更加丰富的真实场景光源下拍摄的图像,可以更加精确地测试出颜色恒常性算法的精确性以及算法在面对复杂环境时的稳健性和泛化性。

3.2 数据集预处理

由于用于训练网络的数据集较小,因此在网络训练之前需要对数据集进行数据增强以及预处理。首先,随机选择原始图像较短边的 10%~100%,之后随机选择左上角图像进行裁剪,将选择到的图像随机旋转 $-30^\circ \sim 30^\circ$,并以 0.6 的概率对图像进行左右旋转,以 0.5 的概率对图像进行上下翻转,将裁剪后的图像调整为 512×512 大小,将数据集中的图像打乱后,选择一个批次的图像输入到网络中,进行训练。将伽马值为 1/2.2 的伽马校正应用于线性 RGB 图像,对图像进行显示,经过预处理后的图像如图 6 所示。

3.3 网络训练

采用二阶段训练法对本文算法进行训练。第一阶段使用 ImageNet 数据集对渐进式多尺度特征级联融合网络结构进行训练,对网络模型中的参数进行初始化。第二阶段使用重处理的 ColorChecker 数据集集中的训练集,进行数据增强后对网络进行迭代训练。第一阶段对大规模数据进



图 6 数据集示例图

Fig. 6 Samples of data set

行训练后,可以为网络结构提供一个好的参数初始值,因此第二阶段使用颜色恒常性数据集训练时网络模型可以更快地收敛,同时可以使网络收敛到一

个泛化能力最高的最优解,这提高了颜色恒常性算法的精确性和泛化性。二阶段训练法流程图如图 7 所示。

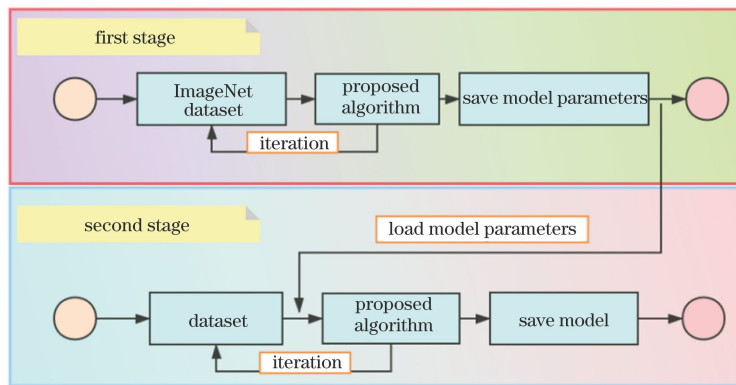


图 7 网络训练步骤图

Fig. 7 Diagram of network training stage

超参数是指在开始训练网络模型之前人为设置其值的参数,如学习率等参数,超参数的设置对于模型性能有着直接影响,在网络训练过程中对网络结构中的超参数调整进行比较,选取可以使得颜色恒常性算法精确度最高和稳健性最好的超参数。将网络训练中的批大小(batch-size)设置为 64,学习率设置为 10^{-3} ,在网络训练过程中逐渐减小学习率,随着循环次数的不断增加,学习率将逐渐按照比例减小。网络训练的前 6000 次使用 SGD 优化函数,之后使用 Adam 优化函数。在网络中的每个卷积层后都对其输出采用 ReLU 激活函数进行处理。对于特征融合之后的第一个卷积层,采用 0.6 的丢弃率以防止网络出现过拟合。将网络训练的迭代次数设置为 15000,在网络训练过程中采用截断训练法。实验结果表明,当网络训练到 12000 次时网络停止训练,即网络在循环迭代训练 12000 次后,损失函数值基本趋向于稳定,网络已经基本接近于完全收敛。训练过程中的损失函数值如图 8 所示。其中横坐标表示训练过程中的迭代次数,纵坐标表示根据角度

误差计算出的损失值。

3.4 评价指标

本文采用角度误差作为光源颜色估计误差的度量参数。网络模型的性能通过比较真实场景光源与本文算法估计得到的光源获得。角度误差的计算式为

$$d_e(\mathbf{e}_g, \mathbf{e}_u) = \arccos\left(\frac{\mathbf{e}_g \cdot \mathbf{e}_u}{|\mathbf{e}_g| \cdot |\mathbf{e}_u|}\right), \quad (4)$$

式中: $\mathbf{e}_g = [e_g^R, e_g^G, e_g^B]$ 表示数据集中的真实场景光源标签; $\mathbf{e}_u = [e_u^R, e_u^G, e_u^B]$ 表示使用本文算法估计出的场景光源。 $\mathbf{e}_g \cdot \mathbf{e}_u$ 表示真实场景光源与估计光源之间的内积操作。角度误差值代表本文算法估计出的场景光源与真实光源之间的接近程度。角度误差值越小则表示估计出的场景光源越接近场景真实光源,角度误差值越大则表示颜色恒常性算法的精确性以及稳定性越差。本文主要对比光源颜色估计角度误差的均值、中位数、三均值、最优 25% 均值、最差 25% 均值和 95th 百分位数。其中角度误差均值用于估计算法的总体性能,中位数

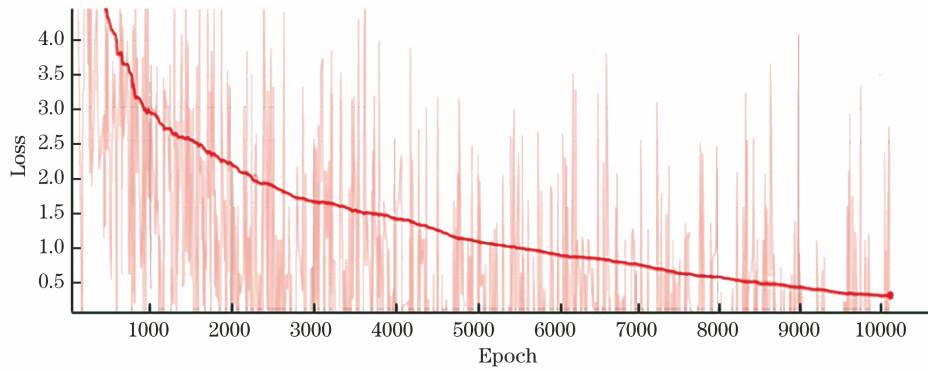


图 8 网络训练损失曲线

Fig. 8 Curve of network training loss

和三均值用于测试离散值对算法的影响,最差 25% 均值和 95th 百分位数用于衡量算法在面对复杂环境时的角度误差,其值越小说明算法的鲁棒性和稳健性越好。

3.5 与其他算法的对比分析

在重新处理的 ColorChecker 数据集上,采用数据集提供的三重交叉验证计算误差,分别对 mean、med、tri、best25%、worst25% 和 95th 进行统计,并与其他方法进行对比,实验结果如表 3 所示。通过对比表格中的数据可知,与传统的颜色恒常性算法相比,本文算法进一步提升了颜色恒常性算法的精

确性与稳定性。C4 算法^[22]是基于深度学习的颜色恒常性算法中效果最好的算法,与 C4 算法相比,本文算法在 mean、med、tri、best25%、worst25% 这 5 个评价指标上分别提升了 10.3%、12.8%、14.8%、12.9%、17.2%。FC4 算法在角度误差 95th 评价指标上是最优的,与 FC4 算法相比,本文算法在 95th 评价指标上提升了 28.1%。在 Worst25% 与 95th 这两个评价指标的提升表明渐进式多尺度特征级融合颜色恒常性算法提高了颜色恒常性算法在面对复杂场景光照或者极端场景光源时光源估计的准确性,增强了算法的鲁棒性和稳健性。

表 3 使用重处理的 ColorChecker 数据集的测试结果

Table 3 Test results obtained by using reprocessed ColorChecker dataset

Method	Mean	Med	Tri	Best25%	Worst25%	95th
Grey-world ^[21]	10.700	10.600	10.700	3.450	12.300	17.400
White-patch ^[25]	9.800	8.000	8.900	3.800	13.600	22.300
Shades-of-grey ^[26]	8.300	7.500	7.800	2.900	11.800	17.000
1-order grey edge ^[4]	5.000	3.700	4.100	3.900	10.100	13.300
2-order grey edge ^[4]	5.400	4.500	4.800	2.600	9.800	12.800
Pixel-based gamut ^[27]	6.900	5.200	5.700	1.800	11.700	18.200
Edge-based gamut ^[27]	6.900	4.600	5.200	2.100	14.600	20.600
Bayesian ^[6]	6.600	4.600	5.200	3.200	10.900	18.400
General grey-world ^[28]	7.600	6.700	7.000	3.800	12.100	16.500
Using CNNs ^[7]	8.200	6.300	6.800	2.600	11.300	20.400
Deep color constancy ^[19]	5.700	4.700	5.000	3.200	8.400	12.400
Exemplar ^[27]	2.890	2.270	2.420	0.820	5.970	6.950
CCC ^[29]	2.000	1.220	1.400	0.350	4.760	5.850
CC-GANs (pix-pix) ^[13]	3.600	2.800	3.100	1.200	7.200	9.400
FC4-AlexNet ^[14]	1.770	1.110	1.290	0.340	4.290	5.440
FC4-SqueezeNet ^[14]	1.650	1.180	1.270	0.380	3.780	4.730
IEN+PSN ^[18]	2.250	1.590	1.730	0.590	5.030	6.080
Multi-hypothesis ^[30]	2.100	1.320	1.530	0.360	5.100	-
C4 ^[22]	1.350	0.880	0.990	0.280	3.210	-
Our method	1.210	0.767	0.843	0.244	2.655	3.043

在 NUS-8 数据集上分别对每个相机的图像子集进行三重交叉验证,得到各个子集光源颜色估计误差的 mean、med、tri、best25%、worst25% 和 95th 的统计结果,将所有子集误差统计结果的几

何平均值作为最终的误差结果,并将本文算法与目前存在的基于统计的颜色恒常性算法和基于学习的颜色恒常性算法进行比较,实验结果如表 4 所示。

表 4 使用 NUS-8 数据集得到的测试结果

Table 4 Test results obtained by using NUS-8 dataset

Method	Mean	Med	Tri	Best25%	Worst25%
Grey-world ^[21]	4.140	3.200	3.390	0.900	9.000
White-patch ^[25]	10.620	10.580	10.490	1.860	19.450
Shades-of-grey ^[26]	3.400	2.570	2.730	0.770	7.410
1-order grey edge ^[4]	3.200	2.220	2.430	0.720	7.360
2-order grey edge ^[4]	3.200	2.260	2.440	0.750	7.270
Pixel-based gamut ^[27]	7.700	6.710	6.900	2.510	14.050
Edge-based gamut ^[27]	8.430	7.050	7.370	2.410	16.080
Bayesian ^[6]	3.670	2.730	2.910	0.820	8.210
Using CNNs ^[7]	7.600	6.900	7.400	3.000	12.400
Deep color constancy ^[18]	6.200	5.000	5.400	3.900	8.600
CCC ^[29]	2.800	1.800	1.900	0.850	6.300
CC-GANs (pix-pix) ^[13]	3.800	3.000	3.700	1.900	8.400
FC4-AlexNet ^[14]	2.120	1.530	1.670	0.480	4.780
FC4-SqueezeNet ^[14]	2.230	1.570	1.270	0.470	5.150
IEN+PSN ^[18]	2.100	1.350	1.510	0.450	5.010
Multi-hypothesis ^[30]	2.350	1.550	1.730	0.460	5.620
C4 ^[22]	1.960	1.420	1.530	0.480	4.400
Our method	1.574	1.027	1.241	0.386	3.512

NUS-8 数据集的场景光源丰富,更加考验模型在面对复杂环境时模型的稳健性。实验结果表明,本文算法在该数据集上的 5 个评价指标上均超过了目前最好的颜色恒常性算法 C4,相对于 C4 算法分别提升了 19.7%、27.7%、18.9%、19.6%、20.1%,其 Worst25% 是目前颜色恒常性算法中较高的。渐进式多尺度特征级联融合颜色恒常性算法可以充分利用图像中不同尺度的特征信息进行融合,得到更加丰富的特征信息,通过级联方法将图像中浅层边缘特征与深层细粒度特征进行融合,以准确地估计光源,渐进式网络结构通过加权累计角度误差损失函数实现了网络结构从粗到细的优化,增强了颜色恒常性算法在面对复杂场景光源或者极端场景光源的鲁棒性与稳健性,使算法在面对复杂场景光照或者极端光源时可以准确估计出场景光源。

本文算法的输出结果如图 9 所示,图 9(a) 表示输入图像,图 9(b) 表示使用本文算法估计得到的经颜色光源校正后的图像,图 9(c) 表示本文算法提取到的特征融合信息图,图 9(d) 表示 C4 算法^[22]校正后的图像,图 9(e) 表示使用 FC4 算法^[14]校正后的图像,图 9(f) 表示经 grey-world 算法^[21]校正后的图

像,图 9(g) 表示经 white-patch 算法^[25]校正后的图像,图 9(h) 表示经 grey-edge 算法^[4]校正后的图像,图 9(i) 表示经 shades-of-grey(SOG)算法^[26]校正后的图像。Grey-world 算法^[21]假设图像中 R、G、B 三个分量的平均值趋于一个灰度常量。White-patch^[25]算法假设图像上最亮的点就是白点,并根据此白点对图像进行色彩校正。SOG 算法^[26]利用闵式距离代替简单求平均的方法对图像进行色彩校正。FC4 算法^[14]通过置信度加权方法并充分利用图像中可以为光源提供更多信息的特征估计光源。C4 算法^[22]在不同数据集之间实现了稳定的泛化能力。将本文算法校正后的图像与其他 4 种传统的颜色恒常性算法进行比较,从主观视觉上可以看出,对于有些场景下的图像(图 9 中第二行和第三行),经 grey-world 算法^[21]和 grey-edge 算法^[4]校正后的图像都出现了偏色。从图 9 中前三行图像中天空的乌云可以看出,与其他算法相比,本文算法对光源的估计更加准确,移除光源后可以准确显示出乌云本身的颜色。本文算法通过将多个尺度提取的图像中的特征信息进行融合,得到更加丰富的特征信息,通过级联方式将图像中浅层边缘特征与深层的细粒度

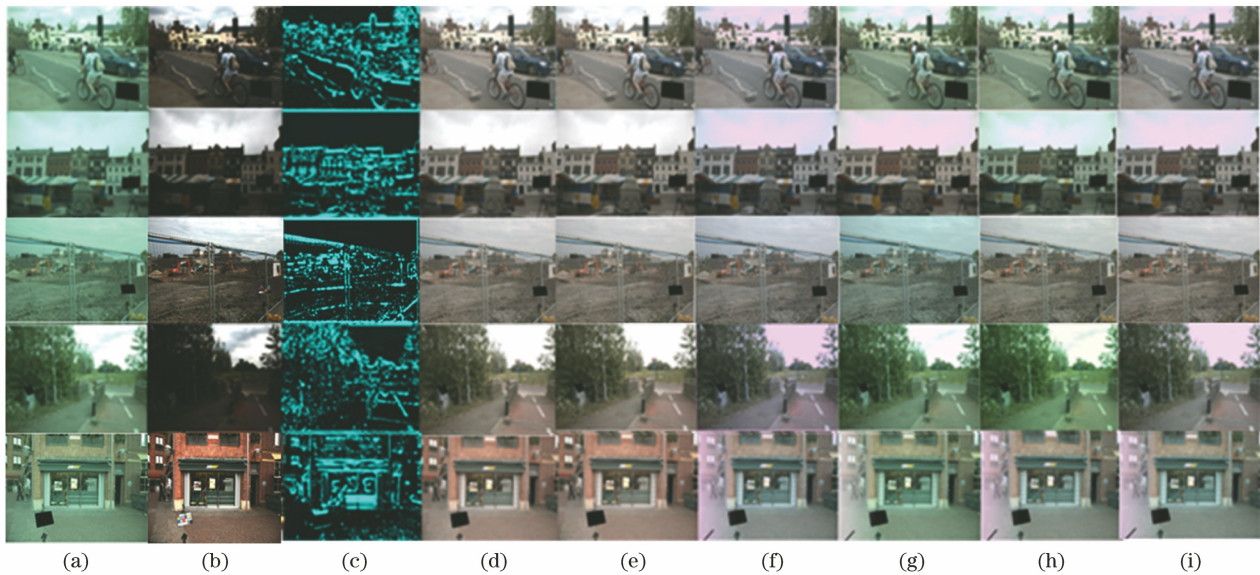


图 9 网络测试可视化结果。(a)输入图像;(b)本文算法估计得到的经光源校正后的图像;(c)本文算法提取到的特征;(d) C4 算法;(e) FC4 算法;(f) grey-world 算法;(g) white-patch 算法;(h) grey-edge 算法;(i) SOG 算法

Fig. 9 Visualization results of network tests. (a) Input images; (b) images after light source correction obtained by proposed algorithm; (c) features extracted by proposed algorithm; (d) C4 algorithm; (e) FC4 algorithm; (f) grey-world algorithm; (g) white-patch algorithm; (h) grey-edge algorithm; (i) SOG algorithm

特征信息相结合,以准确地估计出光源。渐进式网络结构基于加权累计角度误差损失函数实现了网络从粗到细的优化,增强了颜色恒常性算法在面对复杂场景光照或者极端光源时的鲁棒性和稳健性,使其在面对复杂场景光照或者极端光源时可以准确估计出场景光源,这是目前精确度较高、鲁棒性较好的颜色恒常性算法,具有较好的实际应用价值。

4 结 论

为了进一步提高颜色恒常性算法估计光源的精确性、增强颜色恒常性算法在面对复杂场景光照或者极端光源时的鲁棒性及稳健性,提出渐进式多尺度特征级联融合颜色恒常性算法。与现有的颜色恒常性算法相比,本文算法通过将图像中提取的不同尺度的特征进行融合,得到更加丰富的特征,通过级联方式将图像中浅层边缘特征与深层细粒度特征进行融合并估计光源,提高了颜色恒常性算法对光源估计的准确性。渐进式网络结构基于加权累计角度误差损失函数,增强了颜色恒常性算法在面对复杂场景光照或者极端光源时的鲁棒性及稳健性,使算法在面对复杂场景光照环境以及极端光源时可以准确估计出场景光源。实验结果表明,本文算法在颜色恒常性算法角度误差的各项评价指标上均优于目前的颜色恒常性算法。在 NUS-8 数据集的 mean 和 worst25% 评价指标上,本文算法比目前效果最好

的颜色恒常性算法(C4 算法)^[22]提升了 19.7% 和 20.1%,在 med 评价指标上提升了 27.7%。在重处理的 ColorChecker 数据集上,本文算法在 6 个评价指标上均得到了提升,这证明了本文算法的有效性。未来的工作将对颜色恒常性算法的视频序列进行研究,使用颜色恒常性算法对视频序列进行处理,以为计算机视觉的其他任务如无人驾驶、目标跟踪等提供更多的准备工作;同时,会将无监督学习应用于颜色恒常性算法,如自编码网络和对抗神经网络,以此来解决颜色恒常性数据集不充分的问题;在保证颜色恒常性算法稳定性的前提下,寻找更加轻量级的网络模型结构,以进一步提高颜色恒常性算法的稳定性与时效性。

参 考 文 献

- [1] Wu L X, Xu M, Zhu G B, et al. Appearance features in encoding color space for visual surveillance [J]. *Neurocomputing*, 2018, 308: 21-30.
- [2] Lu Z, Jiang X D, Kot A. Color space construction by optimizing luminance and chrominance components for face recognition [J]. *Pattern Recognition*, 2018, 83: 456-468.
- [3] Rabie T. Training-less color object recognition for autonomous robotics [J]. *Information Sciences*, 2017, 418/419: 218-241.
- [4] van de Weijer J, Gevers T, Gijssenij A. Edge-based

- color constancy [J]. *IEEE Transactions on Image Processing*, 2007, 16(9): 2207-2214.
- [5] Finlayson G, Hordley S. Selection for gamut mapping colour constancy[J]. *Image and Vision Computing*, 1999, 17(8): 597-604.
- [6] Brainard D H, Freeman W T. Bayesian color constancy [J]. *Journal of the Optical Society of America A*, 1997, 14(7): 1393-1411.
- [7] Bianco S, Cusano C, Schettini R. Color constancy using CNNs[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 81-89.
- [8] Qiu J Q, Xu H S, Ye Z N. Color constancy by reweighting image feature maps[J]. *IEEE Transactions on Image Processing*, 2020, 29: 5711-5721.
- [9] Afifi M. Semantic white balance: semantic color constancy using convolutional neural network[EB/OL]. (2018-05-31)[2020-11-13]. <https://arxiv.org/abs/1802.00153>.
- [10] Afifi M, Brown M. What else can fool deep learning? addressing color constancy errors on deep neural network performance[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 243-252.
- [11] Laakom F, Raitoharju J, Iosifidis A, et al. Color constancy convolutional autoencoder[C]//2019 IEEE Symposium Series on Computational Intelligence (SSCI), December 6-9, 2019, Xiamen, China. New York: IEEE Press, 2019: 1085-1090.
- [12] Das P, Baslamisli A S, Liu Y, et al. Color constancy by GANs: an experimental survey[EB/OL]. (2018-12-07)[2020-11-20]. <https://arxiv.org/abs/1812.03085>.
- [13] Sidorov O. Conditional GANs for multi-illuminant color constancy: revolution or yet another approach? [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 16-17, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 1748-1758.
- [14] Hu Y M, Wang B Y, Lin S. FC¹: fully convolutional color constancy with confidence-weighted pooling[C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 330-339.
- [15] Li X Y, Yang H J, Yan Z, et al. Image background blurring algorithm based on color constancy[J]. *Laser & Optoelectronics Progress*, 2020, 57(8): 081001. 李晓颖, 杨恒杰, 闫铮, 等. 基于颜色恒常性的图像背景虚化算法[J]. *激光与光电子学进展*, 2020, 57(8): 081001.
- [16] Sidorov O. Artificial color constancy via GoogLeNet with angular loss function [J]. *Applied Artificial Intelligence*, 2020, 34(9): 643-655.
- [17] Košćević K, Subašić M, Lončarić S. Attention-based convolutional neural network for computer vision color constancy [C] // 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA), September 23-25, 2019, Dubrovnik, Croatia. New York: IEEE Press, 2019: 372-377.
- [18] Cui S, Zhang J, Gao J, et al. Illuminant estimation via deep residual learning[J]. *Journal of Image and Graphics*, 2019, 24(12): 2111-2125. 崔帅, 张骏, 高隽. 深度残差学习下的光源颜色估计[J]. *中国图象图形学报*, 2019, 24(12): 2111-2125.
- [19] Lou Z Y, Gevers T, Hu N H, et al. Color constancy by deep learning [C] // Proceedings of the British Machine Vision Conference 2015, September 7-10, Swansea, UK. London: British Machine Vision Association, 2015: 76.
- [20] Shi W, Loy C C, Tang X O. Deep specialized network for illuminant estimation [M] // Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016*. Lecture notes in computer science. Cham: Springer, 2016, 9908: 371-387.
- [21] Buchsbaum G. A spatial processor model for object colour perception[J]. *Journal of the Franklin Institute*, 1980, 310(1): 1-26.
- [22] Yu H L, Chen K, Wang K Q, et al. Cascading convolutional color constancy[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(7): 12725-12732.
- [23] Shi L, Funt B. Re-processed vision of the gehler color constancy dataset of 568 images [EB/OL]. (2014-09)[2020-11-15]. <https://www2.cs.sfu.ca/~colour/data/>
- [24] Cheng D, Prasad D K, Brown M S. Illuminant estimation for color constancy: why spatial-domain methods work and the role of the color distribution [J]. *Journal of the Optical Society of America A*, 2014, 31(5): 1049-1058.
- [25] Brainard D H, Wandell B A. Analysis of the retinex theory of color vision [J]. *Journal of the Optical Society of America A*, 1986, 3(10): 1651-1661.
- [26] Finlayson G D, Trezzi E. Shades of gray and color constancy [C]. In the Twelfth Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications, CIC 2004, Scottsdale, Arizona, USA, November 9-12, 2004.
- [27] Barnard K, Martin L, Coath A, et al. A comparison of computational color constancy algorithms. II.

- Experiments with image data[J]. IEEE Transactions on Image Processing, 2002, 11(9): 985-996.
- [28] Jozsef H R V, Drew M S. Exemplar-based color constancy and multiple illumination[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(5): 860-873.
- [29] Barnard K. Improvements to gamut mapping colour constancy algorithms [M] // Computer vision-ECCV 2000. Lecture notes in computer science. Heidelberg: Springer, 2000, 1842: 390-403.
- [30] Hernandez-Juarez D, Parisot S, Busam B, et al. A multi-hypothesis approach to color constancy [C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 2267-2277.