DOI: 10.13382/j.jemi.B2104908

基于多对一映射生成对抗网络的颜色恒常性算法*

谢林芳 张旭东 孙 锐 范之国 党天一
(合肥工业大学计算机与信息学院 合肥 230009)

摘 要:颜色恒常性是计算机视觉的重要研究方向,但在目前的研究中,大多数算法专注于单光源均匀分布的情况,光源非均匀分布的问题一直没有得到很好的解决。针对该问题,在单光源非均匀分布的情况下,将颜色恒常性转换成一个多对一映射任务,提出了一种基于生成对抗网络直接校正的方法。该方法根据颜色恒常性的特性,将图像拆解成内容编码和光源编码,改变光源编码为目标光源编码,重组生成目标光源下的图像。为了使非标准光源更加多样化,加入光源采样模块,帮助网络学习到更为丰富的光源信息,实现多对一映射。同时,为了在输入不同的光源编码时可以引导图像映射到不同的光源下,加入光源监督模块以区分具有不同光源的图像,帮助光源转换模块更好地将内容编码和特定光源编码结合起来,生成目标图像,实现颜色恒常性。同时,针对本文任务,在现存的数据集上渲染非均匀分布的光源,构造了单光源非均匀分布的数据集。实验结果表明,本文方法较好地解决了光源非均匀分布的问题,在非均匀数据集上超越了其他算法,最终生成的图像也更加接近标准光源下图像。

关键词: 颜色恒常性;单光源;生成对抗网络;图像映射
中图分类号: TP391.41 **文献标识码**: A **国家标准学科分类代码**: 510.4050

MTO-GAN: Learning many-to-one mappings for color constancy

Xie Linfang Zhang Xudong Sun Rui Fan Zhiguo Dang Tianyi
(School of Computer and Information, Hefei University of Technology, Hefei 230009, China)

Abstract: Color constancy is an important research direction in computer vision, but most algorithms focus on uniform distribution of single illuminant, and the problem of non-uniform distribution of illuminant has not been well solved. In order to solve this problem, a direct correction method based on generative adversarial network is proposed to transform color constancy into a many-to-one mapping task under the condition of non-uniform distribution of single illuminant. According to the characteristic of color constancy, the image is divided into content code and illuminant code, and the image under the target illuminant is reconstructed by changing the illuminant code to target illuminant code. At the same time, in order to make non-standard illuminant more diversified, the illuminant sampling module is added to help the network learn more abundant illuminant information and realize many-to-one mapping. In order to guide images to be mapped to different illuminants when different illuminant codes are input, the illuminant supervision module is added to distinguish images with different illuminants, so as to help the illuminant conversion module better combine content coding with specific illuminant coding to generate target images and achieve color constancy. At the same time, aiming at the task of this paper, the non-uniformly distributed illuminant is rendered on the existing dataset, and the dataset with non-uniformly distributed single illuminant is constructed. The experimental results show that the proposed method solves the problem of non-uniform distribution of illuminant well, surpasses other algorithms in non-uniform dataset, and the final image is closer to the image under standard illuminant.

Keywords: color constancy; single illumination; GAN; image mapping

收稿日期: 2021-11-16 Received Date: 2021-11-16
*基金项目:国家自然科学基金(61876057,61971177)、安徽省重点研发计划-科技强警专项(202004d07020012)项目资助

0　引　言

视觉自适应性是视觉系统的最重要的特性之一,使我们能够在复杂的动态环境下保持对外部世界的可靠感知。其中的视觉颜色恒常性,是指不因环境条件变化而保持对客体颜色知觉不变的心理倾向,即当光照条件发生变化时,人类对物体表面颜色的知觉趋于稳定[1]。在计算机视觉领域中,颜色恒常性是一个多学科交叉性很强的研究课题[2-3],是很多图像处理[4-7]的前序工作。例如在图像识别、图像分类、目标检测等图像处理工作中,实验阶段所使用的数据集大多是在良好的光线环境下进行拍摄的,图像均保有原本的色彩;但在实际应用中,受到天气、环境、灯光等因素的影响,所拍摄到的图像往往色彩不佳,不利于进行接下来的图像处理工作。这时就需先对图像进行颜色恒常性的工作,使图像恢复原有色彩,再进行后序处理。因此颜色恒常性仍然是一个值得关注的经典问题。

颜色恒常性任务主要有两大难题,一是存在多个光源,二是光源非均匀分布。研究人员为了利于研究,先将该任务简化成单光源均匀分布的情况,即场景中只有一个均匀分布的光照,图像上每一个像素点的光源值都是相同的。在这个假设前提下大多数算法通常直接把估计的光源向量作为输出,然后利用对角变换校正图像,逐步取得了不错的效果。但对于非均匀情况来说,光源不再是一个向量而是一个分布,因此无法像均匀情况一样估计光源值,导致很难从表现良好的光源均匀分布算法向非均匀迁移,因此光源非均匀分布的问题一直都没有被很好的解决。本文主要针对光源非均匀分布的问题,在单光源非均匀分布的情况下,利用深度学习的框架,提出了一种端到端直接进行校正的方法,省略中间过程,输入源图像输出直接得到目标图像,很好地解决了光源非均匀分布的问题。

本文的贡献如下:

1)提出一种基于生成对抗网络的多对一映射框架 MTO(many-to-one),将颜色恒常性任务转换为多对一映射问题,端到端对非标准光源图像直接进行校正,输入源图像输出目标图像,将中间过程包裹在网络框架中去学习,解决光源非均匀分布的问题。

2) 假设内容空间和光源空间相互独立,把图像分解为内容编码和光源编码,通过内容编码和目标域光源编码输入到光源转换模块进行重组,将图像映射到指定光源下,从而保证图像的内容编码保持不变仅仅在光源空间进行转换,实现颜色恒常性。

3)在 MTO 框架中加入光源采样模块,从目标域中随机抽取光源编码与内容编码重新组合,帮助网络学习到更为丰富的光源信息;同时为了保证输入到光源转换模块的光源编码有示例引导作用,加入光源监督模块以区分具有不同光源的图像,使光源编码可以有效地引导图像映射到指定光源下。

4)由于缺乏单光源非均匀分布的数据集,利用两种公开单光源数据集 cube+和 Gehler-Shi,在其标准光源图像上渲染光源合成了非均匀数据集,用来验证本文方法和以往方法在非均匀情况下的校正能力,最终结果证明本文算法的性能优于其他算法。

1　相关工作

针对单光源颜色恒常性任务,目前的算法可以分为两大类:第 1 类方法估计光源向量,再利用估计值校正图像,第 2 类方法直接进行图像校正,输入源图像输出目标图像。

在估计光源向量的方法中,一部分算法是基于统计学规律提出的:如经典算法中以 Retinex 理论为基础的白板估计算法(white patch)[8]、以灰度世界理论为基础的灰色世界算法(gray world)[9]和 Van 等[10]提出的灰色边缘算法(gray-edge),这些方法的优势是计算简单,但其性能高度依赖于图像的反射率分布。因此研究者引入更复杂的对场景反射率和光源颜色分布的约束以及使用概论推断技术来估计光源的颜色,Hernandez 等[11]提出的多假设策略贝叶斯框架从生成图像的后概率分布中学习估计值,可以较好地适用在多设备拍摄的数据集上;Laakom 等[12]提出了一种基于概率的无监督颜色恒常性方法(PCC),通过使用加权来表示不同图像区域的贡献,该方法可用作集成不同先验信息的通用框架;杨泽鹏等[13]利用多通道置信度加权方法准确估计出全局场景光源,提高了算法的精确性和稳健性。Barron 等[14-15]通过将图像转换到对数色度空间,对光源进行定位,提出了快速傅里叶颜色恒常性算法 FFCC,是近些年来广受关注的一个算法。Hu 等[16]提出的全卷积神经网络 Fc4,将图像块赋予不同的置信度来估计光源,实现了更高的效率和准确性;Lo 等[17]设计了一种基于对比学习的颜色恒常性框架,通过构建对比对来增强光照依赖性特征,在数据稀疏区域也保持了不错的性能;Xiao 等[18]开创了将多域学习引入到颜色恒常性领域的先河,加入相机特定模块适用于不同相机,因此可以利用更大的数据集训练,显著提高了估计精度;Yu 等[19]引入了一种级联卷积颜色恒常性算法(C4),以提高回归学习的鲁棒性,并在一个独特的框架中实现跨数据集的稳定泛化能力。

$$I^{c}(x,y)=\begin{bmatrix} e_R^{-1} & 0 & 0 \\ 0 & e_G^{-1} & 0 \\ 0 & 0 & e_B^{-1} \end{bmatrix} I(x,y) \tag{1}$$

上述算法是在单光源均匀分布的假设下实现的,以非标准光源的 RGB 值作为输出,然后利用式(1)的对角模型校正图像,其中 $I(x,y)$ 是非标准光源下的待校正图像,$I^c(x,y)$ 是校正后的标准光源图像,$[e^R,e^G,e^B]$ 是光源的 RGB 值。但在非均匀分布的情况中,光源是一个三维光源分布矩阵,上述的方法不能很好地适用在单光源非均匀分布的情况上。但对直接校正的方法来说,端到端校正图像,输入源图像直接输出目标图像,黑盒化中间过程,不需要前提假设,更加适用于非均匀情况。Das 等[20]就首次将颜色恒常性任务表述成图像到图像的映射问题,利用 pix2pix[21]、cycleGAN[22] 和 starGAN[23] 这 3 种先进的生成对抗网络,将非标准光源图像直接转换成标准光源图像,实现颜色恒常性。但这 3 种框架都是针对图像风格迁移任务所设计的,它们在单光源均匀分布的颜色恒常性任务上的性能远不如其他先进的方法;Sidorov[24] 则面向单光源和多光源的颜色恒常性任务设计了一种生成对抗网络 AngularGAN,在网络中加入角度损失,帮助网络学习,但角度损失的计算依托均匀分布的假设,无法通过损失的约束解决光源非均匀的问题。本文则专门针对单光源非均匀分布的颜色恒常性问题,设计了一个基于生成对抗网络的多对一映射框架 MTO,通过将图像拆解,加入光源采样模块和光源监督模块,实现非均匀光源下图像的直接校正。

2 基于多对一映射生成对抗网络的颜色恒常性算法

2.1 算法整体框架

颜色恒常性的根本任务是将非标准光源下的图像转换为标准光源下的图像。在图 1 的示例中,此任务就是将左边 3 幅非标准光源下图像分别输入到校正网络中,输出右边对应的标准光源下图像。而每幅图像在输入校正网络前后,图像的内容保持不变仅仅光源发生变化,因此可以用对应光源值代表每一幅图像。对应的任务示意图如图 2 所示,左边一列为 3 种不同的非标准光源,右边一列则为同一种标准光源,因此进一步简化成图 3 所示,此时示例任务可以转换为,将 3 种不同的非标准光源转换到同一种标准光源的三对一映射问题。从而进行推广,最终将颜色恒常性任务转换成多种不同非标准光源映射到同一种标准光源的多对一映射问题。

如上面的分析,颜色恒常性任务已经转换为多对一映射问题。为此,本文设计了一种图 4 所示的多对一映射网络框架 MTO,由编码模块、光源转换模块 Trans 和光源监督模块 Di 三个模块组成。其中编码模块从图像中提取内容编码 C 和光源编码 i;光源转换模块 Trans 则将内容编码和目标光源编码进行重组,生成目标光源下的图像;光源监督模块 Di 可以鉴别每幅图像的光源,帮助光源转换模块 Trans 重组内容编码和指定的光源编码。下面 3 个部分是对各个模块的具体介绍。

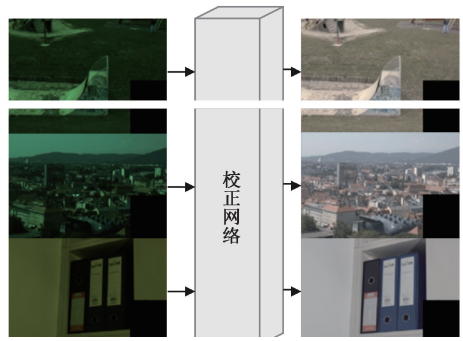

图 1 颜色恒常性任务示意图

Fig. 1 Diagram of color constancy task

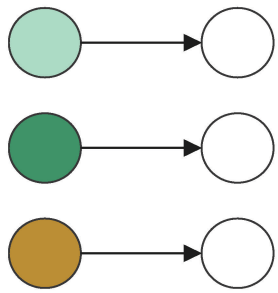

图 2 颜色恒常性任务简化示意图

Fig. 2 Simplification of color constancy task.

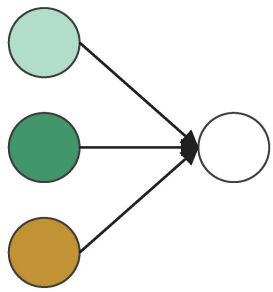

图 3 颜色恒常性任务进一步简化示意图

Fig. 3 Color constancy task further simplifies the schematic

2.2 编码模块

颜色恒常性任务的目的是保持图像内容的不变,改变其光源到标准光源下,因此本文假设图像的内容空间和光源空间相互独立,具有可分离性,对图像的内容和光源分别进行编码,故编码模块由内容编码模块 Ec 和光源编码模块 Ei 两部分组成。其中,内容编码模块 Ec 在输入一幅图像后,利用多个卷积层提取特征,最终得到内容编码 C。光源编码模块 Ei 的输入为光源值$[r,g,b]$,通

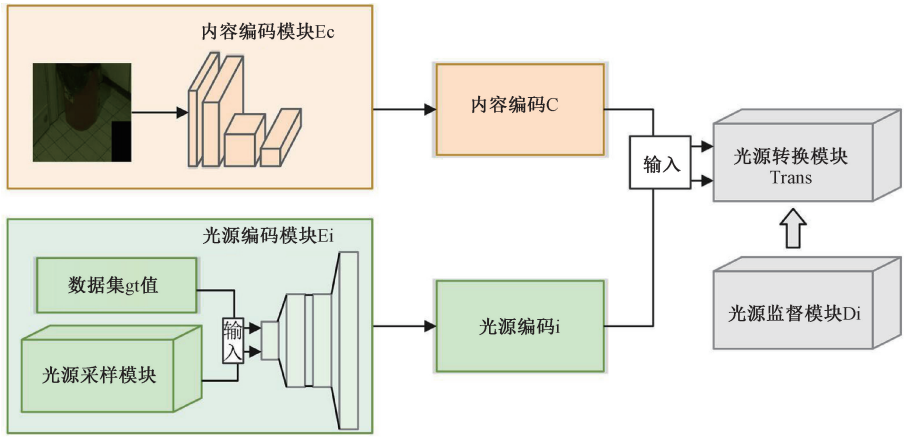

图 4　多对一映射生成对抗网络 MTO 结构示意图

Fig. 4　Many-to-one mapping generative adversarial networks schematic

过多个全连接层输出光源编码 i。

对于光源编码模块来说,它的输入光源值不仅来自于数据集真值,还来自于光源采样模块的采样值。因为在被广泛使用的两个数据集 Gehler-Shi 和 cube+中,其光源真值 gt 的分布情况如图 5 和 6 所示,小球颜色即代表光源颜色。图中显示,两个数据集的 gt 值都主要分布在 $R[0\sim0.5]$ $G[0\sim0.5]$ $B[0\sim0.5]$ 的区间内,它们的分布都近似于一个曲面,没有充斥在整个空间内,虽然其中最大数据集的图像数量已经接近 2 000 张,但光源样本仍然不够丰富,光源分布比较集中,而本文已将颜色恒常性任务转换为多种非标准光源映射到同一种标准光源的多对一映射问题,需要网络学习映射大量的非标准光源到标准光源上,因此本文设计加入图 7 所示的光源采样模块,在 gt 值集中分布的 $R[0\sim0.5]$ $G[0\sim0.5]$ $B[0\sim0.5]$ 立方体区域随机采样一个点,作为光源值 $[r,g,b]$ 输入,最终光源编码模块的输入就由数据集 gt 值和光源采样模块采样得到的光源值两部分组成。

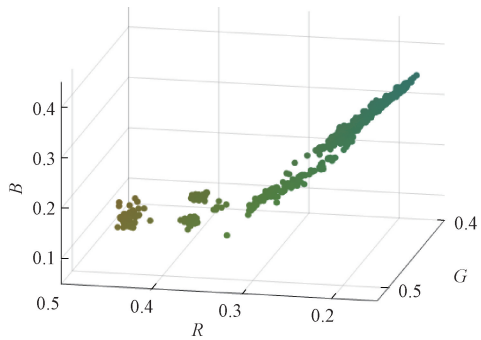

图 5　cube+数据集光源 gt 值分布情况

Fig. 5　Distribution of gt values in cube + dataset

2.3　光源转换模块 Trans

从图像中提取到内容编码和光源编码后,输入到图

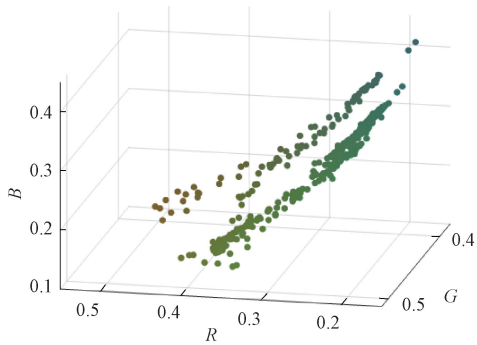

图 6　Gehler-Shi 数据集光源 gt 值分布情况

Fig. 6　Distribution of gt values in Gehler-Shi dataset

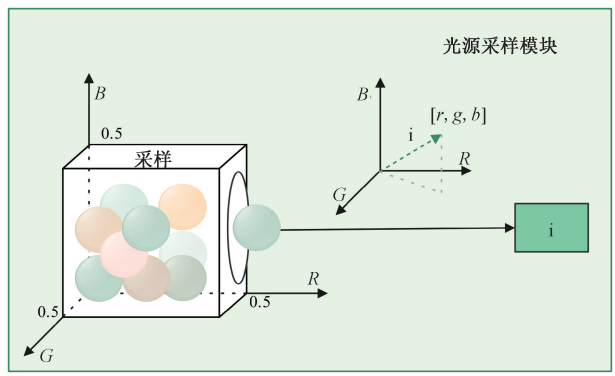

图 7　光源采样模块示意图

Fig. 7　Diagram of illuminant sampling module

8 所示的光源转换模块 Trans 中,内容编码 C 直接输入,目标光源编码 i 作为前 3 层输入充分地与内容编码进行重组结合,在 i 的示例引导下最终生成目标光源下的图像。具体过程为:同时输入标准光源下图像 I_y 和非标准光源下图像 I_x,分别通过内容编码模块得到内容编码 C_y 和 C_x,通过光源编码模块 Ei 得到光源编码 i_y 和 i_x,采样

得到光源编码 i_x1。通过交换非标准和标准光源编码,重组这些内容和光源编码。输入 C_x 和 i_y,光源转换模块 Trans 重组成一个具有光源 i_y 内容 C_x 的图像 I_y';同理,输入 C_y 和 i_x、C_y 和 i_{x1},光源转换模块 Trans 分别把它们重组成一个具有光源 i_x 内容 C_y 的图像 I_x'和一个具有光源 i_{x1} 内容 C_y 的图像 $I_{x1'}$。此时通过两种非标准光源编码 i_x 和 i_{x1} 的引导,将标准光源下的图像 I_y 映射成具有相同内容但在两种不同非标准光源下的图像 I_x 和 I_x',实现了一对多的映射任务,同时将非标准光源图像 I_x 映射成,保持同样内容在标准光源下的图像 I_y'。通过互换光源编码的思想,完成光源转换模块 Trans 自由地将内容编码和光源编码重组的工作,实现了在输入指定光源编码时映射到指定光源下的任务。

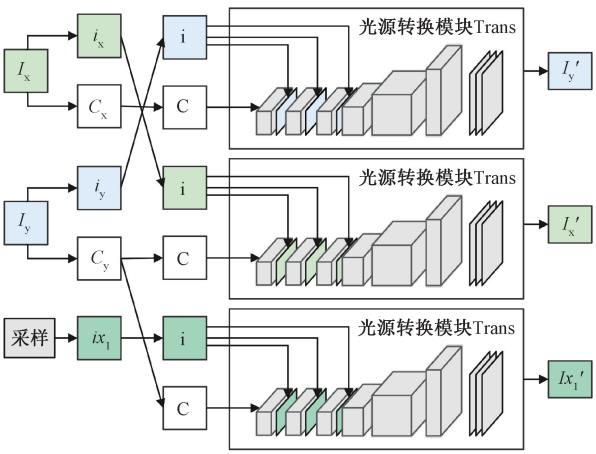

图 8　光源转换模块示意图

Fig. 8　Diagram of illuminant conversion module

2.4　光源监督模块 Di

在光源转换模块 Trans 重组的过程中,会将一个内容编码与多个不同的光源编码进行重组,为了帮助 Trans 模块在输入指定光源编码时生成指定光源下的图像,设计加入图 9 所示的光源监督模块 Di。该模块借助于生成对抗网络判别器的思想,输入为一幅图像,输出为估计的光源值 i′[r′g′,b′]。通过将 Trans 模块生成的图像输入到 Di 模块,再把 Di 模块计算出来的估计值和真值之间的损失回传给 Trans 模块,来帮助 Trans 模块拥有在输入目标光源编码时生成对应光源下图像的能力。而针对光源监督模块的训练,由于估计的光源值实则是 RGB 三维空间中的一个向量,因此将估计光源值 i′和真实光源值 i 的角度误差作为损失函数,如式(2)所示,最终模块的损失为式(3)~(6)所示的光源鉴别损失 $LD_illum(E_c,E_i,Trans,D_i)$。

$$L_fuction(\boldsymbol{a},\boldsymbol{b}) = \cos^{-1}\left(\frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\|\boldsymbol{a}\| \times \|\boldsymbol{b}\|}\right) \tag{2}$$

$$I_y' = Trans(E_c(I_x),E_i(I_x)) \tag{3}$$

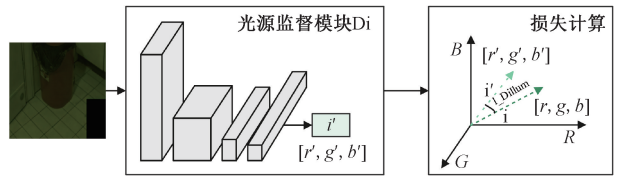

图 9　光源监督模块示意图

Fig. 9　Diagram of illuminant supervision module

$$I_x' = Trans(E_c(I_y),E_i(I_y)) \tag{4}$$

$$I_{x1}' = Trans(E_c(I_y),i_{x1}) \tag{5}$$

$$LD_illum(E_c,E_i,Trans,D_i) = L_fuction(D_i(I_x),label_x) + L_fuction(D_i(I_y),label_y) + L_fuction(D_i(I_x'),label_x) + L_fuction(D_i(I_y'),label_y) + L_fuction(D_i(I_{x1}'),label_{x1}) \tag{6}$$

其中,式(2)为计算向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 的角度误差函数,式(3)~(5)分别为图像 $I_{x'}$、$I_{y'}$、$I_{x1'}$ 的生成过程。在式(3)~(6)中,I_x 代表非标准光源下图像,I_y 代表标准光源下图像,$I_{x'}$ 和 $I_{x1'}$ 代表生成的非标准光源下图像,$I_{y'}$ 代表生成的标准光源下图像,i_{x1} 代表光源采样模块采样得到的非标准光源;$E_c(.)$、$E_i(.)$、$Trans(.)$、$D_i(.)$分别代表内容编码模块、光源编码模块、光源转换模块、光源监督模块的输出结果;label 为光源的 RGB 值,其下标 x、x_1、y 分别代表非标准光源、采样的非标准光源和标准光源,$LD_illum(E_c,E_i,Trans,D_i)$ 表示该损失应用于 E_c、E_i、$Trans$ 和 D_i 模块。

2.5　损失组成

本文借助生成对抗网络的思想,加入了区分图像是真实还是生成的鉴别器 Dr,与整个生成图像的网络框架对抗训练,保证生成图像的真实性,其对抗损失 $L^{adv}_{real_or_fake}(E_c,E_i,Trans,D_r)$ 如式(7)所示。

$$L^{adv}_{real_or_fake}(E_c,E_i,Trans,D_r) = E_{x\sim Pdata(x)}\left[\frac{1}{2}\log D_r(I_x) + \frac{1}{2}\log(1 - D_r(Trans(E_c(I_x),E_i(I_x))))\right] + E_{y\sim Pdata(y)}\left[\frac{1}{2}\log D_r(I_y) + \frac{1}{2}\log(1 - D_r(Trans(E_c(I_y),E_i(I_y))))\right] \tag{7}$$

同时引入循环一致性思想,图像 I 中提取出的内容编码和光源编码,经过光源转换模块 Trans 重组后的图像 I′,应与原图像 I 保持一致,构成图像重构损失 $L_{image_recon}(E_c,E_i,Trans)$,如式(8)所示;同时,对于生成的图像 I′,如果再次输入到内容编码模块 E_c 后,得到的内容编码 C′应与原来的内容编码 C 保持一致,构成内容重构损失 $L_{content_recon}(E_c,E_i,Trans)$,如式(9)所示,最终两个损失构成式(10)所示的重构损失 $L_{recon}(E_c,E_i,Trans)$。

$$L_{image_recon}(E_c,E_i,Trans) = \| I_x - Trans(E_c(I_y'),i_x) \|_1 + \| I_y - Trans(E_c(I_x'),i_y) \|_1 + \| I_y - Trans(E_c(I_{x1}'),i_y) \|_1 \quad (8)$$

$$L_{content_recon}(E_c,E_i,Trans) = \| E_c(I_x) - E_c(I_y') \|_1 + \| E_c(I_y) - E_c(I_x') \|_1 \quad (9)$$

$$L_{recon}(E_c,E_i,Trans) = L_{content_recon} + L_{img_recon} \quad (10)$$

其中，$\| \cdots \|_1$ 代表 L1 loss 损失函数。

因此网络最终使用的总损失 $L(E_c,E_i,Trans,D_r,D_i)$ 包括 3 个部分：光源鉴别损失 $L_{D_illum}(E_c,E_i,Trans,D_i)$、对抗损失 $L^{adv}_{real_or_fake}(E_c,E_i,Trans,D_r)$ 和重构损失 $L_{recon}(E_c,E_i,Trans)$，每个损失再设置适当权重，如式(11)所示。其中对抗损失权重 λadv 取值不宜过大，范围为[1~3]，光源鉴别损失权重 λD_i 取值范围为[3~10]，重构损失权重 λrec 为主要约束，取值范围为[5~20]。

$$L(E_c,E_i,Trans,D_r,D_i) = \lambda_{adv} \cdot L^{adv}_{real_or_fake}(E_c,E_i,Trans,D_r) + \lambda_{Di} \cdot LD_illum(E_c,E_i,Trans,D_i) + \lambda_{rec} \cdot L_{recon}(E_c,E_i,Trans) \quad (11)$$

3　实验细节与配置

3.1　数据集

在单光源的颜色恒常性数据集中，仅仅提供非标准光源下的图像和光源真值 gt，并没有直接给出对应的标准光源下的图像，而本文网络端到端的训练方式需要同时用到非标准光源下和标准光源下图像。因此假设光源均匀分布，利用对角模型得到标准光源下的图像，最终将非标准光源下图像和处理后得到的标准光源下图像作为均匀数据集。而为了验证本文方法在单光源非均匀分布情况下的有效性，利用现有数据集渲染光源非均匀分布的情况，作为非均匀数据集验证本文方法。

1)均匀数据集

本文在两种广泛使用的公开数据集上进行实验：Gehler-Shi 数据集[25]和 cube+数据集[26]。Gehler-Shi 数据集利用 Canon 1D 和 Canon 5D 两种相机收集了 568 张室内和室外场景的图像；cube+数据集仅使用 Canon 550D 相机收集了 1 736 张图像，包括 1 365 张室外图像和 342 张室内图像。每个数据集均提供了真实的光源值 gt，根据提供的非标准光源下图像和 gt 值，利用对角模型校正得到标准光源下的图像。

每个数据集在实验中进行数据增强，包括以 0.5 的概率水平翻转和 0.5 的概率垂直翻转。在消融实验中使用均匀 cube+数据集，对其进行随机划分，80% 的图像作为训练集 20% 作为测试集；在与其他方法对比实验中，数据集采用三折交叉验证训练方式，实验结果是在两个数据集上分别进行得到的。

2)非均匀数据集

在 cube+和 Gehler-Shi 数据集上进行处理，均匀数据集已经获得了全部标准光源下的图像，因此只需在标准光源图像上渲染非均匀分布的光源值，即可得到非均匀图像。对于非标准光源图像来说，相对更接近光源的像素点颜色受光源的影响越大。因此设计非均匀渲染的方式如图 10 所示，随机选择图像上一点为光源中心(center_x,center_y)，将这个点渲染光源值 i_center，剩下所有点的光源值按照高斯分布从光源中心向外逐渐衰减，每一点(x,y)光源值 i=count_i (i_center, x,y)，其中 count_i 函数为以(center_x, center_y)为中心的二维高斯分布函数。

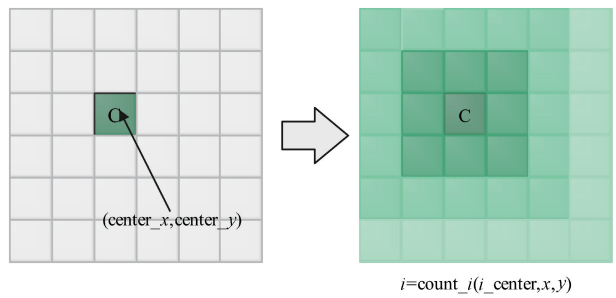

图 10　光源分布示意图

Fig. 10　Illuminant distribution diagram

按照上述方式，对 cube+数据集和 Gehler-Shi 数据集进行渲染，最终得到非均匀 cube+数据集和非均匀 Gehler-Shi 数据集，如图 11 和 12 所示，第 1 列为标准光源下图像，第 2 列为非标准光源下图像，第 3 列为光源分布图。在消融实验和与其他方法对比实验中，对数据集进行随机划分，其中 80% 的图像作为训练集 20% 作为测试集，实验结果是在两个数据集上分别进行得到的。

3.2　实验细节和评价指标

本文采用 Pytorch 框架，以下实验均在 GTX 2070 GPU 的环境下进行。

1)均匀数据集

在均匀数据集下，实验的评价指标为真实光源 i 与估计光源 $\hat{i}$ 之间的角度误差，如式(12)所示。

$$L(i,\hat{i}) = \cos^{-1}\left(\frac{i \cdot \hat{i}}{\| i \| \times \| \hat{i} \|}\right) \quad (12)$$

本文采用的方法为直接对图像进行校正，输出结果为校正后的图像，无法直接得到估计的光源值，而角度误差又是被广泛使用的衡量指标，因此需要利用以下公式得到估计的光源值：

$$\hat{i} = I_x * I_y^{-1} \quad (13)$$

其中，I_x 为输入的非标准光源下图像，I_y 为通过网络

图11　非均匀cube+数据集示意图

Fig. 11　Diagram of non-uniform cube + dataset

图12　非均匀Gehler-Shi数据集示意图

Fig. 12　Diagram of non-uniform Gehler-Shi dataset

校正后的标准光源下图像。再对测试集的所有角度误差计算统计值,包括均值、三均值、中值、最好25%的均值和最差25%的均值。

2)非均匀数据集

由于非均匀数据集光源不均匀分布,无法利用式(11)获得估计的光源值,因此无法计算角度误差,不能利用这个指标。而且在颜色恒常性这个任务中,希望校正后的图像无限接近标准光源下图像,因此采用如下3个评价指标:峰值信噪比(peak signal to noise ratio, PSNR)、均方误差(mean square error, MSE)和结构相似性(structural similarity index, SSIM)。

MSE如式(14)所示,它计算图片张量中对应位置的每个值的平方误差,然后对所有的值求和取平均,MSE值越小,说明预测得到的像素与真实的像素之间的误差越小,也更加相似;PSNR如式(15)所示,是一种评价图像质量的客观标准,PSNR值越大,就代表失真越少;SSIM如式(16)~(19)所示,是一种用于度量两张图片相似度的指标,SSIM的取值范围为[-1,1],它的值越接近于1,图像间的相似度越高。

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} \| X(i,j) - Y(i,j) \|^2 \tag{14}$$

$$PSNR = 10 \cdot \lg\left(\frac{MAX_I^2}{MSE}\right) \tag{15}$$

$$L(X,Y) = \frac{2u_x u_y + C_1}{u_x^2 + u_y^2 + C_1} \tag{16}$$

$$C(X,Y) = \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \tag{17}$$

$$S(X,Y) = \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3} \tag{18}$$

$$SSIM = L(X,Y)^{\partial} \cdot C(X,Y)^{\beta} \cdot S(X,Y)^{\gamma} \tag{19}$$

其中,式(14)为计算图像X和Y的MSE值,$X(i,j)$和$Y(i,j)$代表两幅图像对应位置像素点的值,式(15)用来计算两幅图像的PSNR值,MAX代表两幅图像所有像素点中的最大值,式(16)~(19)为计算图像X和图像Y的SSIM值,其中ux,uy分别为X和Y的均值,σx^2,σy^2分别为X和Y的方差,σxy为X和Y的协方差,C_1,C_2,C_3,∂,β,λ为常数。对所有测试集的MSE、PSNR和SSIM进行统计,计算他们的平均值来比较不同算法之间的优劣。

3.3　消融实验

为证明网络框架加入各个模块的必要性,在均匀的cube+数据集和非均匀的cube+数据集上分别进行如下消融实验,在直接校正的方法中选择cycleGAN作为基线。均匀cube+数据集的结果如表1所示,非均匀cube+数据集的结果如表2所示,用加粗代表最优,定性结果分别如图13和14所示。

表1　均匀cube+数据集上的消融实验结果

Table 1　Ablation results of uniform cube + dataset

实验组	平均值	中值	三均值	最好25%	最差25%
基线	2.76	2.16	2.47	0.62	6.43
实验1	1.34	0.99	1.16	0.28	3.03
实验2	1.36	1.02	1.16	0.28	3.11
实验3	1.10	0.79	0.92	0.22	2.52

表2　非均匀cube+数据集上的消融实验结果

Table 2　Ablation results of nonuniform cube + dataset

实验组	PSNR	MSE	SSIM
基线	29.03	71.23	0.72
实验1	30.87	50.83	0.81
实验2	29.61	53.35	0.78
实验3	32.58	41.91	0.85

●基线　采用 cycleGAN 网络

●实验 1　内容编码模块 Ec+光源编码模块 Ei(其中无光源采样模块)+光源转换模块 Trans

●实验 2　在实验 1 基础上加入光源采样模块

●实验 3　在实验 2 基础上加入光源监督模块 Di,为实验最终框架

根据表中的结果可以看出,无论是在均匀数据集上还是非均匀数据集上,实验 1 的所有量化结果均远远优于基线,证明在假设图像内容空间和光源空间相互独立的情况下,将图像拆分成内容编码加光源编码,再根据目标进行重组确实有效,可以保证图像内容不变的同时仅仅改变图像所处的光源。而实验 2 在加入光源采样模块后,角度误差的平均值、中值和最差 25%均值均不如实验 1 的结果,PSNR 和 SSIM 值低于实验 1,MSE 值高于实验 1,证明实验 2 最终生成的图像不如实验 1 贴近标准光源下图像。但在实验 3 加入光源监督模块后,效果大幅度提升,为所有实验中的最优结果。说明单纯地加入光源采样模块,网络虽然可以学习到更多非标准光源的情况,但也因采样较大的随机性使学习更加复杂,导致效果不提升反下降。因此实验 3 在加入可以区分光源的光源监督模块后,帮助网络更好地学习多种非标准光源,实现多对一映射。充分证明了实验设计加入的模块都有其必要性,共同帮助网络实现最终的最优效果。

图 13　均匀 cube+数据集上的消融实验定性结果对比

Fig. 13　Qualitative results of ablation experiments on a uniform cube + dataset

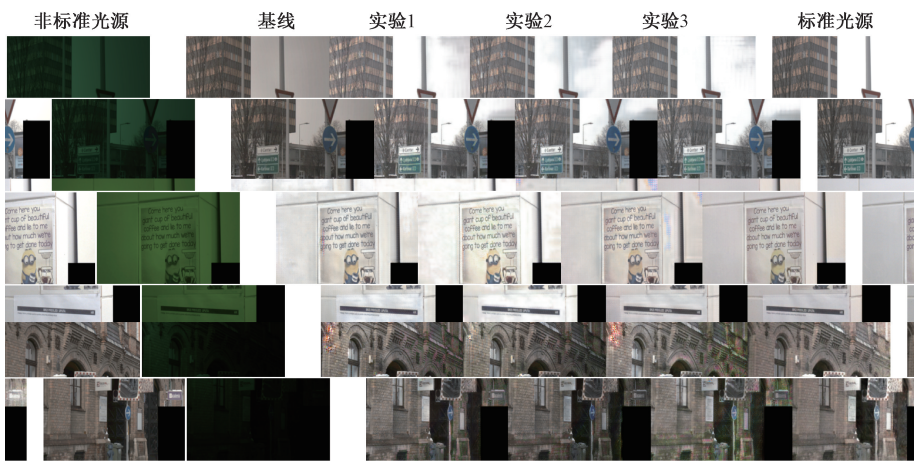

图 14　非均匀 cube+数据集上的消融实验定性结果对比

Fig. 14　Qualitative results of ablation experiments on an inhomogeneous cube + dataset

在如图 13、14 所示的定性结果中,最左边一列为非标准光源下的图像,中间 4 列分别为 4 个实验校正的结果,最右边一列为标准光源下图像。对于均匀 cube+数据集上的实验,在第 1 行的结果中,基线生成的图像模糊,存在大量不平滑区域以及马赛克区域,实验 1、实验 2 和实验 3 相比来说清晰度更好,但实验 1 和实验 2 分别

在椭圆框的车尾灯区域,产生了纹理模糊的情况,只有实验3更接近标准光源下图像,较好地保留了图像的细节。在后两行的实验结果中,基线和实验2在方块区域都产生较大程度的马赛克,实验1也出现少许马赛克,只有实验3几乎没有马赛克区域,与标准光源下图像最为接近。

对于非均匀cube+数据集上的实验,在第1行的结果中,基线相比其他3个实验校正后的图像与真值色差最大,效果最差,但实验1、2和3的结果都在天空区域内出现不同程度的不均匀分布现象,没有完全处理好光源不均匀分布的情况,但相较来说,实验3这种现象更加少,校正结果更加接近标准光源下图像。在第2行室内图像的校正结果中,基线、实验1和实验2校正后的图像表面都出现一些不平滑的分布,实验2还出现少许马赛克区域,只有实验3的结果比较好地处理了光源非均匀分布的问题,图像细节更接近标准光源下图像。同样地,在第3行的结果中,基线、实验1和实验2的结果存在马赛克、色彩不接近真值的问题,实验3的结果更加贴近目标图像。通过以上分析,实验3的最终框架在非均匀和均匀数据集上都达到了最优的效果,再次证明实验加入的每个模块都在帮助网络更好的学习图像校正,缺一不可。

3.4 非均匀数据集下的其他方法对比

在创建的非均匀cube+数据集和Gehler-Shi数据集上分别进行实验,与经典方法和先进方法的对比结果如表3所示,用加粗字体代表最优,下划线代表次优。以横线为分隔符,前面部分为估计光源的方法,后面部分为直接图像校正的方法。

表3 在非均匀cube+和非均匀Gehler-Shi数据集上现有方法的对比实验结果

Table 3 Comparative experimental results of existing methods on non-uniform cube + and non-uniform Gehler-Shi datasets

方法	非均匀cube+			非均匀Gehler-Shi		
	PSNR	MSE	SSIM	PSNR	MSE	SSIM
white-patch[8]	23.18	99.07	0.47	21.91	105.06	0.53
grey-world[9]	23.14	99.97	0.31	21.93	105.74	0.46
shades-of-gray[27]	23.14	100.00	0.38	21.93	105.21	0.42
FFCC[15]	25.23	97.95	0.67	24.76	98.71	0.54
Fc4[16]	23.96	99.12	0.61	22.05	102.39	0.47
MDLCC[18]	24.37	97.65	0.62	23.58	98.60	0.53
C4[19]	24.89	98.13	0.65	23.93	99.87	0.49
pix2pix[21]	29.91	69.19	0.77	29.45	69.89	0.66
cycleGAN[22]	29.03	71.23	0.72	28.64	75.74	0.60
AngularGAN[24]	30.15	66.11	0.80	30.32	54.42	0.71
本文	**32.58**	**41.91**	**0.85**	**31.97**	**47.29**	**0.81**

在两种非均匀数据集上,前3种经典算法仅是基于简单的统计学规律,因此在面对非均匀分布的复杂情况时表现最差。而FFCC、Fc4、MDLCC和C4算法面对非均匀情况,也是先估计光源向量再校正图像,无法适应非均匀的情况,因此定量结果远不如几种直接校正的算法。同时,本文算法在几种直接校正的算法中也表现最为优异,也是所有算法中的最优,PSNR、MSE和SSIM值都优于其他算法,证明本文算法生成的图像失真最小,与真实像素点的误差最小,结构相似度最高。证明了本文的算法较好地解决了光源非均匀分布的问题,超越了之前的算法。

在非均匀cube+数据集上,与其他方法对比的定性结果如图15所示,其中对于估计光源向量的两种算法white-patch和FFCC都没有解决光源非均匀分布的问题,white-patch无法根据简单的规律估计出光源,导致校正后的结果偏暗,与标准光源下图像差距较大;FFCC算法面对非均匀分布的情况,无法准确定位光源,校正后的图像还存在偏绿的情况,没有将图像校正到标准光源下。而在直接校正的算法中,pix2pix和cycleGAN算法没有很好地解决光源非均匀分布的问题,校正后的图像仍旧出现不连续分布的情况,而AngularGAN和本文算法可以很好地解决光源非均匀分布的问题,校正后的图像更加接近标准光源下的图像。

在非均匀Gehler-Shi数据集数据集上,与其他法对比的定性结果如图16所示,先估计光源的两种算法white-patch和FFCC校正后的图像与标准光源下图像的差距最大。而对于直接校正的算法,在第1行的结果中,cycleGAN算法校正的图像与标准光源下图像色差较大,pix2pix和AngularGAN算法校正后的结果也存在一定色差,本文算法的结果最贴近标准光源下图像。而在第2行、第3行光源比较暗的极端情况下,pix2pix、cycleGAN和AngularGAN算法都没有很好解决光源非均匀分布的问题,本文的算法依旧保持了比较好的性能,生成的图像最接近标准光源下图像。通过以上充分说明了在单光源非均匀分布的颜色恒常性问题上本文算法的表现最为优异,超越了之前的算法。

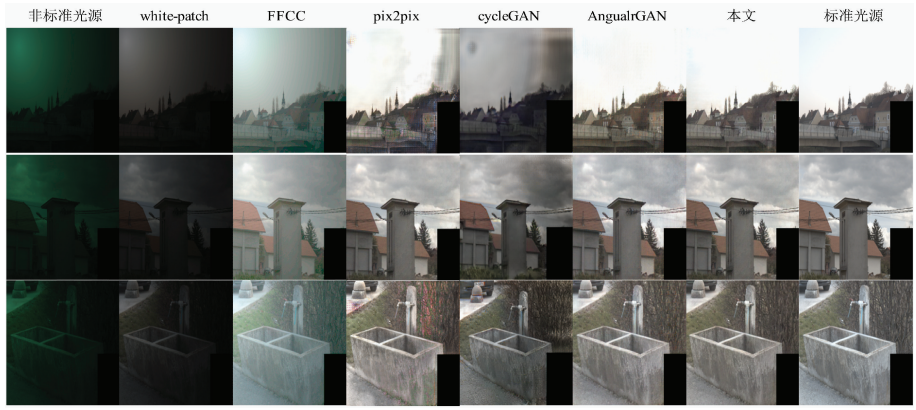

图 15　在非均匀 cube+数据集上域其他方法对比结果

Fig. 15　The results are compared by other methods on the non-uniform cube + dataset

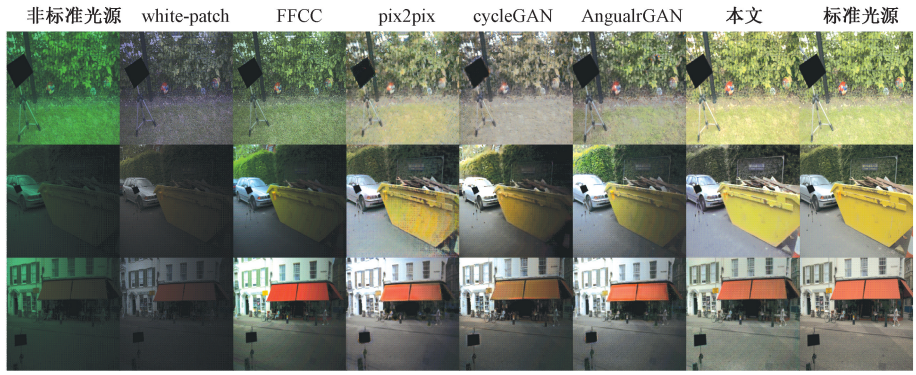

图 16　在非均匀 Gehler-Shi 数据集上域其他方法对比结果

Fig. 16　The results were compared with other methods in the non-uniform Gehler-SHI dataset

3.5　均匀数据集下的其他方法对比

虽然本文主要是针对单光源非均匀分布的情况,但也在两个均匀数据集上分别进行了实验,与经典方法和先进方法的对比结果如表 4 和 5 所示,用加粗字体代表最优,下划线代表次优。以横线为分隔符,前面部分为估计光源的方法,后面部分为直接图像校正的方法。

表 4　在 cube+和 Gehler-Shi 数据集上现有方法的对比实验结果

Table 4　Comparative experimental results of existing methods on cube + and Gehler-Shi datasets

方法	Gehler-Shi					cube+				
	平均值	中值	三均值	最好 25%	最差 25%	平均值	中值	三均值	最好 25%	最差 25%
white-patch	7. 55	5. 68	6. 35	1. 45	16. 12	6. 80	3. 85	5. 21	0. 68	16. 93
grey-world	6. 36	6. 28	6. 28	2. 33	10. 58	3. 52	2. 55	2. 82	0. 60	7. 98
Bayesian[28]	4. 82	3. 46	3. 88	1. 26	10. 49	–	–	–	–	–
General Gray-World	4. 66	3. 48	3. 81	1. 00	10. 09	3. 22	2. 12	2. 44	0. 43	7. 77
Fc4	1. 80	0. 95	1. 18	0. 27	4. 65	1. 35	0. 93	1. 01	0. 30	3. 24
FFCC	1. 61	**0. 86**	1. 02	**0. 23**	4. 27	1. 38	**0. 74**	**0. 89**	**0. 19**	3. 67
MDLCC	1. 58	0. 95	1. 11	0. 37	3. 77	1. 24	0. 83	0. 92	0. 26	2. 91
C4	1. 35	0. 88	**0. 99**	0. 28	3. 21	–	–	–	–	–
pix2pix	2. 84	2. 32	2. 59	0. 71	5. 99	2. 45	2. 01	2. 31	0. 67	5. 05
cycleGAN	2. 90	2. 06	2. 62	0. 62	6. 59	2. 76	2. 16	2. 47	0. 62	6. 43
AngularGAN	2. 56	2. 05	2. 43	0. 58	5. 23	2. 27	1. 46	1. 98	0. 67	4. 88
本文	**1. 31**	1. 21	1. 33	0. 44	**2. 31**	**1. 10**	0. 79	0. 92	0. 22	**2. 52**

在 cube+均匀数据集上,本文方法的平均值和最差 25%的均值是所有方法中的最优,证明了本文方法的平均误差最小,同时在最差的情况下也保持了不错的性能,其他指标也做到了次优的结果,而且本文方法的指标也

远超其他直接校正的方法,整体上取得了较优的结果;在Gehler-Shi的均匀数据集上,本文方法的平均值和最差25%的均值也是所有方法中的最优,再次证明本文的方法可以保证整体的性能。以上结果说明,本文的方法不仅在非均匀情况下达到了最优的效果,同时在均匀数据集上也实现了有竞争力的性能,证明了本文方法在颜色恒常性任务上的有效性,不仅可以较好地解决光源非均匀分布的问题,也可以兼顾均匀情况下的性能。

4 结　论

针对光源非均匀分布的问题,面向单光源颜色恒常性任务,本文设计了一种直接校正的多对一映射网络框架。框架通过端到端输入非标准光源图像输出标准光源图像的方式,省略中间过程解决非均匀分布的问题。算法假设内容空间和光源空间相互独立,分离图像的内容编码和光源编码,保持内容编码不变改变光源编码为标准光源编码,将图像映射到标准光源下。为了丰富光源信息,加入光源采样模块,同时,加入光源监督模块来帮助网络重组特定光源编码和内容编码,实现颜色恒常性。实验结果表明,相较于其他先进算法,本文算法更好地解决了光源非均匀分布的问题。下一步的工作将在本文研究的基础上,聚焦于场景中包含多个光源且光源非均匀分布的情况,同时解决颜色恒常性的两大难题,真正贴近真实场景。

参考文献

[1] 高绍兵. 视觉颜色恒常和自适应性计算模型与应用研究[D]. 成都:电子科技大学,2017.
GAO SH B. Computational models of visual adaptation and color constancy and applications [D]. Chengdu: University of Electronic Science and Technology of China, 2017.

[2] WITZEL C, FLACK Z, SANCHEZ-WALKER E, et al. Colour category constancy and the development of colour naming[J]. Vision Research, 2021, 187: 41-54.

[3] FINLAYSON G, HORDLEY S, SCHAEFER G, et al. Illuminant and device invariant colour using histogram equalisation[J]. Pattern Recognition, 2005, 38(2): 179-190.

[4] 王延年,赵朗月,刘会敏. 基于混合颜色空间的反应孔图像分割算法研究[J]. 国外电子测量技术, 2017, 36(12):20-23.
WANG Y N, ZHAO L Y, LIU H M. Segmentation algorithm of reaction holes image based on lab and HSV color spaces [J]. Foreign Electronic Measurement Technology, 2017, 36(12):20-23.

[5] 王森,伍星,张印辉,等. 基于多尺度降采样规范化割的图像裂纹检测[J]. 仪器仪表学报, 2017, 38(11): 2788-2796.
WANG S, WU X, ZHANG Y H, et al. Image crack detection with multi-scale down-sampled normalized cut[J]. Chinese Journal of Scientific Instrument, 2017, 38(11): 2788-2796.

[6] 罗冬梅,左金水,余文森. 基于双特征融合与自适应提升机制的图像动作识别算法[J]. 电子测量与仪器学报, 2017, 31(12): 1929-1936.
LUO D M, ZUO J SH, YU W S. Motion recognition algorithm based on double feature fusion and adaptive boosting mechanism [J]. Journal of Electronic Measurement and Instrumentation, 2017, 31(12): 1929-1936.

[7] 张璇,郝明磊,行鸿彦,等. 基于三维匹配块的SAR图像相干斑抑制研究[J]. 电子测量与仪器学报, 2017, 31(10): 1621-1626.
ZHANG X, HAO M L, XING H Y, et al. Research on de-speckling SAR image based on 3D matching block [J]. Journal of Electronic Measurement and Instrumentation, 2017, 31(10):1621-1626.

[8] BUNCHSBAUM G. A spatial processor model for object colour perception[J]. Journal of the Franklin Institute, 1980, 310(1): 1-26.

[9] BRAINARD D H, WANDELL B A. Analysis of the retinex theory of color vision [J]. JOSA A, 1986, 3(10): 1651-1661.

[10] VAN DE WEIJER J, GEVERS T, GIJSENIJ A. Edge-based color constancy[J]. IEEE Transactions on Image Processing, 2007, 16(9): 2207-2214.

[11] HERNANDEZ-JUAREZ D, PARISOT S, BUSAM B, et al. A multi-hypothesis approach to color constancy [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,2020: 2270-2280.

[12] LAAKOM F, RAITOHARJU J, IOSIFIDIS A, et al. Probabilistic color constancy [C]. 2020 IEEE International Conference on Image Processing (ICIP). IEEE, 2020: 978-982.

[13] 杨泽鹏,解凯,李桐,等. 多通道置信度加权颜色恒常性算法[J]. 光学学报,2021,41(11):242-252.
YANG Z P, JIE K, LI T, et al. Color constancy with multi-channel confidence-weighted method [J]. Acta Optica Sinica,2021,41(11):242-252.

[14] BARRON J T. Convolutional color constancy [C]. Proceedings of the IEEE International Conference on Computer Vision, 2015: 379-387.

[15] BARRON J T, TSAI Y T. Fast Fourier color constancy [C]. Proceedings of the IEEE Conference on Computer Vision

and Pattern Recognition, 2017: 886-894.
[16] HU Y, WANG B, LIN S. Fc4: Fully convolutional color constancy with confidence-weighted pooling [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4085-4094.
[17] LO Y C, CHANG C C, CHIU H C, et al. CLCC: Contrastive learning for color constancy [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 8053-8063.
[18] XIAO J, GU S, ZHANG L. Multi-domain learning for accurate and few-shot color constancy [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 3258-3267.
[19] YU H, CHEN K, WANG K, et al. Cascading convolutional color constancy [C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12725-12732.
[20] DAS P, BASLAMISLI A S, LIU Y, et al. Color constancy by GANs: An experimental survey [J]. arXiv preprint arXiv: 1812.03085, 2018.
[21] ISOLA P, ZHU J Y, ZHOU T, et al. Image-to-image translation with conditional adversarial networks [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1125-1134.
[22] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 2223-2232.
[23] CHOI Y, CHOI M, KIM M, et al. Stargan: Unified generative adversarial networks for multi-domain image-to-image translation [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8789-8797.
[24] SIDOROV O. Conditional GANs for multi-illuminant color constancy: Revolution or yet another approach? [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2019.
[25] GIJSENIJ A, GEVERS T. Color constancy using natural image statistics and scene semantics [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 33(4): 687-698.
[26] 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]. JOSA A, 2014, 31(5): 1049-1058.
[27] FINLAYSON G D, TREZZI E. Shades of gray and colour constancy [C]. Color and Imaging Conference. Society for Imaging Science and Technology, 2004, 2004(1): 37-41.
[28] GEHLER P V, ROTHER C, BLAKE A, et al. Bayesian color constancy revisited [C]. 2008 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2008: 1-8.

作者简介

谢林芳，2019 年于合肥工业大学获得学士学位，现为合肥工业大学计算机与信息学院硕士研究生，主要研究方向为图像处理。
E-mail：xielinfang@mail.hfut.edu.cn

Xie Linfang received her B. Sc. degree from Hefei University of Technology in 2019. She is currently a M. Sc. candidate in the School of Computer Science at Hefei University of Technology. Her main research interest includes image processing.

张旭东（通信作者），1989 年于合肥工业大学获学士学位 1992 年于合肥工业大学获硕士学位；2005 年于中国科学技术大学获博士学位。现为合肥工业大学教授，主要研究方向为图像处理、模式识别、智能信息处理。
E-mail：xudong@hfut.edu.cn

Zhang Xudong（Corresponding author）received his B. Sc. degree from Hefei University of Technology in 1989, M. Sc. degree from Hefei University of Technology in 1992 and Ph. D. degree from University of Science and Technology of China in 2005, respectively. His main research interests include image processing, pattern recognition and intelligent information processing.