

# 基于图像感知与分割的自动背景虚化算法

刘成民, 叶武剑\*, 刘怡俊

广东工业大学信息工程学院, 广东 广州 510006

**摘要** 针对单目移动设备难以拍摄出背景虚化图像的问题, 结合深度学习的图像感知与分割算法, 提出了一种图像自动背景虚化算法。基于深度学习的图像感知与分割算法得到图像的焦点图、深度图和掩码图三种辅助图。利用辅助图自动确定主体或由用户指定主体, 并计算背景各区域的深度。利用多尺度高斯滤波对背景各区域进行不同程度的虚化处理。将虚化后的各区域与主体融合, 并进行边缘优化, 生成一幅背景虚化图。实验结果表明, 本文算法利用基于深度学习的感知图实现更准确灵活的图像虚化处理, 并能在多种场景下实现对图像的自动定焦或指定区域的重定焦, 产生的虚化效果自然且富有层次感, 能较好地突显主题。

**关键词** 图像处理; 背景虚化; 图像感知; 图像分割; 边缘优化; 自动定焦

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP202259.0410002

## Automatic Background Blurring Algorithm Based on Image Perception and Segmentation

Liu Chengmin, Ye Wujian\*, Liu Yijun

*School of Information Engineering, Guangdong University of Technology,  
Guangzhou, Guangdong 510006, China*

**Abstract** Aiming at the problem that it is difficult for monocular mobile devices to capture background blurring images, this paper proposes an automatic background blurring algorithm based on the image perception and segmentation algorithm of deep learning. We use the image perception and segmentation algorithm of deep learning to obtain three auxiliary images, i. e., the focus map, depth map, and mask map, of the captured image. We use the auxiliary images to automatically determine the subject or specify the subject by the user, and the depth of each area of the background is calculated. Then, the multi-scale Gaussian filter is used to blur each area of the background in different degrees. Finally, the blurred background areas are merged with the subject, and the edges are optimized to finally generate a background blurred image. Experimental results show that this algorithm can realize more accurate and flexible image virtual processing by using perceptual map based on deep learning, and can automatically focus the image subject or refocus the designated area in a variety of scenes. The virtual effect is natural and hierarchical, and can better highlight the image theme.

**Key words** image processing; background blurring; image perception; image segmentation; edge optimization; auto focus

收稿日期: 2021-01-25; 修回日期: 2021-03-02; 录用日期: 2021-03-19

基金项目: 广东省重点领域研发计划(2018B030338001, 2018B010115002)、广东省科技发展专项计划(2016B090904001)、广东省教育厅创新人才项目和广东工业大学青年百人项目(220413548)

通信作者: \*yewjian@126.com

## 1 引言

背景虚化技术又称为浅景深处理技术<sup>[1]</sup>,是摄影技术中常用的表现手法之一。受手机硬件性能的限制,智能手机仅有人像的自动背景虚化处理模式,并且手机获取的背景虚化图效果不佳。此外,目前的主流智能手机多使用双摄像头实现背景虚化,这是因为这类相机利用双目深度估计的方法<sup>[2]</sup>获取图像的深度信息<sup>[3]</sup>,但目前持有单摄像头手机的用户也占有一部分的比例,这类群体同样对照片美化具有很大的需求,而单摄像头手机单次仅能获取一张图像,且无法获取多种图像感知信息。因此,一种应用场景广泛、操作便捷、仅依赖单张图像作为输入的背景虚化算法成为了计算机视觉的研究热点。

随着背景虚化处理技术成为计算机视觉研究领域的热点,图像处理领域设置了一系列的赛事,其中 Advanced Intelligent Mechatronics 成为了该领域最具影响力的赛事之一。如 2019 年的 AIM 散景效果合成的挑战赛中出现了多种优秀的背景虚化处理方案<sup>[4]</sup>。

研究背景虚化的算法主要分为两种。

1) 基于景深渲染的方法。增强景深模拟的真实感是景深渲染算法的主要目的,但这需要巨大的计算开销。为实现算法的高性能,文献[5]利用二次光线相干性的聚类方案,提出了一种景深渲染中的二次光线跟踪方法。文献[6]针对可见性问题,提出了一种从稀疏集中合成自由度图像的算法。文献[7]采用分离式的各向异性高斯滤波快速得到渲染结果,实现算法的高稳定性和高速率。文献[8]提出了径向梯度模糊的图像小景深模拟算法,该方法能精确区分聚焦对象区与非聚焦区,使图像获得了自然、高质的小景深模拟效果。上述工作的结果虽逼近单反相机的背景虚化效果,但在多个物体的场景中,处于同一景深区域的物体都是清晰的,因而达不到单独突显某个物体的效果。

2) 基于图像分割的方法。早期的工作<sup>[9]</sup>提出了自动人像分割图像程式化,它将从人像中提取的空间信息注入到全卷积网络(FCN)中,把肖像区域部分扣取出来,再对其背景进行模糊虚化。文献[10]中也采用了同样的方法,他们通过羽化块学习的滤波器具有边缘保持和抠图自适应能力。另外,文献[11]通过引入背景超像素块优化策略,改

善了显著性检测算法对人像中前景区域的检测效果。但这些工作只关注人像照片,先将肖像从图像中分离出来,然后模糊剩余的图像背景,使得最终得到的背景虚化图像过于简单,同时缺乏高灵活度和较好的应用性。

为不限于人像场景,文献[12]结合图像分割网络和自动对焦技术实现对图像的背景虚化。此外,在对背景的模糊虚化中,利用图像的深度感知信息实现背景的过渡性虚化。文献[13]利用了联合强度估计和动态滤波合成的特定任务效能,将其用于空间感知模糊过程。文献[14]依赖于深度估计网络生成混合了输入图像和输入图像的不同平滑版本的权重图,以生成视觉效果良好的背景虚化图。这些工作扩展了应用场景,同时加入了图像深度信息,使得最终的效果更具有真实感。但是与文献[3-6]一样,当图像中有多个物体位于同一景深区域时,该区域的物体都是清晰的,突显不出某个单独的物体。

针对上述工作的问题,本文提出了一种基于图像感知与分割的自动背景虚化算法。本文算法创新性地将图像感知与图像分割技术相结合,仅依赖单张图像作为输入,可获取图像的景深信息与显著性信息,同时为用户提供了一个灵活的主体选择策略。实验结果表明,本算法能在多种场景下实现对图像主体自动定焦或对指定区域的重定焦,产生的虚化效果富有层次感。相比现有的 SOTA 方法,本文算法更具有优越性。

## 2 基本原理

### 2.1 系统框架

图 1 是本文所提出方法的系统框架图,主要分为 4 个模块。辅助图生成模块以原图作为输入,生成焦点图、景深图以及原图中物体的掩码图。景深计算模块首先根据用户的选择或焦点图的指示,确定原图的主体内容,再以景深图为依据,结合掩码图与确定的主体计算图像中各物体与主体之间的相对景深。图像虚化模块根据相对景深,以不同虚化程度对各物体和背景进行虚化处理。最后,在叠加不同虚化程度的物体到最终输出图时,容易出现物体边缘与背景过于突兀的问题,而均值滤波边缘优化模块解决了这一问题,使得输出的背景虚化图整体更加自然,层次合理。

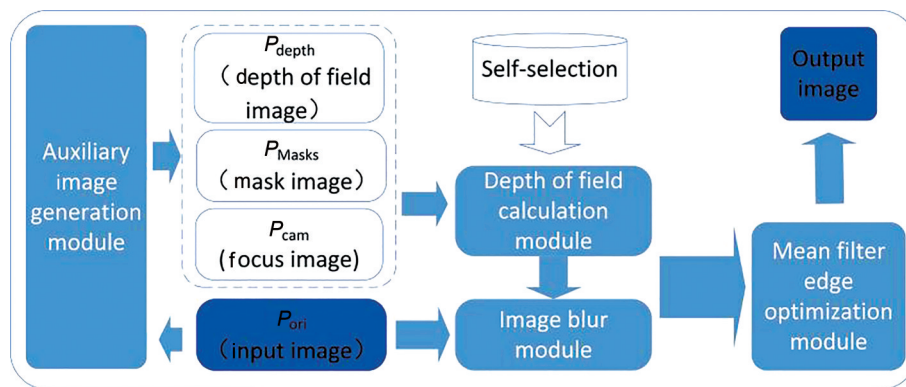


图1 系统框架图

Fig. 1 System framework diagram

### 2.2 辅助图生成模块

辅助图是在背景虚化过程中提供必要信息的图像,主要包括焦点图、景深图,以及图像中各物体的掩码图。焦点图是一张灰度图,在该灰度图上,越靠近图像主体区域,其灰度值越大。而图像主体是含所要表达的语义信息最多、信息量最大的图像

区域。景深图也是一张灰度图,在该灰度图上,距离相机镜头越近的图像区域,其灰度值越小。Mask图也称掩码图,是一张二值图,其中白色区域代表物体所占据的区域,黑色区域代表不属于该物体的区域。原图的各种感知辅助图如图2所示。



图2 辅助图的输出。(a)原图;(b)焦点图;(c)景深图;(d)左边车辆对应的Mask;(e)右边车辆对应的Mask

Fig. 2 Auxiliary image outputs. (a) Original image; (b) focus image; (c) depth of field image; (d) Mask corresponding to vehicle on left; (e) Mask corresponding to vehicle on right

输出所需的三种辅助图是通过将基于深度学习的图像感知算法与分割算法集成于一体,以原图作为输入而生成的。下面介绍辅助图的生成网络。

首先,焦点感知网络主要是使用Zhou等<sup>[15]</sup>提出的类激活映射(CAM)方法获取焦点图,所提出的类激活映射技术结合全局平均池化,能够显示由卷积神经网络(CNN)检测到的语义信息量最多的区域。该算法基于ResNet-18网络,将输入图转化为焦点图,其结构图如图3所示。

深度估计网络是使用Chen等<sup>[16]</sup>提出的野外单

像深度感知网络获得景深图。该算法由一个经过训练的端到端的单一深度网络组成,可以直接估计每像素的深度值。其主要思想是通过构造一个映射原图到像素级深度图上的函数,使得产生的像素级深度值更加精确。

实例分割网络是用加利福尼亚大学提出的一个简单全卷积网络来实现实时实例分割的方法YOLACT<sup>[17]</sup>,该方法的目标是将Mask预测分支添加到单级目标检测网络中,该步骤与Mask R-CNN<sup>[18]</sup>对Faster R-CNN<sup>[19]</sup>所做的工作相似,但减少了定位步骤,使得运行效率更高。为得到Mask图,本文利用YOLACT实例分割算法对原图进行处理。

### 2.3 景深计算模块

本部分主要通过计算图像中各物体的景深,从而有区别地虚化各物体及背景。景深计算模块基于筛选后的Mask图,利用焦点图确定图像主体或由用户自主指定图像主体,再以景深图为基础,结合主体计算出主体与其他物体的相对景深。

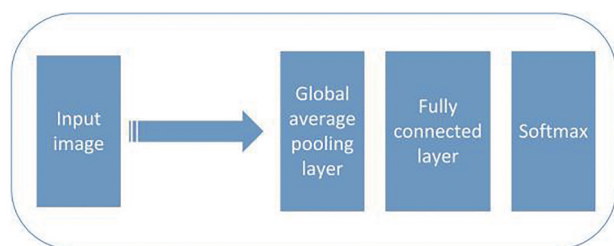


图3 ResNet-18网络结构图

Fig. 3 Diagram of ResNet-18 network structure

### 2.3.1 Mask 图筛选

由于图像分割算法本身的缺陷,生成的Mask中有部分可能是无意义的,如图像中的某个物体离镜头太远,使得Mask面积过小,或者针对同一物体分割出两个重合的Mask。因此,本文提出基于阈值和基于逐像素法的Mask快速筛选方法,用该方法对图像分割结果进行处理,进而筛选出有意义的Mask。

基于阈值进行筛选的步骤:针对面积过小的情

况,计算Mask中白色区域的面积,并设置一个面积阈值  $T$ ,筛除面积小于该阈值的Mask。阈值  $T$  的计算式为

$$T = \frac{\max(S)}{2 \times (W \times H)}, \quad (1)$$

式中:  $T$  为阈值;  $S$  为Mask白色区域面积;  $W$  为图像的宽;  $H$  为图像的长。如图4所示,经过筛选,无意义的Mask被去除。

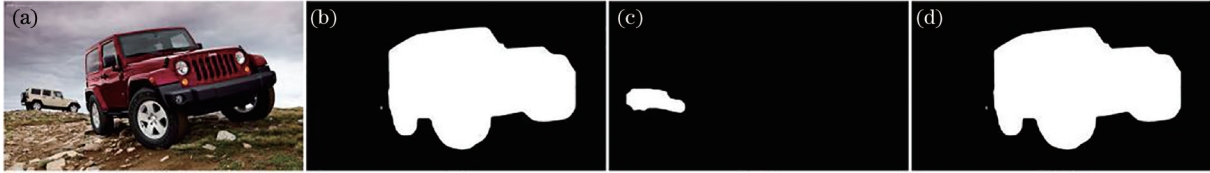


图4 筛除面积过小的Mask。(a)原图;(b)(c)筛选前的Mask;(d)筛选后的Mask

Fig. 4 Filter out Mask that is too small. (a) Original image; (b)(c) Masks before filtering; (d) Mask after filtering

逐像素筛选的步骤:对于重合的Mask,通过采用逐像素法判断两个Mask之间的相似度,检测其是否是同一物体分割出来的两张类似的Mask,并只保留其中面积较小的一个。如图5所示,经过筛选,

只保留一个重合的Mask。

在图像分割之后,将筛选过后的第  $i$  张Mask图记为  $P_{\text{mask}}^i$ ,再令  $K$  等于筛选后的Mask数量,则  $i \in \{0, 1, 2, \dots, K-1\}$ 。

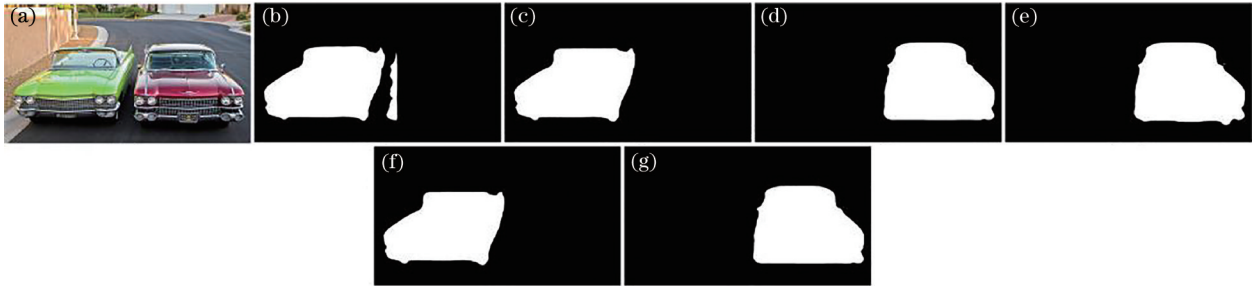


图5 筛除重合的Mask。(a)原图;(b)~(e)筛选前的Masks;(f)(g)筛选后的Masks

Fig. 5 Filter out overlapping Masks. (a) Original image; (b)~(e) Masks before filtering; (f)(g) Masks after filtering

### 2.3.2 确定图像主体

图像主体是背景虚化过程中所要突出的主物体,可能是人物、车辆、动物等。先通过Mask图将原图所包含的所有物体分割开,再通过焦点图确定各物体的焦点权重,最终得到图像的主体。以焦点图为基础,物体的焦点权重计算式为

$$W_{\text{cam}}^i = \frac{\text{sum}(P_{\text{cam}} * P_{\text{mask}}^i)}{\text{sum}(P_{\text{mask}}^i)}, \quad (2)$$

式中:  $W_{\text{cam}}^i$  为第  $i$  个物体的焦点权重;  $P_{\text{cam}}$  为焦点图;  $P_{\text{mask}}^i$  为第  $i$  张Mask图;  $\text{sum}(\cdot)$  为求和运算。

利用(2)式计算出图像的所有Mask所对应物体的焦点权重之后,便可以选择焦点权重最大的物体作为图像的主物体。此外,也可以由用户在筛选后的Mask中直接指定图像的主体。

### 2.3.3 计算相对景深

为了准确计算各物体的景深,在背景虚化过程中对其进行不同程度的虚化,本部分以景深图为基础,得出各物体所对应的景深,景深计算式为

$$D^i = \frac{\text{sum}(P_{\text{depth}} * P_{\text{mask}}^i)}{\text{sum}(P_{\text{mask}}^i)}, \quad (3)$$

式中:  $D^i$  为第  $i$  个物体的景深;  $P_{\text{depth}}$  为景深图;  $P_{\text{mask}}^i$  为第  $i$  张Mask图。

在虚化过程中,以图像主体为原点,离图像主体越远的物体,虚化程度将越大。为此,需计算出图像各物体与主物体的相对景深。本文定义相对景深计算式为

$$W_{\text{depth}}^i = |D^i - D_{\text{chief}}|, \quad (4)$$

式中:  $W_{\text{depth}}^i$  为第  $i$  个物体的相对景深;  $D_{\text{chief}}$  为主物体

的景深。

### 2.4 图像虚化模块

图像虚化模块主要是基于 Mask 图与相对景深,通过控制高斯滤波器的滤波核大小,按不同的虚化程度有序地虚化图像各物体及背景,从而得到层次分明的背景虚化图。具体的实现过程为:先将原图复制多份,将它们记为集合  $T_{ori}$ 。  $T_{ori}$  可表示为

$$T_{ori} \in \{P_{ori}^{bg}, P_{ori}^0, P_{ori}^1, P_{ori}^2, \dots, P_{ori}^{K-1}\}, \quad (5)$$

式中:  $T_{ori}$  为多个原图的集合;  $P_{ori}^{bg}$  为背景区域对应的原图;  $P_{ori}^i$  为筛选后各个 Mask 对应的原图,  $i \in \{0, 1, 2, \dots, K-1\}$ 。

先以一个尺寸较大的滤波核对  $P_{ori}^{bg}$  进行单独处理,得到虚化后的背景图  $P_{blur}^{bg}$ ;再根据  $W_{depth}$  的值逐渐减小滤波核尺寸,对  $P_{ori}^i$  分别进行处理,得到虚化图  $P_{blur}^i$ ;对于主体对应的原图,该模块不对其进行处理,最终得到虚化图像集合  $T_{blur}$ :

$$T_{blur} \in \{P_{blur}^{bg}, P_{blur}^0, P_{blur}^1, P_{blur}^2, \dots, P_{blur}^{K-1}\}, \quad (6)$$

式中:  $T_{blur}$  为处理后的多个原图的集合;  $P_{blur}^{bg}$  为处理后的背景区域对应的虚化图;  $P_{blur}^i$  为处理后的第  $i$  张 Mask 对应的虚化图,  $i \in \{0, 1, 2, \dots, K-1\}$ 。

最后,根据  $P_{mask}^i$  将各个物体扣取出来,再将其融合到最终输出的背景虚化图中,如图 6(a) 所示。融合过程可表示为

$$\begin{cases} P_0 = P_{blur}^{bg} \\ P_{i+1} = P_{blur}^i * P_{mask}^i + P_i * (1 - P_{mask}^i) \end{cases}, \quad (7)$$

式中:  $P_{blur}^{bg}$  为处理后的背景区域对应的虚化图;  $P_{blur}^i$  为处理后的第  $i$  张 Mask 对应的虚化图;  $P_i$  为每一次迭代的输出图。

### 2.5 均值滤波边缘优化模块

由于迭代公式[(7)式]中的  $P_{mask}^i$  作为图像不同部分的加权值,且这个权值只有 0 与 255 两种取值,因此,在由(7)式输出的结果图中,物体出现一个突兀的边缘轮廓。为解决此问题,该模块利用均值滤波器对  $P_{mask}^i$  进行处理,将 0~255 的跃变处理成 0 到 255 渐变的过渡状态。将这样的  $P_{mask}^i$  记为  $P_{mask,b}^i$  并将其应用到迭代公式[(7)式]中时,便能得到各物体边缘自然过渡的输出图  $P_K$ ,如图 6(b) 所示。(7)式可转化为

$$\begin{cases} P_0 = P_{blur}^{bg} \\ P_{i+1} = P_{blur}^i * P_{mask,b}^i + P_i * (1 - P_{mask,b}^i) \end{cases}. \quad (8)$$

对比图 6 中的两个分图,在细节区域 1、2 处可以看到突兀的边缘轮廓得到很好的解决。

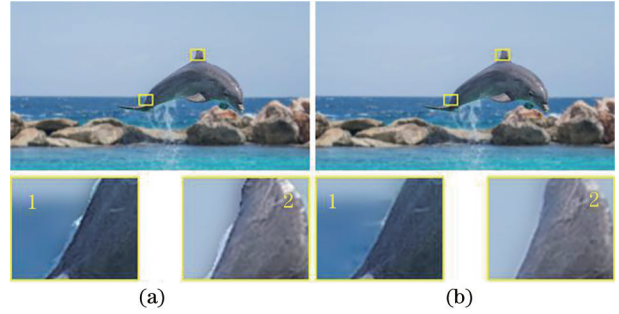


图 6 背景虚化图。(a)简单融合后的背景虚化图;(b)边缘优化后的背景虚化图

Fig. 6 Background blur images. (a) Background blur image after simple fusion; (b) background blur image after edge optimization

## 3 实验结果与分析

本文算法的实验步骤分为 6 个部分:1)不同分割算法对虚化效果的影响分析;2)不同边缘优化方法对虚化效果的影响分析;3)滤波器的核尺寸对虚化效果的影响分析;4)基于自动对焦与选择区域的图像层次性虚化;5)与手机中图像虚化处理方法的对比;6)与 SOTA 方法的对比。以下分别对这 6 部分逐一进行介绍并展示实验结果。

### 3.1 实验一:不同分割算法对虚化效果的影响

实验一对比了不同图像分割算法对虚化效果的影响。如图 7 所示,基于 Mask R-CNN<sup>[18]</sup> 分割算法得到的背景虚化图中,作为前景的主体边缘处被部分分割成背景,主体轮廓被过度地模糊虚化; YOLACT<sup>[17]</sup> 分割算法能更好地分离前后景,边缘处理更精准,能使得最终背景虚化图像效果优于基于 Mask R-CNN 所得到的图像。

从图 7 可看出,分割精度越高,虚化处理后主体区域的信息应保留得越多,则虚化图中的主体区域应与原图的主体更多地保持一致,因此本部分采用了结构相似性(SSIM)指标对使用不同图像分割算法的虚化效果进行评价。YOLACT 分割算法和 Mask R-CNN 分割算法的 SSIM 分别为 0.75 和 0.79,据此得到 YOLACT 分割算法优于 Mask R-CNN 分割算法的结论。因此, YOLACT 被选为本文的图像分割方法。

### 3.2 实验二:不同边缘优化方法对虚化效果的影响

实验二对比了不同边缘优化方法对背景虚化效果的影响,图 8 为效果对比图。如图 8 第 1 列图像所示,虽然 YOLACT 的图像分割效果较优,但主体

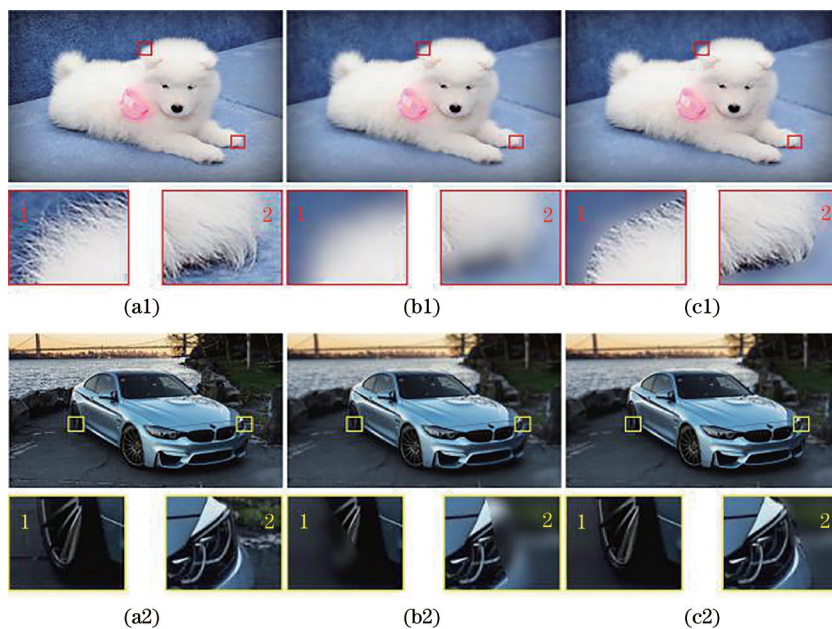


图 7 实验一的效果对比图。(a1)(a2)原图; (b1)(b2) Mask R-CNN; (c1)(c2) YOLACT;  
 Fig. 7 Comparison of results of experiment 1. (a1)(a2) Original image; (b1)(b2) Mask R-CNN; (c1)(c2) YOLACT

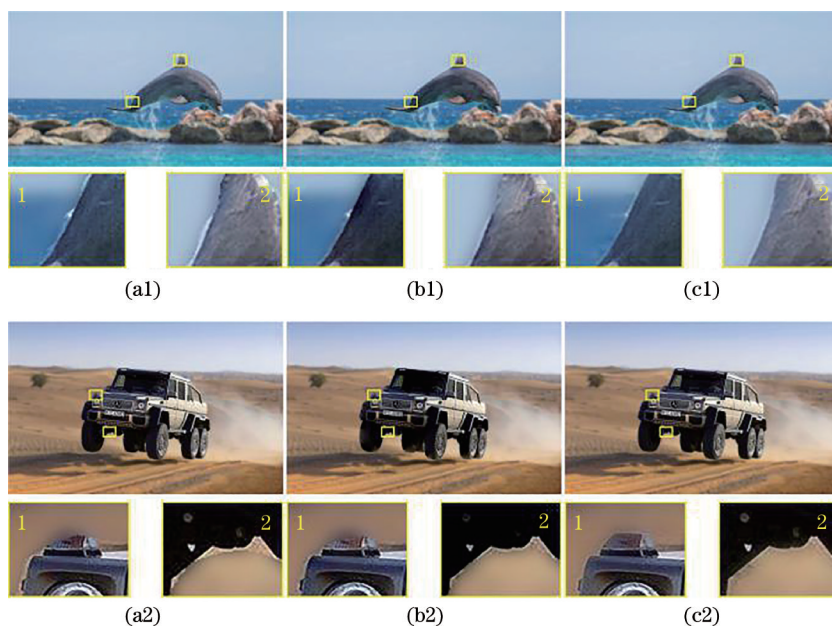


图 8 实验二的效果对比图。(a1)(a2)无边缘优化; (b1)(b2)泊松融合边缘优化; (c1)(c2)均值滤波边缘优化  
 Fig. 8 Comparison of results of experiment 2. (a1)(a2) No edge optimization; (b1)(b2) Poisson fusion edge optimization;  
 (c1)(c2) mean filter edge optimization

与背景的融合图像不可避免地会存在很明显的突兀性边缘。从图 8 的第 2 列图像可看出,利用泊松融合<sup>[20]</sup>算法对边缘部分进行处理之后,主体与背景之间仍存在突兀的边缘;从图 8 的第 3 列图像可以看出,采用简单的均值滤波边缘优化算法能有效地实现虚化背景与清晰主体间的平滑过渡。

由于虚化处理不可避免地改变原图的信息并产

生噪点,之后的边缘优化可能会进一步增加或减少原图的图像信息。因此,本部分采用峰值信噪比 (PSNR) 指标对使用不同边缘优化方法的虚化效果进行衡量。表 1 中的数值显示,与泊松融合边缘优化相比,均值滤波边缘优化方法得到了更高的分值。从图 8 可观察到,图 8(b1)与图 8(a1)、(c1)相比,细节区域 1 处的颜色更深;而图 8(b2)与图 8(a2)、(c2)相

表 1 PSNR 数值对比  
Table 1 Comparison of PSNR

Algorithm	No edge tuning	Poisson fusion edge tuning	Filtering edge tuning
PSNR	22.75	22.43	22.80

比, 细节区域 2 中有两个小亮点被暗化了。出现这些现象的原因是泊松融合边缘优化方法改变原图的信息更多, 而均值滤波边缘优化方法对边缘进行合理调优之后, 恢复了原图的少部分信息。因此, 本文采用均值滤波边缘优化方法进行了边缘优化。

### 3.3 实验三: 滤波器的核尺寸对虚化效果的影响

实验三对比了不同高斯滤波器的核尺寸对背景虚化效果的影响, 图 9 为效果对比图。对滤波核尺寸进行不同程度的微调, 其中, 图 9(b1) 和 (b2)、图 9(c1) 和 (c2) 以及图 9(d1) 和 (d2) 中使用的滤波核尺寸分别为 9, 25, 71。图 9(a1)、(a2) 中细节区域 1 展示的是不同滤波器的核尺寸产生颜色泄漏的程度; 图 9(b1)、(b2) 中细节区域 2 展示的是不同滤波器的核尺寸下背景区域的虚化程度。



图 9 实验三的效果对比图。(a1)(a2)原图; (b1)(b2)偏小的滤波核; (c1)(c2)适中的滤波核; (d1)(d2)偏大的滤波核  
Fig. 9 Comparison of results of experiment 3. (a1)(a2) Original images; (b1)(b2) small filter cores; (c1)(c2) moderate filter cores; (d1)(d2) large filter cores

从图 9 中可看出, 当滤波器的核尺寸过小时, 背景的模糊虚化程度太低, 不能明显地突出主体; 当滤波器的核尺寸过大时, 虽能突显主体, 但是在主体的边缘部分 (细节区域 1) 出现了颜色泄漏, 使得图像变得不够真实; 合适的滤波器的核尺寸既突显了主体, 也减轻了颜色泄漏的程度, 使得图像看起来自然真实。因此, 需要根据需求采用合适的滤波核尺寸对背景进行虚化, 进而实现突显主体、边缘过渡自然的背景虚化效果。

### 3.4 实验四: 基于自动对焦与选择区域的图像层次性虚化

实验四的目的是: 在多个物体的场景中, 通过对图像主体的确定, 实现富有层次感的背景虚化效果。其中, 对图像主体的确定分别有自动对焦识别和用户自主选择两种模式。

图 10 为实验四的效果对比图。本文的层次性指的是当图像中有多个物体时, 根据图 10(a) 的景

深图中的主体与其他物体间的景深信息 (灰度值) 变化调整模糊程度。最终处理的背景虚化图像除主物体之外的背景与其他各物体间有不同的模糊程度, 给人以视觉上的层次感。在无层次感的背景虚化图中, 不同面包区域的虚化程度都是一样的; 在有层次感的背景虚化图中, 不同面包区域的虚化程度是递进的, 而且越远离主体的区域虚化程度越大, 从清晰的面包到其他被虚化的面包, 虚化程度逐渐加大。

### 3.5 实验五: 与手机中图像虚化处理方法的对比

实验五将本文算法得到的背景虚化图与华为 P20 和 iPhone XR 在人像模式下拍摄的背景虚化图进行对比与分析。其中原图是由华为 P20 手机在正常拍照模式 (非人像模式) 下得到的图像。从图 11 中可以看出, 在多人场景下, 华为 P20 与 iPhone XR 拍摄出来的背景虚化图同时突显多人, 本文算法通过自动对焦或者手动选择主体仅能突显单人; 华为

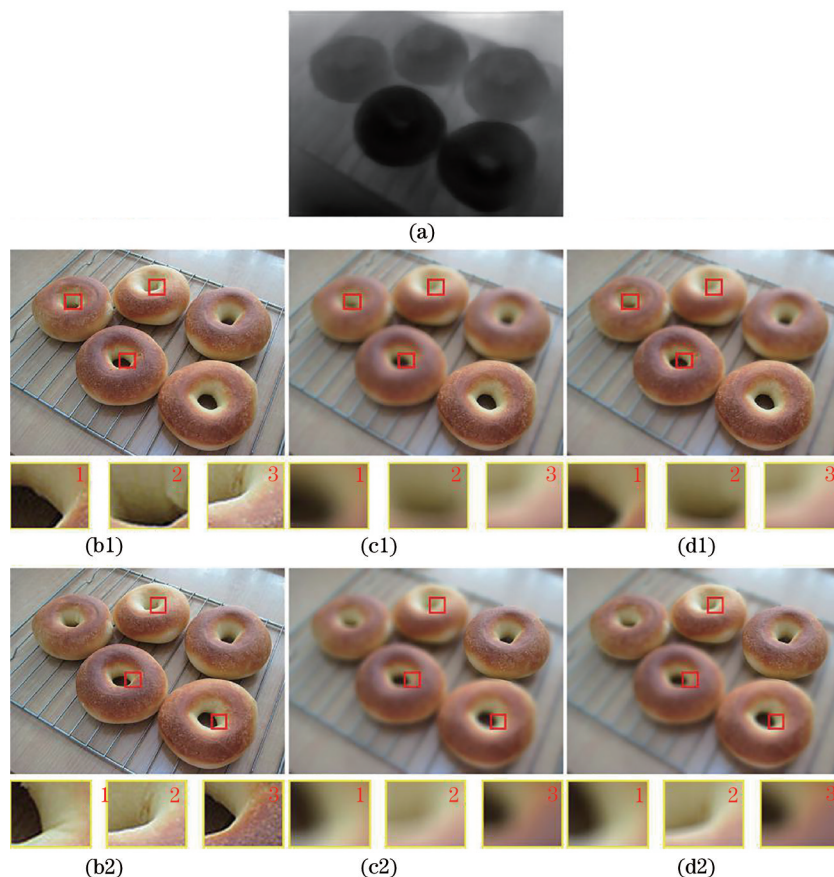


图 10 实验四的效果对比图。(a)景深图;(b1)(b2)原图;(c1)无层次感(自动对焦);(c2)无层次感(自主选择);(d1)有层次感(自动对焦);(d2)有层次感(自主选择)

Fig. 10 Comparison of results of experiment 3. (a) Depth of field image; (b1)(b2) original images; (c1) without sense of hierarchy (auto focus); (c2) without sense of hierarchy (independent selection); (d1) with sense of hierarchy (auto focus); (d2) with sense of hierarchy (independent selection)



图 11 实验五的效果对比图。(a)原图;(b)本方法的自动对焦模式;(c)本方法的自主选择模式;(d)华为 P20;(e) iPhone XR  
Fig. 11 Comparison of results of experiment 5. (a) Original image; (b) autofocus mode of our method; (c) independent selection mode of our method; (d) Huawei P20; (e) iPhone XR

P20 拍摄的图像在主体景深相近的背景部分(细节区域 2)依然是清晰的, iPhone XR 与本文算法则将其模糊虚化处理; 华为 P20 与 iPhone XR 拍摄出来的背景虚化图对两人间的空隙(细节区域 3)处没有进行虚化处理, 而本算法得到的背景虚化图中, 该空隙也被虚化处理了。因此, 本方法效果优于华为 P20 和 iPhone XR 自带的人像拍摄效果。

### 3.6 实验六:与 SOTA 方法的对比

目前关于图像背景虚化效果的评价主要是通过主观指标来实现的<sup>[1]</sup>。本实验主要通过主观调查问卷的方式, 比较本方法与现有 SOTA 方法的视觉效果。调查的对象涉及 43 名参与者, 包括 30 名研究生和 13 名本科生, 男女比例为 4.3:1。用户被要求主观判断图像的背景虚化效果, 并在 1~10 分的



范围内给分,分值越高表示效果越好,然后将用户评分进行统计,以获得最终的平均意见分。将本文解决方案与 Yang、Xiong、Purohit、Dutta、Zheng 等提出的方案<sup>[2]</sup>进行比较,效果对比图及用户平均意见分如表 2 所示。

从图 12 效果图中可以看出 Yang 等的方案在背

景区域的虚化不完全;Xiong、Dutta、Zheng 等的方案在前景人物的处理上存在失真;Purohit 等的方案对背景部分的虚化程度不够,使得图像不够自然;而本文方案在背景区域的处理是最好的,同时前景主体失真度低,在整体性上是最优的。表 2 的数值统计也表明本文方案在 6 种方案中是最好的。



图 12 实验六的效果对比图。(a)原图;(b)本方法;(c)Yang 等所提方法;(d)Xiong 等所提方法;(e)Purohit 等所提方法;(f)Dutta 等所提方法;(g)Zheng 等所提方法

Fig. 12 Comparison of results of experiment 6. (a) Original image; (b) our method; (c) method proposed by Yang *et al.*; (d) method proposed by Xiong *et al.*; (e) method proposed by Purohit *et al.*; (f) method proposed by Dutta *et al.*; (g) method proposed by Zheng *et al.*

表 2 几种算法的虚化效果主观评价比较

Table 2 Comparison of subjective evaluation of blur effects of several algorithms

Algorithm	Yang <i>et al.</i>	Xiong <i>et al.</i>	Purohit <i>et al.</i>	Dutta <i>et al.</i>	Zheng <i>et al.</i>	Our paper
Score	6.94	5.28	6.78	6.08	6.43	<b>7.27</b>

## 4 结 论

针对单目移动设备难以拍摄出背景虚化的图像,提出了一种基于图像感知与分割的自动背景虚化算法,并展示了算法实现的步骤与实验结果。实验结果表明,本文算法利用基于深度学习的感知图结合图像分割技术,实现了准确灵活的图像虚化处理,同时扩展了应用场景的多样性,使最终生成的背景虚化图富有层次感。此外,本文算法降低了搭载平台的硬件要求,单目移动设备仅需拍摄单张图像即可对其进行背景虚化处理。本文算法结合了图像显著性与景深的感知信息,并对图像前景的边缘进行调优,使得其在细节处理、背景区域的虚化、主体的选择等方面优于目前较新的智能手机;同其他算法相比,本文算法的背景虚化效果自然且失真度低。但由于目前应用的三个模型较大,为了使其能在手机端上流畅运行,下一步的研究将优化网络并压缩模型。

## 参 考 文 献

[1] Xiao J S, Du K H, Tu C P, et al. Bokeh display

based on depth information extraction of multi-focus images[J]. Acta Automatica Sinica, 2015, 41(2): 304-311.

肖进胜, 杜康华, 涂超平, 等. 基于多聚焦图像深度信息提取的背景虚化显示[J]. 自动化学报, 2015, 41(2): 304-311.

[2] Xiao J S, Tian H, Zou W T, et al. Stereo matching based on convolutional neural network[J]. Acta Optica Sinica, 2018, 38(8): 0815017.

肖进胜, 田红, 邹文涛, 等. 基于深度卷积神经网络的双目立体视觉匹配算法[J]. 光学学报, 2018, 38(8): 0815017.

[3] 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.

[4] Ignatov A, Patel J, Timofte R, et al. AIM 2019 challenge on bokeh effect synthesis: methods and results[C]//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), October

- 27-28, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 3591-3598.
- [5] Xie G F, Sun X, Wang W C. Interactive depth-of-field rendering with secondary rays[J]. Journal of Computer Science and Technology, 2013, 28(3): 490-498.
- [6] Liu X, Rokne J G. Depth of field synthesis from sparse views[J]. Computers & Graphics, 2016, 55: 21-32.
- [7] Ouyang Z H, Cui G M, Zhao J F, et al. A layered image depth of field rendering algorithm using anisotropic filtering[J]. Optical Technique, 2018, 44(4): 469-475.  
欧阳志恒, 崔光芒, 赵巨峰, 等. 基于分层级各向异性滤波的图像景深渲染算法[J]. 光学技术, 2018, 44(4): 469-475.
- [8] Xu S M, Zhu Y H, Jiang Y Z. Small depth-of-field simulation algorithm based on radial gradient blur[J]. Journal of Hanshan Normal University, 2020, 41(3): 11-15.  
徐颂民, 朱映辉, 江玉珍. 径向梯度模糊的图像小景深模拟算法[J]. 韩山师范学院学报, 2020, 41(3): 11-15.
- [9] Shen X Y, Hertzmann A, Jia J Y, et al. Automatic portrait segmentation for image stylization[J]. Computer Graphics Forum, 2016, 35(2): 93-102.
- [10] Zhu B K, Chen Y Y, Wang J Q, et al. Fast deep matting for portrait animation on mobile phone[C]// Proceedings of the 25th ACM international conference on Multimedia, October 23-27, 2017, Mountain View, California, USA. New York: ACM, 2017: 297-305.
- [11] Su C R, Chen Y Z. Saliency detection based background defocus algorithm for portrait image[J]. Journal of Chinese Computer Systems, 2019, 40(2): 427-431.  
苏超然, 陈羽中. 一种结合显著性检测的肖像照片自动背景虚化算法[J]. 小型微型计算机系统, 2019, 40(2): 427-431.
- [12] Wadhwa N, Garg R, Jacobs D E, et al. Synthetic depth-of-field with a single-camera mobile phone[J]. ACM Transactions on Graphics, 2018, 37(4): 1-13.
- [13] Purohit K, Suin M, Kandula P, et al. Depth-guided dense dynamic filtering network for Bokeh effect rendering[C]//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), October 27-28, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 3417-3426.
- [14] Dutta S. Depth-aware blending of smoothed images for Bokeh effect generation[J]. Journal of Visual Communication and Image Representation, 2021, 77: 103089.
- [15] Zhou B L, Khosla A, Lapedriza A, et al. Learning deep features for discriminative localization[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 2921-2929.
- [16] Chen W F, Fu Z, Yang D W, et al. Single-image depth perception in the wild[C]//30th International Conference on Neural Information Processing Systems, December 5-10, 2016, Barcelona, Spain. Cambridge: DBLP, 2016: 730-738.
- [17] Bolya D, Zhou C, Xiao F Y, et al. YOLACT: real-time instance segmentation[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 9156-9165.
- [18] He K M, Gkioxari G, Dollár P, et al. Mask R-CNN [C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2980-2988.
- [19] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [20] Pérez P, Gangnet M, Blake A. Poisson image editing [J]. ACM Transactions on Graphics, 2003, 22(3): 313-318.