

Image Composition with Depth Registration

基于深度注册的嵌入式图像合成

李赞、王文成、侯飞

International Joint Conference on Artificial Intelligence(IJCAI-2023)

联系人: 李赞, 18810762067, lizhan@ios.ac.cn

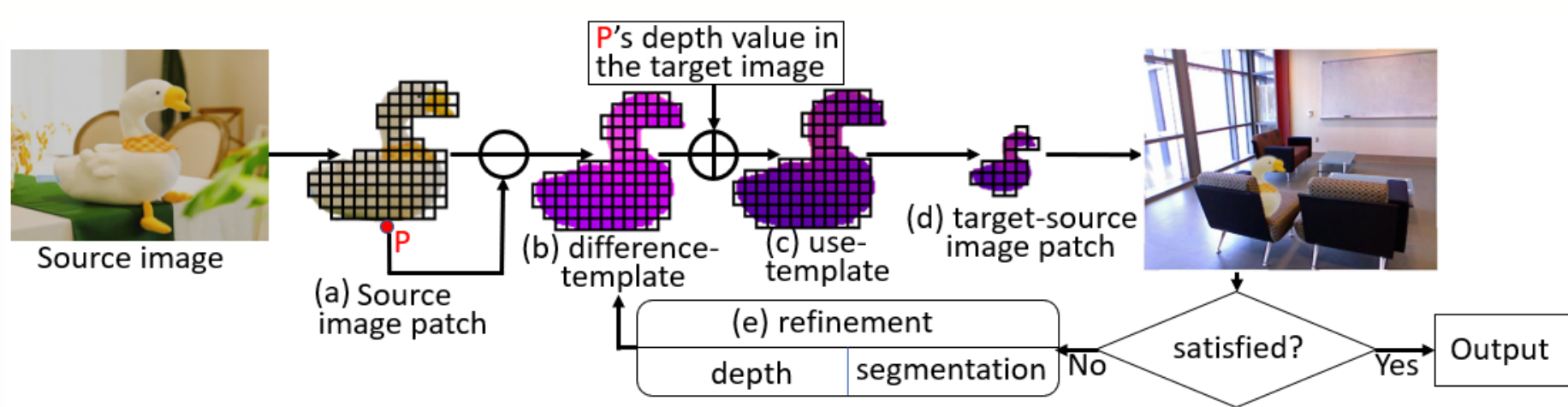
背景

图像合成是指将不同图像中的内容融合成一张看似真实的图像。已有的图像合成方法主要探讨色彩、阴影和亮度的和谐生成,对于图像内容之间的遮挡关系研究不多,往往需要费时的手动处理而效率不高。

方法

本文提出新的方法,从源图像中选择要融入目标图像中的内容(称为源内容)后,通过深度注册,将源内容融入目标图像所表达的三维空间中,由此,用户进行源内容在目标图像中不同位置的摆放时,融合处理可基于图像内容的深度方便判断它们之间的遮挡关系,以得到遮挡关系合理处理的合成图像。其处理步骤如下:

- 基于深度学习方法进行源图像和目标图像的深度估计;
- 基于深度注册初始化,将源内容融入目标图像所表达的三维空间;
- 迭代进行图像分割/深度的调整,以优化深度注册,直至获得满意的合成结果。



深度注册:

- 将源内容放至目标图像中,并调整其大小使得其与目标图像中的一个内容具有相似深度;
- 将目标图像中该内容的锚点像素的深度值赋予源内容的锚点像素(一般来说选择图像内容所有像素中y值最小的点作为锚点);
- 根据相机成像原理,建立源内容深度和大小之间的关系;
- 用户根据编辑意向设定源内容的位置、大小后,根据上述关系计算源内容各像素的深度,完成深度注册。

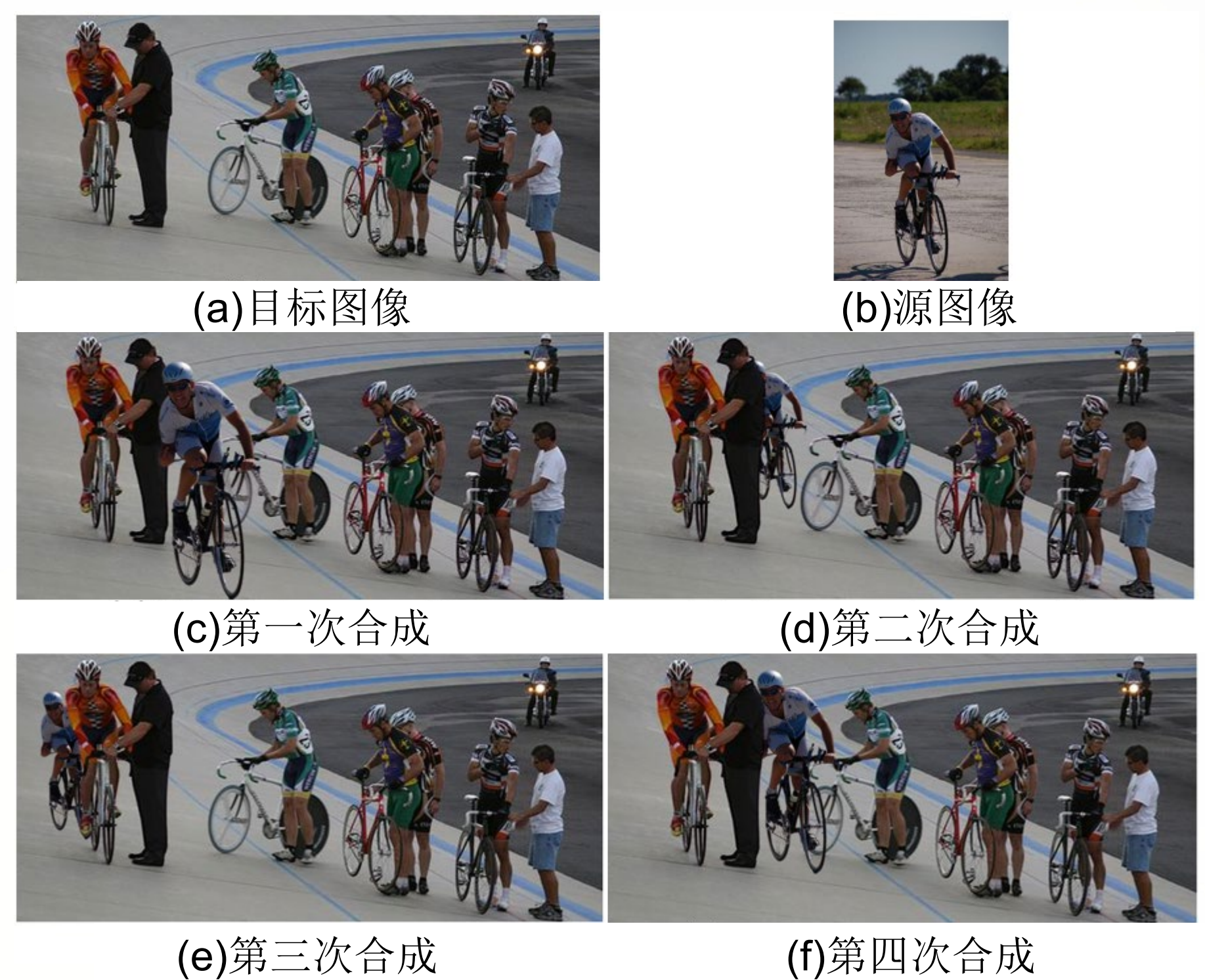
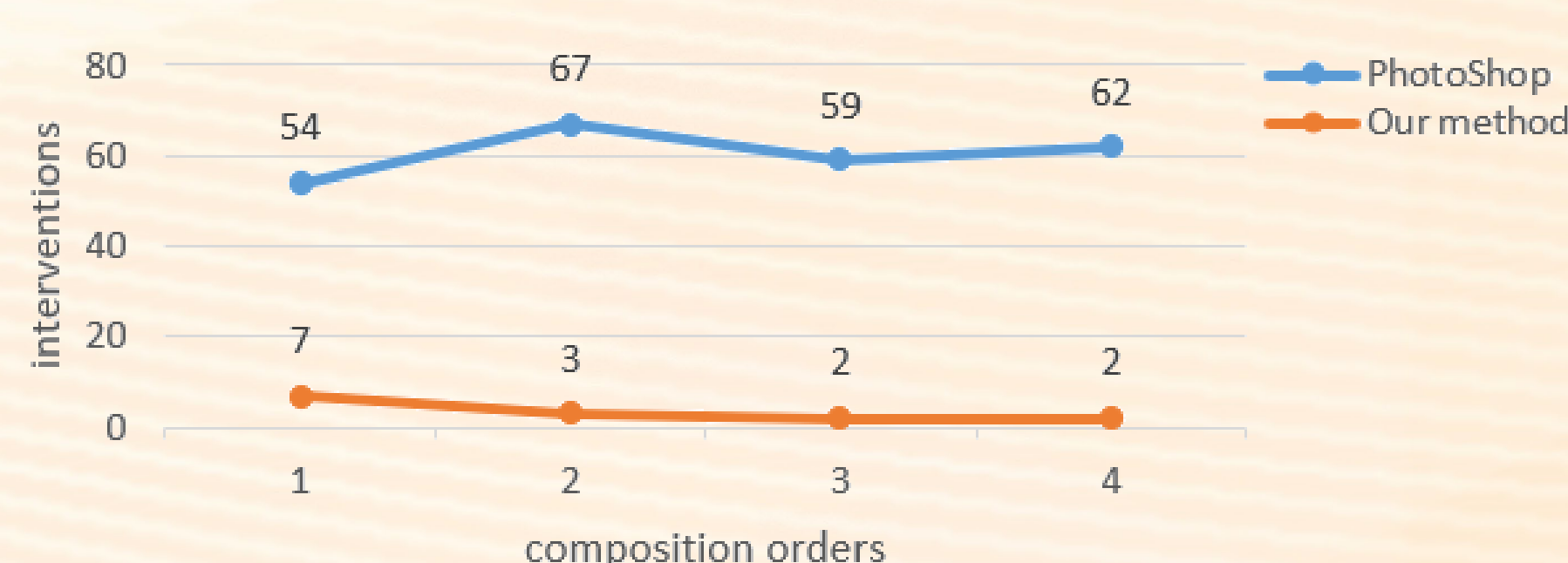


分割/深度的优化调整:

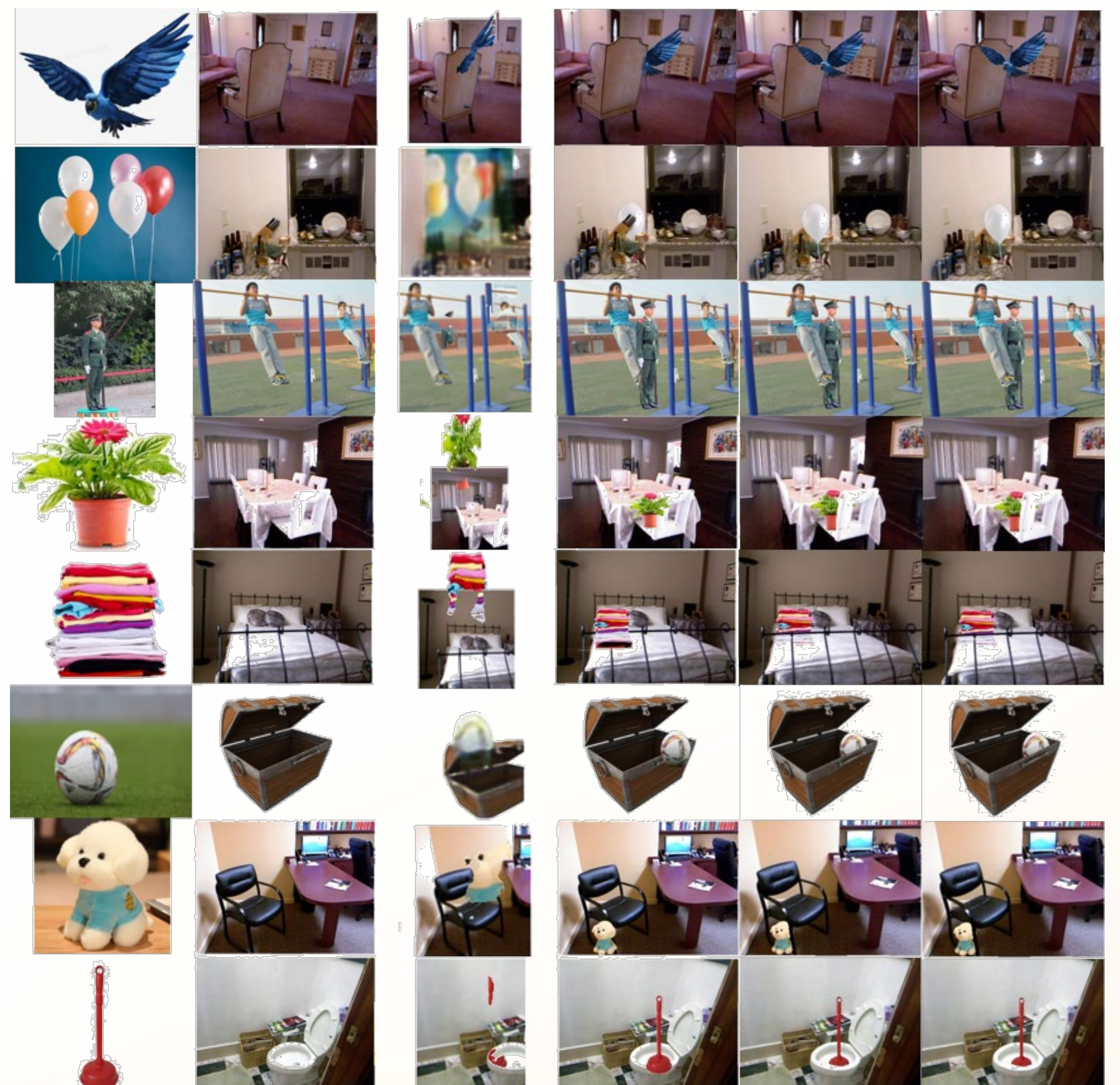
- 分割优化:根据合成图像需求,进行更细致合理的分割,如分割开椅子座上的鸭子等。
- 深度优化:根据编辑操作建立约束条件,更新源内容的深度值,以更适合图像合成处理。

实验结果

本方法可逐步优化深度注册,使源内容更合理融入目标图像对应的三维空间,方便图像合成的遮挡关系处理。如下图所示,在用户频繁改变源图像内容的位置时,只有第一次嵌入需要进行分割优化处理,故交互次数稍高。随后,再次改变源图像内容的位置,则无需更多的调整;反观常用的Photoshop方法,其每次合成处理都需要较多的用户交互。



本方法所得到合成结果,与基于深度学习的CompGAN方法、传统Tan的方法相比,能合理处理遮挡关系;与PhotoShop方法相比,合成结果相当,但我们所需时间开销和交互次数大幅减少。



(a)源图像 (b)目标图像 (c)compGAN (d)Tan的方法 (e)PS (f)我们的方法

对于上图的八个例子,我们用准确率和F1分数来度量源内容在合成图像中具有正确遮挡标签的像素占比,结果如下表所示:

	Accuracy					F1 score				
	CompGAN	Tan's	PS	Ours	Ours*	CompGAN	Tan's	PS	Ours	Ours*
Row1	0.2100	0.7001	1.0000	0.9826	-	0.3472	0.8236	1.0000	0.9912	-
Row2	0.0000	0.7951	1.0000	0.9896	-	0.0000	0.8858	1.0000	0.9948	-
Row3	0.1674	0.8163	1.0000	0.9799	-	0.2694	0.8989	1.0000	0.9875	-
Row4	0.4322	0.6304	1.0000	0.7846	0.9995	0.5186	0.7733	1.0000	0.8326	0.9996
Row5	0.4503	0.7913	1.0000	0.9687	-	0.5910	0.8835	1.0000	0.9798	-
Row6	0.5222	0.5786	1.0000	0.9731	-	0.6557	0.7331	1.0000	0.9762	-
Row7	0.0000	0.9178	1.0000	0.9218	0.9887	0.0000	0.9572	1.0000	0.9586	0.9938
Row8	0.1998	0.8357	1.0000	0.9809	-	0.3363	0.9105	1.0000	0.9884	-
Average	0.2477	0.7582	1.0000	0.9477	0.9941	0.3398	0.8582	1.0000	0.9637	0.9967

可以看到我们的方法在各项指标中都是最好的。上表中,Ours*表示使用优化调整的结果,可知我们只有极少次需要用到优化调整,大多数情况下可以直接得到和PhotoShop相媲美的结果。如下表所示,本方法平均比Photoshop方法快4倍多,所需交互次数也少很多。

	Photoshop		Ours	
	Time	Interventions	Time	Interventions
Row1	142.0s	71	33.0s	4
Row2	158.0s	84	34.0s	5
Row3	139.0s	87	32.0s	4
Row4	160.0s	93	44.0s	5
Row5	144.0s	77	45.0s	5
Row6	137.0s	75	33.0s	4
Row7	133.0s	82	35.0s	4
Row8	152.0s	73	34.0s	4
Average	145.6s	80.3	36.3s	4.4

结论

本文提出的深度注册方法,促进了嵌入式图像合成的方便高效实现,降低了图像内容创作门槛,将提高电影、动画和艺术创作等领域的工作成效。