Deep Rectangling for Image Stitching: A Learning Baseline
SUPPLEMENTARY MATERIAL

Lang Nie[1,2], Chunyu Lin[1,2]*, Kang Liao[1,2], Shuaicheng Liu[3], Yao Zhao[1,2]
[1]Institute of Information Science, Beijing Jiaotong University, Beijing, China
[2]Beijing Key Laboratory of Advanced Information Science and Network, Beijing, China
[3]University of Electronic Science and Technology of China, Chengdu, China

(a) Missing 'surfboard' (b) Missing 'dining table' (c) Missing 'toilet'

Figure 1. Object detection and semantic segmentation results of stitched images and our rectangling results. The arrows highlight the missing parts.

1. Overview

In this material, we first demonstrate the benefits of image rectangling for scene reasoning in Section 2. Then, we illustrate more experimental results of our solution in Section 3, including our rectangling results on the DIR-D dataset and other datasets.

2. Benefits for Scene Reasoning

The proposed rectangling solution offers a nearly perfect visual perception for users by eliminating the irregular boundaries in image stitching. It can also help downstream vision tasks such as object detection and semantic segmentation, which is crucial for scene understanding. As shown in Fig. 1, the detection and segmentation results are derived from Mask R-CNN [1]. We can notice that the objects in the stitched images with irregular boundaries may be missing, such as the surfboard (Fig. 1a), the dining table(Fig. 1b), and the toilet (Fig. 1c). By contrast, our rectangling results 'find' the missing objects. We summarize the improvement as follows:

Almost all existing deep learning models (detection and segmentation) are trained on rectangular images, making them not robust to irregular boundaries in stitched images.

3. More Results

More results on DIR-D are exhibited in Fig. 2, where our solution can deal with variable irregular boundaries and yield perceptually natural rectangular results.

Besides, more cross-dataset results are displayed in Fig. 3, which shows the superiority of rectangling over other solutions such as cropping and completion.

References

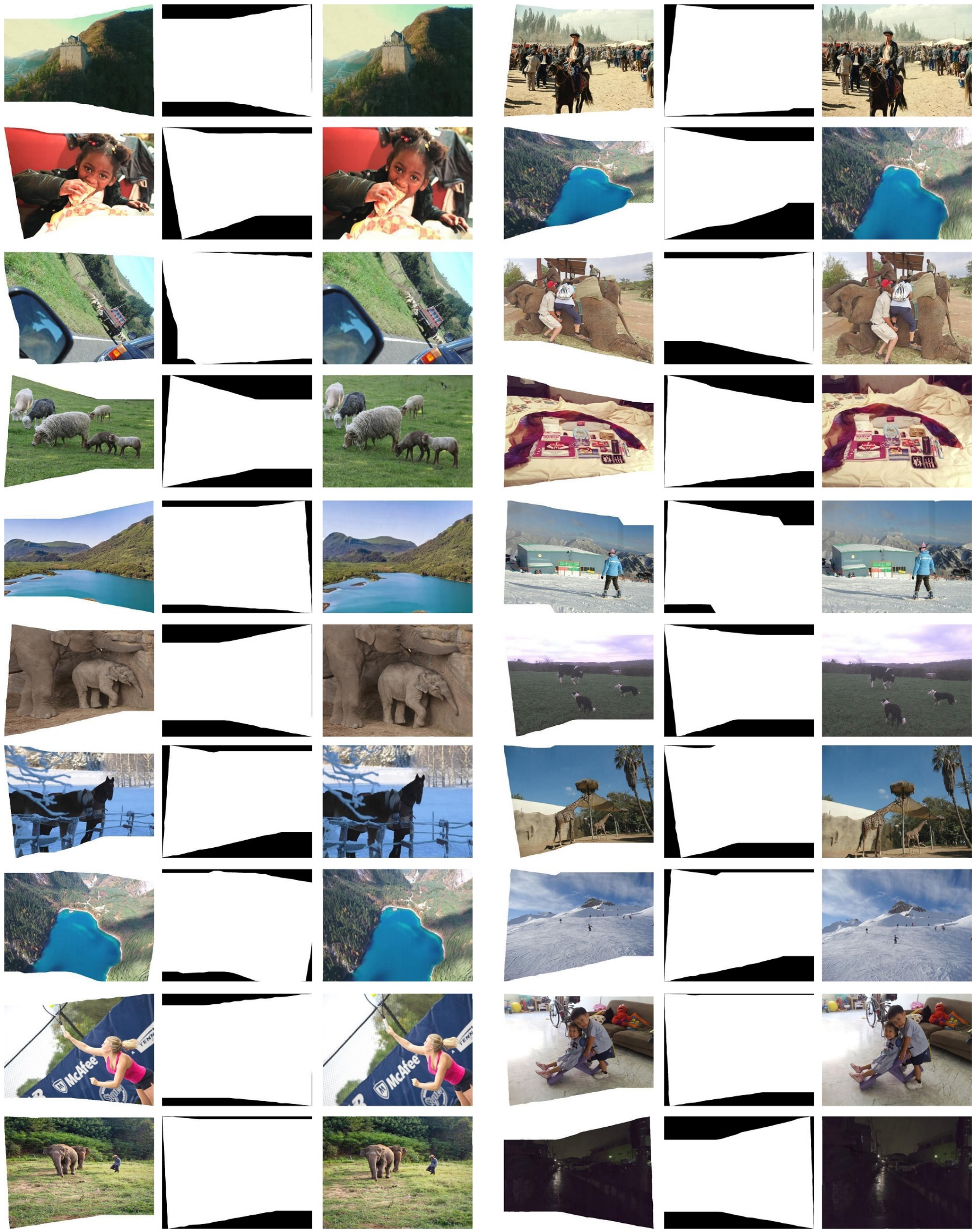
[1] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proc. ICCV*, pages 2980–2988, 2017. 1

[2] Qi Jia, ZhengJun Li, Xin Fan, Haotian Zhao, Shiyu Teng, Xinchen Ye, and Longin Jan Latecki. Leveraging line-point consistence to preserve structures for wide parallax image stitching. In *Proc. CVPR*, pages 12186–12195, 2021. 3

[3] Jing Li, Zhengming Wang, Shiming Lai, Yongping Zhai, and Maojun Zhang. Parallax-tolerant image stitching based on robust elastic warping. *IEEE Trans. on Multimedia*, 20(7):1672–1687, 2017. 3

[4] Tianli Liao and Nan Li. Single-perspective warps in natural image stitching. *IEEE Trans. on Image Processing*, 29:724–735, 2019. 3

[5] Roman Suvorov, Elizaveta Logacheva, Anton Mashikhin, Anastasia Remizova, Arsenii Ashukha, Aleksei Silvestrov, Naejin Kong, Harshith Goka, Kiwoong Park, and Victor Lempitsky. Resolution-robust large mask inpainting with fourier convolutions. *arXiv preprint arXiv:2109.07161*, 2021. 3



Input Mask Ours Input Mask Ours

Figure 2. More results of our solution on DIR-D. Each triplet includes an input, a mask, and our rectangling result from left to right.

Figure 3. More cross-dataset results. The classic image stitching datasets ('crossline' [2], 'eyot' [2] and 'tower' [3]) are stitched by SPW [4]. We adopt LaMa [5] to complete the stitched images, and the rectangling results are generated by the proposed learning baseline.