Refocusing Images Captured from a Stereoscopic Camera

Chia-Lun Ku[1] Yu-Shuen Wang[1] Chia-Sheng Chang[2] Hung-Kuo Chu[2] Chih-Yuan Yao[3]
[1]National Chiao Tung University [2]National Tsing Hua University [3]National Taiwan University of Science and Technology

Figure 1: ***Left:*** *The left image captured from a stereoscopic camera.* ***Middle:*** *The estimated depth map. The dark color indicates the pixel is close to the camera.* ***Right:*** *The synthesized image, in which the focus is at the chair.*

1 Introduction

Traditional photography projects a 3D scene to a 2D image without recording the depth of each local region, which prevents users from changing the focus plane of a photograph once it has been taken. To tackle this problem, Ng et al. [2005] presented light-field cameras that record all focus planes of a scene and synthesized the refocused image using ray tracing. Nevertheless, the captured photographs are of low resolution because the image sensor is divided into subcells. Levin et al. [2007] embedded a coded aperture on the camera lens and recover depth information from blur patterns in a single image. However, the coded aperture blocks around 50% of light. Their system requires longer exposition time when taking pictures. Liang et al. [2008] also embedded a coded aperture on the camera lens to capture the scene but with multiple exposures. It produces high quality depth maps yet is not suitable to hand-held devices. Recently, Microsoft Kinect directly estimates the depth information using infrared light, which works only in a indoor environment.

We introduce a refocus technique that recovers the depth from a pair of images captured from consumer stereoscopic cameras. A disparity map is estimated using semi-global matching (SGM) algorithm [Hirschmuller 2008] to represent the parallax caused by perspective projection. Our key ideas are 1) the sign of disparity indicates whether the pixel is in front of or in rear of the focus plane; and 2) the magnitude of disparity shows the distance to the focus plane. As a result, the depth value of each local region can be obtained easily. Another key observation is that consumer cameras usually carry small lenses to enhance mobility, where their depths of field are deep and the captured photographs are nearly all focused. Therefore, we can simply blur image pixels whenever the focus plane is changed and avoid the challenging deblurring problem. Compared to previous works, our approach does not modify existing cameras, demands short exposition time, and produces high resolution refocus images.

2 Our Approach

The inputs to our system are a pair of stereo images captured from a pre-calibrated stereoscopic camera. The two images are first rectified and the disparity map is then computed using SGM algorithm [Hirschmuller 2008]. As there are some mismatched or occluded pixels within the images, which lead to holes in the disparity map, we fill the holes according to the neighboring disparity values and the gradients of pixel colors. As different objects are usually segmented by sharp boundaries, the regions with similar pixel colors can be assumed to have similar depth values while the others are not. We formulate this idea into the objective function

$$\arg\min_{P} \sum_{\{i,j\}\in\mathbf{E}} w_{ij}|p_i - p_j|^2 + \sum_{i\in\mathbf{C}} |p_i - c_i|^2, \quad (1)$$

and compute for the complete disparity map using a linear solver, where $\mathbf{P} = \{p_0, p_1, ..., p_n\}$ is the disparity map, $\mathbf{E}$ denotes the 4-adjacent neighbors, $\mathbf{C}$ is the set of pixels that have disparity values, and w_{ij} is the Guassian of color difference.

Once the complete disparity map is obtained, the refocused image can be synthesized whenever the focus plane is given. As only blurring is demanded in our system, this process could be accomplished by applying spatial filters with different sizes to pixels. The farther distance to the focus plane, the larger filer size we use. We also blend neighboring pixels that have similar depth values to prevent artifacts induced by fusing foregrounds and backgrounds together.

3 Results and Conclusions

We have presented a system to refocus stereo images. This framework was built on consumer stereo cameras and requires no modification to hardwares. It also enjoys the advantages such as short exposition time and high resolution results that are absent in previous works. The experimental results shown in Figure 1 and our supplemental materials demonstrate the feasibility of our technique.

References

HIRSCHMULLER, H. 2008. Stereo processing by semiglobal matching and mutual information. *IEEE Trans. Pattern Anal. and Mach. Intell. 30*, 2, 328–341.

LEVIN, A., FERGUS, R., DURAND, F., AND FREEMAN, W. T. 2007. Image and depth from a conventional camera with a coded aperture. *ACM Trans. Graph. 26*, 3, 328–341.

LIANG, C.-K., LIN, T.-H., WONG, B.-Y., LIU, C., AND CHEN, H. 2008. Programmable aperture photography: Multiplexed light field acquisition. *ACM Transactions on Graphics 27*, 3, 55:1–55:10.

NG, R., LEVOY, M., BREDIF, M., DUVAL, G., HOROWITZ, M., AND HANRAHAN, P. 2005. Light field photography with a handheld plenoptic camera. *Stanford University Computer Science Tech Report CSTR 2005-02*.

SIGGRAPH 2013, July 21 – 25, 2013, Anaheim, California.

ACM 978-1-4503-2261-4/13/07