Effective Global Prediction for Dense Light-Field Compression by Using Synthesized Multi-focus Images

Takashi SAKAMOTO*† Kazuya KODAMA† Takayuki HAMAMOTO*
*Tokyo University of Science †National Institute of Informatics

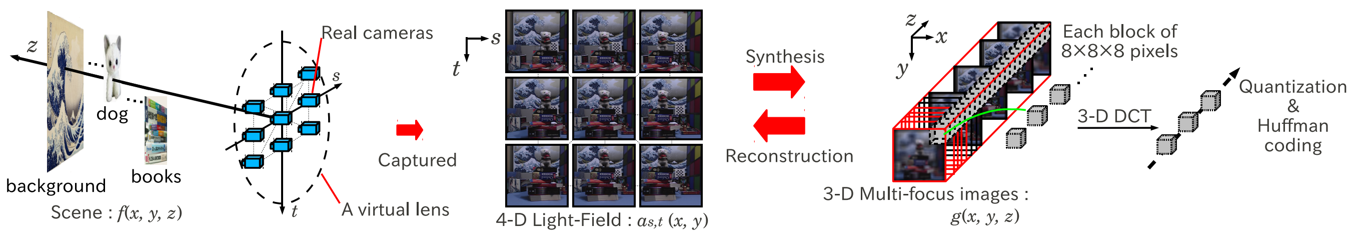


Figure 1: *Light-Field compression using multi-focus images.*

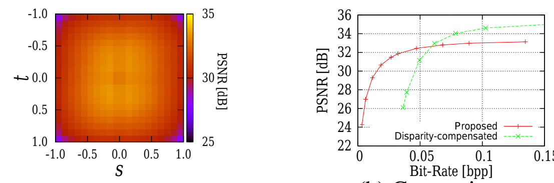
1 Introduction

Light-Field Rendering is a promising technique generating 3-D images from multi-view images captured by dense camera arrays or lens arrays[Isaksen et al. 2000]. However, Light-Field generally consists of 4-D enormous data, that are not suitable for storing or transmitting without effective compression[Magnor and Girod 2000]. We previously derived a method of reconstructing 4-D Light-Field directly from 3-D information composed of multi-focus images without any scene estimation[Kodama et al. 2006]. On the other hand, it is easy to synthesize multi-focus images from Light-Field. Therefore, we can achieve conversion between 4-D Light-Field and 3-D multi-focus images without significant degradation. Recently, researchers in computational photography also study such interesting properties of Light-Field[Levin and Durand 2010]. In this work, based on the conversion, we propose novel global prediction for dense Light-Field compression via synthesized multi-focus images as effective representation of 3-D scenes like Figure 1.

2 Our Approach

We synthesize multi-focus images $g(x, y, z)$ with a virtual lens by merging shifted multi-view images using constant weight function $p(s, t)$ as follows: $g(x, y, z) = \sum_{s,t} p(s, t)a_{s,t}(x+sz, y+tz)$, where a viewpoint is denoted by (s, t) and $a_{s,t}(x, y)$ is the corresponding image. Then, we apply DCT-based coding to synthesized multi-focus images. The compressed 3-D multi-focus images enable us to predict 4-D Light-Field globally. Actually, $g(x, y, z)$ is combined with 3-D scene $f(x, y, z)$ by using a convolution of a 3-D blurring filter $h(x, y, z) = (1/z^2)p(x/z, y/z)$[Kodama et al. 2006]. Based on the relation, Light-Field can be approximately reconstructed by a simple combination of dimension reduction and a 2-D filtering[Ou et al. 2008; Kodama et al. 2010].

We show experimental results using real images [Ota et al. 2009]. 21 × 21 multi-view images are taken from equally-spaced viewpoints, and each image has 512×512 pixels. Disparity between adjacent images is smaller than about 2.0 pixels. Figure 2(a) shows quality of the reconstructed multi-view images from the synthesized 64 multi-focus images compressed at 0.027 bpp. It indicates how the reconstruction quality changes by (s, t), where PSNR is evaluated for 384 × 384 pixels in the center of the image. We obtain 30∼35 dB at almost all the viewpoints. In Figure 2(b), our proposed global prediction using multi-focus images is compared with the conventional local disparity-compensation[Magnor and Girod 2000] by averages of PSNR. Our prediction is robust, especially, at very low bit-rate. Effective Light-Field compression including residuals based on the proposed prediction is expected in future.



(a) Reconstruction quality (b) Comparison

Figure 2: *Evaluation of prediction.*

References

ISAKSEN, A., MCMILLAN, L., AND GORTLER, S. 2000. Dynamically Reparameterized Light Fields. In *Proc. ACM SIGGRAPH*, 297–306.

KODAMA, K., MO, H., AND KUBOTA, A. 2006. Free Viewpoint, Iris and Focus Image Generation by Using a Three-Dimensional Filtering based on Fresquency Analysis of Blurs. In *Proc. IEEE ICASSP*, vol. II, 625–628.

KODAMA, K., IZAWA, I., AND KUBOTA, A. 2010. Robust Reconstruction of Arbitrarily Deformed Bokeh from Ordinary Multiple Differently Focused Images. In *Proc. IEEE ICIP*, 3989–3992.

LEVIN, A., AND DURAND, F. 2010. Linear View Synthesis Using a Dimensionality Gap Light Field Prior. In *Proc. IEEE CVPR*, 1831–1838.

MAGNOR, M., AND GIROD, B. 2000. Data Compression for Light Field Rendering. *IEEE Trans. CSVT 10(3)*, 338–343.

OTA, M., FUKUSHIMA, N., YENDO, T., TANIMOTO, M., AND FUJII, T. 2009. Rectification of Pure Translation 2D Camera Array. In *Proc. IWAIT, 0044*.

OU, X., HAMAMOTO, T., KUBOTA, A., AND KODAMA, K. 2008. Efficient free viewpoint image acquisition from multiple differently focused images. In *Proc. SPIE VCIP*, vol. 6822-73, 1–8.

*e-mail: {sakamoto, hamamoto}@isl.ee.kagu.tus.ac.jp
†e-mail: {sakamoto, kazuya}@nii.ac.jp

SIGGRAPH 2012, Los Angeles, California, August 5 – 9, 2012.
ISBN 978-1-4503-1435-0/12/0008