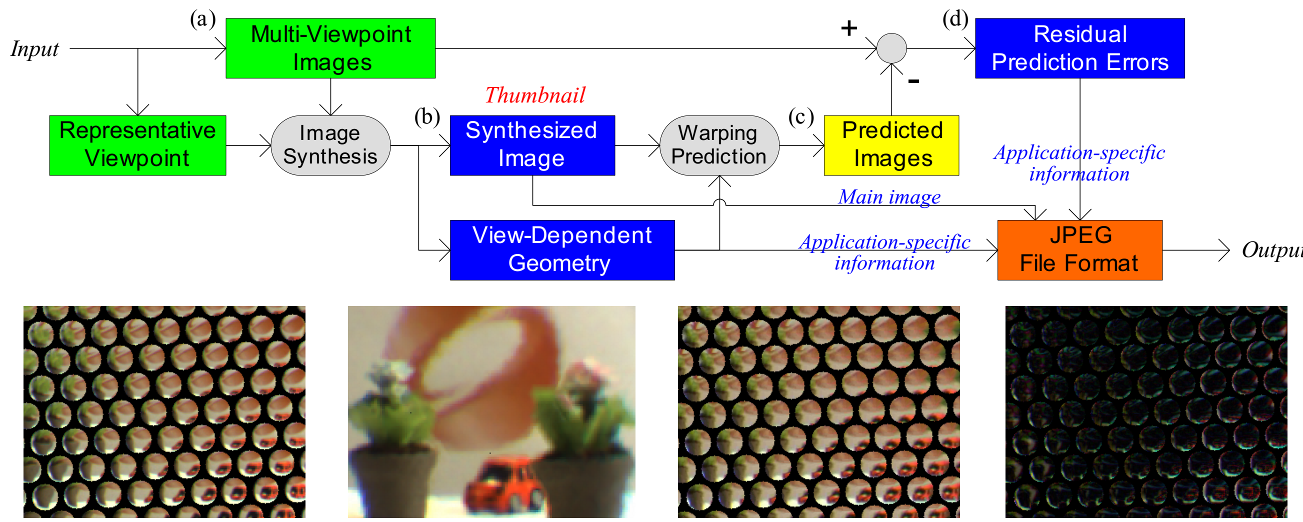
Free-Viewpoint Thumbnail for Light Field Compression

Yuichi Taguchi* Takeshi Naemura
Graduate School of Information Science and Technology, The University of Tokyo



(a) Multi-viewpoint images (b) A synthesized image (*thumbnail*) (c) Predicted images (d) Residual prediction errors

Figure 1: A block diagram of the proposed coder (top) and example images using an integral photography (bottom).

1 Introduction

This paper proposes a new JPEG-compatible coder of light field data. When viewing a 3D scene using light field data, we usually require a specialized light field viewer. However, the proposed coder produces an output file from which we can see a *thumbnail* of the light field by using common JPEG viewers. This *thumbnail* can be a free viewpoint image of the scene which reflects user's preference or creator's idea. We can also achieve good compression efficiency by using this *thumbnail* as a reference image.

In the proposed coder, an image-based rendering technique is introduced at the beginning of coding. The coder first synthesizes a *thumbnail* image at a given viewpoint, and then predicts all input images by warping this *thumbnail*. After dividing the light field data into a 2D *thumbnail* and other information, the coder stores both data together into a standard JPEG file with application-specific information.

2 Proposed Coder

Figure 1 (top) shows a block diagram of the coder. Prior to coding, a viewpoint for the *thumbnail* is selected, which is called a representative viewpoint.

The coder first synthesizes an image at the representative viewpoint using image-based rendering techniques. A view-dependent geometry model is also produced during image synthesis. Then, all of the input images are predicted by warping the synthesized image with the geometry model. The residual prediction error is generated if the predicted image quality is not sufficient. This prediction method is similar to that used in model-aided predictive coding [Magnor et al. 2003], but we take a novel approach that selects a synthesized image at an arbitrary viewpoint as a reference image.

Finally, the synthesized image is encoded using the JPEG compression technique. Other information, i.e. the geometry model and residual prediction errors, are stored in application-specific segments of the JPEG file.

3 Experiments

We encoded an integral photography, which is an example of light field representation, captured by LIFLET system [Yamamoto et al. 2004]. Shown in Fig. 1 (bottom) are example images during the coding process. We selected a representative viewpoint behind the center of the input images' plane. The modeling and rendering technique adopted in LIFLET was used as the geometry reconstruction and image synthesis method. We compressed the geometry model using a fixed-length coder, and the residual prediction errors using a DCT-based coder.

As a preliminary evaluation, we compared the quality of synthesized images using the original and compressed data. We observed view-dependency of the reconstruction quality. The quality of views around the representative viewpoint is kept high even at high compression ratios (see the attached video).

4 Future Work

The presented coding scheme is also useful for dynamic scenes. When we transmit light field video in real time, this view-dependent scheme could work efficiently by selecting the representative viewpoint at the current user's viewpoint. We will extend this scheme to a MPEG-compatible coder of light field video.

Acknowledgment: Special thanks go to Prof. H. Harashima and Mr. K. Takahashi at the University of Tokyo.

References

MAGNOR, M., RAMANATHAN, P., AND GIROD, B. 2003. Multi-view coding for image-based rendering using 3-D scene geometry. *IEEE Trans. CSVT 13*, 11 (Nov.), 1092–1106.

YAMAMOTO, T., KOJIMA, M., AND NAEMURA, T., 2004. LIFLET: Light field live with thousands of lenslets. ACM SIGGRAPH 2004 Emerging Technologies #0130, Aug.

*e-mail: yuichi@hc.ic.i.u-tokyo.ac.jp