Optimized Sampling for View Interpolation in Light Fields Using Local Dictionaries

David C. Schedl
Clemens Birklbauer
Oliver Bimber
firstname.lastname@jku.at
Institute of Computer Graphics, Johannes Kepler University
Linz, Austria

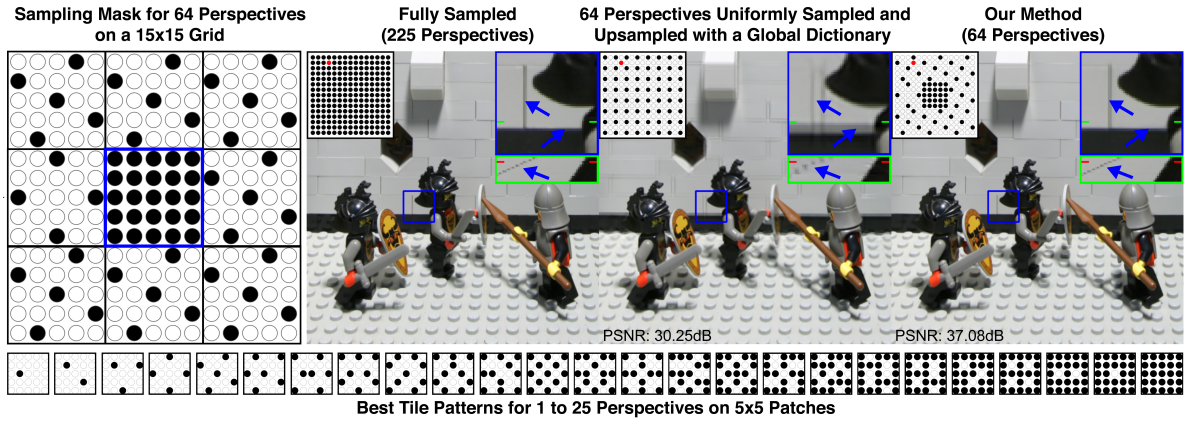

Figure 1: Left: Our sampling mask for 15x15 light fields captured with only 64 cameras. The mask is constructed from best tile patterns for light-field patch sizes of 9x9x5x5 that are determined from a global dictionary (bottom). The center tile, called *guidance area* (blue), is fully sampled to record the local dictionary needed for upsamplig. Center and right: Comparison of reconstruction quality with a fully sampled (225 perspectives) light field and a sparse (64 perspectives) uniformly sampled light field that is upsampled with a related method that applies a global dictionary [Marwah et al. 2013] . Blue rectangles mark close-ups and green rectangles present EPI images. Red dots in the sampling mask indicate the displayed view.

ABSTRACT

We present an angular superresolution method for light fields captured with a sparse camera array. Our method uses local dictionaries extracted from a sampling mask for upsampling a sparse light field to a dense light field by applying compressed sensing reconstruction. We derive optimal sampling masks by minimizing the coherence for representative global dictionaries. The desired output perspectives and the number of available cameras can be arbitrarily specified. We show that our method yields qualitative improvements compared to previous techniques.

CCS CONCEPTS

• **Computing methodologies** → **Computational photography**; **Image-based rendering**; *Image processing*;

KEYWORDS

light fields, upsampling, view interpolation, compressed sensing, computational photography

ACM Reference format:
David C. Schedl, Clemens Birklbauer, and Oliver Bimber. 2017. Optimized Sampling for View Interpolation in Light Fields Using Local Dictionaries. In *Proceedings of SIGGRAPH '17 Posters, Los Angeles, CA, USA, July 30 - August 03, 2017,* 2 pages.
DOI: 10.1145/3102163.3102171

1 INTRODUCTION AND CONTRIBUTION

Compared to standard digital photography, light fields offer various new options, such as refocussing, perspective changes, and 3D filtering as a postprocess. However, capturing them at an adequate resolution remains challenging. Popular approaches typically multiplex the 4D information onto a single 2D sensor, which results either in low spatial resolution, low angular resolution, or both. Multiple sensors (e.g., a camera array) can be used to overcome the aforementioned issue. However, achieving an adequate resolution in the angular domain requires a vast number of cameras, resulting in high construction costs and complexity. In this work, we present an angular superresolution approach for light fields captured with sparse camera arrays. We apply compressed sensing

Project funding: Austrian Science Fund (FWF), contract number P 28581-N33.

SIGGRAPH '17 Posters, Los Angeles, CA, USA

DOI: 10.1145/3102163.3102171

theory for reconstruction and find optimal sampling masks for a desired number of cameras and sampling grid resolution. In contrast to related work, we avoid the need for depth reconstruction, which often fails for non-Lambertian scenes. One of our contributions is the use of online learned local dictionaries extracted directly from the scene sampled with an optimized mask instead of using global dictionaries that are learned offline from a set of representative pre-recorded light fields. Therefore, our method yields superior reconstruction results compared to related techniques. A second contribution is that the number of samples is not constrained to the sampling pattern. Thus, our new approach allows to determine sampling masks for an arbitrary number of cameras. We compute coherence values for representative global dictionaries that provide a formal basis for estimating the reconstruction quality of a given sampling pattern. We find optimal sampling masks by minimizing the coherence.

2 GLOBAL DICTIONARY DRIVEN SAMPLING AND UPSAMPLING WITH LOCAL DICTIONARIES

Our approach can be outlined in four consecutive steps:

(1) finding optimal local sampling masks for computable camera array tile sizes based on a representative global dictionary,
(2) determining an ideal global sampling mask from the determined local tiles,
(3) recording the sparsely sampled light field with the global sampling mask, and
(4) reconstructing missing perspectives using a local dictionary being recorded with the global sampling mask.

Let's assume the example for which we desire a light field with a sampling grid resolution of 15x15 perspectives being recorded with only 64 cameras (cf. Figure 1). Finding the optimal sampling mask in a brute force search by placing 64 samples on a 15x15 grid results in $\approx 10^{57}$ combinatoric possibilities that are infeasible to consider. Therefore, we must reduce the search space by splitting the sampling grid into Q tiles of size U_pxV_p. In our example, we split the grid into 9 tiles of size 5x5.

Each tile can theoretically contain 1 to 25 cameras, and the total number of cameras in all tiles must match 64. Several possibilities exist to distribute cameras within a tile (e.g., 53,120 possibilities to place 5 cameras). To solve this problem, we search for the best pattern for each number of cameras (i.e., 1 to 25) within tile resolution using a global dictionary computed from a set of representative light fields. Here, we assume that coherence values estimate the reconstruction quality sufficiently well.

For capturing the local dictionary that is required for our upsampling, we enforce the center tile to be fully sampled and refer to this tile as *guidance area*. The remaining tiles are selected from the set of best tile patterns while optimizing a global quality metric that is reformulated and solved as a variant of the knapsack problem by mixed-integer linear programming.

A light field recorded with the resulting global sampling mask is upsampled to the full grid resolution by applying sparse reconstruction techniques with the local dictionary extracted from recording the guidance area.

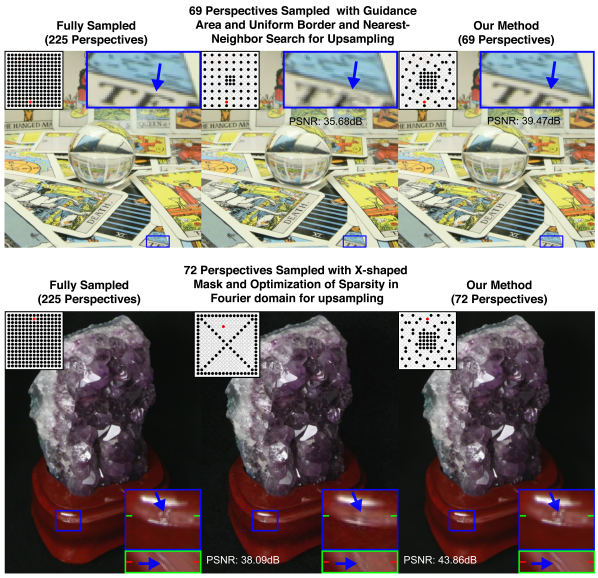

Figure 2: Additional comparisons with related methods [Schedl et al. 2015] and [Shi et al. 2014]. Blue rectangles mark close-ups and green rectangles present EPI images.

3 RESULTS AND OUTLOOK

For evaluation, we used eight light fields to train the global dictionary required for computing the sampling masks, and five test light fields for comparing our results with results of related approaches. The fully sampled light fields serve as ground truth, and we apply PSNR as a measure of reconstruction quality. We compare our approach with the three most related methods that also do not rely on depth reconstruction for upsampling (cf. Figures 1 and 2): Marwah et al. 2013 (we use a uniform mask for sampling and a global dictionary for upsampling), Shi et al. 2014 (uses an X-shaped mask for sampling and an optimization of sparsity in Fourier domain for upsampling), and our previous method [Schedl et al. 2015] (uses a guidance area and uniform border as a sampling mask, and nearest-neighbor search for upsampling). Either, our approach delivers a better quality (higher PSNR) with the same number of cameras, or requires a lower number of cameras for the same quality (similar PSNR). Furthermore, it is much more flexible since it supports an arbitrary number of cameras for determining an optimal sampling mask. Although our experiments indicate that computed sampling masks lead to improved upsampling results even for light fields that differ much from those used in the global dictionary, a comprehensive study is required that leads to general representative global dictionaries for various applications and light field cameras. This will be part of future work.

REFERENCES

K. Marwah, G. Wetzstein, Y. Bando, and R. Raskar. 2013. Compressive Light Field Photography using Overcomplete Dictionaries and Optimized Projections. *ACM Trans. Graph.* 32, 4 (2013), 1–11.

D. C. Schedl, C. Birklbauer, and O. Bimber. 2015. Directional Super-Resolution by Means of Coded Sampling and Guided Upsampling. In *Computational Photography (ICCP), 2015 IEEE International Conference on.* 1–10. DOI:https://doi.org/10.1109/ICCPHOT.2015.7168365

Lixin Shi, Haitham Hassanieh, Abe Davis, Dina Katabi, and Fredo Durand. 2014. Light Field Reconstruction Using Sparsity in the Continuous Fourier Domain. *ACM Trans. Graph.* 34, 1, Article 12 (Dec. 2014), 13 pages. DOI:https://doi.org/10.1145/2682631