Vision, Modeling, and Visualization (2013)
Michael Bronstein, Jean Favre, and Kai Hormann (Eds.)

Poster

Datasets and Benchmarks for Densely Sampled 4D Light Fields

Sven Wanner, Stephan Meister and Bastian Goldluecke

Heidelberg Collaboratory for Image Processing

Abstract

We present a new benchmark database to compare and evaluate existing and upcoming algorithms which are tailored to light field processing. The data is characterised by a dense sampling of the light fields, which best fits current plenoptic cameras and is a characteristic property not found in current multi-view stereo benchmarks. It allows to treat the disparity space as a continuous space, and enables algorithms based on epipolar plane image analysis without having to refocus first. All datasets provide ground truth depth for at least the center view, while some have additional segmentation data available. Part of the light fields are computer graphics generated, the rest are acquired with a gantry, with ground truth depth established by a previous scanning of the imaged objects using a structured light scanner. In addition, we provide source code for an extensive evaluation of a number of previously published stereo, epipolar plane image analysis and segmentation algorithms on the database.

1. Introduction

The concept of a light field was originally used mainly in computer graphics as a powerful tool to describe scene appearance [AB91, LH96], but recently it is also getting more and more attention from the computer vision community. One of the likely reasons is the availability of cheap recording devices. While the first light field capturing techniques used large camera arrays [WJV*05] which are expensive and not very practicable, hand-held light field cameras [Ng06, PW10] are now available on the consumer market.

However, the driving force for successful algorithm development is the availability of suitable benchmark datasets with ground truth data in order to compare results and initiate competition. The current public light field databases we are aware of are the following.

- **Stanford Light Field Archive**
  http://lightfield.stanford.edu/lfs.html
  The Stanford Archives provide more than 20 light fields sampled using a camera array [WJV*05], a gantry and a light field microscope [LNA*06], but none of the datasets includes ground truth disparities.
- **UCSD/MERL Light Field Repository**
  http://vision.ucsd.edu/datasets/lfarchive/lfs.shtml
  This light field repository offers video as well as static light fields, but there is also no ground truth depth available, and the light fields are sampled in a one-dimensional domain of view points only.
- **Synthetic Light Field Archive**
  http://web.media.mit.edu/~gordonw/SyntheticLightFields/index.php
  The synthetic light field archive provides many interesting artificial light fields including some nice challenges like transparencies, occlusions and reflections. Unfortunately, there is also no ground truth depth data available for benchmarking.
- **Middlebury Stereo Datasets**
  http://vision.middlebury.edu/stereo/data/
  The Middlebury Stereo Dataset includes a single 4D light field which provides ground truth data for the center view, as well as some additional 3D light fields including depth information for two out of seven views. The main issue with the Middlebury light fields are that they are designed with stereo matching in mind, and thus the baselines are quite large and thus not representative for plenoptic cameras and unsuitable for direct epipolar plane image analysis.

While there is a lot of variety and the data is of high quality, we observe that all of the available light field databases either lack ground truth disparity information or exhibit large camera baselines and disparities, which is not representative for plenoptic camera data. Furthermore, we believe that a

DOI: 10.2312/PE.VMV.VMV13.225-226

large part of what distinguishes light fields from standard multi-view images is the ability to treat the view point space as a continuous domain. There is also emerging interest in light field segmentation [KSS12, SHH07, EM03, WSG13], so it would be highly useful to have ground truth segmentation data avaliable to compare light field labeling schemes. The above datasets lack this information as well.

Contributions. To alleviate the above shortcomings, we present a new benchmark database which consists at the moment of 13 high quality densely sampled light fields. The database offers seven computer graphics generated datasets providing complete ground truth disparity for all views. Four of these datasets also come with ground truth segmentation information and pre-computed local labeling cost functions to compare global light field labeling schemes. Furthermore, there are six real world datasets captured using a single Nikon D800 camera mounted on a gantry. Using this device, we sampled objects which were pre-scanned with a structured light scanner to provide ground truth ranges for the center view, see figure 1. An interesting special dataset contains a transparent surface with ground truth disparity for both the surface as well as the object behind it - we believe it is the first real-world dataset of this kind with ground truth depth available.

We also contribute a CUDA C library with complete source code for several recently published algorithms to demonstrate a fully scripted evaluation on the benchmark database and find an initial ranking of a small subset of the available methods on disparity estimation. We hope that this will ease the entry into the interesting research area which is light field analysis, and are fully commited to increasing the scope of the library in the future.

The light field archive as well as the full version of this paper, which describes the archive structure and benchmarks in detail, can be downloaded on

http://lightfield-analysis.net.

The CUDA library `cocolib` for continuous convex optimization and light field analysis is available on

http://cocolib.net.

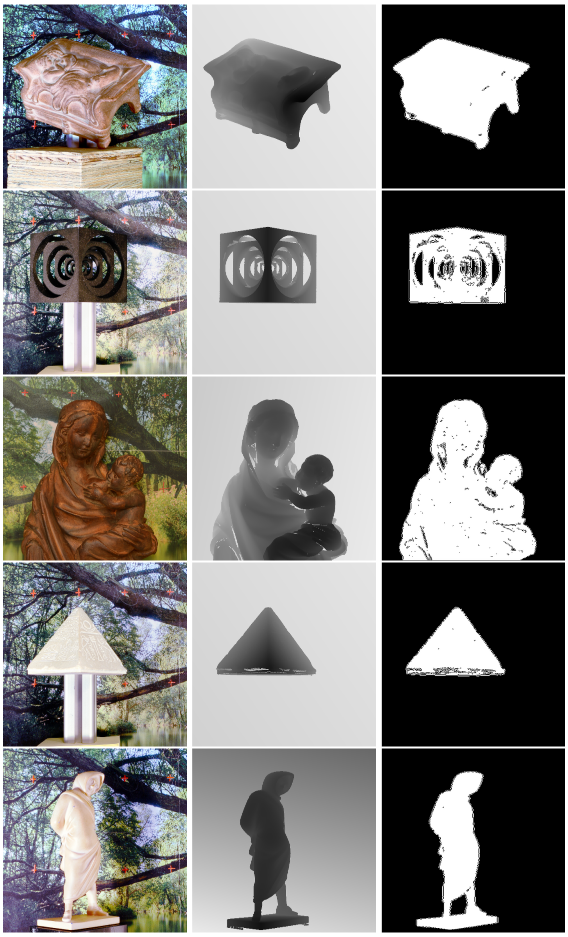

Figure 1: *Datasets in the category Gantry. From left to right: center view, depth channel, mask which indicates regions with valid depth information.*

References

[AB91] ADELSON E., BERGEN J.: The plenoptic function and the elements of early vision. *Computational models of visual processing 1* (1991). 1

[EM03] ESEDOGLU S., MARCH R.: Segmentation with Depth but Without Detecting Junctions. *Journal of Mathematical Imaging and Vision 18*, 1 (2003), 7–15. 2

[KSS12] KOWDLE A., SINHA S., SZELISKI R.: Multiple View Object Cosegmentation using Appearance and Stereo Cues. In *Proc. European Conference on Computer Vision* (2012). 2

[LH96] LEVOY M., HANRAHAN P.: Light field rendering. In *Proc. SIGGRAPH* (1996), pp. 31–42. 1

[LNA*06] LEVOY M., NG R., ADAMS A., FOOTER M., HOROWITZ M.: Light field microscopy. *ACM Transactions on Graphics (TOG) 25*, 3 (2006), 924–934. 1

[Ng06] NG R.: *Digital Light Field Photography*. PhD thesis, Stanford University, 2006. Note: thesis led to commercial light field camera, see also `www.lytro.com`. 1

[PW10] PERWASS C., WIETZKE L.: The Next Generation of Photography, 2010. `www.raytrix.de`. 1

[SHH07] STEIN A., HOIEM D., HEBERT M.: Learning to Find Object Boundaries Using Motion Cues. In *Proc. International Conference on Computer Vision* (2007). 2

[WJV*05] WILBURN B., JOSHI N., VAISH V., TALVALA E.-V., ANTUNEZ E., BARTH A., ADAMS A., HOROWITZ M., LEVOY M.: High performance imaging using large camera arrays. *ACM Transactions on Graphics 24* (July 2005), 765–776. 1

[WSG13] WANNER S., STRAEHLE C., GOLDLUECKE B.: Globally Consistent Multi-Label Assignment on the Ray Space of 4D Light Fields. In *Proc. International Conference on Computer Vision and Pattern Recognition* (2013). 2