Adopting an Unconstrained Ray Model in Light-field Cameras for 3D Shape Reconstruction

Filippo Bergamasco, Andrea Albarelli, Luca Cosmo, Andrea Torsello[1], Emanuele Rodolá, Daniel Cremers[2]

[1]DAIS Università Ca' Foscari Venezia, Venice Italy [2]Department of Computer Science Technische Universität München, Garching, Germany

Due to their recent availability as off-the-shelf commercial devices, light-field cameras have attracted increasing attention from both scientific and industrial operators.

Traditional cameras are designed to capture the amount of light radiation directed toward an image plane. The captured rays can converge to a common projection point (as for the pinhole model), could go through a common axis (as for the models including radial distortion), or can follow any other distribution, even ditching any parametric model. However, regardless of the camera model, the mechanics remains basically projective, and the result of the imaging process is a 2D image.

Light-field cameras pursue a different goal: to capture the full plenoptic function generated by each observed material point [1], which includes the intensity of the light radiating from each point along all the directions over the sphere. Of course, this goal is not practically achievable by any physical sensor, due to the technical and theoretical problem involved. In practice, most, if not all, the light-field devices ever built are made up of an array (explicit or implicit) of traditional cameras, each one contributing to capture a portion of the plenoptic function. An example can be seen in Figure 1, where we show a detail of the composite image captured by a Lytro light-field camera [5]. The number, type and arrangement of such cameras, as well as their calibration, has been a very active topic in recent research.

One of the main hurdles in plenoptic photography derives from the composite imaging formation process which limits the ability to exploit the well consolidated stack of calibration methods that are available for traditional cameras. While several efforts have been done to propose practical approaches, most of them still rely on the quasi-pinhole behaviour of the single microlens involved in the capturing process. This results in several drawbacks, ranging from the difficulties in feature detection, due to the reduced size of each microlens, to the need to adopt a model with a relatively small number of parameters.

This paper makes two main contributions, that we feel to be important to the light-field community.

First, we analyze the use of a calibration method that escapes the need to adopt a parametric model by exploiting dense correspondences generated using phase coding technique [2]. While dense calibration has been already used in literature, this is the first time that it is attempted with light-field cameras and its correct behaviour is not granteed. In fact both the initialization hurdles and the sparsity of the micro-lenses pixel could hinder the process, leading to unsatisfactory results. To this end, we perform an in-depth study of the different aspect of calibration under various conditions.

The second contribution is related to the recovery of 3D shapes. This is a common application of light-field cameras, especially using multiway-stereo algorithms.Unfortunately, most triangulation methods are based on epipolar geometry constraints, that can only be exploited if pinhole cameras are available. To this end, we propose a triangulation step suitable for any camera model and we use it to compare our method with the recent state-of-the-art.

Calibrating Light-field Cameras using a Parameter-Free Model

Following [2], we adopt a non-parametric camera model where each ray is modeled as an independent line within a common reference frame. Such reference frame is not directly related to the physical sensor. In fact, according to this model, image coordinates can be considered just labels for the imaging rays, which are not related to them by means of any analytic function.

A solution space this large needs an exceptional number of observations, and this can only be obtained using a dense coding strategy, which assigns to each image pixel (i.e. to each ray) a pair of coordinates on the calibration

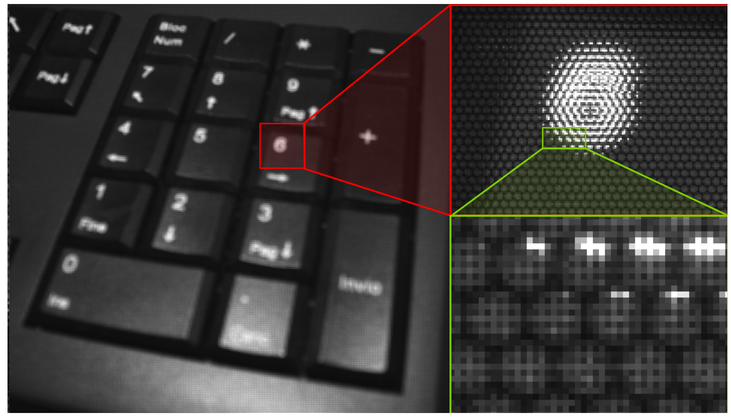

Figure 1: A real-word light-field camera is no more than a tightly packed array of very small (distorted) pinhole cameras sharing the same imaging sensor.

target. There are several ways to do this, in this paper we follow [2] adopting a flat monitor as the target [3, 6] and using a multi period phase shift coding [4] in order to obtain dense target coordinates on a Lytro camera sensor. The coding has been performed both horizontally and vertically. The dense correspondences acquired over several poses of the target can be used to feed the iterative optimization method presented in [2] obtaining the characterization of each ray that has been correctly codified within a large enough number of different poses. Such method, however, has been designed to work on quasi-pinhole cameras and there is no guarantee that it works with a plenoptic camera. Neither it is obvious that the dense coding would work well with the considered imaging process, especially for the higher camera zoom levels.

With our experimental evaluation we show that both target coding and rays calibration work well, to the extent that the complete camera sensor can be calibrated for moderated zoom levels. Additionally, we propose specially crafted techniques for interpolating, selecting and triangulating the several views of the same material point that occurs in different microlens.

[1] Edward H. Adelson and James R. Bergen. The plenoptic function and the elements of early vision. In *Computational Models of Visual Processing*, pages 3–20. MIT Press, 1991.

[2] F. Bergamasco, A. Albarelli, E. Rodola, and A. Torsello. Can a fully unconstrained imaging model be applied effectively to central cameras? In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 1391–1398, June 2013. doi: 10.1109/CVPR.2013.183.

[3] M.D. Grossberg and S.K. Nayar. A general imaging model and a method for finding its parameters. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, pages 108–115 vol.2, 2001. doi: 10.1109/ICCV.2001.937611.

[4] Erik Lilienblum and Bernd Michaelis. Optical 3d surface reconstruction by a multiperiod phase shift method. *Journal of Computers (JCP*, 2:73–83, 2007.

[5] Ren Ng, Marc Levoy, Mathieu Brédif, Gene Duval, Mark Horowitz, and Pat Hanrahan. Light field photography with a hand-held plenoptic camera. *Computer Science Technical Report CSTR*, 2(11), 2005.

[6] Peter Sturm and Srikumar Ramalingam. A generic concept for camera calibration. In *Computer Vision - ECCV 2004*, volume 3022 of *Lecture Notes in Computer Science*, pages 1–13. Springer Berlin Heidelberg, 2004. ISBN 978-3-540-21983-5.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.