Computer Vision and Image Understanding 145 (2016) 140–147

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Stepwise calibration of focused plenoptic cameras

Klaus H. Strobl*, Martin Lingenauber

Robotics and Mechatronics Center, German Aerospace Center (DLR), D-82230 Wessling, Germany

ARTICLE INFO

Article history:
Received 15 December 2014
Accepted 22 December 2015

Keywords:
Focused plenoptic camera
Light-field
Plenoptic 2.0
Metric calibration

ABSTRACT

Monocular plenoptic cameras are slightly modified, off-the-shelf cameras that have novel capabilities as they allow for truly passive, high-resolution range sensing through a single camera lens. Commercial plenoptic cameras, however, are presently delivering range data in non-metric units, which is a barrier to novel applications e.g. in the realm of robotics. In this work we revisit the calibration of focused plenoptic cameras and bring forward a novel approach that leverages traditional methods for camera calibration in order to deskill the calibration procedure and to increase accuracy. First, we detach the estimation of parameters related to either brightness images or depth data. Second, we present novel initialization methods for the parameters of the thin lens camera model—the only information required for calibration is now the size of the pixel element and the geometry of the calibration plate. The accuracy of the calibration results corroborates our belief that monocular plenoptic imaging is a disruptive technology that is capable of conquering new markets as well as traditional imaging domains.

1. Introduction

A considerable number of machine vision users think that multi-view triangulation is required in order to retrieve accurate 3-D information using cameras without a priori information about the scene. This is predominantly accomplished either by passively taking images from different vantage points (stereo vision) or by actively projecting a known pattern from a separate location (Kinect). In other words, the notion that 3-D information is lost when light rays traverse the front lens of the camera is widespread. Experts know, however, that this is not the case as light rays are differently diffracted by a lens depending on the distance to the emitting object [1]. In fact, one of the potential outputs of monocular plenoptic cameras is range sensing.

1.1. Plenoptic imaging

Plenoptic (also light-field) imaging is about measuring light in a higher dimensionality than in standard 2-D imaging. In fact, light transmission can be contemplated in a higher-dimensional space, the so-called plenoptic function [2]. Current plenoptic imaging samples the plenoptic function in 4-D , viz. 2-D projection position on the chip together with the direction of incoming light rays. The quest for this extra information is anything but new, but advances in parallel computation and modern workmanship of microlens arrays (MLAs) have recently made commercial products possible [3–5]. In a nutshell, monocular plenoptic cameras capture the 3-D image produced by the main lens within the camera by using an MLA in front of the sensor chip. By capturing the whole 3-D image, classical habits when using planar sensors like keeping the aperture size small in order to increase depth of field are lifted and more light can be gathered from the scene. Different camera designs open up new possibilities to trade off lateral precision against angular resolution of the reprojected ray directions. The original monocular plenoptic cameras focus the image on the MLA, achieving a limited spatial resolution at that particular depth equal to the number of valid microlenses. These microlenses produce defocused images that sample the ray direction at the position of the microlens. In 2009 Lumsdaine and Georgiev introduced the focused plenoptic camera (or plenoptic camera 2.0), which makes it possible to adapt this rather rigid trade-off between angular and spatial resolutions towards more spatial resolution [6]. This is performed by a modification of the focus distance to the main lens with the result that microlenses produce focused images that, on the other hand, more loosely sample ray direction.

Many characteristics of plenoptic cameras are in conformity with the standard reference on disruptive technologies in Ref. [7]. For instance, plenoptic cameras initially produce a deficient standard output (fair images) at a higher cost, which makes them of no interest to the average consumer. They, however, clearly have the potential to improve and open up new markets while sharply reducing costs. Their current applications are offline refocusing and

* Corresponding author.
E-mail address: klaus.strobl@dlr.de (K.H. Strobl).

http://dx.doi.org/10.1016/j.cviu.2015.12.010
1077-3142/

total focusing (*i.e.*, increased depth of field) of 2-D images. More relevant potential applications are passive, 3-D video recording, 3-D modeling, range-based segmentation and tracking, industrial inspection (e.g. in narrow cavities), and imaging in challenging, low-light environments (e.g. underwater or in space). Most of these applications rely on the capability of plenoptic cameras to provide *metric* information of the scene in the form of 2.5-D depth images. This is, however, not yet commercially available as depth is currently being delivered in internal units related to image processing (disparities). We address the metric calibration of focused, monocular plenoptic cameras in order to transform their depth output into metric space.

1.2. Related work

There is less research on the metric calibration of focused plenoptic cameras and the works on the calibration of the original, unfocused ones are only of partial use [8]. Next we review the only available approaches in Refs. [9–12]. They all have in common that they start out from synthetic images generated by the RxLive software of Raytrix GmbH (viz. the total focus image and the depth image), not from the raw images of the camera. The conformity of the generated synthetic images with the camera models used for calibration is of course critical. It is a judicious decision to rely on the manufacturers, however, since (i) they are most qualified to do that job, (ii) they still keep individual design details in secret, and (iii) in order to avoid mismatching between our potential reconstruction attempts and the eventual operation on GPGPUs. In addition, the calibration process is simplified by using the synthetic images because we leverage established methods for pinhole camera calibration [13,14]. It is worth noting that the geometry of the MLA is not included in the calibration process as it can be estimated in a separate procedure using the Raytrix software. The best-known work in Ref. [9] details the modeling and calibration of the focused plenoptic camera, failing to obtain absolute range accuracy. Further the automatic initialization of calibration parameters is not addressed; the authors make use of privileged information from the manufacturer. The recent Master's thesis in Ref. [12], however, does achieve superior results by largely implementing the above approach. Still, considerations on the initialization of calibration parameters are not being addressed. The author makes strong use of filtering approaches to wipe out peripheral artifacts in depth estimation, which might constrain the general applicability of the approach. Zeller et al. in Ref. [11] perform calibration by minimizing the reprojection residuals with respect to (w.r.t.) a set of measured calibration points for which the object distance is known—at least for the initialization of their method. In addition, the method requires assumed intrinsic values. In the same spirit, Luhmann et al. in Ref. [10] opt for measuring ranges of a planar calibration object, which is error-prone and inconvenient [15,16]. Incidentally, the current internal approach for metric calibration at Raytrix GmbH also relies on a linear actuator in order to produce a polynomial that directly converts virtual depths into metric distances [12]. This type of empirical models, however, is only applicable within the scope of the calibration data.

1.3. Contributions

In this work we revisit the type of calibration approaches that are based on the standard camera calibration method described in Refs. [13,14]. We suggest modifications to particular modeling details and present justifications. Special care has been taken to keep the approach in the spirit of the standard method, *i.e.*, to take images of a known planar calibration pattern in unknown pose and to facilitate automatic bootstrapping of the parameters prior to nonlinear optimization. This keeps the amount of required prior knowledge (e.g. specifications by the manufacturer) to a minimum, making the whole calibration process more generic and easier. More importantly, we introduce a novel approach for stepwise calibration by alternate use of total focus and depth (synthetic) images. Our motivation is to avoid the impact of higher levels of noise in the depth images on a large part of the intrinsic parameters like the focal length and the radial lens distortion. These parameters can be estimated in advance by exclusively using total focus images, as in traditional camera calibration. After that, the optimization of the remaining parameters using the depth images and the results of the first optimization takes place, see Fig. 1. By doing so, calibration accuracy is increased and the optimization robustness is promoted as the formulations of both optimizations become better conditioned compared with joint optimization methods [9]. In addition, the optimization of the lens distortion model will not get entangled with the optimization of the (potentially very similar) depth distortion model.

Overall, we produce an easy-to-use, automatic method for metric plenoptic camera calibration. The only required data are synthetic total focus and virtual depth images of a planar calibration plate of known geometry and the metric size of their virtual sensor elements.

2. Proposed method

2.1. The thin lens camera model and the focused plenoptic camera

The pinhole camera model is a valid approximation for most cameras and applications. It relates projection directions in the camera reference frame S_C with projection positions in the sensor reference frame S_S. This projection is independent of the actual range to the scene, which makes it unsuitable for modeling plenoptic cameras aimed at inferring the depth of the scene out of inner camera projections. The pinhole camera model is derived from the thin lens camera model in the case of smaller aperture sizes. The thin lens camera model embraces the thin lens approximation of light rays passing through a thin lens, which states that ray directions are still projected following the pinhole camera

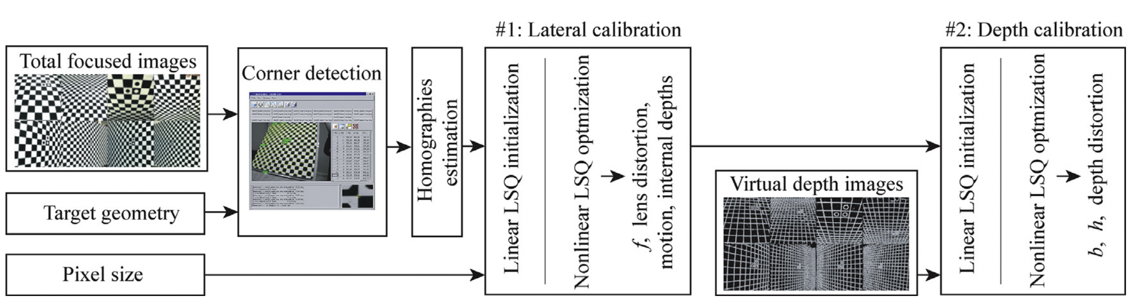

Fig. 1. Diagram representing the information flow in the presented method.