ISPRS Journal of Photogrammetry and Remote Sensing 118 (2016) 83–100

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Depth estimation and camera calibration of a focused plenoptic camera for visual odometry

Niclas Zeller [a,b,*], Franz Quint [a], Uwe Stilla [b]

[a] Faculty of Electrical Engineering and Information Technology, Karlsruhe University of Applied Sciences, Moltkestraße 30, 76133 Karlsruhe, Germany
[b] Department of Photogrammetry and Remote Sensing, Technische Universität München, Arcisstraße 21, 80333 München, Germany

ARTICLE INFO

Article history:
Received 27 July 2015
Received in revised form 15 March 2016
Accepted 27 April 2016
Available online 18 May 2016

Keywords:
Accuracy analysis
Camera calibration
Focused plenoptic camera
Light-field
Probabilistic depth map
Visual odometry

ABSTRACT

This paper presents new and improved methods of depth estimation and camera calibration for visual odometry with a focused plenoptic camera.

For depth estimation we adapt an algorithm previously used in structure-from-motion approaches to work with images of a focused plenoptic camera. In the raw image of a plenoptic camera, scene patches are recorded in several micro-images under slightly different angles. This leads to a multi-view stereo-problem. To reduce the complexity, we divide this into multiple binocular stereo problems. For each pixel with sufficient gradient we estimate a virtual (uncalibrated) depth based on local intensity error minimization. The estimated depth is characterized by the variance of the estimate and is subsequently updated with the estimates from other micro-images. Updating is performed in a Kalman-like fashion. The result of depth estimation in a single image of the plenoptic camera is a probabilistic depth map, where each depth pixel consists of an estimated virtual depth and a corresponding variance.

Since the resulting image of the plenoptic camera contains two plains: the optical image and the depth map, camera calibration is divided into two separate sub-problems. The optical path is calibrated based on a traditional calibration method. For calibrating the depth map we introduce two novel model based methods, which define the relation of the virtual depth, which has been estimated based on the light-field image, and the metric object distance. These two methods are compared to a well known curve fitting approach. Both model based methods show significant advantages compared to the curve fitting method.

For visual odometry we fuse the probabilistic depth map gained from one shot of the plenoptic camera with the depth data gained by finding stereo correspondences between subsequent synthesized intensity images of the plenoptic camera. These images can be synthesized totally focused and thus finding stereo correspondences is enhanced. In contrast to monocular visual odometry approaches, due to the calibration of the individual depth maps, the scale of the scene can be observed. Furthermore, due to the light-field information better tracking capabilities compared to the monocular case can be expected.

As result, the depth information gained by the plenoptic camera based visual odometry algorithm proposed in this paper has superior accuracy and reliability compared to the depth estimated from a single light-field image.

1. Introduction

The concept of a plenoptic camera has been developed more than hundred years ago by Ives (1903) and Lippmann (1908). However, only for the last few years the existing graphic processor units (GPUs) are capable to evaluate the recordings of a plenoptic camera with acceptable frame rates ($\geqslant$25 fps).

To gather the whole 4D light-field (hence the name "plenoptic"), a microlens array (MLA) is placed in front of the sensor. Today there exist basically two main concepts of MLA based plenoptic cameras: The unfocused plenoptic camera proposed by Adelson and Wang (1992) and developed further by Ng (2006) and the focused plenoptic camera, which was described for the first time by Lumsdaine and Georgiev (2008). Compared to the unfocused plenoptic camera, the focused plenoptic camera has a higher spatial resolution but lower angular resolution. This high spatial resolution is especially beneficial for estimating depth out

* Corresponding author.
E-mail addresses: niclas.zeller@hs-karlsruhe.de (N. Zeller), franz.quint@hs-karlsruhe.de (F. Quint), stilla@tum.de (U. Stilla).

http://dx.doi.org/10.1016/j.isprsjprs.2016.04.010
0924-2716/

of the recorded raw image (Perwaß and Wietzke, 2012). In our research we are using a focused plenoptic camera.

The accuracy of the depth information gathered by a focused plenoptic camera is rather low for a distance of a few meters compared to other depth sensors, like Time-of-Flight (TOF) cameras or stereo camera systems with a large baseline, at least at a comparable field of view (FOV). Besides, the depth accuracy of a focused plenoptic camera strongly decays when reducing the focal length. Thus, a trade-of between wide FOV and acceptable accuracy has to be found.

On the other hand as shown by Perwaß and Wietzke (2012), the plenoptic camera offers a much larger depth of field (DOF) compared to a monocular camera at the same aperture. Thus, a plenoptic camera has a much shorter close range limit than e.g. a stereo camera system.

Another plus for plenoptic cameras are their small dimensions, which are similar to those of a conventional camera. In future there will also be miniaturized light-field sensors available, which can be assembled in smartphones (Venkataraman et al., 2013).

In many navigation applications such small sensors are profitable, for example on unmanned aerial vehicles (UAVs), where space and weight is limited. But also for indoor navigation or blind people assistance, where bulky sensors can be annoying, such small and light sensors are beneficial.

For this kind of applications today mostly monocular visual odometry (or Simultaneous Localization and Mapping (SLAM)) systems are used, which gain depth information from camera motion. However, such monocular systems come with some drawbacks. One drawback of a monocular visual odometry system is its scale ambiguity. Thus, especially in navigation applications additional sensors are needed to gather metric dimensions. Another disadvantage of the monocular system is that no depth is obtained without any motion of the camera or for rotations around the camera's optical center.

Thus, a plenoptic camera seems to be a good compromise between a monocular and a stereo camera for a visual odometry system. Since for a plenoptic camera rough depth information is available for each single frame, it is to be expected that tracking will become much more robust compared to a monocular system.

Fig. 1 shows two typical scenarios for indoor navigation recorded by a plenoptic camera. Here, far as well as very close objects with less than one meter distance to the camera are present in the same scene. In such scenes a plenoptic camera benefits from its large DOF. Even though the scene has a high variation in depth the camera is able to record the whole scene in focus.

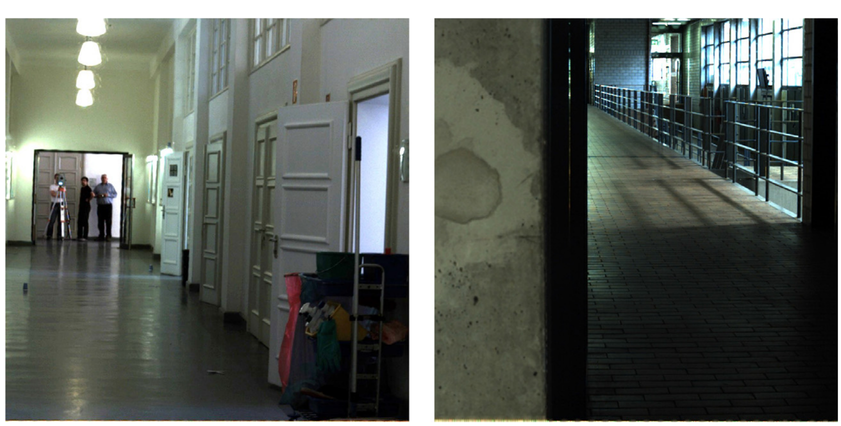
The presented work unifies and extends the previous publications (Zeller et al., 2014, 2015a,b). Thereby, this paper provides the complete work-flow of a focused plenoptic camera based visual odometry. We firstly present the concept of the focused plenoptic camera and how depth can be estimated in principle out of the recorded light-field (Section 2). Additionally, we derive the theoretically achievable depth accuracy of the camera (Section 3). Afterwards, we propose a new depth estimation algorithm which estimates a probabilistic depth map from a single recording of a focused plenoptic camera (Section 4). The probabilistic depth map will be beneficial for the visual odometry system. Like any camera system that is used for photogrammetric purposes, a plenoptic camera has to be calibrated. We present a complete framework to calibrate a focused plenoptic camera, especially for an object distance range of several meters (Section 5). For this we develop three different depth models and compare them to each other. Finally we incorporate the depth estimation as well as the camera calibration into a focused plenoptic camera based visual odometry system (Section 6). All proposed methods are extensively evaluated (Section 7).

1.1. Related work

1.1.1. Depth estimation

For the last years various algorithms for depth estimation based on the recordings of plenoptic cameras or other light-field representations have been developed. First methods were published even more than 20 years ago (Adelson and Wang, 1992).

Since light-field based depth estimation represents a multidimensional optimization problem, always a trade-off between low complexity and high accuracy or consistency has to be chosen. Wanner and Goldluecke (2012, 2014) for instance present a globally consistent depth labeling which is performed directly on the 4D light-field representation and results in a dense depth map. Jeon et al. (2015) make use of the phase-shift theorem of the Fourier transform to calculate a dense, light-field based disparity map with sub-pixel accuracy, while Heber and Pock (2014) use principal component analysis to find the optimum depth map. Some approaches make use of geometric structures like 3D line segments (Yu et al., 2013) to improve the estimate and to reduce complexity. Tosic and Berkner (2014) present a so called scale-depth space which provides a coarse depth map for uniform regions and a fine one for textured regions. Other methods reduce complexity by the use of local instead of global constraints and thus result in a sparse depth map. Such sparse maps supply depth only for textured



(a) (b)

Fig. 1. Two scenarios typical for indoor navigation recorded by a focused plenoptic camera. Here, far as well as very close objects with less than one meter distance to the camera are present in the same scene. Due to the large DOF of a plenoptic camera such scenes with a high variation in depth still can be recorded completely in focus by the plenoptic camera.