ISPRS Journal of Photogrammetry and Remote Sensing 114 (2016) 1–9

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

On the accuracy potential of focused plenoptic camera range determination in long distance operation

Hannes Sardemann [*], Hans-Gerd Maas

TU Dresden, Institute of Photogrammetry and Remote Sensing, Helmholtzstr. 10, 01069 Dresden, Germany

ARTICLE INFO

Article history:
Received 31 July 2015
Received in revised form 19 January 2016
Accepted 19 January 2016
Available online 12 February 2016

Keywords:
Plenoptic camera
Light field
Range image
Accuracy

ABSTRACT

Plenoptic cameras have found increasing interest in optical 3D measurement techniques in recent years. While their basic principle is 100 years old, the development in digital photography, micro-lens fabrication technology and computer hardware has boosted the development and lead to several commercially available ready-to-use cameras. Beyond their popular option of a posteriori image focusing or total focus image generation, their basic ability of generating 3D information from single camera imagery depicts a very beneficial option for certain applications.

The paper will first present some fundamentals on the design and history of plenoptic cameras and will describe depth determination from plenoptic camera image data. It will then present an analysis of the depth determination accuracy potential of plenoptic cameras. While most research on plenoptic camera accuracy so far has focused on close range applications, we will focus on mid and long ranges of up to 100 m. This range is especially relevant, if plenoptic cameras are discussed as potential mono-sensorial range imaging devices in (semi-)autonomous cars or in mobile robotics.

The results show the expected deterioration of depth measurement accuracy with depth. At depths of 30–100 m, which may be considered typical in autonomous driving, depth errors in the order of 3% (with peaks up to 10–13 m) were obtained from processing small point clusters on an imaged target. Outliers much higher than these values were observed in single point analysis, stressing the necessity of spatial or spatio-temporal filtering of the plenoptic camera depth measurements. Despite these obviously large errors, a plenoptic camera may nevertheless be considered a valid option for the application fields of real-time robotics like autonomous driving or unmanned aerial and underwater vehicles, where the accuracy requirements decrease with distance.

1. The plenoptic camera

With plenoptic cameras, a new sensor technology has evolved in the last decade, that is capable of determining distances and 3D positions within one camera. An array of thousands of microscopic small lenses is placed in front of the image plane. Each lens generates a small image on the imaging sensor. In addition to the popular option of a posteriori image focussing, the multiple images of the scene form a basis for the determination of 3D information from single camera imagery. The compact design makes it interesting for real-time robotic applications like autonomous driving or unmanned aerial and underwater vehicles. In these kind of applications the accuracy requirements decrease with distance. Accuracy is rather uncritical at large distances, while higher requirements are posed at smaller distances. This matches the accuracy behaviour of plenoptic cameras.

With the goal of applying plenoptic cameras in the fields of real-time robotics or (semi-)autonomous cars in mind, the paper will analyse the accuracy of distance determination in mid and long range operation of up to 100 m. Following fundamental considerations on the history and design of plenoptic cameras and the principles of depth determination from plenoptic camera image data, experiments on depth accuracy will be evaluated. In a depth range of 1–100 m, accuracy analysis was conducted using a focused plenoptic camera of the type R42 from the manufacturer Raytrix.

1.1. Historical background

The idea of a plenoptic camera or light field camera goes back to the beginning of the 20th century. Frederic E. Ives (1903)

---

[*] Corresponding author.
E-mail addresses: hannes.sardemann@tu-dresden.de, hannes.sardemann@gmx.de (H. Sardemann), hans-gerd.maas@tu-dresden.de (H.-G. Maas).

http://dx.doi.org/10.1016/j.isprsjprs.2016.01.012

registered an US patent describing a camera and projector for recording and observing stereograms using a pinhole array. A few years later, Lippmann (1908) published his concept of integral photography, using an array of micro-lenses instead of pinholes placed in front of the photographic plate. In the following decades the concept of integral and lenticular photography evolved (Roberts, 2003) until the 1990s, when modern computer hardware, digital photography and the ability of building microscopic small lenses has brought a new boost to cameras with micro-lens arrays (MLA).

Adelson and Wang (1992) described a camera, that can achieve stereo photography with a single lens and were the first to name it plenoptic camera. They put an array of 100×100 micro-lenses in front of the imaging sensor. Each micro-lens is focused to infinity and projects an image of the aperture to a 5×5 pixel area on the sensor. This technique is known as the standard plenoptic camera. Ng et al. (2005) realised a hand-held plenoptic camera, that works with that principle. It has a 16 megapixel sensor and 87,000 micro-lenses each projecting onto 14×14 pixels. Since 2012, Lytro,[1] a company founded by Ng, offers plenoptic cameras with similar specifications for the consumer market.

Lumsdaine and Georgiev (2009) presented a different light field camera approach. By using micro-lenses that are focused to the projected main image inside the camera, they gather micro-images with only slightly overlapping fields of view. Each micro-lens captures a small part of the main image. This technique is known as the focused plenoptic camera or plenoptic camera 2.0 and is superior in lateral resolution compared to the standard plenoptic camera. To improve the depth of field, micro-lenses with different focal length can be used (Georgiev and Lumsdaine, 2012). Raytrix[2] offers multi-focus plenoptic cameras (Perwass and Wietzke, 2012) in different sensor and MLA configurations for industrial use.

1.2. Functional principle of a focused plenoptic camera

The outer design of a plenoptic camera is the same as for a standard digital camera. The main lens with focal length f_L projects an image of an object at distance a_L at the position b_L behind the lens. The projection follows the lens equation:

$$\frac{1}{f_L} = \frac{1}{a_L} + \frac{1}{b_L} \tag{1}$$

Inside the plenoptic camera is a micro-lens array (MLA) with $n_1 \times n_2$ hexagonally packed micro-lenses. Each micro-lens has a diameter D and projects onto an area of $N \times N$ pixels.

While in a standard plenoptic camera (Adelson and Wang, 1992) the MLA is placed at the focal plane b_L, in focused plenoptic cameras it is placed at a small distance in front of or behind the focal plane b_L. The imaging sensor is placed at a distance b behind the MLA (Fig. 1). The micro-lenses are focused on the main image and project it on the sensor.

There are two versions of focused plenoptic cameras (Georgiev and Lumsdaine, 2009): The Keplerian mode, where the MLA and the image sensor are placed behind the main image (Fig. 1) and the Galilean mode with MLA and sensor placed between the main lens and the main image (Fig. 2). In the Galilean mode, the main image would be formed behind the MLA and the image sensor. Therefore, it will not actually appear and is only virtual, but can still be focused on the sensor. This version is applied in Raytrix cameras.

1.3. Rendering a total focus image

In standard plenoptic cameras, each micro-image can be seen as a N^2 big pixel. Thus, a fully generated picture has a lateral resolution of $n_1 \times n_2$ pixels. Another method for rendering images is to choose one pixel at the same position $(i,j) \in (N \times N)$ of each micro-image and stack them together to one image. By this, N^2 images from different points of view can be rendered. There are also techniques to render images of higher resolution (Lumsdaine and Georgiev, 2008; Bishop et al., 2009; Bishop and Favaro, 2011).

In focused plenoptic cameras, each micro-lens projects only a part of the image. The MLA is placed in a way that each object in the depth of field is seen by at least two micro-lenses. Objects close to the camera appear in more micro-images than objects far away (Fig. 3). By connecting unique patches of the micro-images, a focused image can be rendered (Georgiev and Lumsdaine, 2010). The size of a patch depends on its depth. Thus, depth has to be known for rendering. The effective lateral resolution of the rendered image is proportional to depth and has a maximum of $1/4$ of the imaging sensor resolution for far away objects. Another method for generating a total focus image is to average the grey values of corresponding pixels in two or more micro-images (Perwass and Wietzke, 2012).

1.4. Calculating a depth map

The focused plenoptic camera measures depth not directly from the scene, but from a (virtual) 3D main image inside the camera. The main lens projects the scene inside the camera. The decreased, flipped and depth-crushed image is then imaged on the sensor by the micro-lenses. By image matching, homologous points can be detected in two or more micro-images and their depth in the main image can be calculated. Thus, the distance a between MLA and image point can be determined. Since b differs with camera models, a relative depth map unit, the virtual depth, is defined as

$$v = \frac{a}{b} \tag{2}$$

where b is the distance between MLA and imaging sensor. All parts of the main image, that are at a distance $a \geqslant 2b$ from the MLA appear in at least two micro-images. Thus, $v \geqslant 2$ applies for virtual depth.

Since the depth map represents distances in the flipped main image, objects that are far away from the camera appear close to the micro-lenses (Fig. 2). Thus, small virtual depths mean large object distances and vice versa. It can be shown, that the virtual depth accuracy decays approximately proportional with v (Zeller et al., 2014): The depth accuracy of any point in the main image, that is measured with two neighbouring micro-lenses, decays proportional to the square of the virtual depth. This effect is the same as in stereo photogrammetry. Furthermore, accuracy is proportional to the base length. The relation can be described as

$$\sigma_v \sim \frac{v^2}{B} \tag{3}$$

with base length B. In plenoptic cameras, the base is not constant, but depends on the number of micro-lenses used for depth determination:

$$B = k \cdot D \tag{4}$$

with $k = 1, \sqrt{3}, 2, \ldots$ for a hexagonally packed MLA (Fig. 4). Parts of the main image appear in more than two micro-images, the further away they are from the MLA (i.e. the closer the object is to the camera). Fig. 3 shows that behaviour for an object close and far from the camera. For an object, which is very close to the camera (Fig. 3a),

---

[1] www.lytro.com.

[2] www.raytrix.de.