

Geometry based Three-Dimensional Image Processing Method for Electronic Cluster Eye

Sidong Wu^a Gexiang Zhang^{a,b}, Ming Zhu^c, Tao Jiang^c, and Ferrante Neri^{d,1}

^a*School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, People's Republic of China*

^b*Robotics Research Center, Xihua University, Chengdu 610039, People's Republic of China*

^c*Control Engineering College, Chengdu University of Information Technology, Chengdu 610225, People's Republic of China*

^d*Centre for Computational Intelligence, School of Computer Science and Engineering, De Montfort University, The Gateway House, LE1 9BH, Leicester, United Kingdom*

Abstract.

In recent years, much attention has been paid to the electronic cluster eye (eCley), a new type of artificial compound eyes, because of its small size, wide field of view (FOV) and sensitivity to motion objects. An eCley is composed of a certain number of optical channels organized as an array. Each optical channel spans a small and fixed field of view (FOV). To obtain a complete image with a full FOV, the images from all the optical channels are required to be fused together. The parallax from unparallel neighboring optical channels in eCley may lead to reconstructed image blurring and incorrectly estimated depth. To solve this problem, this paper proposes a geometry based three-dimensional image processing method (G3D) for eCley to obtain a complete focused image and dense depth map. **In G3D, we derive the geometry relationship of optical channels in eCley to obtain the mathematical relation between the parallax and depth among unparallel neighboring optical channels. Based on the geometry relationship, all of the optical channels are used to estimate the depth map and reconstruct a focused image. Subsequently, by using an edge-aware interpolation method, we can further gain a sharply focused image and a depth map.** The effectiveness of the proposed method is verified by the experimental results.

Keywords. Image processing, Electronic cluster eye, Artificial eyes, Mobile robotics, Geometric algorithms

1. Introduction

Modern robots are progressively becoming more realistic and similar to humans [23] being asked to solve tasks such as path planning [58], path following [18], trajectory tracking [49], reproducing the behavior of human limbs [24].

A fundamental role in robotics is played by Artificial Intelligence [55], Computational Modelling [32, 57], and Metaheuristic Optimisation. Algorithms be-

longing to the last category are usually inspired by nature. Some recently proposed examples include inspiration by music [41,40], physics [43,42,44], swarm behavior [34,37,36]. Artificial Intelligence is often integrated within robot controllers [45,28] and combined to neural systems for Machine Intelligence [19,2,13].

The eye of the robot is very important and complex device since it is one of the sensors that allow a robot to make decisions and thus being autonomous. Several problems arise in relation to the eyes of the robots. Some are related design of the eye itself, e.g. retina modelling [25] while others are related to the interpretation of what the cameras incorporated the eye's capture. Some examples include image segmentation [3], feature extraction [47,35], image recognition

¹Corresponding Author: Ming Zhu, Control Engineering College, Chengdu University of Information Technology, Chengdu 610225, People's Republic of China, E-mail: zhuming@cuit.edu.cn; Ferrante Neri, School of Computer Science and Engineering, De Montfort University, The Gateway House, LE1 9BH, Leicester, United Kingdom, E-mail: fneri@dmu.ac.uk.

[7,26], and image processing in medical applications [16,17].

The present paper focuses on robots' eyes and, more specifically on artificial compound eyes. Inspired by natural compound eyes, an artificial compound eye is manufactured to be an alternative imaging system to miniaturize an optical system [48]. Due to many merits, such as small size, wide field of view (FOV) and sensitivity to motion objects, much attention has been paid to artificial compound eye systems in recent years [51]. Inspired by a wasp parasite called *Xenos Peckii*, an electronic cluster eye (eCley) [5] was introduced by Brückner *et al.* in 2010 as a new type of artificial compound eyes. The details about the optical module, the working principle and the manufacturing process can be referred to [5]. The camera module of eCley is shown in Fig. 1 and its basic system parameters are listed in Table 1.

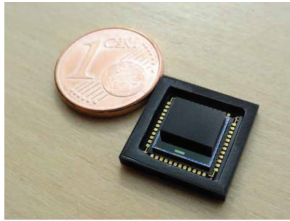


Figure 1. The eCley module (The one cent coin is used for the size comparison [5]).

Table 1. Parameters of the eCley system in [5]

Property	Value
Size of eCley ($L \times W \times H$)	$6.8mm \times 5.2mm \times 1.4mm$
FOV ($H \times V$)	$58^\circ \times 46^\circ$
F-number	3.7
Focal length	$0.778 mm$
No. of channels	13×17
pixel pitch	$3.2 \mu m$
diameter of microlens	$0.375-0.386 mm$
angular shift	4°

An eCley is composed of a certain number of optical channels organized as an array. The array of microlenses is used for imaging. As shown in Fig 2, each imaging channel can capture only a small part of the full FOV, so the FOV of each imaging channel is small and the FOVs of neighboring channels are partial over-

lapped. By using this structure, the total track length is only about $1.4mm$, which is much shorter than single-aperture camera [1]. In addition, since the focal length of eCley is only $0.778mm$, it has a merit that the eCley focuses on very closed objects easily. Therefore, eCley can be used for the navigation of micro-robot and the internal examination based on micro capsule robot. While the captured image is composed of a series of separated, small FOV and low-resolution images. It is hard for robots to make decision based on the small FOV and low-resolution image of a single channel or even several channels, thus the post-process to obtain a complete high-resolution image is required.

In the eCley system, a complete image with a full FOV is obtained by fusing the images from all the optical channels. The principal optic axes of optical channels are considered to be unparallel so as to gain a larger FOV than the parallel case on condition that the overall size of eCley is not changed. Each channel has its inherent offset against its neighboring channel. Thus, the parallaxes of objects at the same distance are not identical in different channels. The parallax from unparallel neighboring optical channels in eCley may lead to reconstructed image blurring and incorrectly estimated depth.

To solve this problem, a geometry based three-dimensional image processing method (G3D) for eCley is proposed to obtain a complete focused image and dense depth map. The main idea of G3D is the introduction of the geometry relationship of optical channels for eCley. Based on the geometry relationship, the mathematical relation between the parallax and depth among unparallel neighboring optical channels can be derived. Further, a sharply focused image and a depth map can be gained. Compared with the methods in [30,31,59], G3D has three main advantages:

1. G3D can reconstruct a complete all-in-focus image.
2. Based on the geometry relationship derived, all channel images can be directly used to estimate the depth of an object.
3. G3D can reach sub-pixel precision in principle, because the matching is based on the space coordinates, instead of pixel shift.

The rest of this paper is organized as follows. Section 2 reviews the related work. Section 3 presents the

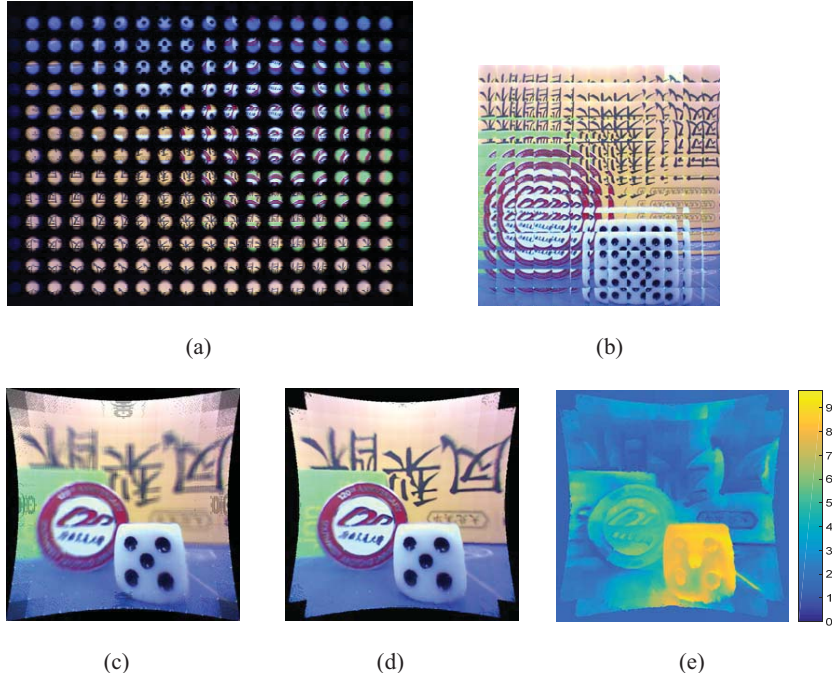


Figure 2. (a) The captured image. (b) The extracted images after flat field correction. (c) Reconstructed image focused on a single distance. (d) The reconstructed image of our method. (e) The depth map of our method.

geometry based three-dimensional image processing method. Experimental results are shown in Section 4. Finally, conclusions are drawn in Section 5.

2. Related Work

This study discusses the three-dimensional image processing method with respect to eCley covering object depth estimation and image reconstruction. So this section starts from brief introductions of object depth estimation and image reconstruction in the area of computer vision, and then moves to the survey of the related work on artificial compound eyes.

Object depth estimation is one of the most active research topics in computer vision, see [56,6]. According to the different setting of system parameters, two different methods, depth from defocus (DFD) [52] and stereo matching [38,11], are widely studied. DFD attempts to extract depth information from multiple images (at least two) captured at different focus setting [4], while the stereo matching method tries to obtain depth from multiple images captured

at different positions. The performance analysis and comparisons between DFD and stereo matching algorithms was presented in [39]. Moreover, many public test images can be found in the web site by the link: <http://vision.middlebury.edu/>, and plenty of stereo matching and multi-view stereo algorithms have been listed and evaluated.

In the meantime, super-resolution (SR) image reconstruction is also a very attractive research area over the past two decades. An SR reconstruction algorithm is to obtain one or multiple high-resolution images from one or multiple low-resolution images. In general, SR algorithms consist of four main categories: nonuniform interpolation approach, frequency domain approach, regularized SR reconstruction approach and projection onto convex sets approach [33,27].

In the area of artificial compound eyes, except for the system design, image reconstruction and depth estimation are two active research areas [51]. In SR reconstruction, Kitamura *et al.* proposed a pixel rearrange method to reconstruct the high-resolution image based on Thin Observation Module by Bound Optics (TOMBO) [21]. In [29], an iterative back projec-

tion method (IBP) was presented to further improve the reconstructed results. El-Sallam *et al.* introduced a spectral-based blind method for TOMBO [9,10], but the point spread function (PSF) of sub-images should be known before. As for depth estimation, most algorithms were designed for TOMBO. In [53], a multiple base-lines method was discussed. In [14], a sum of squared difference (SSD) evaluation based back-project method was described based on the fact that a back-project method will result in poor reconstruction performance when the object is located at the wrong position. Especially for eCley, a braiding algorithm was introduced in [30,31] to obtain the reconstructed image with the final resolution 640*480. The final image can only focus on a specified distance at a time. If the object is not in the specified distance, the image blurring appears. In [59], Ziegler *et al.* presented an algorithm based on stereo matching and fusion method to obtain the final depth map and complete image. The depth of each channel image was obtained by fusing the results of stereo matching of four neighbouring images. In [20], an approach based on the intensity transitional area was discussed. The approach can achieve subpixel precision, but the depth information can be obtained only at the object edge.

In summary, the work on image processing for eCley stays at the initial stage and lots of problems are required to be solved before it goes to practical applications.

3. G3D

To eliminate the reconstructed image blurring and incorrectly estimated depth resulting from the parallax coming from unparallel neighboring optical channels in eCley, this section discusses the proposed geometry based three-dimensional image processing method (G3D) for eCley. The main idea of G3D is that the geometry relationship of optical channels in eCley is derived to obtain the mathematical relation between the parallax and depth among unparallel neighboring optical channels. The geometry relationship is very useful for gaining a sharply focused image and a depth map.

The main steps of G3D are shown in Fig. 3. In the eCley system, the captured image goes through several processes: image extraction, flat field correction,

spatial coordinate calculation, matching cost computation, matching cost aggregation, depth selection, focused image reconstruction, image segmentation, and depth refinement. Finally, the dense depth map and focused image can be obtained. Each step will be described in detail as follows.

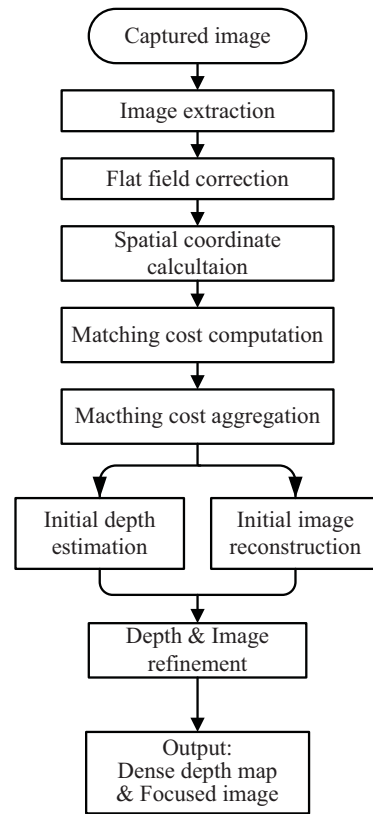


Figure 3. Schematic diagram of G3D.

3.1. Image extraction and flat field correction

An eCley system consists of multiple optical channels, but the images of all the channels are captured by one image sensor. The image of each channel occupies a small circular region on the image sensor. Assume that the eCley system has $m \times n$ channels, where m and n are odd numbers, and the resolution of the image sensor is $IW \times IH$ pixels. The central pixel $(\frac{m+1}{2}, \frac{n+1}{2})$ of the central channel is located at the center $(\frac{IW}{2}, \frac{IH}{2})$ of the image sensor. The intervals between images of ad-

adjacent channels are equivalent. In our eCley system, there are 13×17 channels, and the image captured by each channel is a circular region with a diameter of 79 pixels. The total resolution of a raw image captured by the image sensor is 1536×2048 pixels. The interval between images of adjacent channels with regard to the central pixels is 111 pixels. Therefore, the channel image is extracted as a squared image centered at the central pixel of channel with a resolution of 55×55 pixels. Because the large oblique angle of incidence leads to high distortion, in this paper, only 13×13 channels are considered.

Due to the border effect, the amount of incident light at the pixels near the border of image channel is less than the rest pixels', so the border of each image is darker than the rest of the image, which is shown on the left-up corner of the left image (pixels in the red box) in Figure 4. The pixels located in the left-up corner of this red box are darker than the rest pixels in the red box. With the effect of intensity attenuation, the depth of each pixel can not be accurate estimate because the brightness consistency assumption does not hold and the influence is more severe than the random sampling noise and the system noise of eCley. To reduce these effects, the flat field correction [21,50] is performed. The steps are as follows:

1. Reference image capture. Two types of reference images are captured by using the eCley. One image is the white reference image, the other is the black reference. The black reference image is the Dark Current (DC) offset signal. To reduce the sampling noise, multiple reference images are captured continuously, in our case, 5 images of each reference type are enough to obtain a smooth reference image, then the average intensity of each pixel is calculated as the pixel intensity. $W(x,y)$ and $B(x,y)$ are the averaged responsivity of reference white image and black image at pixel (x,y) , respectively.
2. Pixel response calculation. The pixel responsivity of camera sensor $Pr(x,y)$ is calculated according to equation $Pr(x,y) = W(x,y) - B(x,y)$. Due to the white reference image is the image of an uniform white object. The $Pr(x,y)$ can reflect the distortion ratio of camera sensor. To further reduce the noise effect, a Least

Square (LS) polynomial fitting method is used to fit the curve of pixel responsivity in each channel. The pixel responsivity after LS fitting is $P(x,y)$.

3. Flat field correction. For a captured object image $T(x,y)$, the corrected image $T'(x,y)$ is computed as follows:

$$T'(x,y) = \frac{T(x,y) - B(x,y)}{P(x,y)} \quad (1)$$

The result of the flat field correction is shown in Fig.4(b). Comparing with the pixels in the red box of both left and right images, the light intensity of the border area is reduced in the original image after flat field correction. So the effect of intensity attenuation in the red box is reduced significantly.



Figure 4. The flat field correction of a channel image. (a) Before the flat field correction. (b) After the flat field correction.

3.2. Spatial coordinate calculation

This step is the key point of G3D and is also the foundation of the subsequent steps of G3D. The aim of spatial coordinate calculation is to handle the problem of nonuniform disparities at the same distance due to the nonuniform spatial sampling interval in the eCley system. Concretely, on the basis of several depth plane candidates, the pixel is back-projected to each depth plane to obtain its corresponding spatial coordinates, thus, several processes in the local stereo matching method [38], such as matching cost computation, matching cost aggregation and depth selection, can be used to estimate the correct depth by using the spatial coordinates.

Differing from a single traditional lens camera, an array of microlenses is placed on the eCley, where each microlens has a small diameter, up to $400\mu m$, and the spacing between neighboring microlens is less

than $10\mu m$. With such a small diameter, the FOV of a microlens is quite small. Thus, all microlenses are considered to jointly gain a large FOV. However, the principal optic axes of adjacent channels in eCley are not parallel and there is a certain degree of angular shift, as shown in Fig. 5. The large oblique angle of incidence results in the image astigmatism. To alleviate the astigmatism, some measures, such as the introduction of a well designed chirped array in [8], were studied. But the oblique incidence will lead to the nonuniform spatial sampling of the object. As shown in Fig. 6, the sampling interval of the central channel is nearly uniform at the same distance, but the sampling interval of the marginal channel with an oblique incidence is nonuniform and larger than the central channel's. Even if the distances to the object are the same, the disparities of center image pair and marginal image pair are not identical. Consequently, the stereo matching method based on the disparity of pixels [38] can not yield accurate results.

Differing from the spatial sampling interval, the angular sampling interval of channel images is uniform. So this study proposes to directly use the spatial coordinates of pixels to obtain the corresponding pixels among neighboring channels so as to estimate the correct depth of object. The pixel is back-projected into space based on the geometry relationship of optical channels and the incidence angle of the pixel, thus, the spatial coordinates of pixels captured the same object will be nearly the same at the correct depth plane. The calculation of spatial coordinates is described in detail as follows.

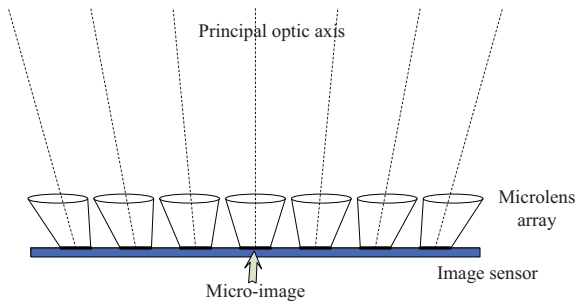


Figure 5. The simplified graph of the principal optic axis in eCley (only one dimension). The principal optic axes are not parallel with each other, and the angular deviation between adjacent channels is about 4°

In a camera system, suppose that W , Δx and f are the width of an image sensor, the pixel pitch and the focal length, respectively, which are shown in Fig. 7. The angular resolution $\Delta\theta$ of this camera system is

$$\Delta\theta = \arctan \frac{\Delta x}{f} \quad (2)$$

and the FOV θ is

$$\theta = \arctan \frac{W}{f} \quad (3)$$

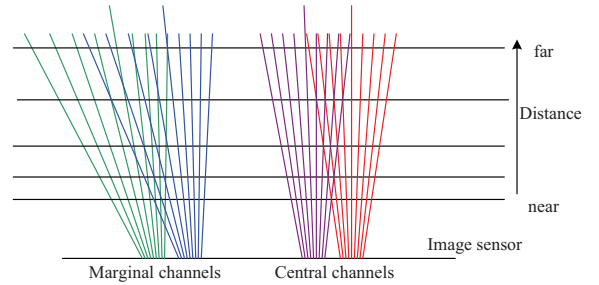


Figure 6. Interleaving pixels from neighboring channels in the eCley. Left: two marginal channels. Right: two central channels. For simplicity, each channel has nine pixels in one dimension. The central channels nearly have the uniform spatial sampling interval, while the marginal channels do not have. At the same distance, the central channels have exactly the same parallax for all pixels, but the parallax of the pixels in the marginal channels is not uniform.

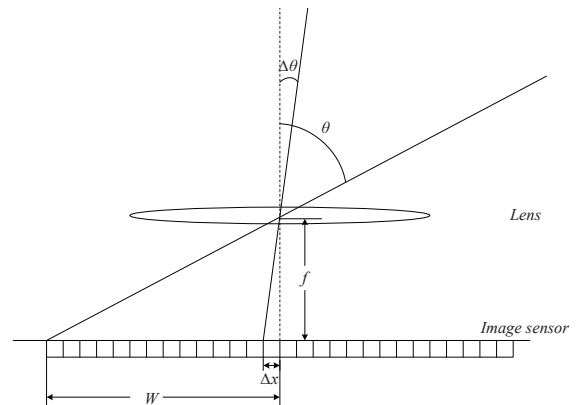


Figure 7. The simplified graph showing the relationship of camera system parameters (only in one dimension).

In our eCley system, the pixel pitch is $3.2\mu\text{m}$; the channel image has a size of 55×55 pixels; and the focal length f is $778\mu\text{m}$. According to (2) and (3), the angular resolution $\Delta\theta=0.235^\circ$ and the FOV $\theta=6.345^\circ$ are obtained.

For a micro-lens channel shown in Fig. 8, the principal optic axis has a horizontal offset angle $\Delta\phi_x$ and a vertical offset angle $\Delta\phi_y$. Let $Ic(Ic_x, Ic_y)$ be the center pixel of an image. An object point $T(D_x, D_y, Z)$ is projected into the image of this channel at $I(I_x, I_y)$. Thus, the horizontal and vertical incident angles, $\Delta\psi_h$ and $\Delta\psi_v$, of object T , are

$$\begin{aligned}\Delta\psi_h &= \Delta\phi_x + (I_x - Ic_x) * \Delta\theta \\ \Delta\psi_v &= \Delta\phi_y + (I_y - Ic_y) * \Delta\theta\end{aligned}\quad (4)$$

The spatial coordinates of the object have the following relationship

$$\begin{aligned}D_x &= \tan(\Delta\psi_h) * Z \\ D_y &= \tan(\Delta\psi_v) * Z\end{aligned}\quad (5)$$

Once the depth Z is known, the spatial coordinate of an object can be obtained. But the spatial coordinate of each channel image is not based on a unified coordinate system. The origins of the coordinate systems of different channel images are distinct, so a coordinate transformation is required.

In a unified coordinate system, let the optical center of the central channel $c(c_x, c_y)$ be the original point $O(O_x, O_y) = (0, 0)$, and B be the interval between adjacent channels. The unified coordinate $O_i(O_{ix}, O_{iy})$ of the optical center in the i^{th} channel can be computed as

$$\begin{aligned}O_{ix} &= O_x + (i_x - c_x) * B \\ O_{iy} &= O_y + (i_y - c_y) * B\end{aligned}\quad (6)$$

where $i_x \in [-N, N]$ and $i_y \in [-M, M]$ indicate the horizontal and vertical positions of channels. The total number of channels is $(2 * N + 1) \times (2 * M + 1)$, and the central channel $c(c_x, c_y)$ is located at $(i_x, i_y) = (0, 0)$.

Consequently, for a specified depth Z , according to (5) and (6), the unified spatial coordinate

$T'(D'_x, D'_y, Z)$ of a pixel $I(I_x, I_y)$ in the i^{th} channel is calculated as

$$\begin{aligned}D'_x &= D_x + O_{ix} \\ D'_y &= D_y + O_{iy}\end{aligned}\quad (7)$$

where the superscript i is added to (7) to indicate the i^{th} channel.

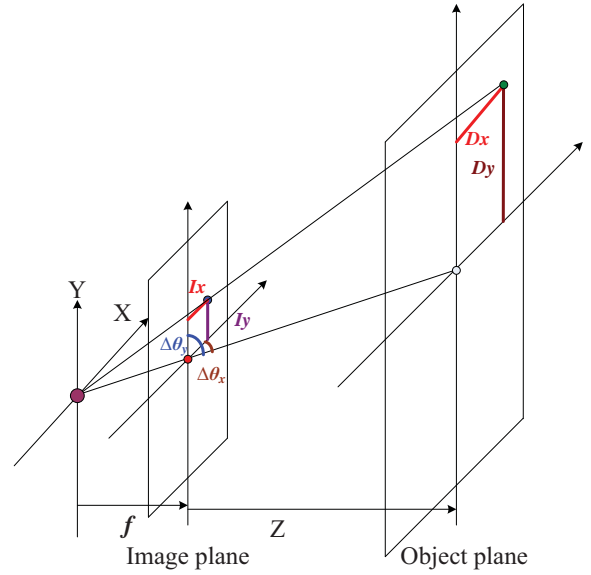


Figure 8. Schematic graph for imaging process. Brown point: optical center; red point: image center; green point: object position; blue point: the projection position of the object on the image plane. The principal optic axis of the image plane and the image plane are not orthogonal and it has a horizontal offset angle $\Delta\phi_x$ and a vertical offset angle $\Delta\phi_y$.

3.3. Matching cost computation

Matching cost computation is to calculate the intensity difference between a pixel and its corresponding pixels at adjacent channels. Like stereo matching, the depth is estimated based on several depth plane candidates. With the spatial coordinates of each pixel in each depth plane obtained in (7), the matching cost of these pixels with similar spatial coordinates can be obtained. If the pixel is placed at the proper distance, the

matching cost would be small due to the similar intensities of its corresponding pixels.

According to the eCley parameters, there is a slight offset between the principal optic axes of adjacent channel images, though an object in the FOV of eCley can still be captured by at least four channels (except for the object located at the marginal of the FOV), as show in Fig. 9. Due to the fact that the direction of FOV between adjacent channels has a shift and the width of each image is only 55 pixels, which lead to an inherent parallax between two adjacent channel images and little usefully contextual information for depth matching, respectively, the traditional two-image stereo correspondence algorithms easily cause mismatching. Although the method proposed by Ziegler *et al.* [59] estimates the disparity of one channel image with four neighboring images, essentially, this method is based on two image stereo corresponding algorithm.

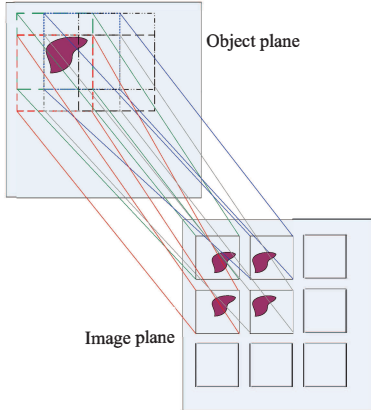


Figure 9. Schematic graph showing the object captured by 4 adjacent channels.

In order to gather as much information as possible for depth estimation, this study uses the pixels from different images located at similar spatial coordinate to compute the matching cost simultaneously. For a specified depth plane, all pixels are projected to this plane based on eq. (7) in Section 3.2. As shown in Fig. 10, the black dots are the spatial points at this depth plane. The spatial coordinates of several pixels from different images may not be consistent and have a small deviation. According to the expected resolution of the reconstructed image, the depth plane is

divided into multiple small and non-overlapped windows, each of which indicates a sampling area of a re-sampling point. Thus, the pixel points within the small window are considered as the supporting points of the re-sampling point. As shown in Fig. 10(b), the orange dot is a re-sampling point; the pixel points located in the blue box are the supporting points of the re-sampling point; the intensity of the orange dot is the average intensity of the supporting points. Moreover, the sampling interval is related to the depth. The farther the depth plane is, the larger the sampling interval will be. For simplicity, those points in all candidate depth planes are re-projected to a fixed depth plane Z_f , thus the sampling intervals at different distances are consistent. The coordinates of each point are obtained by using (8)

$$\begin{aligned} X_{zf} &= D'_x * \frac{Z_f}{Z} \\ Y_{zf} &= D'_y * \frac{Z_f}{Z} \end{aligned} \quad (8)$$

where Z is a known depth and $S(X_{zf}, Y_{zf}, d)$ is the spatial coordinate of a pixel, where we suppose all pixels are projected to the plane at depth d . The corresponding intensity is a 3-dimensional vector $I(X_{zf}, Y_{zf}, d)$ of the RGB space. The depth plane is re-sampled with a resolution of $W * H$. The re-sampling interval Rw is

$$Rw = \min\left\{\frac{\max(X_{zf}) - \min(X_{zf})}{W}, \frac{\max(Y_{zf}) - \min(Y_{zf})}{H}\right\} \quad (9)$$

The sampling area of a re-sampling point $S'(rx, ry, d)$ is a small window centered at $S'(rx, ry, d)$ with a size of $Rw \times Rw$. Thus the intensity $I'(rx, ry, d)$ of the re-sampling point and the matching cost $C(rx, ry, d)$ of this point is

$$\begin{aligned} I'(rx, ry, d) &= \frac{1}{Ni} \sum_k I(k) \\ C(rx, ry, d) &= \sqrt{\frac{1}{Ni} \sum_k \sum_{m \in r, g, b} (I_m(k) - I'_m(rx, ry, d))^2} \end{aligned} \quad (10)$$

where k is a 3-dimensional vector that indicates the supporting pixel points of $S'(rx, ry, d)$; $I(\cdot)$ is a inten-

sity function; N_i is the number of supporting points; and m indicates the color channels.

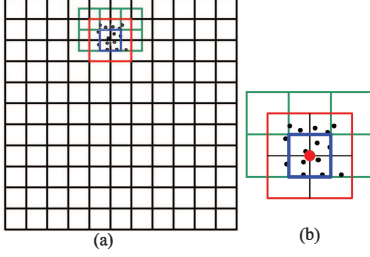


Figure 10. An example for showing the pixels projected to the spatial plane, where the black points denote the projected point of pixels; blue and green boxes indicate the sampling regions of one pixel. (a) The spatial plane with several pixels projected to this plane. (b) A magnified figure of the green box in (a).

3.4. Matching cost aggregation

Matching cost aggregation is widely used in local stereo matching methods to aggregate the local matching information and improve the matching results. In this study, due to the matching ambiguity, the depth estimation results based on the point-wise matching cost may be noisy and less accurate. So a local window-based edge preserved method based on the bilateral filter is used for cost aggregation [54,12]. Assume that a window P with a size of $[WL, WL]$ is centered at point $S'(rx, ry, d)$. Thus its aggregated cost $Cs(rx, ry, d)$ is computed as

$$Cs(rx, ry, d) = \sum_{(qx, qy) \in P} \frac{\omega(qx, qy, d) \times C(qx, qy, d)}{\omega(qx, qy, d)} \quad (11)$$

where (qx, qy) indicates the coordinate of the pixel points located in the window P ; $\omega(qx, qy, d)$ is the weight of point $S'(qx, qy, d)$. The weight is based on an edge-preserved filter and can be written as

$$\omega(qx, qy, d) = \exp\left(-\left(\frac{\Delta cl(qx, qy, d)}{\gamma_a} + \frac{\Delta gl(qx, qy, d)}{\gamma_b}\right)\right)$$

$$\Delta cl(qx, qy, d) = \sqrt{\sum_{m \in r, g, b} (I'_m(qx, qy, d) - I'_m(rx, ry, d))^2}$$

$$\Delta gl(qx, qy, d) = \sqrt{\sum_{cg \in x, y} (S'_{cg}(qx, qy, d) - S'_{cg}(rx, ry, d))^2} \quad (12)$$

where Δcl measures the color difference between point $S'(qx, qy, d)$ and point $S'(rx, ry, d)$; Δgl measures the Euclidean distance between two points; γ_a and γ_b are weighting parameters and determined empirically; cg indicates the horizontal and vertical coordinates.

3.5. Initial depth estimation and image reconstruction

In this section, the depth with the minimal matching cost is selected as the initial estimated depth. For different values of distance d , the depth information $d'(rx, ry)$ of a point can be selected by using Winner-Takes-All (WTA) method [38].

$$d'(rx, ry) = \arg \min_d Cs(rx, ry, d) \quad (13)$$

In addition, it is well known that the parallax of an object captured by neighboring images is related to the depth of the object. If an object is placed in an improper depth plane, the obtained image of the object will be blurred. So the image reconstruction is based on depth information. Concretely, the intensity $I'(rx, ry, d)$ of a point $S'(rx, ry, d)$ at specified distance d is the average intensity of pixels in the sampling area. Multiple intensities of $S'(rx, ry, d)$ are obtained based on multiple values of d . Among these intensities, the intensity corresponding to the minimal matching cost is considered as the intensity of this point. If the difference of these intensities is small, the matching cost will be low, which indicates that these points corresponding to these intensities are more likely to be the same spatial point or they are different points with similar intensities. While the different points will not lead to the reduction of the image quality. So this method is simple and effective. The intensity of the reconstructed image is

$$In(rx, ry) = I'(rx, ry, d'(rx, ry)) \quad (14)$$

3.6. Depth and image refinement

However, the depth obtained may still suffer mismatching. Therefore, a reliable depth map is constructed by comparing the minimal cost $C(rx, ry, d')$ with the second minimal cost $C(rx, ry, d'_1)$. If the remainder that the difference between the two costs is divided by $C(rx, ry, d'_1)$ is larger than a threshold λ , the depth is marked as a reliable depth, otherwise, it is not a reliable depth. The λ is determined empirically, the larger value indicates a more reliable depth. Thus, the reliable depth map $D(rx, ry)$ is computed by eq.(15).

$$\begin{aligned} D(rx, ry) &= d'(rx, ry), \quad re(rx, ry) > \lambda \\ D(rx, ry) &= 0, \quad re(rx, ry) \leq \lambda \\ re(rx, ry) &= \frac{C(rx, ry, d') - C(rx, ry, d'_1)}{C(rx, ry, d'_1)} \end{aligned} \quad (15)$$

Now a sparse reliable depth map and an initial reconstructed image $In(px, py)$ are obtained. For simplicity, two dimensional $p = (px, py)$ is used. To fill the depth in those invalid pixels, based on the initial reconstructed image, an edge-aware interpolation method called matting Laplacian [22] is performed. The depth refinement problem of D can be formulated as minimizing the follow function,

$$E(d_f) = d_{f_v}^\top M d_{f_v} + \lambda_d (d_{f_v} - D_v)^\top O (d_{f_v} - D_v) \quad (16)$$

where d_{f_v} and D_v are the vector forms of the final depth map d_f and D , respectively. O is a diagonal matrix whose O_{pp} is 1 if pixel p is a valid pixel, and 0 otherwise. λ_d balance between fidelity to the consistent depth map and smoothness. M is the matting Laplacian matrix and the (p, q) element of M is,

$$\begin{aligned} \sum_{g|s \in \vartheta_g} (\delta_{pq} - \frac{1}{|\vartheta_g|}) (1 + (In(p) - \mu_g) \\ (\Sigma_g + \frac{\epsilon_g}{|\vartheta_g|} U_3)^{-1} (In(q) - \mu_g)) \end{aligned} \quad (17)$$

where U_3 is a 3×3 identity matrix, μ_g and Σ_g are 3×1 mean vector and 3×3 covariance matrix of colors in a window ϑ_g , respectively. $|\vartheta_g|$ is the size of ϑ_g and ϵ_g is a regulation parameter. $In(p)$ and $In(q)$ are the colors of input image In at pixel p and q respectively.

Therefore, the final depth map d_{f_v} of image In can be obtained by solving the follow linear system,

$$(M + \lambda_d O) d_{f_v} = \lambda_d O D_v \quad (18)$$

Then based on the final depth map d_f , a more smooth image $I_s(p)$ is obtained,

$$I_s(p) = I'(p, d_f(p)) \quad (19)$$

4. Experimental results

In this section, the feasibility of the proposed method is tested and the comparisons with the two methods in [59,46] are made in terms of the quality of reconstructed images and the depth map. Actually, the depth estimation method in [59] is replaced by the cost filter technique [15] in this paper. Both the proposed method and the other algorithms for comparison were implemented in MATLAB 2016b.

Since there is no ground truth for our image sets, the quantitative comparisons are performed by comparing the reconstruction errors between the final image and the original channel images. Suppose a reconstructed image $I_s(p)$ with a resolution $[W_s, H_s]$, pixel p is obtained from a supporting pixel sets $I^s(p)$ of original channels images. Then the reconstructed mean squared error (MSE) is,

$$mse = \frac{1}{W_s \times H_s} \sum_p (\frac{1}{N_s} \sum_{k \in N_s} (I_s(p) - I^s(k))^2) \quad (20)$$

where N_s is the number of supporting pixels.

The captured image is extracted from all channels and shown in Fig. 11, where the left is the original image without flat field correction and the right is the image after flat field correction. In this paper, due to the large oblique distortion, only 13×13 channels are considered. **The most time-consuming step is the cost**

aggregation. The larger the window size of P is, the more time is needed for cost aggregation. Figure 12 shows the executing time at the cost aggregation step with different window sizes and the MSE of the final reconstructed image. As shown in the figure, the computing time increases as the window size goes larger. In the meantime, the MSE decreases. Therefore, to balance the time consumption and the result accuracy, the window size with 11×11 is used in the study. The parameters used in this paper are listed in Table 2.

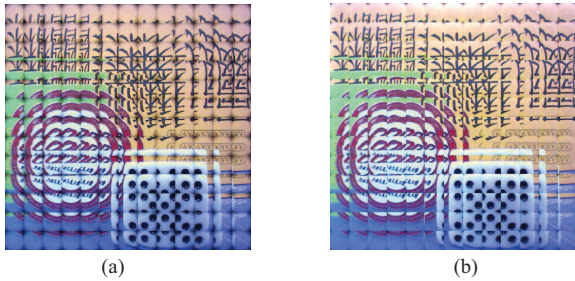


Figure 11. Input image captured by the eCley camera system. (a) Original image before flat field correction. (b) Image after flat field correction.

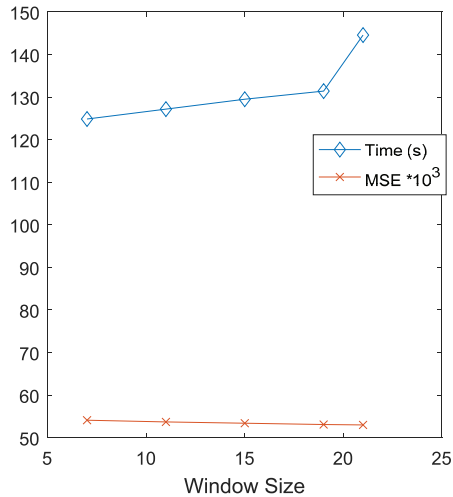


Figure 12. The executing time at the cost aggregation step with different window sizes P. The X-axis is the width of the windows.

Figure 13 shows five test images and the corresponding reconstructed images. The first, the second and the fifth images both have three fronto-parallel ob-

jects placed in front of the eCley system, the third image has two objects, different from the above two images, they are placed on a textured slanted plane, and the fourth image is a pure slanted plane, all of the objects shown in image are placed up to 200mm away from the eCley system. As shown in Fig. 13, the proposed method has a better performance than the one gained by method in [59] and by the method in [46], especially for the third and the fourth image, where the words on the slanted plane have less artifact. In addition, the MSE listed in Table 3 also indicates that the proposed method has the minimal reconstructed error among the three methods.

Fig. 14 shows the final depth map obtained by the proposed method and the methods in [46,59]. Since the principal optical axis among eCley apertures is not parallel, therefore, the estimated depth obtained by directly using the stereo method on the eCley images will suffer from the problem that the depth of an object will decrease as the object approaches to the image border. As shown in the right two columns in Fig. 14, especially for the images of the first two rows, the objects are fronto-parallel to the eCley system. While in the proposed method, this problem is largely avoided, and the depth of an object is more accurate.

Figs. 15 - 19 show the 3D point clouds of the reconstructed images based on the depth maps. As shown in the Figs. 15, 16 and 19, the objects obtained by the proposed method are located at the same plane,

Table 2. Parameters used in this study

reconstructed image resolution	640 × 640
angular resolution $\Delta\theta$	0.235°
pixel pitch	3.2 μ m
γ_a	0.1
γ_b	3
window size of P	11 × 11
λ	0.05
λ_d	0.005

Table 3. MSE of the reconstructed image

image	G3D	[59]	[46]
image 1	0.048	0.0823	0.0843
image 2	0.0536	0.0593	0.0654
image 3	0.0535	0.0766	0.0829
image 4	0.045	0.0464	0.0523
image 5	0.0374	0.0465	0.0439

and the objects are accurately recovered, while based on the methods in [46,59], the objects near the border of the image are not located at the same plane. In addition, for the slanted plane shown in Figs. 17 and 18, the proposed method can also accurately recover the patterns on the slanted plane, while in the methods [46,59], the reconstructed images are much worse and the patterns on the slanted plane are hardly recognised. Therefore, based on the depth maps and the 3D point cloud maps, the reconstructed images obtained by the proposed method and the methods in [46,59] are not much different, while the stereo images have much difference. The distance of an object in the proposed method has not dramatic change, but for the methods in [46,59], the distance of the object located in the corner of the reconstructed image always has a dramatic change.

It is worth pointing out that there is not much difference in the aspect of computing time, several minutes on Matlab 2016b, between the proposed method and the one in [59], the computing times are listed in Table 4. The two methods use the same cost aggregation approach. As for the method in [46], since this is a global method, and a lot of iterations is needed to obtain a good result, the computing time is about two time longer than the proposed method to obtain a sufficient result. In the proposed method, all neighboring images of the reference image are used for cost matching, while the methods in [59,46] use a two-view stereo matching method and fusion. In this study, no fusion process is required.

Table 4. Average computing time

method	G3D	[59]	[46]
time (sec)	169.53	143.34	302.78

5. Conclusions

In this work, a three-dimensional image processing method is presented by using the multiple apertures images of eCley to reconstruct a focused image and to estimate depth information. Based on the projection geometry, the proposed method directly obtains an initial depth map and a reconstructed image according to the spatial coordinate of pixels. The recon-

structed image and the initial depth map are further used to obtain a dense and consistent depth map by using an edge-aware interpolation method. Subsequently, the reconstructed image is further refined based on the final dense and consistent depth map. Therefore, the three-dimensional structure of the scene can be recovered. Differing from the depth map obtained by stereo matching and fusion method, the proposed method back-projects the pixel into space and directly uses the images from all channels to estimate the depth map. The information from multiple channels is merged into a single matching step. The reconstructed image and the depth map are compared with the methods in literature. Experimental results show that the proposed method can correctly estimate the reconstructed image and the depth map, and achieve a better reconstructed image and depth results than two state-of-the-art stereo methods. Future work will focus on the auto calibration method to rectify the channel image and then further improve the depth result and the use of super-resolution image reconstruction because the results still contain some errors.

6. Acknowledgements

This work is supported by the National Natural Science Foundation of China (61373047, 61672437, 61702428 and 51641506), Sichuan Science and Technology Program (18ZDYF2877, 18ZDYF1985), the State Key Laboratory of Robotics Program (2014-O09), Scientific Research Foundation of CUIT (J201508), Scientific Research Fund of Sichuan Provincial Science & Technology Department (2015GZ0304); Fund of Robot Technology Used for Special Environment Key Laboratory of Sichuan Province (14zxtk04) and China Scholarship Council.

References

- [1] G. Adams, "Hand held dyson relay lens for anti-counterfeiting," in *Proc. IEEE International Conference on Image Systems and Technologies*, 2010, pp. 273–278.
- [2] S. Aydin, S. Demirtas, K. Ates, and M. A. Tunga, "Emotion recognition with eigen features of frequency band activities embedded in induced brain oscillations mediated by affective pictures," *Int. J. Neural Syst.*, vol. 26, no. 3, 2016.

- [3] J. Baumgartner, A. G. Flesia, J. Gimenez, and J. A. Pucheta, "A new image segmentation framework based on two-dimensional hidden markov models," *Integrated Computer-Aided Engineering*, vol. 23, no. 1, pp. 1–13, 2016.
- [4] R. Ben-Ari, "A unified approach for registration and depth in depth from defocus," *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, no. 6, pp. 1041–1055, 2014.
- [5] A. Brückner, J. Duparré, R. Leitel, P. Dannberg, A. Bräuer, and A. Tünnermann, "Thin wafer-level camera lenses inspired by insect compound eyes," *Optics Express*, vol. 18, no. 24, pp. 24 379–24 394, 2010.
- [6] M. Curkovic, D. Vucina, and A. Curkovic, "Enhanced 3d parameterization for integrated shape synthesis by fitting parameter values to point sets," *Integrated Computer-Aided Engineering*, vol. 24, no. 3, pp. 241–260, 2017.
- [7] K. K. Delibasis, S. V. Georgakopoulos, K. Kottari, V. P. Plagianakos, and I. Maglogiannis, "Geodesically-corrected zernike descriptors for pose recognition in omni-directional images," *Integrated Computer-Aided Engineering*, vol. 23, no. 2, pp. 185–199, 2016.
- [8] J. Duparré, F. Wippermann, P. Dannberg, and A. Reimann, "Chirped arrays of refractive ellipsoidal microlenses for aberration correction under oblique incidence," *Optics express*, vol. 13, no. 26, pp. 10 539–10 551, 2005.
- [9] A. A. El-Sallam and F. Boussaid, "Spectral-based blind image restoration method for thin tombo imagers," *Sensors*, vol. 8, no. 9, pp. 6108–6124, 2008.
- [10] —, "A high resolution color image restoration algorithm for thin tombo imaging systems," *Sensors*, vol. 9, no. 6, pp. 4649–4668, 2009.
- [11] Y. Furukawa and C. Hernández, "Multi-view stereo: A tutorial," *Found. Trends. Comput. Graph. Vis.*, vol. 9, no. 1-2, pp. 1–148, Jun. 2015.
- [12] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397–1409, June 2013.
- [13] T. J. Hirschauer, H. Adeli, and J. A. Buford, "Computer-aided diagnosis of parkinson's disease using enhanced probabilistic neural network," *J. Medical Systems*, vol. 39, no. 11, p. 179, 2015.
- [14] R. Horisaki, S. Irie, Y. Nakao, Y. Ogura, T. Toyoda, Y. Masaki, and J. Tanida, "3d information acquisition using a compound imaging system," in *Proc. SPIE*, 2007, p. 66950F.
- [15] A. Hosni, C. Rhemann, M. Bleyer, C. Rother, and M. Gelautz, "Fast cost-volume filtering for visual correspondence and beyond," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 2, pp. 504–511, 2013.
- [16] W. Hsu, "Automatic atrium contour tracking in ultrasound imaging," *Integrated Computer-Aided Engineering*, vol. 23, no. 4, pp. 401–411, 2016. [Online]. Available: <https://doi.org/10.3233/ICA-160517>
- [17] —, "A hybrid approach for brain image registration with local constraints," *Integrated Computer-Aided Engineering*, vol. 24, no. 1, pp. 73–85, 2017.
- [18] G. Iacca, F. Caraffini, and F. Neri, "Memory-saving memetic computing for path-following mobile robots," *Appl. Soft Comput.*, vol. 13, no. 4, pp. 2003–2016, 2013.
- [19] S. Iliya and F. Neri, "Towards artificial speech therapy: A neural system for impaired speech segmentation," *Int. J. Neural Syst.*, vol. 26, no. 6, pp. 1–16, 2016.
- [20] T. Jiang, M. Zhu, K.-D. Kuhnert, and L. Kuhnert, "Distance measuring using calibrating subpixel distances of stereo pixel pairs in artificial compound eye," in *Proc. International Conference on Informative and Cybernetics for Computational Social Systems*, 2014, pp. 118–122.
- [21] Y. Kitamura, R. Shogenji, K. Yamada, S. Miyatake, M. Miyamoto, T. Morimoto, Y. Masaki, N. Kondou, D. Miyazaki, J. Tanida *et al.*, "Reconstruction of a high-resolution image on a compound-eye image-capturing system," *Applied Optics*, vol. 43, no. 8, pp. 1719–1727, 2004.
- [22] A. Levin, D. Lischinski, and Y. Weiss, "A closed-form solution to natural image matting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 228–242, 2008.
- [23] M. Li, W. Li, and H. Zhou, "Increasing N200 potentials via visual stimulus depicting humanoid robot behavior," *Int. J. Neural Syst.*, vol. 26, no. 1, 2016.
- [24] Z. Lu, X. Chen, X. Zhang, K. Tong, and P. Zhou, "Real-time control of an exoskeleton hand robot with myoelectric pattern recognition," *Int. J. Neural Syst.*, vol. 27, no. 5, pp. 1–11, 2017.
- [25] P. Martínez-Cañada, C. A. Morillas, B. Pino, E. Ros, and F. J. Pelayo, "A computational framework for realistic retina modeling," *Int. J. Neural Syst.*, vol. 26, no. 7, pp. 1–18, 2016.
- [26] R. G. Mesquita and C. A. B. Mello, "Object recognition using saliency guided searching," *Integrated Computer-Aided Engineering*, vol. 23, no. 4, pp. 385–400, 2016.
- [27] K. Nasrollahi and T. B. Moeslund, "Super-resolution: a comprehensive survey," *Machine vision and applications*, vol. 25, no. 6, pp. 1423–1468, 2014.
- [28] F. Neri and E. Mininno, "Memetic compact differential evolution for cartesian robot control," *IEEE Comp. Int. Mag.*, vol. 5, no. 2, pp. 54–65, 2010.
- [29] K. Nitta, R. Shogenji, S. Miyatake, and J. Tanida, "Image reconstruction for thin observation module by bound optics by using the iterative backprojection method," *Applied optics*, vol. 45, no. 13, pp. 2893–2900, 2006.
- [30] A. Oberdörster, A. Brückner, F. Wippermann, and A. Bräuer, "Correcting distortion and braiding of micro-images from multi-aperture imaging systems," in *Proc. SPIE*, 2011, p. 78750B.
- [31] A. Oberdörster, A. Brückner, F. Wippermann, A. Bräuer, and H. P. Lensch, "Digital focusing and refocusing with thin multi-aperture cameras," in *Proc. SPIE*, 2012, p. 829907.
- [32] L. Pan, G. Paun, G. Zhang, and F. Neri, "Spiking neural p systems with communication on request," *Int. J. Neural Syst.*, vol. 27, 2017, to appear.
- [33] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE signal processing magazine*, vol. 20, no. 3, pp. 21–36, 2003.
- [34] I. Poikolainen, F. Neri, and F. Caraffini, "Cluster-based pop-

- ulation initialization for differential evolution frameworks,” *Inf. Sci.*, vol. 297, pp. 216–235, 2015.
- [35] A. Rigos, G. E. Tsekouras, M. I. Vousedoukas, A. Chatzispavlis, and A. F. Velegarakis, “A chebyshev polynomial radial basis function neural network for automated shoreline extraction from coastal imagery,” *Integrated Computer-Aided Engineering*, vol. 23, no. 2, pp. 141–160, 2016.
- [36] S. Rostami, F. Neri, and M. Epitropakis, “Progressive preference articulation for decision making in multi-objective optimisation problems,” *Integrated Computer-Aided Engineering*, vol. 24, 2017, to appear.
- [37] S. Rostami and F. Neri, “Covariance matrix adaptation pareto archived evolution strategy with hypervolume-sorted adaptive grid algorithm,” *Integrated Computer-Aided Engineering*, vol. 23, no. 4, pp. 313–329, 2016.
- [38] D. Scharstein and R. Szeliski, “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms,” *International journal of computer vision*, vol. 47, no. 1-3, pp. 7–42, 2002.
- [39] Y. Y. Schechner and N. Kiryati, “Depth from defocus vs. stereo: How different really are they?” *International Journal of Computer Vision*, vol. 39, no. 2, pp. 141–162, 2000.
- [40] N. H. Siddique and H. Adeli, “Harmony search algorithm and its variants,” *IJPRAI*, vol. 29, no. 8, 2015.
- [41] —, “Hybrid harmony search algorithms,” *International Journal on Artificial Intelligence Tools*, vol. 24, no. 6, 2015.
- [42] —, “Gravitational search algorithm and its variants,” *IJPRAI*, vol. 30, no. 8, pp. 1–22, 2016. [Online]. Available: <https://doi.org/10.1142/S0218001416390018>
- [43] —, “Physics-based search and optimization: Inspirations from nature,” *Expert Systems*, vol. 33, no. 6, pp. 607–623, 2016.
- [44] —, “Simulated annealing, its variants and engineering applications,” *International Journal on Artificial Intelligence Tools*, vol. 25, no. 6, pp. 1–24, 2016.
- [45] M. G. Soto and H. Adeli, “Recent advances in control algorithms for smart structures and machines,” *Expert Systems*, vol. 34, no. 2, 2017.
- [46] J. Sun, N.-N. Zheng, and H.-Y. Shum, “Stereo matching using belief propagation,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 25, no. 7, pp. 787–800, 2003.
- [47] V. Tirronen, F. Neri, T. Kärkkäinen, K. Majava, and T. Rossi, “An enhanced memetic differential evolution in filter design for defect detection in paper production,” *Evolutionary Computation*, vol. 16, no. 4, pp. 529–555, 2008.
- [48] R. Volkkel, H. P. Herzig, P. Nussbaum, R. Dandliker, and W. B. Hugle, “Microlens array imaging system for photolithography,” *Optical Engineering*, vol. 35, no. 11, pp. 3323–3330, 1996.
- [49] X. Wang, G. Zhang, F. Neri, T. Jiang, J. Zhao, M. Gheorghe, F. Ipate, and R. Lefticaru, “Design and implementation of membrane controllers for trajectory tracking of nonholonomic wheeled mobile robots,” *Integrated Computer-Aided Engineering*, vol. 23, no. 1, pp. 15–30, 2016.
- [50] T. C. Williams and C. R. Shaddix, “Simultaneous correction of flat field and nonlinearity response of intensified charge-coupled devices,” *Review of Scientific Instruments*, vol. 78, no. 12, p. 123702, 2007.
- [51] S. Wu, T. Jiang, G. Zhang, B. Schoenemann, F. Neri, M. Zhu, C. Bu, J. Han, and K.-D. Kuhnert, “Artificial compound eye: a survey of the state-of-the-art,” *Artificial Intelligence Review*, vol. 48, no. 4, pp. 573–603, 2017.
- [52] T. Xian and M. Subbarao, “Performance evaluation of different depth from defocus (dfd) techniques,” in *Proc. SPIE*, 2005, p. 600009.
- [53] K. Yamada, H. Mitsui, T. Asano, J. Tanida, and H. Takahashi, “Development of ultrathin three-dimensional image capturing system,” in *Proc. SPIE*, 2006, p. 60560V.
- [54] K.-J. Yoon and I.-S. Kweon, “Locally adaptive support-weight approach for visual correspondence search,” in *Proc. CVPR IEEE*, June 2005, pp. 924–931.
- [55] Y. Zeinali and B. A. Story, “Competitive probabilistic neural network,” *Integrated Computer-Aided Engineering*, vol. 24, no. 2, pp. 105–118, 2017.
- [56] D. Zhang, F. He, S. Han, L. Zou, Y. Wu, and Y. Chen, “An efficient approach to directly compute the exact hausdorff distance for 3d point sets,” *Integrated Computer-Aided Engineering*, vol. 24, no. 3, pp. 261–277, 2017.
- [57] G. Zhang, H. Rong, F. Neri, and M. J. Pérez-Jiménez, “An optimization spiking neural P system for approximately solving combinatorial optimization problems,” *Int. J. Neural Syst.*, vol. 24, no. 5, 2014.
- [58] Z. Zhu, J. Xiao, J. Li, F. Wang, and Q. Zhang, “Global path planning of wheeled robots using multi-objective memetic algorithms,” *Integrated Computer-Aided Engineering*, vol. 22, no. 4, pp. 387–404, 2015.
- [59] M. Ziegler, F. Zilly, P. Schaefer, J. Keinert, M. Schöberl, and S. Foessel, “Dense lightfield reconstruction from multi aperture cameras,” in *Proc. International Conference on Image Processing*, Oct 2014, pp. 1937–1941.

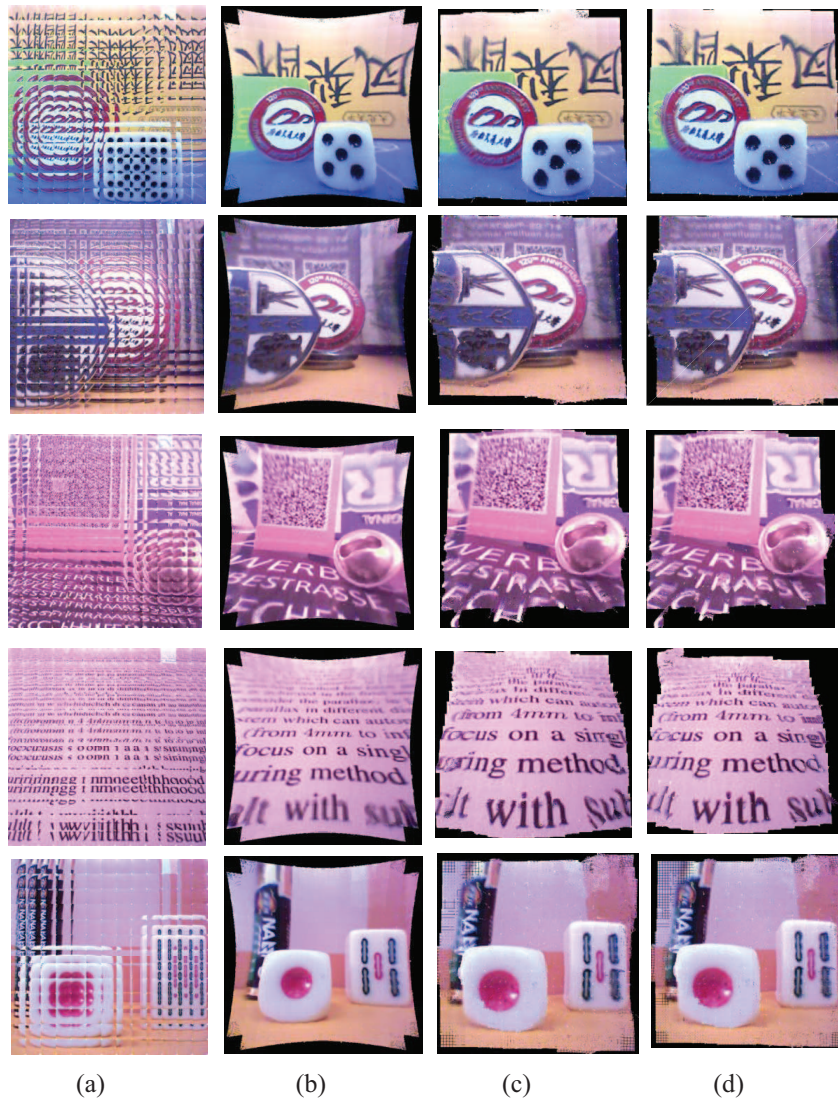


Figure 13. Test images and the reconstructed image. (a) Four test images. (b) Reconstructed image based on the depth of proposed method. (c) Reconstructed image based on [46]. (d) Reconstructed image based on [59].

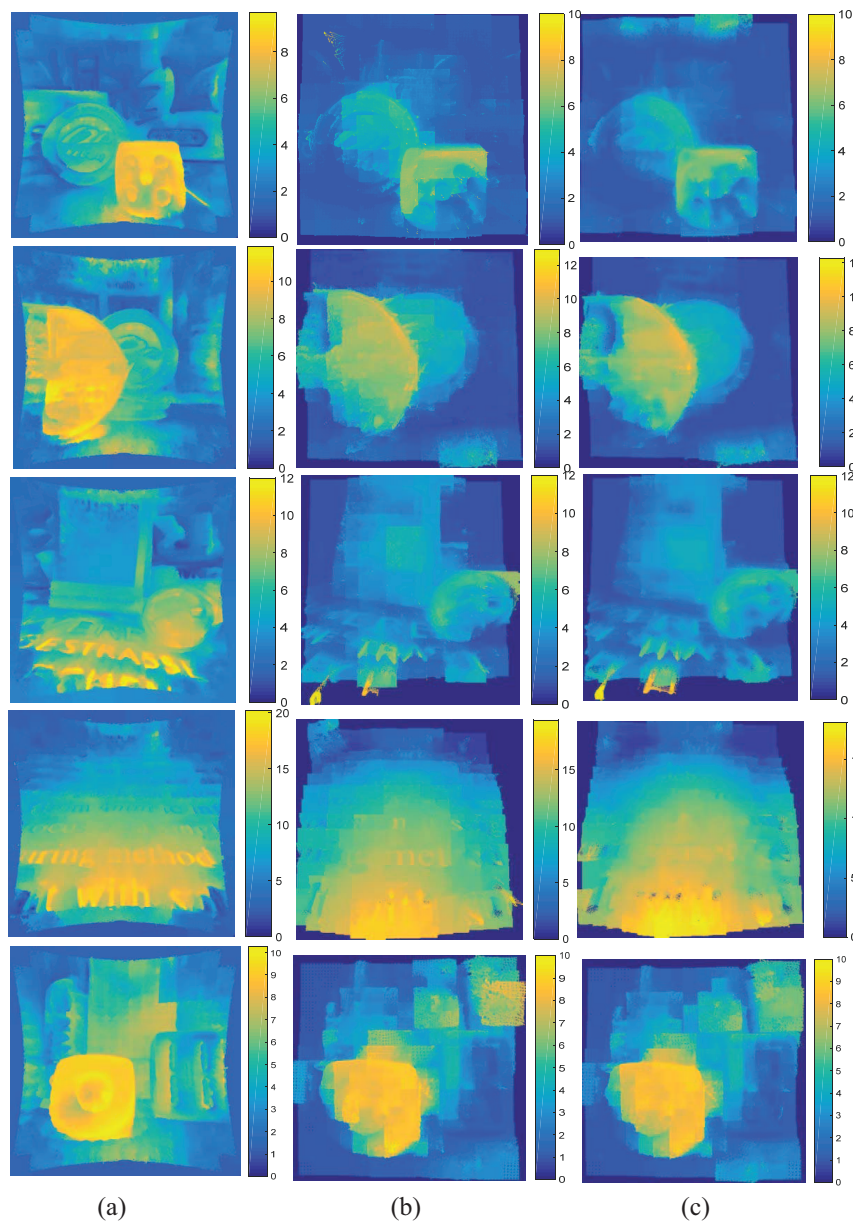


Figure 14. The final depth map. (a) The proposed method. (b) Depth in [46]. (c) Depth in [59].

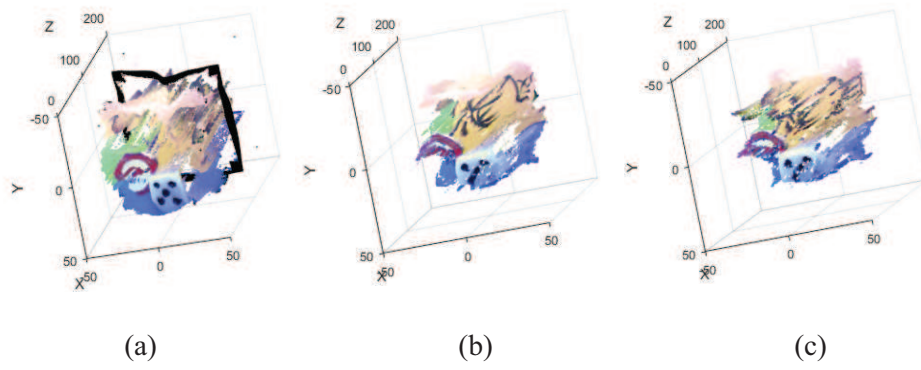


Figure 15. The 3D point cloud of image 1. (a) The proposed method. (b) Result in [46]. (c) Result in [59].

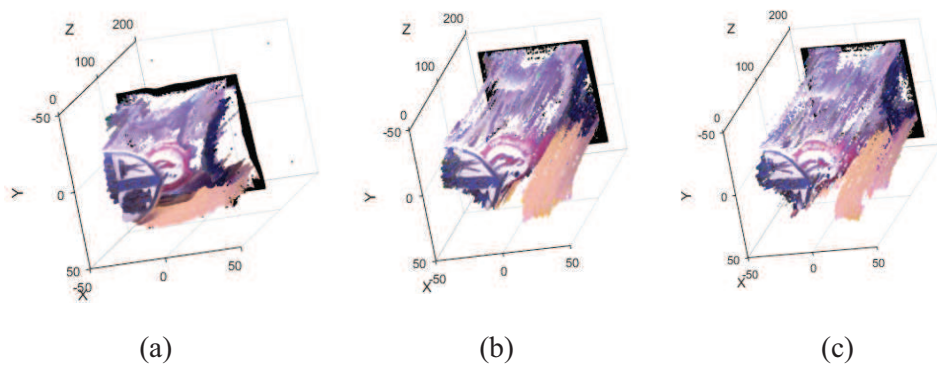


Figure 16. The 3D point cloud of image 2. (a) The proposed method. (b) Result in [46]. (c) Result in [59].

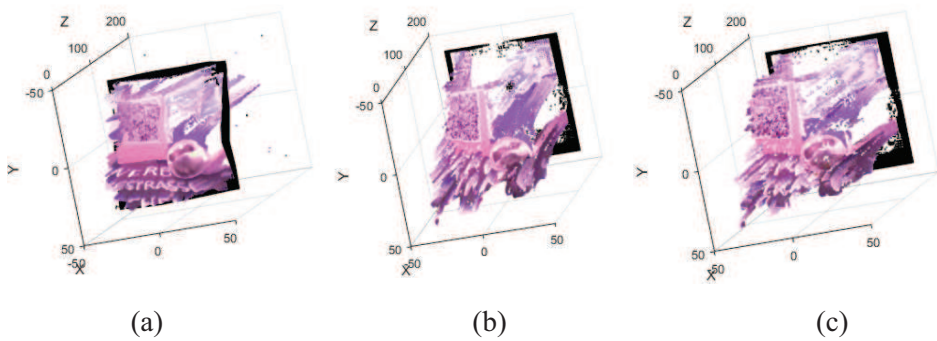


Figure 17. The 3D point cloud of image 3. (a) The proposed method. (b) Result in [46]. (c) Result in [59].

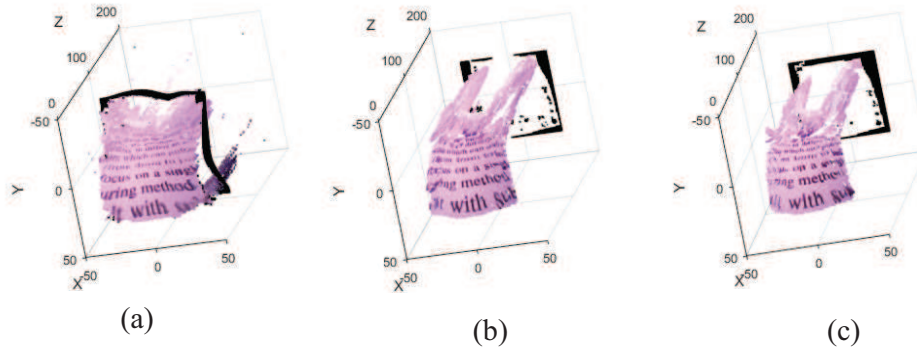


Figure 18. The 3D point cloud of image 4. (a) The proposed method. (b) Result in [46]. (c) Result in [59].

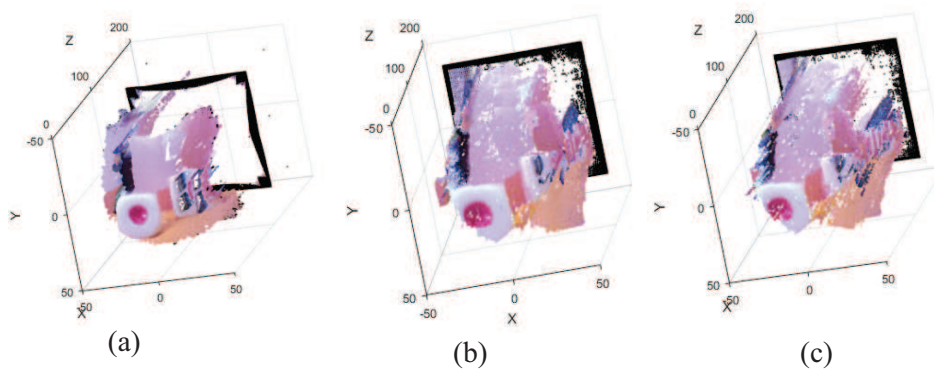


Figure 19. The 3D point cloud of image 5. (a) The proposed method. (b) Result in [46]. (c) Result in [59].