

Article

All-in-Focus Three-Dimensional Reconstruction Based on Edge Matching for Artificial Compound Eye

Sidong Wu * , Liuquan Ren and Qingqing Yang

School of Automation, Chengdu University of Information Technology, Chengdu 610225, China;
3220502005@stu.cuit.edu.cn (L.R.); 3230504004@stu.cuit.edu.cn (Q.Y.)

* Correspondence: wsd@cuit.edu.cn

Abstract: An artificial compound eye consists of multiple apertures that allow for a large field of view (FOV) while maintaining a small size. Each aperture captures a sub-image, and multiple sub-images are needed to reconstruct the full FOV. The reconstruction process is depth-related due to the parallax between adjacent apertures. This paper presents an all-in-focus 3D reconstruction method for a specific type of artificial compound eye called the electronic cluster eye (eCley). The proposed method uses edge matching to address the edge blur and large textureless areas existing in the sub-images. First, edges are extracted from each sub-image, and then a matching operator is applied to match the edges based on their shape context and intensity. This produces a sparse matching result that is then propagated to the whole image. Next, a depth consistency check and refinement method is performed to refine the depth of all sub-images. Finally, the sub-images and depth maps are merged to produce the final all-in-focus image and depth map. The experimental results and comparative analysis demonstrate the effectiveness of the proposed method.

Keywords: computer vision; artificial compound eye; 3D reconstruction; stereo vision; edge matching



Citation: Wu, S.; Ren, L.; Yang, Q. All-in-Focus Three-Dimensional Reconstruction Based on Edge Matching for Artificial Compound Eye. *Appl. Sci.* **2024**, *14*, 4403. <https://doi.org/10.3390/app14114403>

Academic Editor: Antonio Fernández-Caballero

Received: 9 April 2024

Revised: 16 May 2024

Accepted: 20 May 2024

Published: 22 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Artificial compound eye (ACE) is a type of camera with a miniature volume and large FOV. Different kinds of artificial compound eye systems have been proposed [1–3]. The ACE-captured image is an array of sub-images with slightly different viewpoints. All the sub-images can be reconstructed into one large FOV image, but depth information is required for the all-in-focus image reconstruction. As one of the most famous and typical ACE systems, the eCley (Electronic Cluster Eye [1]) (shown in Figure 1) uses a lens array of 17×13 channels, and adjacent channels have an inherent offset. The height of its optical module, about two times shorter than a comparable single-aperture camera with the same resolution, is only 1.4 mm. Oberdörster et al. present a stitching method that merges all the sub-images and reconstructs an image with a resolution of 700×550 pixels [4]. This method does not consider the effect of parallax, and the images are merged at a specified distance. Consequently, the object located at another distance will be blurred, as shown in the left image of Figure 2. The reconstructed image with all 17×13 channels without depth information is shown in Figure 2a, and our reconstructed image with central 13×13 channels is shown in Figure 2b. To address this problem, the object distance (depth) information is essential.

Stereo matching, one of the fundamental tasks in computer vision, can estimate the depth of objects from at least two cameras with overlapped FOV [5]. Many stereo matching methods have been proposed in recent years (BP [6], SGBM [7], cost-volume [8], Lac-GwcNet [9], PCW-Net [10], IGEV-Stereo [11], etc.) and have achieved state-of-the-art results. Although the adjacent images of ACE have overlapped FOV, traditional stereo matching methods cannot obtain satisfied depth results directly. As shown in Figure 3e,f, the depth map of the traditional method is noisy and inconsistent. This is mainly because of the following:

1. The low-resolution and small FOV. In this paper, the rectified sub-image has only 101×101 pixels, and the FOV of the sub-image is about 9.18° , which will lead to a large textureless area in the sub-image.
2. The large FOV offset. For the eCley, the inherent offset between adjacent sub-images is about 34 pixels, and the overlap area between adjacent images is less than 70%. As shown in Figure 4, Figure 4a gives the target sub-image located in the center and its four neighboring sub-images placed at the corresponding location. Figure 4b–d show the overlap between the target sub-image and each adjacent sub-image, where the green area indicates the non-overlapped area. This will result in much less contextual information being used.
3. The photo-consistency may not hold for all pixels. Due to the low resolution and noise, the boundary of objects will be blurred, and the point-wise depth map will be very noisy.

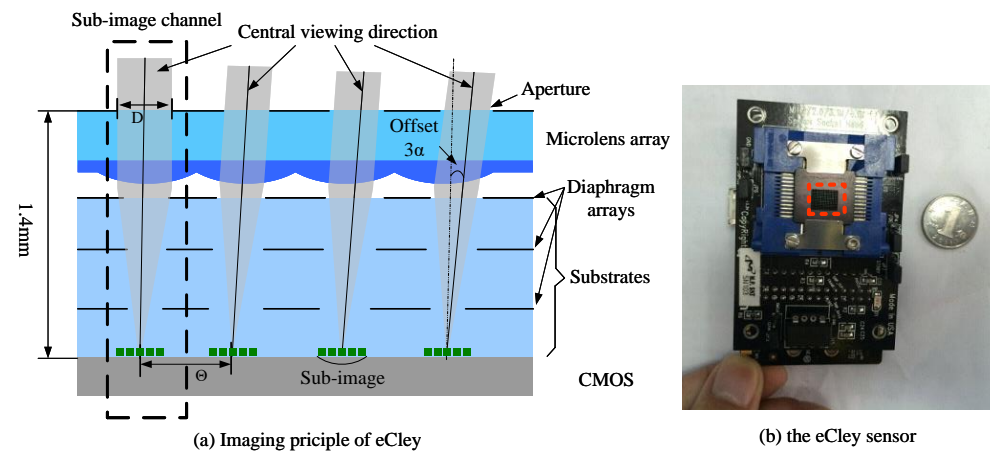


Figure 1. The eCley [1]. (a) The imaging principle. (b) The eCley sensor, the red box is the lens array.

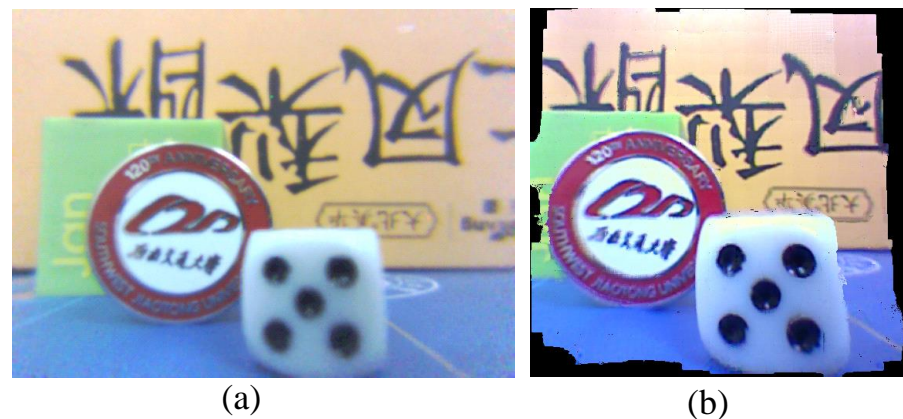


Figure 2. The reconstructed image. (a) The reconstructed image without depth information. (b) The reconstructed image with our method.

To obtain accurate depth for the ACE, an edge matching and propagation (EdMP) method is proposed. The method takes inspiration from local feature matching and propagation methods [12]. Since there is less texture in the sub-image, the edge feature is used to indicate the salient information. To accurately match the edge features between adjacent images, the edge matching cost is computed by combining a modified census transform [13] and shape context method [14]. Additionally, to eliminate the no-match problem caused by the inherent offset, the edge matching results between a sub-image and its four adjacent sub-images are fused to obtain the complete and robust sparse edge depth. The edge depth is then propagated to the whole sub-image to obtain the dense

depth map. Finally, the depth of the sub-image is refined by consistency checking with its four neighboring sub-images. Using the depth maps of all the sub-images, it is easy to obtain an all-in-focus reconstructed image. The proposed method is validated by the real captured images, and the experimental results show superior performance compared to other methods. The sub-image depth is shown in Figure 3. For the sub-image in Figure 3a, the edge depth map, dense depth map, and the refined depth map are given in Figure 3b–d, respectively. And our method achieves much better depth results compared with Figure 3e,f.

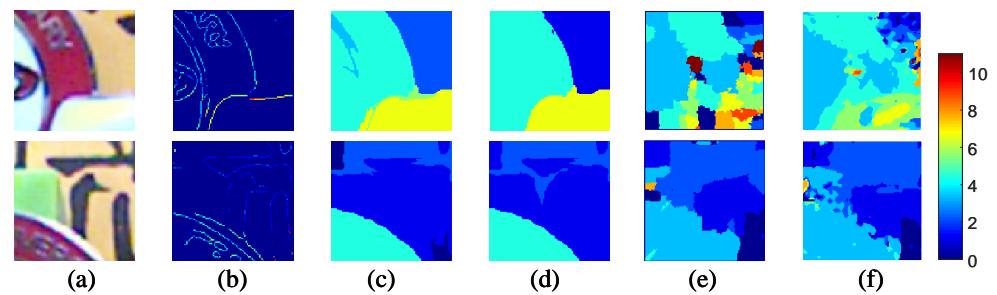


Figure 3. The depth map of sub-images with different methods, Different depths are shown in different colors, the color bar is shown on the far right of the image. (a) The sub-image for depth estimation. (b) The edge depth map. (c) The dense depth propagated from edge depth. (d) The refined depth map. (e) The depth map based on BP [6]. (f) The depth map based on Cost Filter [8].

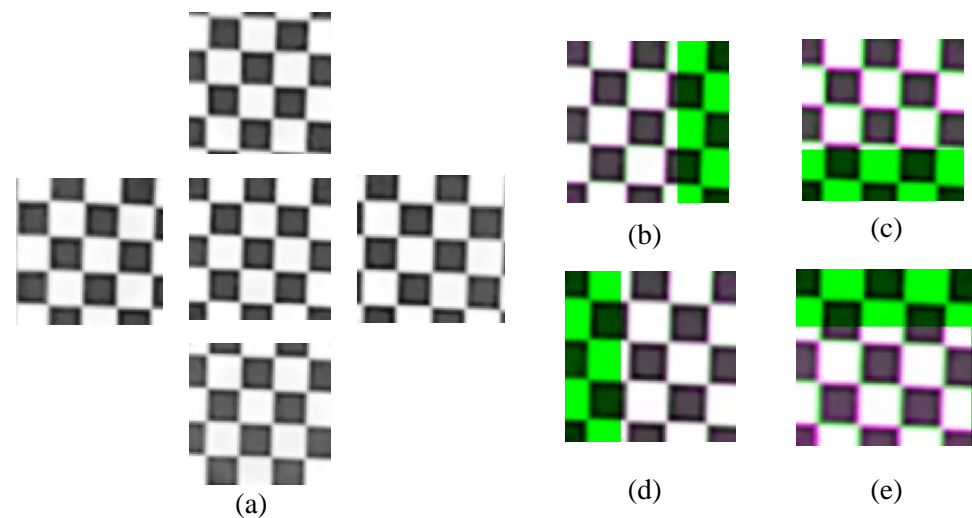


Figure 4. The overlapped area with adjacent sub-images. (a) A sub-image with its four neighboring sub-images. (b) The overlapped area with left adjacent sub-image (the white area indicates the overlapped area, and the green area shows the non-overlapped area). (c) With upper adjacent sub-image. (d) With right adjacent sub-image. (e) With bottom adjacent sub-image.

The rest of this paper is organized as follows: Section 2 introduces the related work. The overview of the proposed method is presented in Section 3. The details of EdMP are given in Section 4. Section 5 shows the experimental results, and the conclusions are provided in Section 6.

2. Related Work

In 2001, Tanida et al., inspired by the dragonfly’s apposition compound eye, proposed a compact imaging system called TOMBO (Thin Observation Module by Bound Optics) [15]. In 2004, Dupparé et al. proposed another artificial apposition compound eye called APCO (Apposition Compound Eye Objective) [16]. Although APCO has multiple imaging channels, each channel only contains one pixel, which leads to an overall resolution of only

130 × 130. Dupparé et al. proposed a superposition compound eye system called Cley (Cluster Eye) [17]. In 2007, Brückner et al. proposed an artificial neural superposition eye to mimic its bio-prototype [18]. Druart et al. was inspired by a parasite of wasps called *Xenos Peckii* and proposed the MULTICAM [19]. Each channel of MULTICAM can be seen as a small camera, and the full FOV image of MULTICAM can be reconstructed by the image processing method. In 2010, Brückner et al. followed this idea and proposed the eCley system [1]. eCley can achieve a VGA resolution while keeping the overall thickness at 1.4 mm. In 2011, Meyer et al. proposed an oCley system [20]. Apart from ACE with plane structure, many curved ACE systems were also proposed to achieve a large FOV. Jeong et al. proposed a hemispherical compound eye [21]. Song et al. proposed an arthropod-inspired camera [22]. Floreano et al. proposed the CurvACE (Curved Artificial Compound Eye) [23]. Many other ACE systems can be found in [3,24–26].

Although a lot of ACE systems were proposed and researched, the research is mainly focused on the optical imaging, and the related image processing method is not well studied. Kitamura et al. [27] rearranged pixels of ACE aperture images to reconstruct the high-resolution image for TOMBO. Nitta et al. proposed an iterative back-projection method to improve the result [28]. However, these methods reconstruct the image only on a specified distance. Horisaki et al. proposed a three-dimensional estimation method [29] and then combined the depth map into image reconstruction [30]. Dobrzynski et al. [31] proposed a flexible compound-eye-like sensor to detect motion and estimate the proximity. Luke et al. proposed compound-eye sensor with motion hyperacuity in the same year [32], and then compared it with a CCD Camera [33]. Gao et al. used stereo matching and the SIFT feature to estimate the depth of a TOMBO image [34]. Park et al. proposed a time-stamp-based optic flow ACE sensor for motion detection [35]. Agrawal et al. proposed an edge detection method for ACE to detect horizontal, vertical, forward diagonal, and backward diagonal edges [36]. Lee et al. proposed a COMPU-EYE system and designed a depth estimation method to reconstruct a high-resolution and sharp image [37]. For eCley, Oberdörster proposed a braiding method to reconstruct the image with a final resolution of 640 × 480 [4,38], but the braiding method can only focus on a specified distance. Ziegler et al. introduced a depth-related image reconstruction method [39], where the depth of each channel image was obtained by fusing the matching results with four neighboring images. Jiang et al. considered the intensity transitional area and proposed a distance measurement method at the object edge and achieved subpixel precision [40]. Wu et al. considered the oblique incidence of the apertures and proposed a geometry-based 3D reconstruction method named G3D [41]. Here, the geometry relationship of optical channels in eCley is derived and then the mathematical relation between the parallax and depth among unparallel neighboring optical channels is obtained to reconstruct the final image. In addition, Wu et al. also proposed an optical flow method for eCley to estimate the object movement [42], and multi-aperture images of eCley provided good constraints on the consistency of optical flow estimates.

Although the ACE system has a similar structure with other imaging systems, like the integral imaging system [43] and light field system [44], and both can use the stereo principle to obtain the depth of objects, for eCley, the inherent offset between adjacent channels leads to an object that can only span at most three channels in one direction. For light field reconstruction, this is quite sparse light field information; therefore, the light field based image reconstruction method is not considered in this paper.

In summary, some progress has been made in recent work on ACE image reconstruction, but further enhancement and research in areas such as accurate depth estimation and high-resolution image reconstruction are still needed.

3. Overview of Edge Matching and Propagation

The workflow of the proposed method is given in Figure 5, and our method mainly consists of the following five processing modules: (1) image correction; (2) edge depth

estimation; (3) edge depth propagation; (4) inconsistency check and depth refinement; and (5) all-in-focus reconstruction.

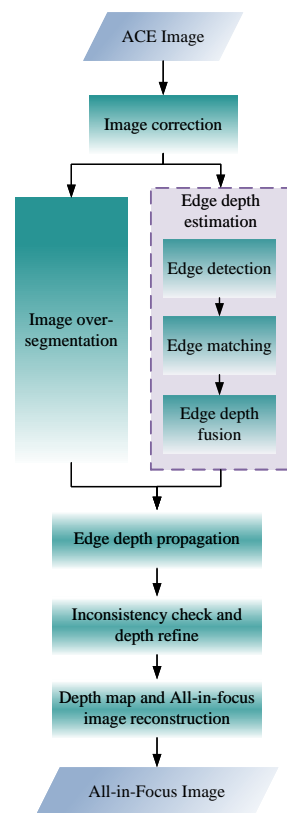


Figure 5. Overview of Edge Matching and Propagation.

The five modules are briefly explained below, and the following section will give each module in detail:

1. Image correction. Because the ACE image suffers distortion, the first step is to rectify sub-images. This will greatly reduce the complexity of the following depth estimation.
2. Edge depth estimation. The inputs to the edge depth estimation module consist of a reference image and four adjacent images, which have been rectified already. The main purpose of this module is to obtain a consistent sparse edge depth map of the reference image.
3. Edge depth propagation. The purpose of this module is mainly to propagate the edge depth to the whole image. Associated with over-segmentation results and four neighbor images, this matching cost of the segmented area is much more robust to obtain the correct depth.
4. Inconsistency checking and refinement. The main purpose of this module is to identify segmented areas which have unreliable depth. Instead of the left–right check used in many stereo matching algorithms, matching consistency with four neighboring images is used.
5. All-in-focus reconstruction. Finally, based on the refined depth map of each sub-image, the sub-image arrays are rendered to reconstruct the all-in-focus image, and the depth maps are also merged into a whole depth map.

4. Edge Matching and Propagation Method

4.1. Image Correction

The incidence angles of the eCley sub-apertures are not parallel but have an inherent offset angle. This results in an oblique incidence of the fringe aperture, which further leads to distortion of the image. As shown in Figure 6a, the chessboard is captured by eCley,

and the central sub-image (cyan box) is less distorted, but the fringe sub-image suffers from oblique distortion (red box). The oblique distortion requires a complex process for accurate image reconstruction [41]. Therefore, this paper performs the image correction first, and the correction experimental schematic is shown in Figure 6b. The chessboard is large enough to span all the sub-images, and the image correction method proposed in [42] is used in this paper.

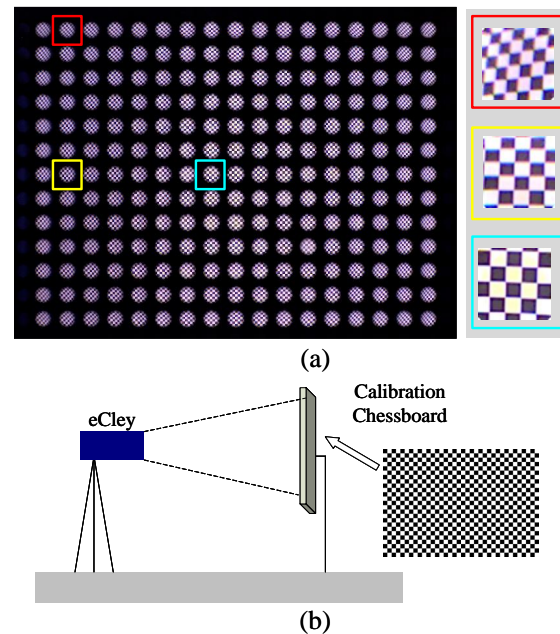


Figure 6. The image correction experiment setup. (a) The original captured image. Left, the image array; right, the enlarged sub-images in the corresponding color boxes (red, yellow, and cyan). (b) The experiment setup.

4.2. Edge Depth Estimation

4.2.1. Sub-Image Matching Feature Selection

The resolution of the rectified sub-image is 101×101 pixels, as directly matching the adjacent sub-image may cause a mismatch. To obtain accurate depth, the local feature matching-based method inspired by [12] is proposed. However, the resolution of the sub-image is low and the FOV is small, which leads to few feature points by using many robust feature extraction methods. Figure 7 shows six different features (SURF [45], FAST [46], Harris [47], BRISK [48], MSER [49], and Canny [50]) extracted for the rectified sub-image, and Table 1 gives their counts. The sub-image has fewer SURF, FAST, Harris, BRISK, and MSER feature points, whereas Canny feature points reach 554, much more than the other features. Considering the total pixels of the sub-image is 101×101 , the Canny feature is also much less than the pixels, which can largely speed up the computation. Therefore, the edge feature is adopted to obtain the sparse depth.

Table 1. Feature count in one sub-image

Feature	SURF	FAST	Harris	BRISK	MSER	Canny
Count	6	1	18	3	32	554

In addition to the large number of edge features, another reason for choosing edge features is that all-in-focus image reconstruction is mainly meant for obtaining the object depth information, and the depth estimation algorithms are essential to finding the matching points of the reference image points from the corresponding image. In the original information of the image, the edge points are the most informative features and can be used to describe

the magnitude and direction of the image gradient. The multiple consecutive edge points can also be concatenated to carve the pattern and describe the shape of the pattern, etc.

Then the sparse edge depth of one sub-image can be estimated by matching it with its four neighboring sub-images. The depth estimation diagram is shown in Figure 8.

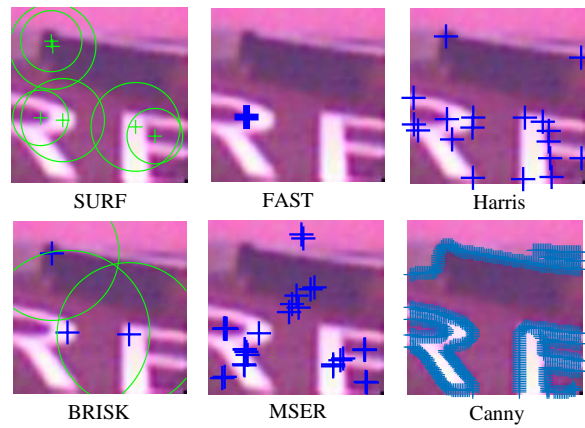


Figure 7. The extracted features in one sub-image with different methods, the green and blue cross in the figure show these features.

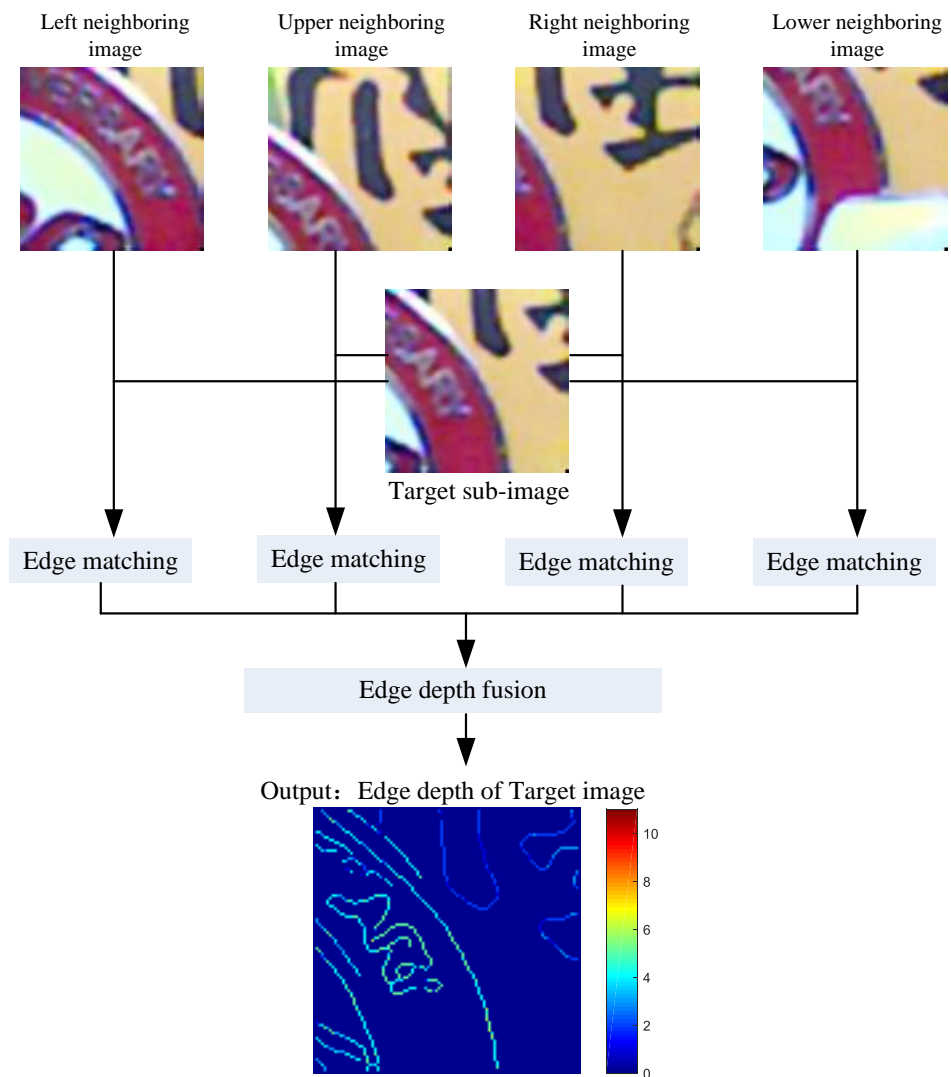


Figure 8. The edge depth estimation diagram.

4.2.2. Edge Feature Matching

In the eCley, the focal length is $f = 0.778$ mm, the pixel size is $3.2 \mu\text{m}$, the rectified image is upsampled by a factor of 2 from the original image, and there are 111 pixels between centers of the rectified adjacent sub-images, so Δx of the rectified sub-image is $1.6 \mu\text{m}$. According to the relationship between disparity and distance of the stereo image in Equation (1),

$$Z = \frac{B \cdot f}{d \cdot \Delta x} \quad (1)$$

where Z is the distance of the object, and d is the estimated disparity (or object depth). Therefore, when $d = 1$, then $Z = 172$ mm; when $d = 10$, then $Z = 17.2$ mm; and when $d = 15$, then $Z = 11.47$ mm. The distance of the object is only 1.1 cm when the disparity reaches 15, so, in the edge matching, the disparity is assumed to be below 15.

Although eCley images can extract a large number of Canny edge features, the edge features only respond to the location where the image color changes suddenly, and edge matching is prone to mismatching if only the color information is used. To obtain robust edge depths, two different types of operators (improved Census operator [13] and SC (Shape Context) operator [14]) are fused for edge feature matching.

Supposed for the sub-image I , $I(p)$ indicates the intensity of pixel p . The improved Census operator of p is given in Equation (2),

$$T(p) = \bigotimes_{q \in W'_p} \zeta(I(p), I(q)), \quad (2)$$

$$\zeta(I(p), I(q)) = \begin{cases} 1, & b + \alpha < c \\ 0, & b - \alpha \leq c \leq b + \alpha \\ -1, & b - \alpha > c \end{cases}$$

where \bigotimes denotes concatenation, W'_p is the local window centered in pixel p , and α denotes noise tolerance. Due to the low resolution of the sub-image, the system noise may lead to fluctuations in the edge transition region. Hence, the noise tolerance is introduced to improve the matching stability.

The SC operator is widely used in object matching with similar shapes such as digit recognition. Therefore, the SC operator introduces the geometry similarity for edge matching. For pixel p , the SC operator obtains the edge content information from statistical histogram $h_p(k)$ of window W_p centered in p using Equation (3),

$$h_p(k) = \#\{q \neq p : (q - p) \in \text{bin}(k)\} \quad (3)$$

where $\text{bin}(\cdot)$ uses the log-polar coordinate system and uniformly partitions the coordinate system into K regions. $\text{bin}(k)$ is the k_{th} part. q is the edge point in W_p , and $h_p(k)$ indicates the number of q in $\text{bin}(k)$.

The Census and SC operators are used to match the edge point. Suppose I is the target sub-image, $I_i \in N_I$ is the adjacent sub-image, and N_I denotes the four neighboring sub-images of I . For sub-image pair I and I_i , the matching cost of edge point $I(p)$ to have a disparity of d is

$$C^i(p, d) = (1 - \beta) \cdot C_s(p, p'(d)) + \beta \cdot C_c(p, p'(d)) \quad (4)$$

where $C_s(p, p'(d))$ and $C_c(p, p'(d))$ are the cost of SC and Census operators (given in Equation (5)), respectively. $p'(d)$ indicates the edge point with disparity d in the matching sub-image I_i .

$$C_c(p, p'(d)) = \text{Hamm}(T(p), T_i(p'(d)))$$

$$C_s(p, p'(d)) = \frac{1}{2} \sum_{k=1}^K \frac{[h_p(k) - h_{p'(d)}(k)]^2}{h_p(k) + h_{p'(d)}(k)} \quad (5)$$

where $\text{Hamm}(\cdot)$ denotes Hamming distance. If no matching point exists in the matching sub-image with disparity d for pixel p , $C^i(p, d) = \infty$, then the winner-takes-all (WTA) method is used to select the edge depth map $D^i(p)$.

4.2.3. Edge Depth Fusion

$D^i(p)$ obtains the depth map of I with an adjacent sub-image I_i , but the inherent offset between adjacent sub-images leads to an incomplete depth map, as shown in Figure 9. To estimate the complete depth map of I , depth maps of all four neighboring sub-images are fused.

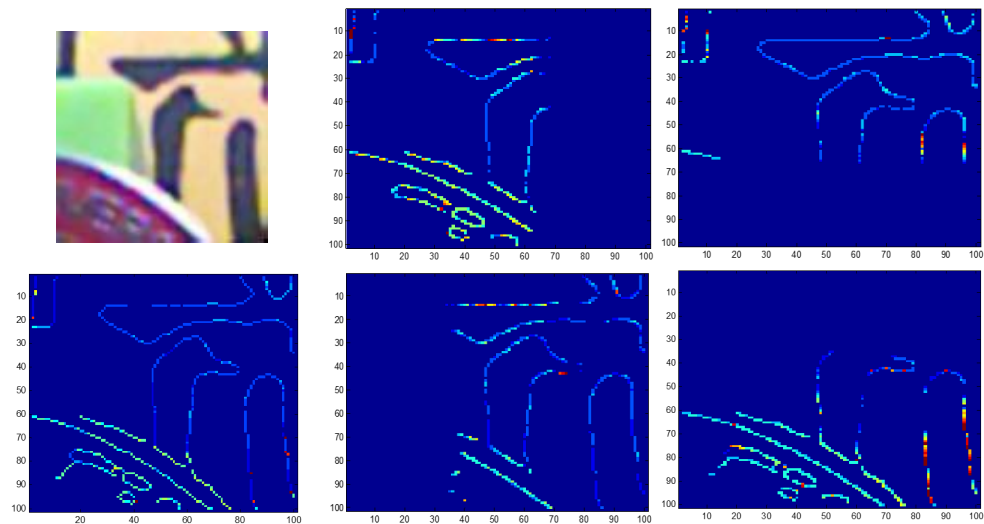


Figure 9. The edge depth map the color bar of the depth map keep the same as Figure 3. From left to right, first row: target sub-image, depth map matched with left-neighbor sub-image, and with top-neighbor sub-image; second row: fused depth map, depth map matched with right-neighbor sub-image, and with bottom-neighbor sub-image

In the depth fusion step, the depth reliability of $D^i(p)$ is first detected using Equation (6). Then the edge direction of pixel p is utilized, and if the offset between edge direction and depth search direction is smaller than 5° , the depth of pixel p is also considered unreliable, and set to 0.

$$D^i(p) = \begin{cases} d, & C^i(p, d) \leq \delta \\ 0, & C^i(p, d) > \delta \end{cases} \quad (6)$$

After that, all four edge depths are fused to obtain the final edge map $D(p)$, which is fused in the following three cases:

1. The pixel p in the four $D^i(p)$ has the same depth. The depth of $D(p)$ remains the same.
2. The pixel p in the four $D^i(p)$ has different depths, and the difference is smaller than 2 pixels. The smaller depth is chosen.
3. The pixel p in the four $D^i(p)$ has different depths, and the difference is not smaller than 2 pixels. The depth of the pixel is considered unreliable, and $D(p)$ is set to 0.

4.3. Edge Depth Propagation

Since the edge depth map is only a sparse depth map, it is necessary to propagate the sparse depth map to obtain a dense map. Based on the assumption that the depth will not abruptly change in a similar color region, the propagation process mainly utilizes the sub-eye image brightness information to segment the image into regions and then propagates the edge depth to the regions. The pseudo-code is shown in Algorithm 1. The sub-images I , I_i , and the edge depth $D(p)$ are the inputs to Algorithm 1; the output is the dense depth map $D_d(p)$. The target image I is first over-segmented by the meanshift method [51] in this paper (step 1), then each segmented area Am and segmented information S are obtained.

The depth value and the depth number for each area $S.area$ are saved (step 2–step 5). The depth of the neighboring area is used (step 9–step 11). The matching cost of area $S.area$ is then computed according to the candidate depth (step 12–step 14) in Equation (7). The area depth is then selected by WTA.

Algorithm 1 Edge Depth Propagation

Input: Edge depth map $(D(p))$ and color images (I, I_i)

Output: Dense depth map $D_d(p)$

```

1: Image over-segmentation,  $[Am, S] = \text{meanshift\_segmentation}(I)$ 
2: for each segmented area  $S.area$  do
3:    $d\_can \leftarrow$  find edge depth values in this area;
4:    $d\_num \leftarrow$  the number of depth in this area;
5: end for
6:  $S.area \leftarrow$  sort  $S.area$  based on  $d\_num$ ;
7:  $d_s = 0$ 
8: for each segmented area  $S.area$  do
9:   for each neighboring area  $Am$  do
10:     $d\_can \leftarrow$  add the estimated depth  $d_s$ ;
11:   end for
12:   for each candidate depth  $d\_can$  do
13:     $c(d) \leftarrow$  compute the matching cost;
14:   end for
15:    $d_s = \text{argmin}(c(d))$ 
16:    $D_d(S.area) = d_s$ 

```

$$c(d) = \frac{1}{num} \sum_{num} \frac{\sum_i \delta_i(p) c^i(p, d)}{\sum_i \delta_i(p) + \epsilon}$$

$$c^i(p, d) = (1 - \lambda) \cdot Mt^i(p, d) + \lambda \cdot Mg^i(p, d) \quad (7)$$

$$Mt^i(p, d) = \min\{|I(p) - I_i(p'(d))|, \tau_t\}$$

$$Mg^i(p, d) = \min\{|G(p) - G_i(p'(d))|, \tau_g\}$$

where num is the pixel number of the corresponding area, $\delta_i(p) = 1$ if pixel p in I has a corresponding edge pixel with depth d in sub-image I_i , ϵ is a small constant to avoid dividing by zero. $Mt^i(p, d)$ and $Mg^i(p, d)$ are intensity cost and gradient cost, respectively. λ is weight. $G(p)$ computes the gradient of p ; $p'(d)$ denotes the corresponding pixel in sub-image I_i to have the depth of d , τ_t ; and τ_g are the truncate threshold.

4.4. Inconsistency Check and Depth Refinement

The propagated depth map of sub-image I only uses the color and gradient information of itself. Therefore, the depth of the image array can be obtained independently. However, the estimated depth may be wrong due to the low context in the sub-image. Considering the adjacent sub-images have a common FOV, the depth information of the neighboring sub-image is also utilized to refine the estimated depth map $D_d(p)$. The consistency check and depth refinement module is mainly divided into two steps. The first step is to check the consistency and fill the depth in each segmented area, and the second step is to check the depth consistency for each pixel and then the depth of the inconsistent pixel is interpolated by using the matting Laplacian method [52].

4.5. All-in-Focus Reconstruction

After the inconsistency check and depth refinement, the depth map of each sub-image is obtained. Due to the small FOV of the sub-image in the eCley, the sub-image can only capture a small part of the entire FOV. The full FOV image is obtained by fusing all sub-images.

The rectified sub-image can be considered spatially uniformly sampled, and the spatial coordinates (x, y, z) of each pixel can be obtained by projecting the sub-image array into space. To reconstruct the image with the full FOV, the image range is obtained based on the x, y upper and lower limits. Then the position of the pixel points in the reconstructed image and the sampling interval are determined based on the reconstructed image resolution. There are multiple pixels in the resampling interval. When the z values of all pixels are close, the intensity of the resampled pixel is the average intensity of all pixels, and the depth of the resampled pixel is the average depth of all pixels. When the z values of all pixels are not close enough, occlusion may occur, the intensity of the resampled pixel is assigned with the average intensity of the pixels with the smaller z , and the depth of the resampled pixel is assigned with the average depth of the pixels with the smaller z . After iterating all the resample pixels, the reconstructed image I' and fused depth map D' are obtained.

5. Experiments

In this section, the proposed method is validated and compared quantitatively and qualitatively with three methods [6,8,41]. All algorithms are implemented in Matlab on the same laptop. The effectiveness of the EdMP method is mainly verified by the depth estimation results of the sub-image, the definition of the depth-based 2D reconstructed images, and the 3D point cloud images.

5.1. Experimental Settings

eCley has 17×13 sub-images, but only the central 13×13 sub-images are used for all-in-focus reconstruction. Each sub-image has a resolution of 101×101 pixels after rectification. For the parameters, the noise tolerance $\alpha = 0.05$, the weight $\beta = 0.07$, $\gamma = 0.5$, $\delta = 0.08$ empirically. By comparing the MSE of a test image, the different values of α , β , γ , and δ have similar MSE. The size of local window influence the computation time and depth estimation result, as illustrated in [41]. To balance the efficiency and effectiveness, $W(p)$ and $W'(p)$ are both 15×15 in this paper. K is set as the same as in [13] and consists of 3 radius regions and 12 angle regions for a total of 36 regions.

The qualitative evaluation metrics include the reconstruction error and the reconstructed image definition. The reconstruction error is measured by MSE (mean squared error) in [41]. The image definition is mainly measured by comparing the gradient of the image; therefore, four commonly used gradient functions (Brenner gradient D_b , Tenengrad gradient D_t , SMD method D_s , and Product of intensity variance D_p) for evaluating image definition are used [53].

$$D_b = \sum_x \sum_y |I(x+2, y) - I(x, y)|^2 \quad (8)$$

$$D_t = \sum_x \sum_y |G(x, y)|^2, G(x, y) > T \quad (9)$$

$$G(x, y) = \sqrt{G_x^2(x, y) + G_y^2(x, y)}$$

$$D_s = \sum_x \sum_y (|I(x+1, y) - I(x, y)|^2 + |I(x, y+1) - I(x, y)|^2) \quad (10)$$

$$D_p = \sum_x \sum_y (|I(x+1, y) - I(x, y)| * |I(x, y+1) - I(x, y)|) \quad (11)$$

where $G_x(x, y)$ and $G_y(x, y)$ are the Sobel gradient. The higher the value of these four functions, the more drastic the changes in the image edge and the higher the image definition.

5.2. Evaluation

5.2.1. Image Correction

To achieve accurate depth estimation, image correction plays a crucial role. To verify the effect of image correction, in Figure 10, the calibration board and another test pattern are

placed parallel to eCley, and the sub-images are fused and reconstructed at a certain distance. If the sub-images are corrected and reconstructed at the right distance, the reconstructed image will be all-in-focus. As shown in Figure 10a,b, the reconstructed image without correction is blurry at the border of the image, while the image is all-in-focus after correction. Figure 10c,d show the MSE of each pixel in reconstructed image of (a) and (b), where the white area has larger MSE than the black area. As one can see, the reconstructed image after correction has larger MSE in the transition area. The MSE of the whole image of (a) and (b) are 0.0427 and 0.0205, respectively. The sub-images of the calibration board before and after correction are shown in (e) and (f), respectively. The sub-images of different channels may have different distortions, which are corrected with our method. Moreover, with the test pattern, the sub-images are fused and reconstructed in two resolutions— 480×480 and 640×640 —and the reconstructed images in both resolutions are blurry before correction, while all-in-focus after correction.

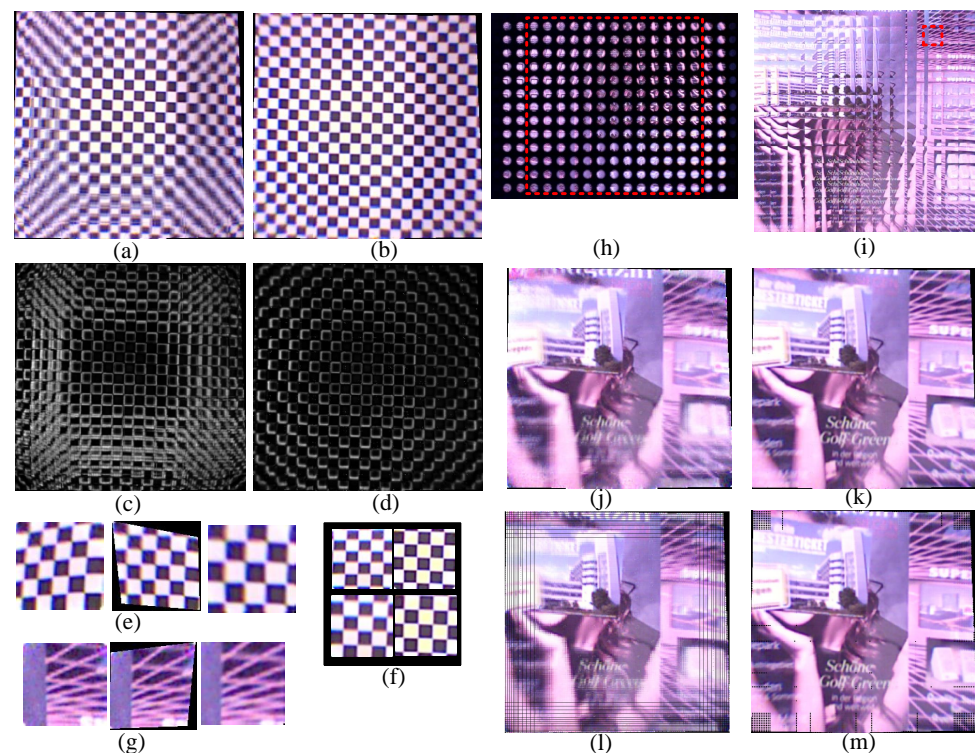


Figure 10. Image correction results. (a,b) The fused calibration board image before and after correction, respectively. (c,d) The MSE of the fused image of (a,b) respectively. (e) Sub-image of calibration board before and after correction. (f) Another two sub-images. (g) Sub-image (red box in (i)) of a test pattern before and after correction. (h) Image of the test pattern. (i) Corrected sub-images array of red box in (h). (j) Fused image without correction (resolution 480×480). (k) Fused image after correction (resolution 480×480). (l) Fused image without correction (resolution 640×640). (m) Fused image after correction (resolution 640×640).

5.2.2. Edge Depth Estimation

In Figure 11, a comparison of the edge-matching operators is presented. The results show that both the Census and SC operators are capable of producing accurate edge depth maps. However, using a single operator can result in depth instability for some edge points and be discarded. By combining the two operators, more edge points on the map can obtain depth values and have greater support in the subsequent depth propagation process.

5.2.3. Sub-Image Depth

Figure 12 shows the depth map of the eCley sub-image array after edge depth propagation and optimization. As seen in the figure, the depth map becomes denser after edge depth

propagation. However, it suffers from a large number of inconsistencies and depth abruptness due to the process not using the depth information from neighboring images. After the consistency check and depth optimization, the image array's depth becomes highly consistent.

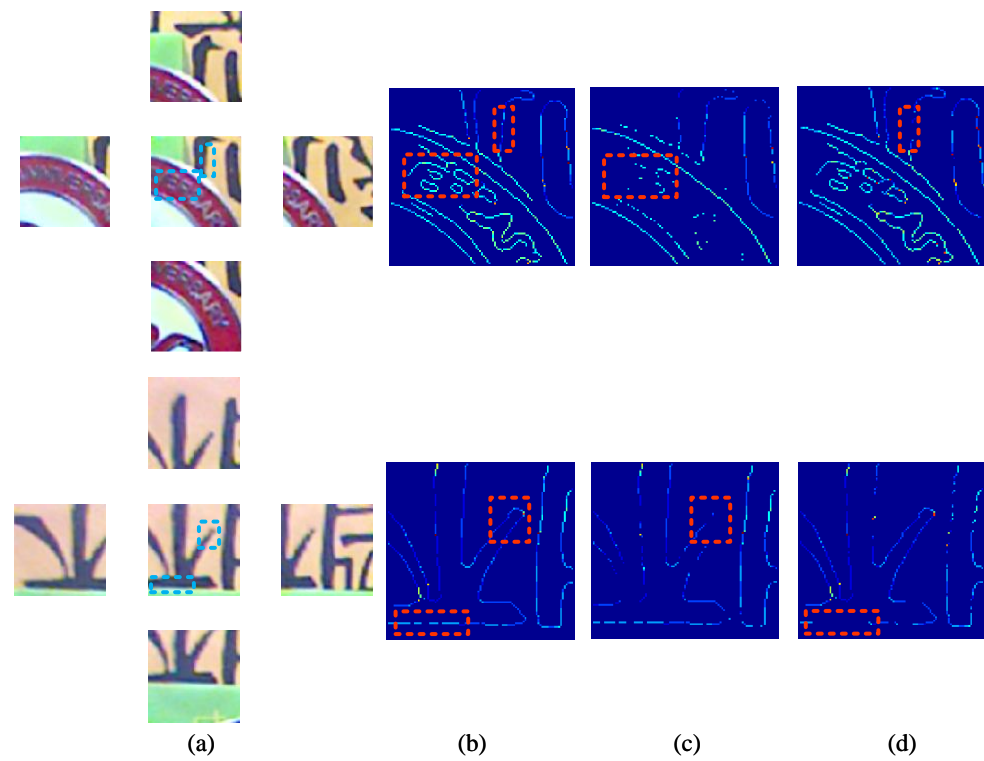


Figure 11. Edge depth comparison the depth of the blue box in (a) is highlighted in the red box of (b–d). Upper: sub-image of channel (6,6), Bottom: sub-image of channel (3,2). (a) Target sub-image for matching located in the central, and four neighboring sub-images, (b) edge depth map with Census+SC, (c) edge depth map with only the Census operator, (d) edge depth map with only the SC operator.

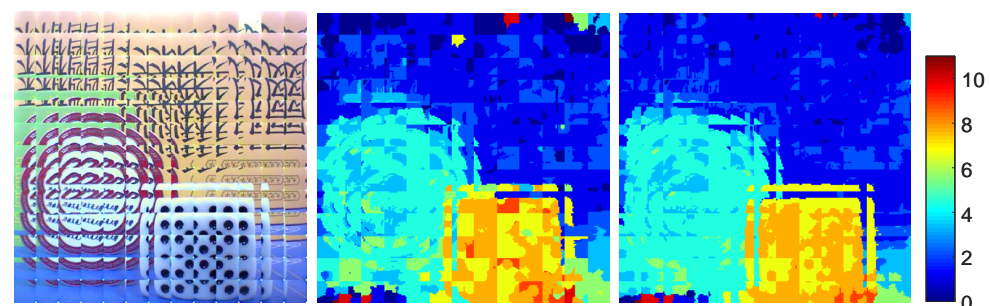


Figure 12. eCley sub-image array depth map. (Left) Sub-image array. (Middle) Depth map after edge depth propagation. (Right) Depth map after depth refinement.

5.2.4. Image Reconstruction

The reconstructed image is evaluated in this section. The test sub-image arrays after correction are shown in Figure 13, and the objects in the experiment are located at different distances. Figure 14 compares the all-in-focus reconstructed image of our method with images that are focused on a specific distance. As shown in the figure, if the reconstruction is focused only on a certain distance, objects located at other distances will appear out of focus and become blurry. However, by combining the distances during reconstruction, the reconstructed image will be all-in-focus.

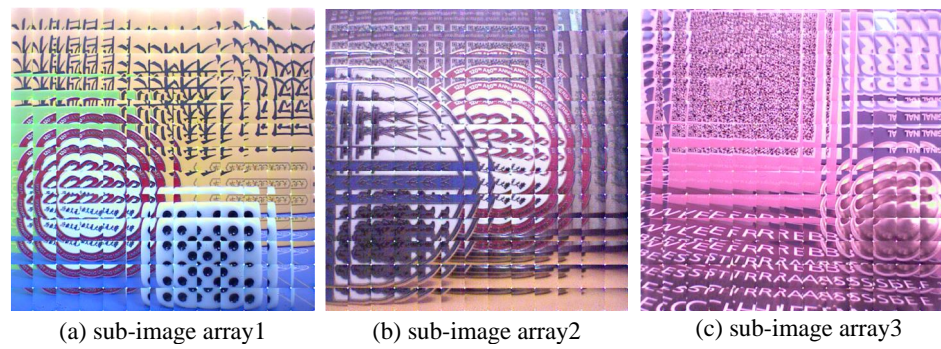


Figure 13. Test sub-image arrays captured by the eCley.

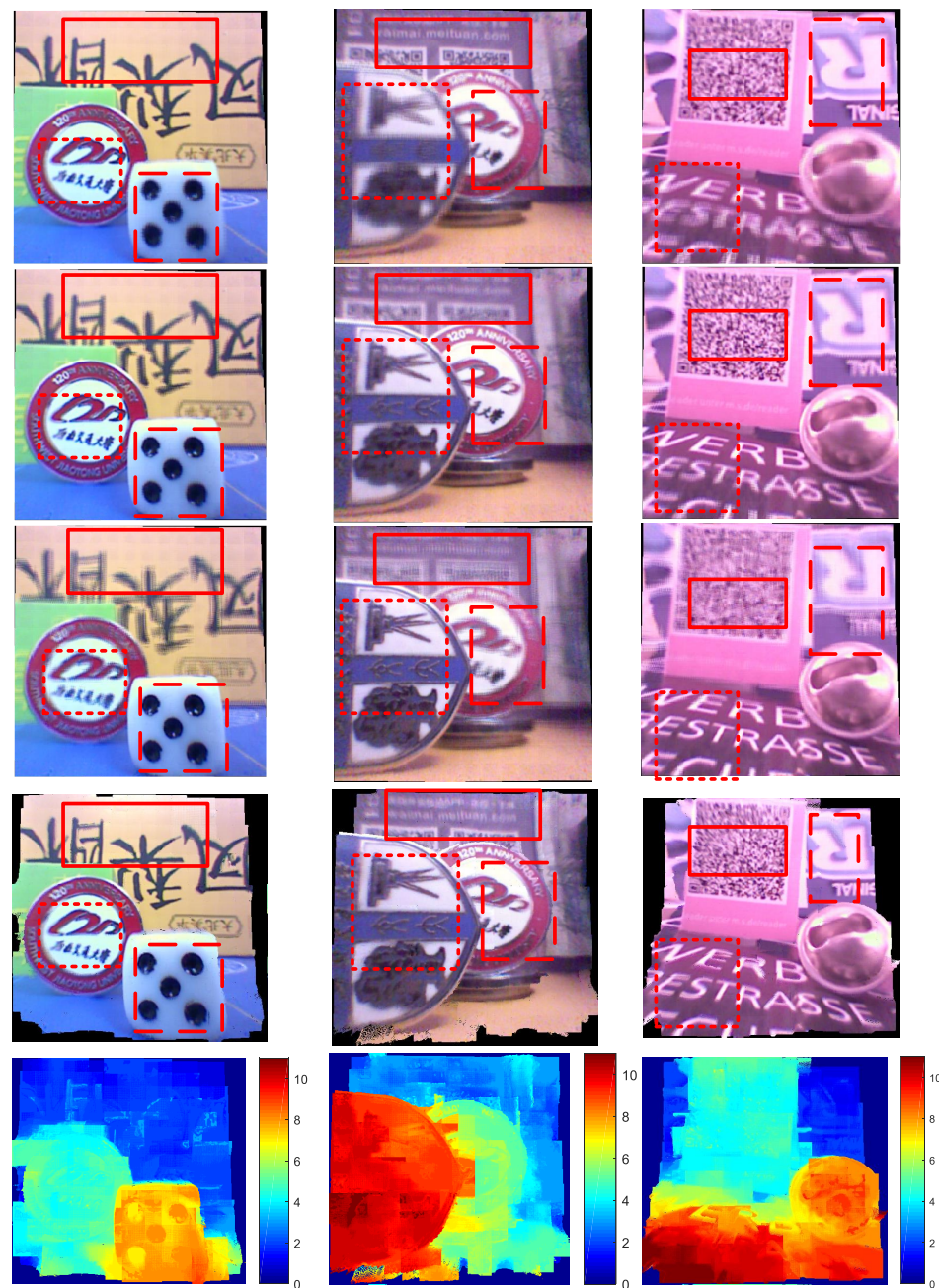


Figure 14. Image reconstruction, the red solid, dashed, and dot boxes are objects or patterns located at different distances. From top to bottom: Focus on far object, focus on middle distance object, focus on near object, all-in-focus image by our method, depth map of the reconstructed image.

The reconstructed images are compared with three methods (BP [6], Cost Filter [8], and G3D [41]). The images and corresponding quantitative evaluation results are shown in Figure 15 and Table 2, respectively. To better visualize the result of Table 2, Figure 16 shows the result in a histogram. It is evident from the results that EdMP outperforms the other methods in terms of all the metrics, except for the metrics D_b , D_s , and D_p of sub-image 1. This indicates that EdMP is capable of reconstructing images with the best definition. The reconstructed image in Figure 15, especially the details shown in the bottom row, also demonstrates that EdMP yields the sharpest images. To evaluate the computational complexity of the proposed method, all the methods are tested on a HP laptop with Intel Core i5-3210M 2.5 GHz CPU (Intel, Santa Clara, CA, USA) and Matlab2018b (version 9.5.0.944444), and the average execution times of these four methods are listed in Table 3. G3D projects the sub-image directly into space to obtain pixel depth information, so the G3D method is the most efficient. EdMP is slightly slower than G3D due to the need for neighboring sub-image matching, but the method is significantly more efficient compared to Cost Filter and BP.

Table 2. Reconstructed image evaluation.

eCley Image	Metrics	Sub-Image 1	Sub-Image 2	Sub-Image 3
BP [6] 2003	MSE	0.0823	0.0593	0.0766
	D_b	2732	1910	1783
	D_t	11682	7427	11386
	D_s	665	367	489
	D_p	3982	2417	3142
Cost Filter [8] 2012	MSE	0.0843	0.0654	0.0829
	D_b	2662	1821	1704
	D_t	10821	7309	10923
	D_s	587	321.4	501
	D_p	3811	2398.2	3028.7
G3D [41] 2018	MSE	0.048	0.0536	0.0535
	D_b	2883 ¹	1987.2	1847
	D_t	11861	7654.5	11556
	D_s	684.4	385	519.8
	D_p	4099	2501.8	3179.8
EdMP [this work] 2024	MSE	0.0412	0.0507	0.0476
	D_b	2520	2288.1	2045
	D_t	16,446	11,983	17,116
	D_s	463.6	622.7	648.9
	D_p	3105	3273.8	3522.7

¹ The bold means the best.

Table 3. The average execution time.

Method	EdMP	Cost Filter	BP	G3D
Time (s)	180.24	197.92	312.77	169.53

After the image is reconstructed, it is re-projected using the depth map to create a 3D point cloud map. The point cloud maps are shown in Figure 17. Based on the red box of images, the 3D point cloud created by EdMP is the most complete among the four methods for different objects. This indicates that the reconstructed images of EdMP have the most accurate depth information.



Figure 15. Reconstructed image comparison with different depth estimation methods. From top to bottom: BP-based method, Cost Filter-based method, G3D-based method, EdMP, details of the corresponding box in the reconstructed images (dashed green box: BP, solid green box: Cost Filter, dashed red box: G3D, solid red box: EdMP).

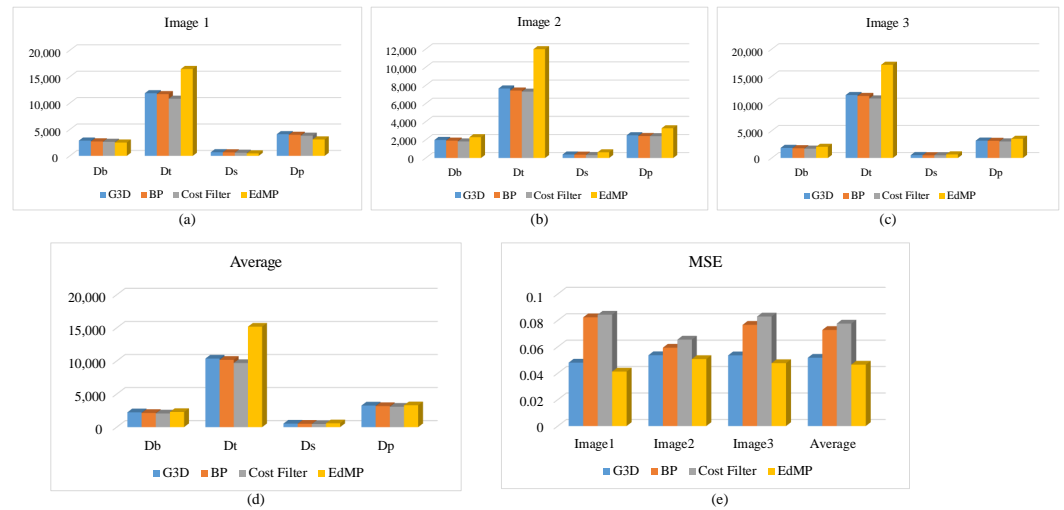


Figure 16. The metrics. (a) The image definition metrics of sub-image 1. (b) The metrics of sub-image 2. (c) The metrics of sub-image 3. (d) The average metrics of sub-image 1–3. (e) The MSE of three sub-images and their average.

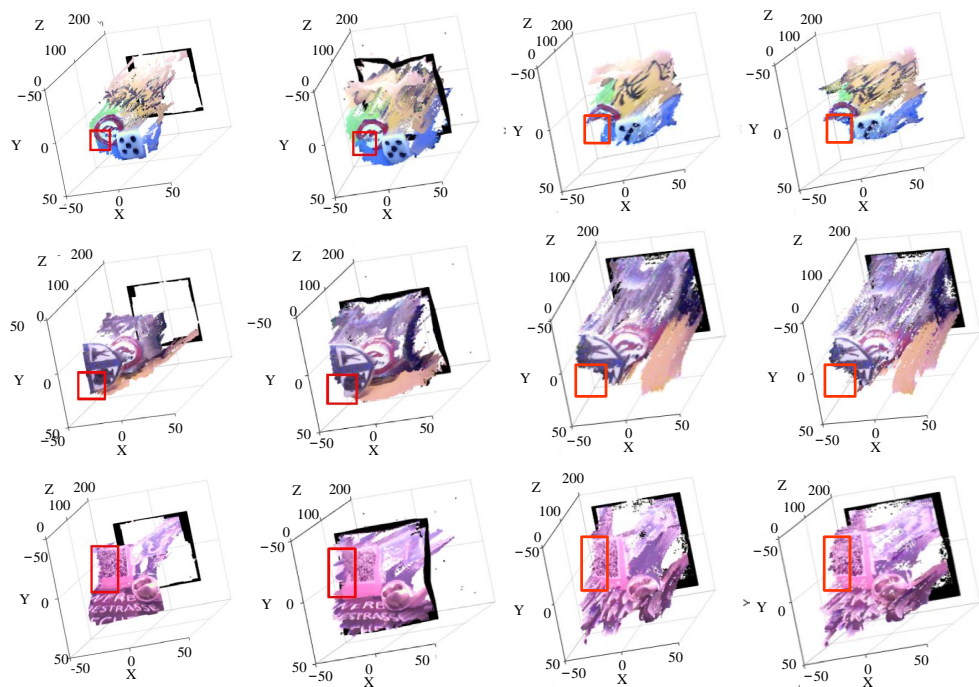


Figure 17. Reconstructed 3D point cloud map. First column: EdMP; second column: G3D; third column: BP; fourth column: Cost filter.

5.2.5. Failure Cases

It must be noted that the estimated depth map is propagated based on the image over-segmentation result, and the segmented area is set to an identical depth. When the scenarios contain a sloping plane, the proposed method will fail and the reconstructed image along the sloping plane will be blurry and aliased. As shown in Figure 18, the sloping plane is segmented with different sizes. Although finer segmentation granularity improves depth map results and reconstructed images to some extent, slope plane depth discontinuities persist, and the reconstructed images are still aliased and blurred.

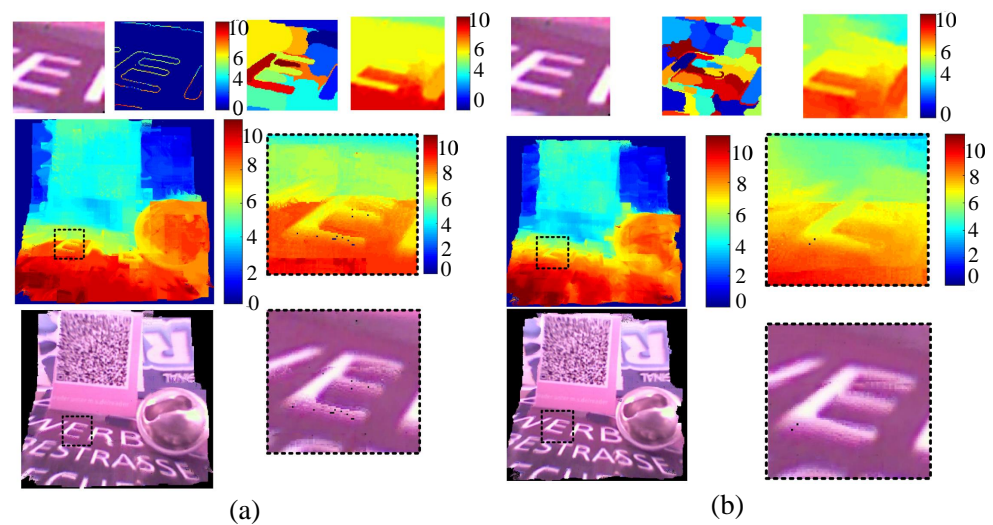


Figure 18. The failure cases. (a) The result of slope plane area, (b) the result of slope plane with finer segmentation. The first row include the sub-image of the slope plane, edge depth map, image over-segmentation result, and depth map of the sub-image. The second row include the depth map of the reconstructed image, and detailed depth maps in the black dotted box. The third row include the reconstructed all-in-focus image, and details of the black dotted boxes.

By comparison with another three methods, EdMP achieved the best result for eCley-captured ACE images. This is mainly because the BP and Cost Filter methods are directly modified from the stereo matching method, but the ACE image has a lower resolution, smaller FOV, and also a large offset angle. Without incorporating multiple adjacent sub-images, the depth estimation will be influenced by the noise and image quality. G3D utilized all adjacent sub-images to estimate the depth. However, G3D calculated the incidence angle of each pixel based on the theoretical design values. The depth estimation performance is influenced by the image distortion. Our method, EdMP, corrected the image distortion parameters and utilized four adjacent sub-images for depth estimation. From the failure case, we can see that, since the EdMP uses depth propagation to estimate the depth of the non-edge area, this mechanism will fail in the sloping plane, so a sub-pixel depth estimation method is needed. In addition, the depth obtained by depth propagation can also be influenced by large textureless areas, and the depth estimation on textureless areas is also a fundamental problem in stereo depth estimation.

6. Conclusions

In this paper, we propose an all-in-focus reconstruction method EdMP for eCley. Based on the rectified eCley image, EdMP fully considers the eCley camera and imaging characteristics, and maintains the parallax and depth consistency of multiple neighboring channels by compensating for the inherent offset between sub-images. We propose to use edge information for sparse depth acquisition, obtain reliable depth information by matching four neighboring channels simultaneously, and then propagate the depth to the full image using the depth plane consistency assumption, and finally further optimize the depth by performing a uniform consistency detection for all sub-image depths. In the experiments, EdMP has the smallest MSE among four methods and achieves an average of 0.0465, compared to 0.0517, 0.0727, and 0.0775 for the G3D, BP, and Cost Filter methods, respectively. The effectiveness of EdMP is validated in the experiments and the disadvantages of EdMP is also analyzed. In the future, the depth estimation and reconstruction of the slope plane and the robust depth estimation in the textureless area will be further researched.

Author Contributions: Conceptualization, methodology, writing—review and editing, S.W.; validation and visualization, L.R.; writing—original draft preparation, L.R. and Q.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by the National Natural Science Foundation of China, under Grant NO. 62103064; the Sichuan Science and Technology Program, under Grant NO. 2023YFG0196, NO. 2022YFNO020, NO. 2023YFN0077, NO. 2023JDZH0023; the Opening Project of Unmanned System Intelligent Perception Control Technology Engineering Laboratory of Sichuan Province, under Grant NO. WRXT2020-005; the Key Laboratory of Lidar and Device, P.R.China, under Grant NO. LLD2023-411010; and the Scientific Research Foundation of CUIT under Grant NO. KYTZ202109.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Acknowledgments: The authors would like to thank Gexiang Zhang for their helpful suggestions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Brückner, A.; Duparré, J.; Leitel, R.; Dannberg, P.; Bräuer, A.; Tünnermann, A. Thin wafer-level camera lenses inspired by insect compound eyes. *Opt. Express* **2010**, *18*, 24379–24394. [[CrossRef](#)]
- Yamada, K.; Mitsui, H.; Asano, T.; Tanida, J.; Takahashi, H. Development of ultra thin three-dimensional image capturing system. In *Proceedings of the Three-Dimensional Image Capture and Applications VII*; SPIE: San Jose, CA, USA, 2006; Volume 6056, pp. 287–295.
- Wu, S.; Jiang, T.; Zhang, G.; Schoenemann, B.; Neri, F.; Zhu, M.; Bu, C.; Han, J.; Kuhnert, K.D. Artificial compound eye: A survey of the state-of-the-art. *Artif. Intell. Rev.* **2017**, *48*, 573–603. [[CrossRef](#)]
- Oberdörster, A.; Brückner, A.; Wippermann, F.C.; Bräuer, A. Correcting distortion and braiding of micro-images from multi-aperture imaging systems. In *Proceedings of the Sensors, Cameras, and Systems for Industrial, Scientific, and Consumer Applications XII*; SPIE: San Francisco, CA, USA, 2011; Volume 7875, pp. 73–85.
- Scharstein, D.; Szeliski, R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Comput. Vis.* **2002**, *47*, 7–42. [[CrossRef](#)]
- Sun, J.; Zheng, N.N.; Shum, H.Y. Stereo matching using belief propagation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2003**, *25*, 787–800.
- Hirschmuller, H. Stereo processing by semiglobal matching and mutual information. *IEEE Trans. Pattern Anal. Mach. Intell.* **2007**, *30*, 328–341. [[CrossRef](#)] [[PubMed](#)]
- Hosni, A.; Rhemann, C.; Bleyer, M.; Rother, C.; Gelautz, M. Fast cost-volume filtering for visual correspondence and beyond. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *35*, 504–511. [[CrossRef](#)] [[PubMed](#)]
- Liu, B.; Yu, H.; Long, Y. Local similarity pattern and cost self-reassembling for deep stereo matching networks. *Proc. AAAI Conf. Artif. Intell.* **2022**, *36*, 1647–1655. [[CrossRef](#)]
- Shen, Z.; Dai, Y.; Song, X.; Rao, Z.; Zhou, D.; Zhang, L. PCW-Net: Pyramid combination and warping cost volume for stereo matching. In *Proceedings of the European Conference on Computer Vision*, Tel Aviv, Israel, 23–27 October 2022; pp. 280–297.
- Xu, G.; Wang, X.; Ding, X.; Yang, X. Iterative geometry encoding volume for stereo matching. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Vancouver, BC, Canada, 17–24 June 2023; pp. 21919–21928.
- Lhuillier, M.; Quan, L. A quasi-dense approach to surface reconstruction from uncalibrated images. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 418–433. [[CrossRef](#)]
- Zabih, R.; Woodfill, J. Non-parametric local transforms for computing visual correspondence. In *Proceedings of the European Conference on Computer Vision*, Stockholm, Sweden, 2–6 May 1994; pp. 151–158.
- Belongie, S.; Malik, J.; Puzicha, J. Shape matching and object recognition using shape contexts. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 509–522. [[CrossRef](#)]
- Tanida, J.; Kumagai, T.; Yamada, K.; Miyatake, S.; Ishida, K.; Morimoto, T.; Kondou, N.; Miyazaki, D.; Ichioka, Y. Thin observation module by bound optics (TOMBO): Concept and experimental verification. *Appl. Opt.* **2001**, *40*, 1806–1813. [[CrossRef](#)]
- Duparré, J.; Dannberg, P.; Schreiber, P.; Bräuer, A.; Tünnermann, A. Artificial apposition compound eye fabricated by micro-optics technology. *Appl. Opt.* **2004**, *43*, 4303–4310. [[CrossRef](#)] [[PubMed](#)]
- Duparré, J.; Wippermann, F. Micro-optical artificial compound eyes. *Bioinspir. Biomim.* **2006**, *1*, R1. [[CrossRef](#)] [[PubMed](#)]
- Brückner, A.; Duparré, J.; Dannberg, P.; Bräuer, A.; Tünnermann, A. Artificial neural superposition eye. *Opt. Express* **2007**, *15*, 11922–11933. [[CrossRef](#)] [[PubMed](#)]
- Druart, G.; Guérineau, N.; Haïdar, R.; Lambert, E.; Tauvy, M.; Thétas, S.; Rommeluère, S.; Primot, J.; Deschamps, J. MULTICAM: A miniature cryogenic camera for infrared detection. In *Proceedings of the Micro-Optics 2008*; SPIE: Strasbourg, France, 2008; Volume 6992, pp. 129–138.
- Meyer, J.; Brückner, A.; Leitel, R.; Dannberg, P.; Bräuer, A.; Tünnermann, A. Optical cluster eye fabricated on wafer-level. *Opt. Express* **2011**, *19*, 17506–17519. [[CrossRef](#)] [[PubMed](#)]

21. Jeong, K.H.; Kim, J.; Lee, L.P. Biologically inspired artificial compound eyes. *Science* **2006**, *312*, 557–561. [[CrossRef](#)] [[PubMed](#)]
22. Song, Y.M.; Xie, Y.; Malyarchuk, V.; Xiao, J.; Jung, I.; Choi, K.J.; Liu, Z.; Park, H.; Lu, C.; Kim, R.H.; et al. Digital cameras with designs inspired by the arthropod eye. *Nature* **2013**, *497*, 95–99. [[CrossRef](#)] [[PubMed](#)]
23. Floreano, D.; Pericet-Camara, R.; Viollet, S.; Ruffier, F.; Brückner, A.; Leitel, R.; Buss, W.; Menouni, M.; Expert, F.; Juston, R.; et al. Miniature curved artificial compound eyes. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 9267–9272. [[CrossRef](#)] [[PubMed](#)]
24. Lee, G.J.; Choi, C.; Kim, D.H.; Song, Y.M. Bioinspired artificial eyes: Optic components, digital cameras, and visual prostheses. *Adv. Funct. Mater.* **2018**, *28*, 1705202. [[CrossRef](#)]
25. Cheng, Y.; Cao, J.; Zhang, Y.; Hao, Q. Review of state-of-the-art artificial compound eye imaging systems. *Bioinspir. Biomim.* **2019**, *14*, 031002. [[CrossRef](#)]
26. Kim, M.S.; Kim, M.S.; Lee, G.J.; Sunwoo, S.H.; Chang, S.; Song, Y.M.; Kim, D.H. Bio-inspired artificial vision and neuromorphic image processing devices. *Adv. Mater. Technol.* **2022**, *7*, 2100144. [[CrossRef](#)]
27. Kitamura, Y.; Shogenji, R.; Yamada, K.; Miyatake, S.; Miyamoto, M.; Morimoto, T.; Masaki, Y.; Kondou, N.; Miyazaki, D.; Tanida, J.; et al. Reconstruction of a high-resolution image on a compound-eye image-capturing system. *Appl. Opt.* **2004**, *43*, 1719–1727. [[CrossRef](#)] [[PubMed](#)]
28. Nitta, K.; Shogenji, R.; Miyatake, S.; Tanida, J. Image reconstruction for thin observation module by bound optics by using the iterative backprojection method. *Appl. Opt.* **2006**, *45*, 2893–2900. [[CrossRef](#)] [[PubMed](#)]
29. Horisaki, R.; Irie, S.; Ogura, Y.; Tanida, J. Three-dimensional information acquisition using a compound imaging system. *Opt. Rev.* **2007**, *14*, 347–350. [[CrossRef](#)]
30. Horisaki, R.; Nakao, Y.; Toyoda, T.; Kagawa, K.; Masaki, Y.; Tanida, J. A thin and compact compound-eye imaging system incorporated with an image restoration considering color shift, brightness variation, and defocus. *Opt. Rev.* **2009**, *16*, 241–246. [[CrossRef](#)]
31. Dobrzynski, M.K.; Pericet-Camara, R.; Floreano, D. Vision Tape—A flexible compound vision sensor for motion detection and proximity estimation. *IEEE Sens. J.* **2011**, *12*, 1131–1139. [[CrossRef](#)]
32. Luke, G.P.; Wright, C.H.; Barrett, S.F. A multiaperture bioinspired sensor with hyperacuity. *IEEE Sens. J.* **2010**, *12*, 308–314. [[CrossRef](#)]
33. Prabhakara, R.S.; Wright, C.H.; Barrett, S.F. Motion detection: A biomimetic vision sensor versus a CCD camera sensor. *IEEE Sens. J.* **2010**, *12*, 298–307. [[CrossRef](#)]
34. Gao, Y.; Liu, W.; Yang, P.; Xu, B. Depth estimation based on adaptive support weight and SIFT for multi-lenslet cameras. In *Proceedings of the 6th International Symposium on Advanced Optical Manufacturing and Testing Technologies: Optoelectronic Materials and Devices for Sensing, Imaging, and Solar Energy*; SPIE: Xiamen, China, 2012; Volume 8419, pp. 63–66.
35. Park, S.; Lee, K.; Song, H.; Cho, J.; Park, S.Y.; Yoon, E. Low-power, bio-inspired time-stamp-based 2-D optic flow sensor for artificial compound eyes of micro air vehicles. *IEEE Sens. J.* **2019**, *19*, 12059–12068. [[CrossRef](#)]
36. Agrawal, S.; Dean, B.K. Edge detection algorithm for Musca-Domestica inspired vision system. *IEEE Sens. J.* **2019**, *19*, 10591–10599. [[CrossRef](#)]
37. Lee, W.B.; Lee, H.N. Depth-estimation-enabled compound eyes. *Opt. Commun.* **2018**, *412*, 178–185. [[CrossRef](#)]
38. Oberdörster, A.; Brückner, A.; Wippermann, F.; Bräuer, A.; Lensch, H.P. Digital focusing and refocusing with thin multi-aperture cameras. In *Proceedings of the Digital Photography VIII*; SPIE: Bellingham, CA, USA, 2012; Volume 8299, pp. 58–68.
39. Ziegler, M.; Zilly, F.; Schaefer, P.; Keinert, J.; Schöberl, M.; Foessel, S. Dense lightfield reconstruction from multi aperture cameras. In *Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP)*, Paris, France, 27–30 October 2014; pp. 1937–1941.
40. Jiang, T.; Zhu, M.; Kuhnert, K.D.; Kuhnert, L. Distance measuring using calibrating subpixel distances of stereo pixel pairs in artificial compound eye. In *Proceedings of the 2014 International Conference on Informative and Cybernetics for Computational Social Systems (ICCSS)*, Qingdao, China, 9–10 October 2014; pp. 118–122.
41. Wu, S.; Zhang, G.; Zhu, M.; Jiang, T.; Neri, F. Geometry based three-dimensional image processing method for electronic cluster eye. *Integr. Comput. Aided Eng.* **2018**, *25*, 213–228. [[CrossRef](#)]
42. Wu, S.; Zhang, G.; Neri, F.; Zhu, M.; Jiang, T.; Kuhnert, K.D. A multi-aperture optical flow estimation method for an artificial compound eye. *Integr. Comput. Aided Eng.* **2019**, *26*, 139–157. [[CrossRef](#)]
43. Javidi, B.; Carnicer, A.; Arai, J.; Fujii, T.; Hua, H.; Liao, H.; Martínez-Corral, M.; Pla, F.; Stern, A.; Waller, L.; et al. Roadmap on 3D integral imaging: Sensing, processing, and display. *Opt. Express* **2020**, *28*, 32266–32293. [[CrossRef](#)] [[PubMed](#)]
44. Wu, G.; Masia, B.; Jarabo, A.; Zhang, Y.; Wang, L.; Dai, Q.; Chai, T.; Liu, Y. Light field image processing: An overview. *IEEE J. Sel. Top. Signal Process.* **2017**, *11*, 926–954. [[CrossRef](#)]
45. Bay, H.; Tuytelaars, T.; Van Gool, L. Surf: Speeded up robust features. In *Proceedings of the European Conference on Computer Vision*, Graz, Austria, 7–13 May 2006; pp. 404–417.
46. Rosten, E.; Drummond, T. Fusing points and lines for high performance tracking. In *Proceedings of the IEEE International Conference on Computer Vision*, Beijing, China, 17–21 October 2005; Volume 2, pp. 1508–1515.
47. Harris, C.; Stephens, M. A combined corner and edge detector. In *Proceedings of the Alvey Vision Conference*, Manchester, UK, 31 August–2 September 1988; Volume 15, p. 147–152.
48. Leutenegger, S.; Chli, M.; Siegwart, R.Y. BRISK: Binary robust invariant scalable keypoints. In *Proceedings of the IEEE International Conference on Computer Vision*, Barcelona, Spain, 6–13 November 2011; pp. 2548–2555.

49. Mikolajczyk, K.; Tuytelaars, T.; Schmid, C.; Zisserman, A.; Matas, J.; Schaffalitzky, F.; Kadir, T.; Gool, L.V. A comparison of affine region detectors. *Int. J. Comput. Vis.* **2005**, *65*, 43–72. [[CrossRef](#)]
50. Canny, J. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **1986**, *PAMI-8*, 679–698. [[CrossRef](#)]
51. Comaniciu, D.; Meer, P. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 603–619. [[CrossRef](#)]
52. Levin, A.; Lischinski, D.; Weiss, Y. A closed-form solution to natural image matting. *IEEE Trans. Pattern Anal. Mach. Intell.* **2007**, *30*, 228–242. [[CrossRef](#)] [[PubMed](#)]
53. Zuo-Lin, L.; Xiao-hui, L.; Ling-Ling, M.; Yue, H.; Ling-li, T. Research of definition assessment based on no-reference digital image quality. *Remote Sens. Technol. Appl.* **2011**, *26*, 239–246.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.