Faculty & Staff Scholarship

2016

Analysis of Camera Arrays Applicable to the Internet of Things

Jiachen Yang

Ru Xu

Zhihan Lv

Houbing Song

2 of 13

In this paper, a camera array is designed by us, which can be used as both a parallel array and a converged array. We find out the threshold of shooting distance for the converged camera array. When the shooting distance is shorter than the threshold, converged camera array is preferred, while parallel camera array is more useful for longer shooting distance.

The rest of this paper is organized as follows. Section 2 introduces the camera array model. Then, we analyze the camera arrays in Section 3. In Section 4, we design a camera array for stereo imaging. Finally, the conclusion is given in Section 5.

2. Camera Array Model

Cameras are modeled based on the theory of pinhole camera. The pinhole camera model describes the mathematical relationship between the coordinates of a three-dimensional point and its projection onto the image plane of an ideal pinhole camera. The model does not include, for example, geometric distortions or blurring of unfocused objects caused by lenses and finite sized apertures.

2.1. Pinhole Camera Model

The geometry related to the mapping of a pinhole camera is illustrated in Figure 1.

The figure describes the relationship between the coordinates in space and those in the image. There are two coordinate systems. One is the 3D orthogonal coordinate system with its origin at the point O, which is also the location of the camera aperture. The three axes of the coordinate system are X1, X2, X3. The other coordinate system is the 2D coordinate system with the axes Y1, Y2, whose origin is at R. The model is shown as follows.

The display of an actual 3D scene by parallel camera arrays is only taken outdoors, which is not sunk concavely [15,16]. Therefore, it is unsatisfactory. However, the display scene taken with converged arrays is produced outside and concavely sunk. This makes the scene look alive; but at the same time, there is a flaw that effects the keystone, and vertical parallax will have a negative impact on viewing quality [15,16].

The comparison of parallel cameras and converged cameras has received much attention, and their pros and cons have been discussed greatly [17]. Most of the relative analyses are based on experiment results, and it is always hard to make the experimental values as precise as the theoretical ones. Although some camera array models are built in theory [18–20], the fact that only one direction or one plane is considered and other human factors make the models imperfect [21,22]. Our previous work [23] checked the stereoscopic distortion with varied parameters, including inter-camera distance, shooting distance and camera focal length when acquiring images. In this paper, we focus on the shooting distance and try to find the effect of shooting distance on parallel and converged camera arrays precisely. As the base of the analysis, a parallel camera array model and a simplified converged camera array model are built. Based on these models, we analyze their horizontal and vertical parallax and find the relationship between the parallax and shooting distance threshold of the converged cameras. According to the relationship, we can get the best shooting distance of converged cameras, which is also a threshold to distinguish the differences of parallel and converged cameras. It is helpful to make the choice of which kind of cameras to take the stereo-images. To identify the threshold, we design a camera array that can be used as a parallel camera and a converged camera, then carry out some experiments, using them as the proof of our findings.

In Section 2, we build a parallel array model and a simplified converged array model, then analyze the relation between the convergence distance of the converged array and the parallax. The analysis shows that the converged camera array is appropriate for short-distance shooting. In Section 3, we design a converged camera array, which can be also used as a parallel camera array through angle tuning. Then, we carry out some experiments using the designed camera array to identify the pros and cons of the parallel and converged camera array in Section 4. These experiments also identify that the auto-converged camera array outperforms the parallel one if the shooting distance varies in a short-distance range. In Section 5, we give the conclusion and an outlook for future work. Generally speaking, the main contribution of our paper is that we did a detailed analysis about the two kinds of camera arrays' performance over different shooting distances. The work can benefit the application of the camera arrays in different environments. Another point is the realization of the camera array, which can be used as a parallel camera array, as well as a converged camera array. This is convenient for the different choices according to the real needs.

2. Camera Array Models

2.1. Converged Camera Array Model

We build a simplified converged model, as Figure 1 shows. A and B are the optical center of the two cameras. C is a reference object, which lies on the z axis, and AC, BC are the cameras' optical axes. D is the target object, which we want to display. F is the projection of object D on the $x - z$ plane. E is the point that F projects on the z axis, so $EF//OA$. According to the geometry, Equation (1) can be obtained, where L indicates convergence distance, t indicates the distance between A and B and β represents the converged deflection angle.

$$\tan \beta = \frac{t}{2L} \qquad (1)$$

There are two kinds of coordinate systems used in Figure 1, the world coordinate system and the image coordinate system. (x_0, y_0, z_0) in the world coordinate system is a point in an object plane.

Its projections in the right and left camera are correspondingly (x_r, y_r) and (x_l, y_l) in the image coordinate system. θ_1 is the angle between AE and AF; θ_2 is the angle between BE and BF. Then, we can get:

$$\begin{cases} \theta_1 = \beta - \arctan(\frac{t+2x_0}{2z_0}) \\ \theta_2 = \arctan(\frac{t-2x_0}{2z_0}) - \beta \\ \frac{x_l}{f} = \tan\theta_1 \\ \frac{x_r}{f} = \tan\theta_2 \end{cases} \tag{2}$$
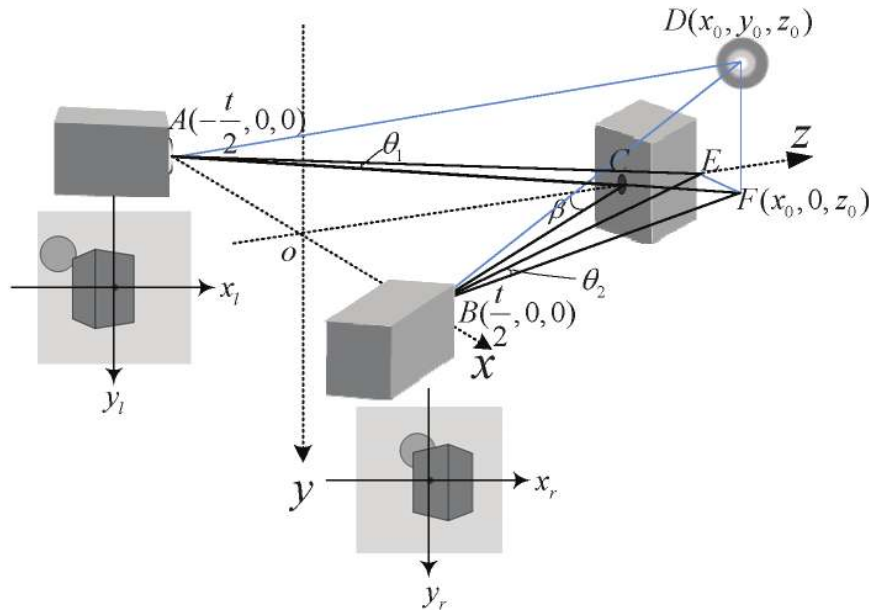
Figure 1. Camera array schematic diagram.

x_l and x_r can be obtained according to Equation (2),

$$\begin{cases} x_l = f\tan[\beta - \arctan(\frac{t+2x_0}{2z_0})] \\ x_r = f\tan[\arctan(\frac{t-2x_0}{2z_0}) - \beta] \end{cases} \tag{3}$$

AF is obtained after extending AC; then, we can use the geometric similarity to get:

$$\frac{y_l}{f} = -\frac{y_0}{AF} \tag{4}$$

$$AF = z_0\cos\beta + (x_0 + \frac{t}{2})\sin\beta \tag{5}$$

According to Equations (4) and (5),

$$y_l = -\frac{y_0 f}{z_0\cos\beta + (x_0 + \frac{t}{2})\sin\beta} \tag{6}$$

$$y_r = -\frac{y_0 f}{z_0\cos\beta - (x_0 - \frac{t}{2})\sin\beta} \tag{7}$$

To match with the display coordinate system, the image system needs to be expanded M times. For points in the display screen (x_{sl}, y_{sl}) and (x_{sr}, y_{sr}), we could get:

$$x_{sl} = Mx_l, y_{sl} = My_l, x_{sr} = Mx_r, y_{sr} = My_r \tag{8}$$

Then, we get the horizontal and vertical parallax, h and v, respectively, with (x_{sr}, y_{sr}) and (x_{sl}, y_{sl}),
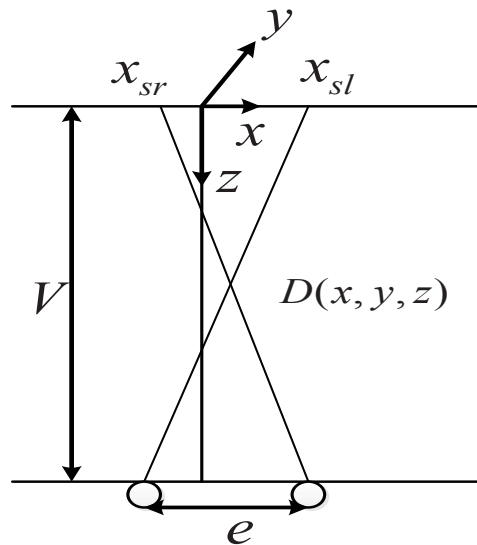
$$h = x_{sr} - x_{sl}, v = y_{sr} - y_{sl} \tag{9}$$

Figure 2. Stereopsis diagram.

For a point (x, y, z) on a stereo image in the image coordinate system and a view distance V (seen in Figure 2), we can obtain the following relationship based on the human eye imaging characteristics.

$$\begin{cases} \frac{h}{e} = \frac{z}{V-z} \\ \frac{V}{V-z} = \frac{x_{sl}+x_{sr}}{2x} \\ \frac{V}{V-z} = \frac{y_{sl}+y_{sr}}{2y} \end{cases} \tag{10}$$

According to Equations (3)–(10), we can get h, v, x, y, z as follows.

$$h = Mf\{\tan[\arctan(\frac{t-2x_0}{2z_0}) - \beta] - \tan[\beta - \arctan(\frac{t+2x_0}{2z_0})]\} \tag{11}$$

$$v = \frac{My_0 f}{z_0 \cos\beta + (x_0 + \frac{t}{2})\sin\beta} - \frac{My_0 f}{z_0 \cos\beta - (x_0 - \frac{t}{2})\sin\beta} \tag{12}$$

$$x = \frac{eMf\{\tan[\beta - \arctan(\frac{t+2x_0}{2z_0})] + \tan[\arctan(\frac{t-2x_0}{2z_0}) - \beta]\}}{2(e+h)} \tag{13}$$

$$y = -\frac{eMfy_0}{2(e+h)}[\frac{1}{z_0 \cos\beta + (x_0 + \frac{t}{2})\sin\beta} + \frac{1}{z_0 \cos\beta - (x_0 - \frac{t}{2})\sin\beta}] \tag{14}$$

$$z = \frac{VMf}{e+h}\{\tan[\arctan(\frac{t-2x_0}{2z_0}) - \beta] - \tan[\beta - \arctan(\frac{t+2x_0}{2z_0})]\} \tag{15}$$

2.2. Parallel Camera Array Model

In the converged model, if the convergence distance L approaches infinity, the converged deflection angle β will approach zero. That means the camera arrays can be turned into parallel cameras. Then, the relationship in the parallel model can be easily obtained based on what we have in the converged model.

$$\begin{cases} h = \frac{Mft}{z_0} \\ v = 0 \\ x = -\frac{eMfx_0}{ez_0+Mft} \\ y = -\frac{eMfy_0}{ez_0+Mft} \\ z = -\frac{VMft}{ez_0+Mft} \end{cases} \tag{16}$$

2.3. Model Analysis

In stereoscopic observation, vertical parallax is unfavorable. In Equation (16), we can see that in images taken by parallel camera arrays, the vertical parallax is zero theoretically, even though physical camera errors and many other kinds of external factors may change the ideal value; while for converged arrays, the vertical parallax is:

$$v = \frac{2tfM\sqrt{4L^2+t^2}x_0y_0}{\left(2Lz_0 + \frac{t^2}{2}\right)^2 - (tx_0)^2} \tag{17}$$

As we consider that $t << L$,

$$v = \frac{tfMx_0y_0}{Lz_0^2} \tag{18}$$

For the converged camera array model, the vertical parallax v diminishes as convergence distance L increases. As a matter of fact, the viewing angle of the camera lens usually meets $x_0 < z_0$, $y_0 < z_0$, and L is usually greater than 1 m. If we set t as 70 mm, f as 6.5 mm, M as 50 and assume the range of L is from 1 m to 10 m, v will be:

$$0.00228m \le v = \frac{tfm}{L} \le 0.02275m \tag{19}$$

This means the converged camera arrays can have a vertical parallax of more than 1 cm, which is enough to have a negative effect on stereoscopy in theory.

In terms of horizontal parallax h, it only depends on z_0 with an inverse relationship for parallel camera arrays, as Equation (16) shows; while for converged camera arrays, h is related to x_0, z_0 and L, as Equation (11) shows. The corresponding relations for parallel and converged camera arrays are illustrated in Figure 3. We set t as 70 mm, f as 6.5 mm and M as 50. L is a variable. We choose four values of L, 1 m, 4 m, 7 m and 10 m as samples.

From Figure 3, we can see that for the converged array, the horizontal parallax h is more and more similar to the h of parallel camera arrays in three conditions: when the convergence distance L gradually becomes greater, when $|x_0|$ becomes smaller or when z_0 is larger.

In another view, the differences of the parallax effect between these two camera arrays can be analyzed, as z_0 is set as a fixed value. Figure 4 shows corresponding relations, and the four chosen values of z_0 are 1 m, 4 m, 7 m and 10 m. From these relation curves, we can see in the same condition,

as the shooting distance z_0 becomes longer, the h of converged camera arrays is more similar to the h of parallel ones.
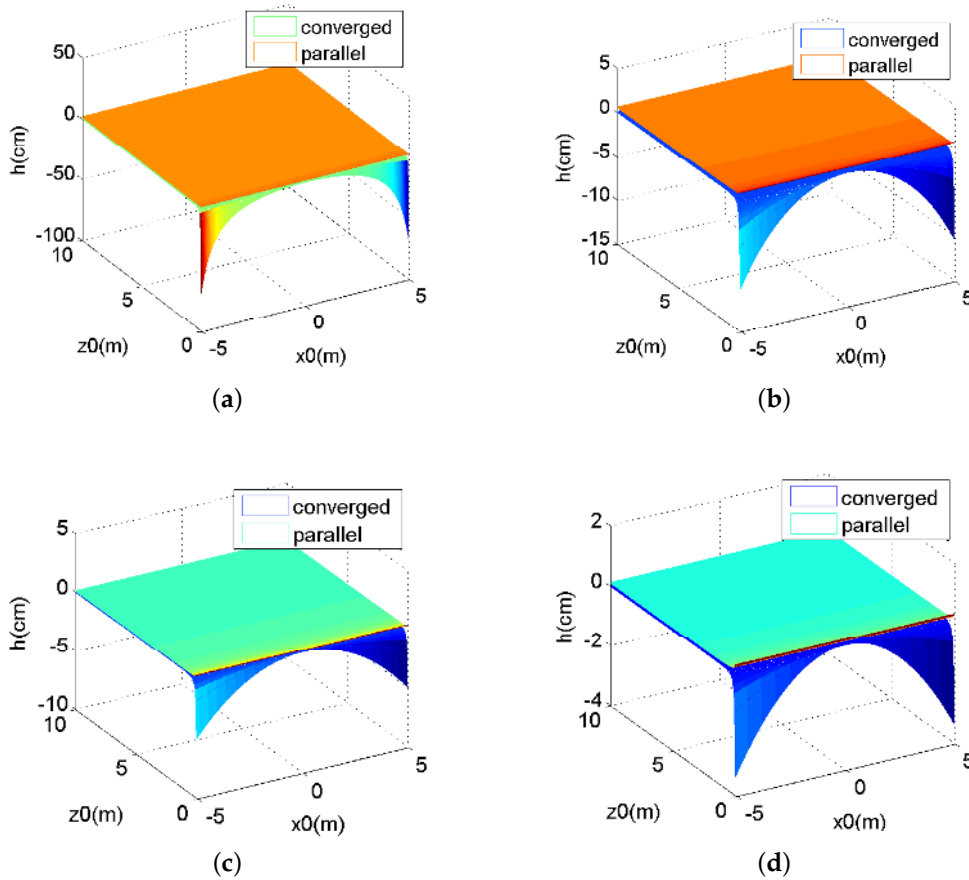
Figure 3. The curve of horizontal parallax h and x_0, z_0. (**a**) L= 1 m; (**b**) L = 4 m; (**c**) L = 7 m; (**d**) L = 10 m.

As shown in Figures 3 and 4, when the convergence distance L is small, converged arrays outperform parallel arrays because the negative and positive parallax existing in converged stereo images makes depth sensation more obvious. This is very important in a 3D display. On the other hand, when L is large, there is little difference between these two arrays, and both of their depth sensations are unimpressive.

Based on the horizontal parallax threshold, when L is longer than 7 m, the horizontal parallax h is so small, that the stereo sense dies away. On the other hand, when shooting distance z_0 is greater than 7 m, converged and parallel camera arrays are almost the same. In conclusion, when the shooting distance is shorter than 7 m, converged arrays are preferable for the scenes taken and 3D display, but when it is greater than 7 m, parallel arrays will stand out as a more suitable option.
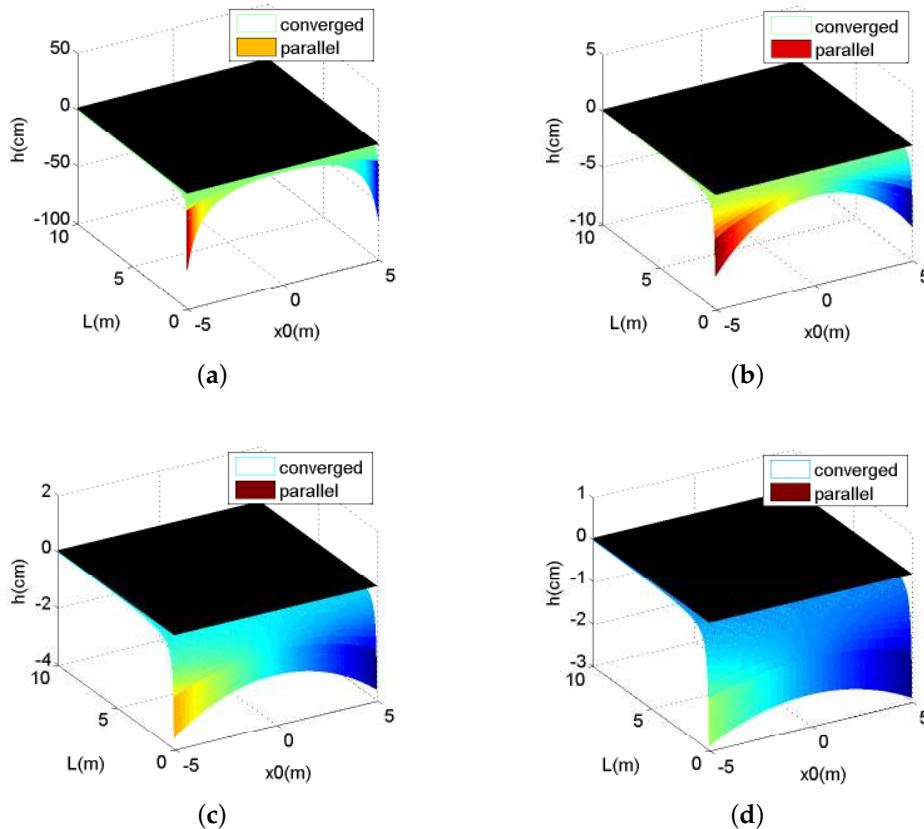
Figure 4. The curve of horizontal parallax h and x_0, L. (**a**) $z_0 = 1$ m; (**b**) $z_0 = 4$ m; (**c**) $z_0 = 7$ m; (**d**) $z_0 = 10$ m.

3. Auto-Converged Camera Array Realization

To identify the analysis in Section 2, we design an auto-converged camera array. The configuration in Figure 5 is made up of two CCD sensors spaced 7 cm apart, pan-tilt-zoom, four stepping motors, rotate axis and rotary bracket. There are four stepping motors that are used to control the four degrees of free rotary movement to simulate the configuration turning left or right and the cameras' switching angle of view. For accurate fusion of spatial sensor information, proper calibration among the cameras is required. The entire sensor is mounted on a tripod beside the computer. In this paper, we adopt the Kalman filter to track the target [24]. The designed cameras can converge and focus on moving targets automatically based on the measured distance.

An experimental test-bed was configured using multimedia PC components, Microcontroller Unit (MCU) and Digital Signal Processor (DSP). A diagram of the sensor, date acquisition and processing components is shown in Figure 6. The multimedia 1394 video capture card can be used to grasp 1024×768 pixel color images from either camera. Image processing and camera control are performed in real-time on the DSP and MCU.

The general working principle of the system is: firstly, digital images captured by cameras are processed with the DSP chip; secondly, the result of processing is transmitted to the microcomputer through the bus; thirdly, the movement of the camera system is controlled with a stepper motor driven by the microcomputer. As shown in Figure 7, the system consists of the video capture module, the data processing module and the communication and control module. Figure 8 shows the physical frame of the designed auto-converged camera array.
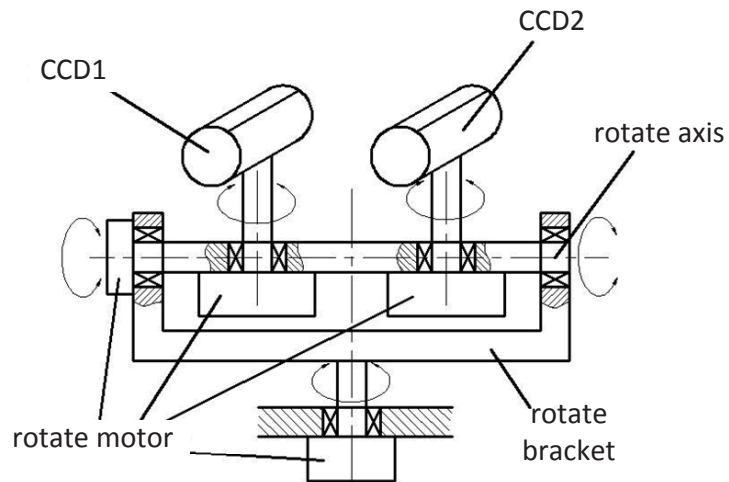
Figure 5. The sensor configuration of the auto-converged camera array.
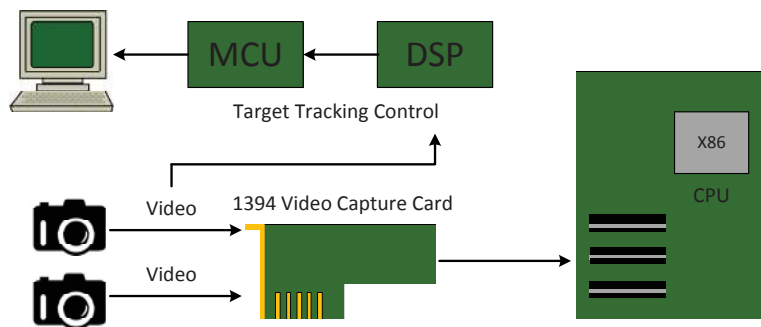
Figure 6. Component interconnections.

Figure 7. The physical diagram of the auto-converged camera array.

Auto-converged arrays are widely used in three-dimensional reconstructions, face recognition, gesture recognition and image assessment based on stereo vision, and so on. In image assessment, stereo images are necessary for the experiments. Auto-converged camera arrays can be a perfect tool to take stereo images for the study of image assessment, as Yang *et al.* do in [25].

Figure 8. Structure of the system hardware.

4. Auto-Converged Experiment and Measure

In order to verify the analysis above, we use a parallel and a converged camera array to shoot images of a moving man as the target in the same condition, then contrast the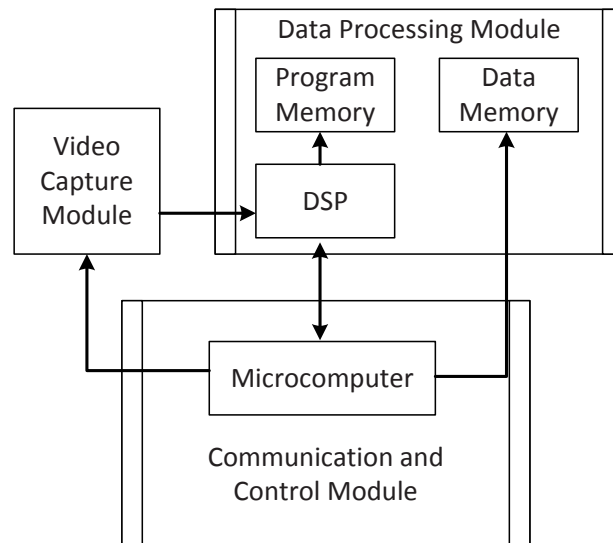se two groups of images, as shown in Figure 9. The resolution of the camera array used is 1024×768. The convergence distance varies in the range of 1 m to 10 m.

We study the stereo images taken by the parallel array and find that the stereo sense is obvious if the shooting distance is large, but for a short distance, the images cannot be fused. To be more precise, if the shooting distance is less than 3 m, it is difficult for the obtained parallax images to be fused; if it is between 3 m and 4 m, fusing is realized, but there exists much ghosting; if it goes beyond 4 m and up to 10 m, better fusing is realized with very little ghosting, but it is still not enough to realize the perfect life-like effect of the actual three-dimensional scene.

Equation (16) shows that horizontal parallax from parallel images decreases as shooting distance increases. If shooting distance z_0 is less than 3 m, the fusion of parallax images cannot be realized, as the horizontal parallax goes beyond the allowed range; if z_0 varies from 3 m to 4 m, the images can be fused, but there exists much ghosting; if z_0 is greater than 4 m, the parallax is in the allowed range, and the corresponding images fuse perfectly. However, only positive parallax appears in fused images, so only a protruding visual effect can be obtained.

We study converged stereo-images and find that if the shooting distance is less than 5 m, stereo images taken by the converged array have a good stereo sense; if it changes from 5 m to 7 m, the visual stereo perception decreases; if it is greater than 7 m, the converged camera array works almost the same as the parallel camera array. It makes no perceptual difference between them. Therefore, 7 m is the threshold to distinguish the difference of the parallel and converged camera arrays. The three-dimensional scene can be displayed in a life-like manner with the images taken by the converged array, which is protruding outside and concavely sunk.

We use the designed camera array to take two groups of pictures of a man moving slowly toward the camera. One group is taken by a parallel camera array, and the other is taken by a converged one. The shooting distance for both cameras in the experiment varies in the range from 1 m to 10 m. We take integers of the distance as our experiment setting and take five pictures at every distance. Then, we get 50 pictures for the parallel array and 50 pictures for the converged one. Figure 9 gives some examples of the pictures taken. The observers feel dizzy when they look at the parallel stereo images. Their eyes will be tired if the scene taken in the pictures is close to the cameras; while stereoscopic images taken by the parallel camera array offer us a comfortable depth sensation, and the target

keeps in near the center of the imaging plane. The converged camera array reproduces the true three-dimensional scene.

(**a**)   (**b**)
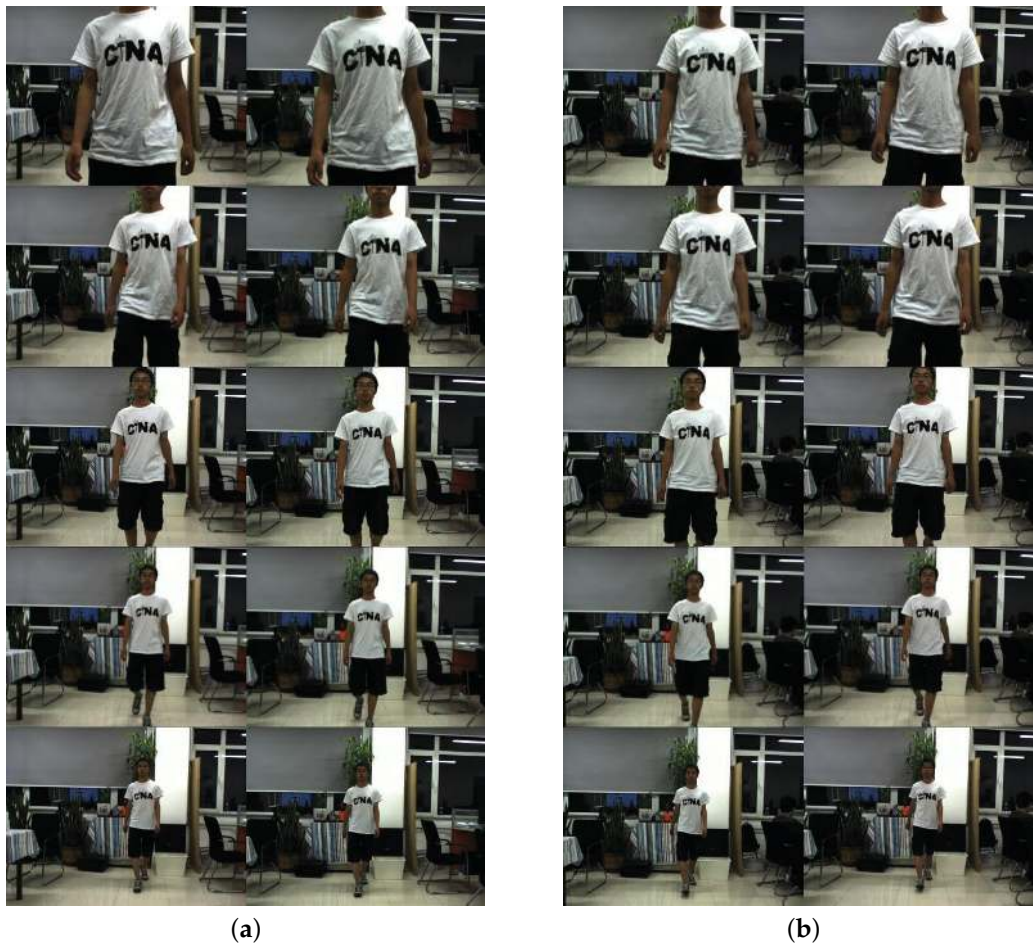
Figure 9. Images taken by using the converged camera array and the parallel camera array. (**a**) Images taken by using the converged camera array; (**b**) images taken by using the parallel camera array.

Table 1. DIIVINE quality of the images taken by the converged and the parallel camera array.

Shooting Distance	1 m	2 m	3 m	4 m	5 m
converged	0.75	0.79	0.82	0.84	0.86
parallel	0.60	0.67	0.75	0.79	0.81
shootingdistance	6 m	7 m	8 m	9 m	10 m
converged	0.86	0.83	0.81	0.82	0.80
parallel	0.83	0.83	0.82	0.83	0.81

We also adopt the objective evaluation method for stereo image quality to evaluate the stereo images based on the algorithm called DIIVINE in [26]. Table 1 gives the corresponding results. We can see from the figures in the table that when the shooting distance is less than 7 m, the overall performance of the converged cameras is better than the parallel cameras, while when the shooting distance is more than 7 m, they are comparable. Combined with the analysis above, we can conclude that converged camera arrays outperform parallel camera arrays when the shooting distance is less

than 7 m; however, if we take the keystone distortion into consideration, parallel camera arrays will be better in the condition of a shooting distance greater than 7 m.

5. Conclusions

In this paper, we build a parallel camera array model and a simplified converged camera array model as the base of subsequent analysis. A physical camera array is developed, which can be used as a parallel array, as well as a converged array. This is convenient for the different choices according to the real needs. The pros and cons of these two kinds of cameras are discussed. A conclusion is that converged arrays are more suitable for short-distance shooting and parallel arrays for long-distance shooting. Our work can be guidance to the application of the camera arrays in different environments. For future work on the auto-converged camera array, we need to focus on camera calibration and visual stereo-video evaluation.

Author Contributions: Jiachen Yang and Ru Xu did the overall analysis. Zhihan Lv and Houbing Song helped carry out the experiment and data statistics.

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

References

1. Lv, Z.; Halawani, A.; Feng, S.; Réhman, S.; Li, H. Touch-less Interactive Augmented Reality Game on Vision Based Wearable Device. *Pers. Ubiquitous Comput.* **2015**, *19*, 551–567.
2. Monares, Á.; Ochoa, S.F.; Santos, R.; Orozco, J.; Meseguer, R. Modeling IoT-based solutions using human-centric wireless sensor networks. *Sensors* **2014**, *14*, 15687–15713.
3. Tiete, J.; Domínguez, F.; Silva, B.D.; Segers, L.; Steenhaut, K.; Touhafi, A. SoundCompass: A distributed MEMS microphone array-based sensor for sound source localization. *Sensors* **2014**, *14*, 1918–1949.
4. Sun, T.; Fang, J.Y.; Zhao, D.; Liu, X.; Tong, Q.X. A Novel Multi-Digital Camera System Based on Tilt-Shift Photography Technology. *Sensors* **2015**, *15*, 7823–7843.
5. Lv, Z.; Tek, A.; Da Silva, F.; Empereur-Mot, C.; Chavent, M.; Baaden, M. Game on, science-how video game technology may help biologists tackle visualization challenges. *PLoS ONE* **2013**, *8*, 57990.
6. Kim, D.; Kim, K.; Lee, S. Stereo camera based virtual cane system with identifiable distance tactile feedback for the blind. *Sensors* **2014**, *14*, 10412–10431.
7. Yamaguchi, M.; Wakunami, K.; Inaniwa, M. Computer generated hologram from full-parallax 3D image data captured by scanning vertical camera array. *Chin. Opt. Lett.* **2014**, *12*, 060018-1–060018-6.
8. Jovanov, L.; Pižurica, A.; Philips, W. Fuzzy logic-based approach to wavelet denoising of 3D images produced by time-of-flight cameras. *Opt. Express* **2010**, *18*, 22651–22676.
9. Bradshaw, M.F.; Hibbard, P.B.; Parton, A.D.; Rose, D.; Langley, K. Surface orientation, modulation frequency and the detection and perception of depth defined by binocular disparity and motion parallax. *Vis. Res.* **2006**, *46*, 2636–2644.
10. Shim, H.; Lee, S. Hybrid exposure for depth imaging of a time-of-flight depth sensor. *Opt. Express* **2014**, *22*, 13393–13402.
11. Son, J.Y.; Javidi, B.; Kwack, K.D. Methods for displaying three-dimensional images. *Proc. IEEE* **2006**, *94*, 502–523.
12. Jung, K.; Park, J.I.; Choi, B.U. Interactive auto-stereoscopic display with efficient and flexible interleaving. *Opt. Eng.* **2012**, *51*, 027402-1–027402-12.
13. Shin, D.; Daneshpanah, M.; Javidi, B. Generalization of three-dimensional N-ocular imaging systems under fixed resource constraints. *Opt. Lett.* **2012**, *37*, 19–21.
14. Son, J.Y.; Javidi, B. Three-dimensional imaging methods based on multiview images. *Disp. Technol. J.* **2005**, *1*, 125–140.

15. Zhang, L.; Tam, W.J. Stereoscopic image generation based on depth images for 3D TV. *Broadcast. IEEE Trans.* **2005**, *51*, 191–199.

16. Kooi, F.L.; Toet, A. Visual comfort of binocular and 3D displays. *Displays* **2004**, *25*, 99–108.

17. Yamanoue, H. The differences between toed-in camera configurations and parallel camera configurations in shooting stereoscopic images. In Proceedings of the 2006 IEEE International Conference on Multimedia and Expo, Toronto, ON, Canada, 9–12 July 2006; pp. 1701–1704.

18. Yoon, K.H.; Ju, H.; Park, I.; Kim, S.K. Determination of the optimum viewing distance for a multi-view auto-stereoscopic 3D display. *Opt. Express* **2014**, *22*, 22616–22631.

19. Son, J.Y.; Yeom, S.; Lee, D.S.; Lee, K.H.; Park, M.C. A stereoscopic camera model of focal plane detector array. *Disp. Technol. J.* **2011**, *7*, 281–288.

20. Fazio, G.; Hora, J.; Allen, L.; Ashby, M.; Barmby, P.; Deutsch, L.; Huang, J.S.; Kleiner, S.; Marengo, M.; Megeath, S.; *et al*. The infrared array camera (IRAC) for the spitzer space telescope. *Astrophys. J. Suppl. Ser.* **2004**, *154*, 10–16.

21. Sawhney, H.S.; Kumar, R. True multi-image alignment and its application to mosaicing and lens distortion correction. *Pattern Anal. Mach. Intell. IEEE Trans.* **1999**, *21*, 235–243.

22. Yoneyama, S.; Kikuta, H.; Kitagawa, A.; Kitamura, K. Lens distortion correction for digital image correlation by measuring rigid body displacement. *Opt. Eng.* **2006**, *45*, 023602.

23. Yang, J.; Gao, Z.; Chu, R.; Liu, Y.; Lin, Y. New stereo shooting evaluation metric based on stereoscopic distortion and subjective perception. *Opt. Rev.* **2015**, *22*, 459–468.

24. Amirzadeh, A.; Karimpour, A. An interacting Fuzzy-Fading-Memory-based Augmented Kalman Filtering method for maneuvering target tracking. *Digit. Sig. Process.* **2013**, *23*, 1678–1685.

25. Yang, J.; Liu, Y.; Meng, Q.; Chu, R. Objective Evaluation Criteria for Stereo Camera Shooting Quality under Different Shooting Parameters and Shooting Distances. *IEEE Sens. J.* **2015**, *15*, 4508–4521.

26. Moorthy, A.K.; Bovik, A.C. Blind image quality assessment: From natural scene statistics to perceptual quality. *IEEE Trans. Image Process.* **2011**, *20*, 3350–3364.