

Research on Individual Differences in Ability to Recognize Obstacles Employing Binocular Stereopsis and Development of Stereoscopic Test Video Sequences for Analyses

Emma Miwa, Michihiro Hayashi, Justin Han, Sakuichi Ohtsuka
International College of Technology, Kanazawa, Kanazawa, Japan

Abstract

Final goal of this study is to analyze the side effects of the development of binocular stereopsis when children become habituated with the use of flat panel displays such as smart phones or tablets in their daily environment from infants. To this end, experiments are conducted employing a head-mounted display (HMD). As an initial step, this study proposes stereoscopic test video sequences.

Author Keywords

Stereopsis, individual differences, flat panel display, peripheral vision, development, screen time.

1. Introduction

In daily life, we are surrounded by numerous obstacles. When moving around the home, people must avoid tables, chairs, and in hallways, houseplants may share the floor space. In our robotics laboratory (See Figure 1), there are tables, tool cabinets, and parts racks, and a number of obstacles block people's movements as they attempt to reach their work area. As one good example, a broom handle often was suddenly perceived in the field of view forcing the walker to make a sudden evasive maneuver (See Figure 2). Many lab members as well as visitors were late in perceiving the handle and found it to be a disturbing obstacle.

It is suggested that people employed the peripheral vision created by binocular stereopsis to recognize the intruding handle and judge it to be an obstacle [1]. Regardless of their level of technical expertise, all who recognized the handle as an obstacle were interested in painting, photography, etc., and thus were keen observers of real 3D objects and the real 3D surrounding space.

Therefore, it is thought that they had acquired the ability to estimate the position of objects and predict collisions by the peripheral vision of their binocular stereopsis, and many

exhibited spatial recognition and shape grasping abilities. In addition, it is known that personally acquired trajectory prediction strategies strongly depend on training in sports [2][3]; tests have found that skill in sports, especially table tennis, leads to differences in the impact of optical illusions [3].

It has been suggested that binocular stereoscopic abilities may be altered in the generation that grew up with the presence of electronic displays in their daily environment from an early age [4]. The World Health Organization (WHO) established "Guidelines on physical exercise and sleep for infants under 5 years of age from a physical development perspective"; it recommends that children under 2 years of age should not watch electronic displays [5].

In fact, a common topic of discussion is that some engineers of the younger generation can design complex mechanical structures using 3D-CAD, but cannot judge the spatial interference between objects or the front and back without rotating the objects on the screen (e.g., see ref. [4]). Note that, even though titled 3D (three-dimensional), a three-dimensional object displayed on an electronic display is a 2-D rendering of a 3D-CAD model inside the computer. Essentially, there is no 3D parallax information, and shape recognition using binocular stereo vision functions is not required.

This study introduces preliminary experiments employing a 3D stereoscopic video camera capable of capturing two images with both motion and binocular parallax information, and presenting the images in an HMD to confirm the ability of individuals to recognize spatial obstacles and collision avoidance behavior. In addition, we examine what types of images should be presented and the method of capturing images. Here, we also need to pay attention to the activity of standardizations committee concerning to HMD ergonomics [6].



Figure 1. View of authors' laboratory with desks and shelves



Figure 2. Laboratory staff avoiding the broom handle (Demonstration)



Figure 3. Appearance of the 3D camera as part of HMD

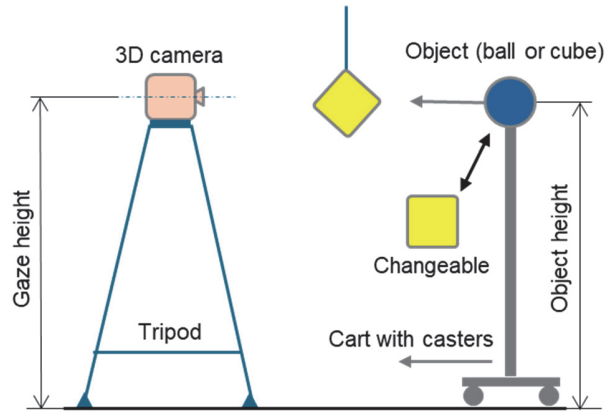


Figure 4. Schematic illustration of video capturing. Note that the camera was fixed and only the object was moved.

Table 1. Specifications of 3D camera

Distance between the two lenses	65 mm
Aperture	F1.8
Focal length	28 mm
FOV	66° (H) 52° (V) 79° (D)
Sensor size	1/2 inch
Video resolution	3840 × 1080
Video format	MP4
Dimensions (camera)	94 × 52.8 × 22
Weight (camera)	160 g

2. Preliminary experimental result

We captured a representative stereoscopic video clip by employing a stereoscopic video camera “QooCam EGO” made by ShenZhen KanDao Technology Co., Ltd. See Table 1 for its specifications.

This video camera was combined with a personal full view stereoscopic 3D display (See Figure 3). Therefore, subjects using this display are able to have the experience of moving around the laboratory.

A total of five subjects participated. Four lab-members, two teachers and two students, and one visitor who spontaneously reported that she found the handle to be annoying.

All five subjects found the handle to be annoying in the 3D display as in the real 3D space.

Therefore, in summary, it is suggested that the above evaluation system (the stereoscopic 3D video camera to capture environmental images, and HMD to reproduce the environment) is valid for our purpose.

3. Proposal for experimental capture system for human visual space perception analyses

3.1 Perspectives on creating evaluation images

The first viewpoint is that the current location of the object of interest (OoI), should be one that permits easy detection by central vision, or hard peripheral vision. The second viewpoint is that the estimated trajectory of the OoI, should be such that the object could damage the observer directly.



Figure 5. Equipment setting in the video shooting environment

3.2 Equipment for capturing test video sequences

Figure 4 is a schematic illustration of the video capturing system. Figure 5 shows a photograph of the actually constructed experimental setup; the grid on the floor has 1m intervals. It was placed in an expansive studio and the background wall was colored uniform gray by matte paint in order not to give additional information to the observer. As you can see, the setup was constructed (1) stereoscopic video camera (abovementioned QooCam EGO) fixed on a tripod, and (2) a matte yellow cube (side length of 200 mm) for the obstacle was hung from the ceiling by a transparent thread, and (3) a movable glossy blue ball, with diameter of 200mm, mounted on a mobile trolley.

One of our experimental objectives was to force the observer to judge the depth relationship (i.e., collision, or go through front or back) between the two objects (yellow cube and blue ball); with this arrangement the observer could acquire basically acquired no additional depth information from the surroundings.

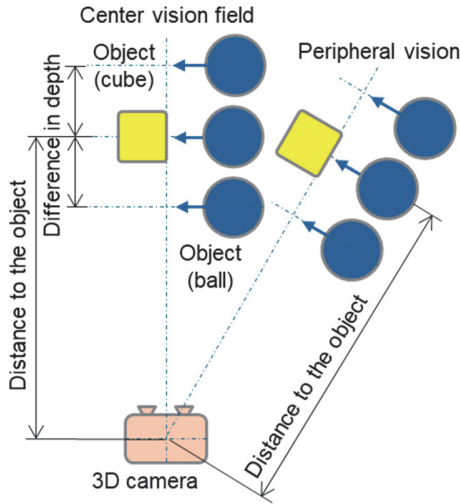


Figure 6. Prediction of collisions between objects and observer

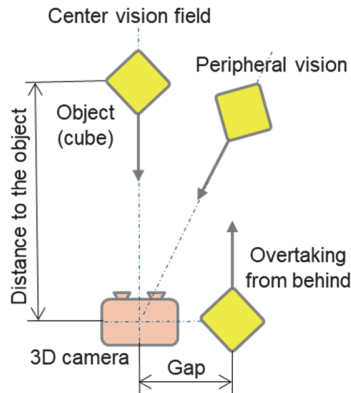


Figure 7. Prediction of obstacle collision as the goal

3.3 Examples of arrangement of camera and objects, and sample stereoscopic images

Figures 6 and 7 show schematic illustrations of examples of the arrangements of camera and objects.

First experimental task was judge the collision of the two objects, neither object was estimated to be unable to damage the observer directly (i.e., safe), as shown in Figure 6. The distance from the camera to the cube was about 1400 mm, and we captured videos of the ball colliding with the cube from the right or crossing in front of and behind it. We employed the images captured with the cube in front of the camera for the central vision experiment, and the images captured with the cube at an angle of about 16 degrees to the right for the peripheral vision experiment. Figures 8 and 9 show examples for both cases; the stereoscopic images are produced by free fusing.

In the second experiment, one object was taken as being in that it could collide with the observer, as shown in Figure 7. The videos were captured assuming that the same cube used in the previous experiment approached the observer or that it was overtaking the observer.

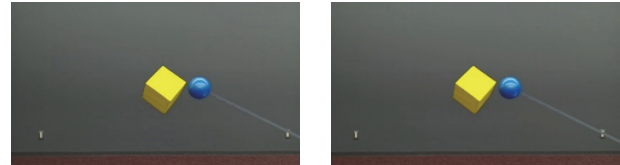


Figure 8. Stereo images for collision prediction in the central vision field

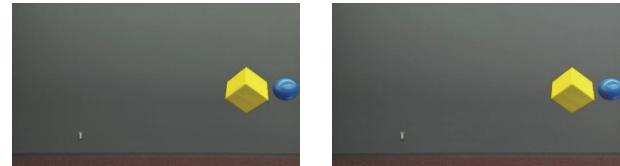


Figure 9. Stereo images for collision prediction in the peripheral vision



Figure 10. Stereo images for collision prediction in the central vision field

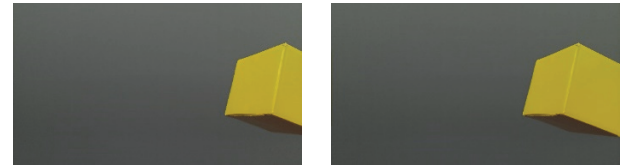


Figure 11. Stereo images for collision prediction in the peripheral vision

Number of conditions was four: two collision cases and two non-collision cases. In the former cases, there were included that (a1) the height of the obstacle was set at the line of sight, and (a2) to collide with the chest (300 mm lower than the camera). In the latter, (b1) the height of the obstacle was set to approach and pass overhead (300 mm higher than the camera), and (b2) the obstacle was set to overtake from the right rear. Figures 10 and 11 show examples (a1) and (b2), respectively.

3.4 Preliminary evaluation of captured stereoscopic video images

For the first step, we confirmed the video sequences in the safe condition, i.e., the judgment of the collision of two objects. Observer viewed the stereoscopic video clips by employing the replay function of the video camera while seated in a chair. Video camera was mounted on a tripod.

Number of observers were three; one teenager (student), one in their twenties (technical staff), and one in their thirties (teacher); none were the authors. Number of experimental conditions were three as shown in Figure 6, only the case of central vision condition was examined.

All subjects judged all depth relationships correctly when viewing three stereoscopic video clips; they had actually visited the real experimental setup before viewing the video.

Therefore, it was suggested that our method using stereoscopic video taken in the studio was effective in simulating the observer’s judgment in real environments.

Table 2. Evaluation sheet for situation between the cube and ball

Movie [1]	Movie [2]	Movie [3]	Movie [4]	Movie [5]	Situation
+	+	+	+	+	Overtaking
*	*	*	*	*	Collision with the box
-	-	-	-	-	Passing by the front
5	5	5	5	5	Surely
4	4	4	4	4	Perhaps
3	3	3	3	3	
2	2	2	2	2	
1	1	1	1	1	Not confident

4. Example of preliminary subjective test

4.1 Experimental conditions

Observers’ task was collision judgment of two objects, i.e., still yellow cube and moving blue ball, in conditions of central- and peripheral- vision. The depth conditions were that the ball depth was located (1) 400 mm farther, (2) 200 mm farther, (3) collision (0 mm), (4) 200 mm nearer, and (5) 400 mm nearer. The videos were played twice in each condition at random. Observers were required to answer ball location and confidence of judgement by using the evaluation sheet shown in Table 2. Number of observers were two; one twenties (technical staff) and one thirties (teacher), both participated in the previous experiment.

4.2 Results and discussions

Figure 12 shows the results. Positive/negative value in horizontal axis shows that the ball was farther/closer than the cube. First, all judgments of collision in central vision condition were correct, whereas the peripheral vision condition yielded incorrect judgements (two in #2). Second, the confidence scores were scattered, except for the case of +400 mm. Two reasons are advanced: one was a visual field effect, the other was a distance effect. Therefore, we need to collect more data and the analysis should be from multiple viewpoints.

The above results are generally consistent with common sense. However, it is known that some people, such as athletes and professional drivers, have excellent ability to estimate the position of objects and predicting their trajectories in their peripheral vision. Therefore, more precise experimental planning is needed to obtain fuller knowledge for developing that ability in others.

5. Conclusions

In recent a few years, AR/VR technologies, such as head-mounted displays (HMDs), have been rapidly advancing and are starting to enter the consumer market. This makes it easier to analyze the characteristics of human visual perception, as related to real 3D space.

This study demonstrated preliminary experiments and results

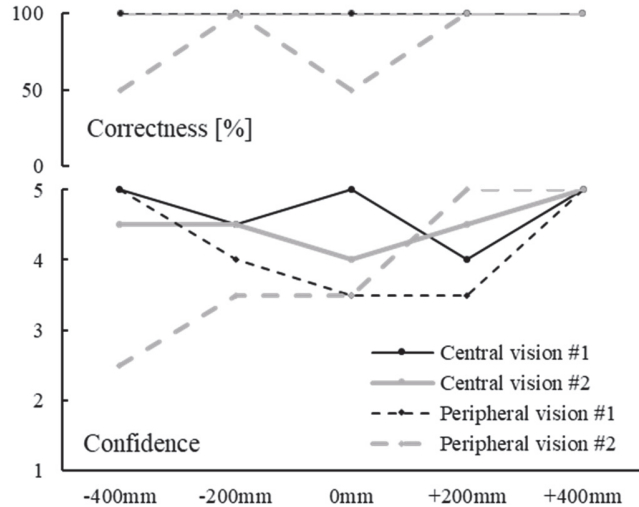


Figure 12. Correctness and confidence in the positional relationship between the cube and the ball

attained by reproducing an actual laboratory situation. Next, we analyzed some requirements for producing useful test video sequences. A part of the experimental setup suitable for video data capture was also described.

The final goal of this study is to analyze the side effects on the development of binocular stereopsis of children when they grow up with in the ubiquitous presence of flat panel displays such as smart phones or tablets; for this, experiments employing HMDs will be conducted.

6. References

- Fademrecht L., B’ulthoff I., and de la Rosa S. (2016). Action recognition in the visual periphery. *Journal of Vision*, 16(3):33, 1–14, <https://doi.org/10.1167/16.3.33>
- Ohtsuka, S., Nakagawa, H., Fukaya, Y., Torisato, N. and Hira, S. (2019). 3-2: Importance of Object Contour Retention in 3D Space Based on “The Law of Inertia” Hypothesis in Human Perception, with Analysis of 2D- and 3D- Footsteps Illusion. *SID Symposium Digest of Technical Papers*, 50: 5-8. <https://doi.org/10.1002/sdtp.12841>
- Lee J.-S. et al. (2021). Stereopsis and Response Times between Collegiate Table Tennis Athletes and Non-Athletes, *Int. J. Environ. Res. Public Health*, 18: 6287. <https://doi.org/10.3390/ijerph18126287>
- Ohtsuka S. (2023). Discussion on latent atypical development of both visual perception and its interaction with tactile perception in ubiquitous artificial visual environment — Differences in fundamentals between three- and two- dimensional percepts —, *IEICE Tech. Rep.*, 123(320) HIP2023-87: 53-58. <https://ken.ieice.org/ken/paper/20231222KcAi/eng/> (Abstract in English)
- World Health Organization (WHO) (2019). Guidelines on physical activity, sedentary behaviour and sleep for children under 5 years of age. <https://iris.who.int/handle/10665/311664>
- Hiroyasu, Ujike; Kei, Hyodo; Mitsunori, Tada; Hiroshi, Watanabe; “Standardizations of Ergonomics for Head Mounted Displays (HMDs),” *SID Symposium Digest of Technical Papers*, Vol.55, 2024, <https://doi.org/10.1002/sdtp.17682>