# Stereoscopy in Cinematographic Synthetic Imagery

Jonathan Eisenmann and Rick Parent

Department of Computer Science and Engineering
The Ohio State University, Columbus, Ohio, USA



http://dx.doi.org/10.1117/12.805994

# Stereoscopy in Cinematographic Synthetic Imagery

Jonathan Eisenmann* and Rick Parent+

Department of Computer Science and Engineering
The Ohio State University, Columbus, Ohio, USA

## ABSTRACT

In this paper we present experiments and results pertaining to the perception of depth in stereoscopic viewing of synthetic imagery. In computer animation, typical synthetic imagery is highly textured and uses stylized illumination of abstracted material models by abstracted light source models. While there have been numerous studies concerning stereoscopic capabilities, conventions for staging and cinematography in stereoscopic movies have not yet been well-established. Our long-term goal is to measure the effectiveness of various cinematography techniques on the human visual system in a theatrical viewing environment. We would like to identify the elements of stereoscopic cinema that are important in terms of enhancing the viewer's understanding of a scene as well as providing guidelines for the cinematographer relating to storytelling.

In these experiments we isolated stereoscopic effects by eliminating as many other visual cues as is reasonable. In particular, we aim to empirically determine what types of movement in synthetic imagery affect the perceptual depth sensing capabilities of our viewers. Using synthetic imagery, we created several viewing scenarios in which the viewer is asked to locate a target object's depth in a simple environment. The scenarios were specifically designed to compare the effectiveness of stereo viewing, camera movement, and object motion in aiding depth perception. Data were collected showing the error between the choice of the user and the actual depth value, and patterns were identified that relate the test variables to the viewer's perceptual depth accuracy in our theatrical viewing environment.

**Keywords:** Stereoscopy, Computer Animation, Depth Perception, Cinematography

## 1. INTRODUCTION

### 1.1 Stereoscopic Cinema

Motivated by the recent activity on the part of the movie industry in adopting stereo technology, we are interested in looking at how stereo viewing impacts the theater experience. This is particularly interesting in the context of computer animation in which the production process has control over every aspect of the resulting imagery. We would like to move towards developing guidelines for filmmakers on the effective use of stereo technology in conjunction with computer animation.

### 1.2 Problem Statement

While there have been numerous studies concerning stereoscopic capabilities,[1–3] viewer comfort, and perception,[4–7] conventions for staging and cinematography in stereoscopic movies have not yet been well-established. We wish to measure the effectiveness of various cinematographic techniques on the human visual system in a theatrical viewing environment, which we define to consist of a large, flat screen viewed from a static, relatively distant position without head-tracking. Toward this end, we present our experiments and the results which lay the foundation for further exploration of relevant concepts in the study of stereoscopic viewing of synthetic imagery.

* eisenmaj@cse.ohio-state.edu
+ parent@cse.ohio-state.edu

### 1.3 Cinema vs. Virtual Reality

There has been much work done on perception of stereo in theme park rides and virtual reality,[8–11] but relatively less so on the theater experience. The theater experience differs from theme park rides and virtual reality because the viewer is only occasionally traversing through the environment. More typically, the viewer is a stationary observer of action or dialogue. We would like to explore these differences in this paper and in subsequent research.

### 1.4 Scope of our study

In this paper, we report our initial experiments that consider the effect of camera and object movements on the accuracy of the viewers' perception of depth with and without stereo viewing.

## 2. EXPERIMENTS

### 2.1 Apparatus

For our experiments, we utilized a custom polarized stereoscopic arrangement provided by Alan Price of the Advanced Computing Center for the Art and Design (ACCAD) at the Ohio State University. The setup consisted of two DLP projectors with circularly polarized lenses which reflected two images off of a mirror and onto the rear side of the screen. We used a rear-projected system due to space limitations and also to keep the viewers from blocking the projection of the image.

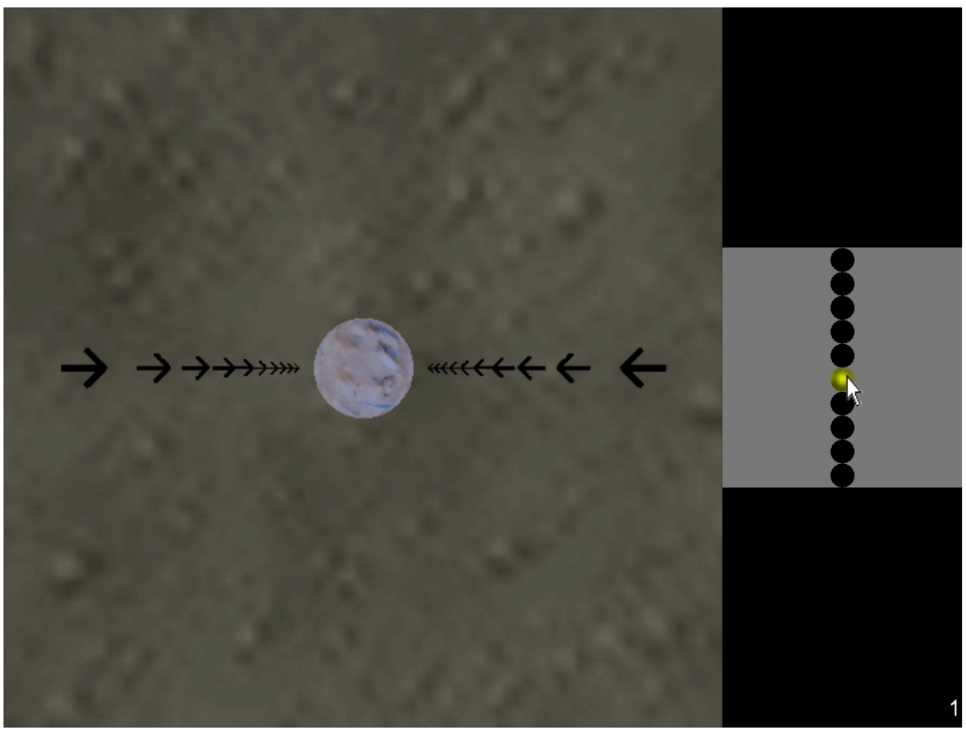


Figure 1. Our test environment

### 2.2 User Interface

We created the virtual environment we used for our subject tests with the Virtools development toolkit from Dassault Systemes. The two images necessary for binocular viewing were created side by side for projection as a horizontal span with pixel resolution of 2048 by 768, and each image was displayed by a single projector with a pixel resolution of 1024 by 768. While this resolution is less than that of a movie theater projection system, we were limited by the specifications of the available hardware, and we felt that this would not greatly impact

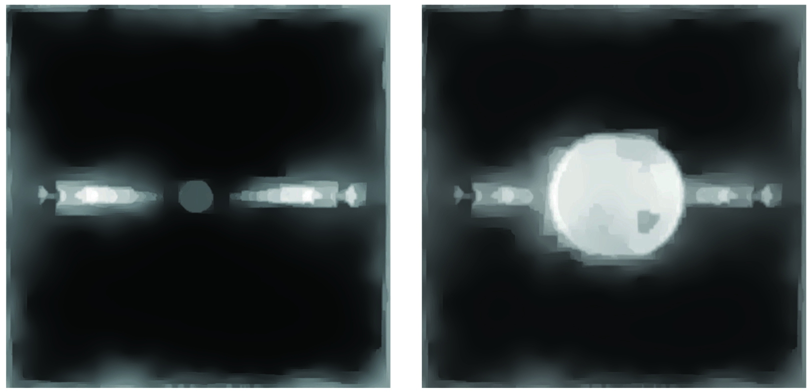

Figure 2. These images show the optical flow of interocular disparity when the target sphere is located at the most distant position (left) and the nearest position (right) to the subject.

our testing. The two images were then polarized and projected on top of one another. Figure 1 shows our test environment.

The 3D portion of our environment occupies a square area on the left side of the projected image. It consists of the unshaded target sphere with a noisy texture, an unshaded background with a different noisy texture, and two rows of 2D arrows placed on either side of each of the ten possible positions that can be occupied by the target sphere.

The 2D portion of our environment, the subject depth-picking interface, is on the right. It contains ten circles that represent the ten possible positions from a top-down, orthographic view. These ten positions were the same in each different test scenario. The subject indicates his or her choice by clicking on one of the circles. There is a counter in the bottom right hand corner for indicating how many responses have been recorded for the current test.

The subjects wore circularly polarized glasses and were seated 10.5 feet away from the projection screen which measures 10 by 5.5 feet. Our virtual cameras had a 35mm standard field of view. The relatively small eye separation we used in our 3D tests was 2.32 inches, the effect of which is shown in the images in Figure 2 which demonstrate the interocular disparity between the left and right images.

### 2.3 Preliminary Experiments

An initial synthetic scene was constructed with various depth and motion cues; we conducted preliminary testing on three subjects to establish consistencies and validate our stereoscopic display. These results served to direct our subsequent experiments where the objective was to isolate stereoscopic effects by eliminating as many other visual cues as possible. Our basic scene consisted of a synthetic camera placed in the center of a large, uniformly lit, neutral color, matte sphere which served as the background. The camera could see the interior of the sphere, equally distant in all directions. A smaller, uniformly lit, matte sphere of a neutral but contrasting color was positioned in front of the camera at random depths. In some tests constant retinal size of the target sphere was enforced so it always occupied the middle third of the field of view. The subjects were asked to select the depth of the target. Several tests were run with varying conditions, and data was collected showing how the choices of the subjects matched the correct responses.

The human visual system relies in part on the relative retinal size of an object to discern depth.[12] Our initial studies indicated that, with this cue absent, depth perception accuracy was reduced by about 50% regardless of the presence or absence of textures or stereoscopy. Stereoscopic viewing provided about 15% more depth information than monoscopic viewing in cases where the images were textured. In untextured tests, stereoscopy actually gave much less depth information. Subtle camera movements seemed to provide a remarkable amount of depth perception, but it was not clear under what scenarios this movement helped or distracted. For example, when it was used in conditions where the field of view was narrow it seems to be more distracting than helpful, perhaps because the parallax effect was exaggerated in such a compressed frustum. Larger fields of view seemed to yield greater depth perception accuracy. After these initial tests, we revised our test environment as discussed below.

### 2.4 Experimental Design

As a result of our initial tests, we sought to balance the competing principles of identifying and isolating the test variables with presenting the imagery in a relevant context. We did not want to isolate the variables at the cost of losing relevancy so we only eliminated the following depth cues: shading, shadows, and comparison with sizes of familiar objects. We intentionally made the objects in the scene unfamiliar in terms of size so that the only size information a viewer could utilize in judging depth came from relative size comparisons. We included static reference objects in the scene (the arrows) to provide the viewers with better spatial orientation and to help the viewers judge the depth of the target relative to fixed positions in space. We adjusted the convergence of the 3D camera configurations so that the nearest five positions appeared in front of the screen, and the furthest five positions appeared behind the screen plane.

### 2.5 Subjects

We asked students and professors at our research facility to participate in the subject testing process. We tested subjects whose ages ranged from 20 to 60 with the majority of the participants' ages ranging between 20 and 30. There were 3 females and 9 males. In addition, 8 of our subjects had corrected vision.

### 2.6 Test Cases

| **Camera Configuration** | **Camera Movement** | **Object Movement** |
|---|---|---|
| Monoscopic | None | None |
| Toe-in Stereoscopic | Trucking Oscillation | Distal Oscillation |
| Parallel Stereoscopic | Swaying Oscillation | Lateral Oscillation |

Table 1. Test variables

The test cases were composed of all possible combinations of the variables shown above. The resulting test sequence contained 27 different scenarios. Even though it has been widely accepted in the literature[1,13] that the parallel camera configuration is better for reducing distortion (specifically keystone distortion) in stereoscopic applications, we included the toe-in configuration to see how the two would compare in our results. Trucking camera movement consisted of small oscillating movements into and out of the scene along the line of sight. Swaying camera movement was implemented as a small rotation about the vertical axis to the left and right about a pre-defined center point in the scene.

### 2.7 Testing Procedure

The subjects were briefed on the environment, the depth judgment task, and the user interface. They were then asked to complete ten depth judgments for each of the 27 test scenarios. In each test, the target sphere appeared in the 3D environment for one second at one of the ten possible depths and then disappeared. The placement of the target sphere was determined by a uniformly random selection process. The subjects were asked to click on the circle that seemed to the viewer to most accurately represent the depth position of the sphere. There was no time limit for the depth judgment response; however, we only allowed the subjects to view the sphere in 3D space for a short time so that they would respond based on pure perceptual data and not be tempted to geometrically analyze spatial relationships for extended periods of time. At the beginning of the testing process and each time the camera configuration changed, the ten possible positions were demonstrated (from far to near) to the subject to give them an idea of both the range of depth values possible and the relationship between each of the adjacent positions. The subjects were given the opportunity to practice clicking with the mouse pointer on the 2D depth picking interface if desired.

We tested subjects with the 2D camera configuration first, followed by the toe-in 3D configuration, and finally the parallel 3D configuration. The target sphere oscillations were slight in order to avoid giving too many occlusion depth cues to the subjects by letting the sphere get too close to the arrows on either side. Subjects were asked to locate the target by making judgments about its average depth during its distal oscillatory movement, and it never moved more than half a position in either direction.

## 3. RESULTS AND DISCUSSION

Our findings are detailed in the charts below. As we expected, the toe-in configuration proved to be sufficiently similar to the parallel configuration so we have omitted the data in our charts for the sake of readability. The three charts in Figure 3 show the total error, averaged over all subjects, for each of the 27 test scenarios. We define error to be the absolute integer distance between the true position of the target sphere and the subject's response, and total error represents the sum of the absolute errors for the ten trials of each test scenario. The ranges shown with each bar represent the standard deviation. The first graph represents the scenarios with no object movement, the second graph shows total errors for the distal object movement scenarios, and the third graph displays the scenarios with lateral object motion.

Figure 4 compares the errors between the scenario with a stationary camera and lateral target movement and the scenario with swaying camera movement and a stationary target sphere. Figure 5 shows the total error (averaged over all subjects and all motion combinations) based on whether the target sphere was placed in front of the screen plane or behind it.

We statistically analyzed a few of the affects on stereoscopic depth accuracy using a Student's T-test. As we expected, the addition of binocular disparity during subject testing decreased the error by an average of 54% ($p < 0.001$), a significant improvement. This can be seen in Figure 3 by the significant decrease in error between the monoscopic and the stereoscopic scenarios.

The most interesting insight we gained dealt with the improvements in accuracy gained by camera sway in stereoscopic viewing conditions. A visual comparison of the three charts in Figure 3 will reveal the accuracy gains made under this condition as opposed to both static and trucking stereoscopic camera conditions. It is interesting that swaying camera movement yielded more error in monoscopy, but increased accuracy in stereoscopy. In the static object scenarios, the swaying camera movement caused an 18% decrease in error as compared with the static camera stereoscopic viewing conditions ($p = 0.08$). Moreover, in figure 4 a 35% decrease in error can be seen when comparing the stereoscopic static camera, laterally moving object configuration and the stereoscopic swaying camera, static object scenarios ($p = 0.10$). This suggests to us that subtle camera movements to the left or right can reveal a great deal to the viewer about depth in a scene. This most likely comes from the added parallax induced by the laterally moving camera. On the other hand, in the static object scenarios, the trucking camera motion yielded 52% more error than the static camera ($p = 0.08$) which indicates that this type of camera movement should typically be avoided if accurate depth perception is desired.

## 4. FUTURE WORK

In retrospect, we realize that our target sphere without shading might have made the task more difficult than if we had simply replaced it with a flat disc. Additionally, the depth picking interface required excessive focus and fusion adjustments due to the constant alternation between looking at a 2D image on the screen plane and looking at the scene in perspective. Also, using a test (such as the Titmus Stereo Test) to determine if stereopsis is present in our subjects would have been helpful in the subject screening process, and we plan to use such a test in the future.

In future investigations, we also plan to randomize the order of tests to avoid sequencing effects, particularly with respect to subject learning during the testing process. In addition, we would like to record the response time of each test so that we can gauge how easy or difficult the subjects are finding the task at hand. We also intend to use the region of interest method of Froner and Holliman, 2005,[14] so that we avoid the problems we noticed in these experiments with lack of depth resolution behind the screen, as seen in Figure 5.

The p-values resulting from our statistical analysis provide a satisfactory level of statistical significance given our small sample size (12 subjects), however, in future experiments we would like to increase our number of test subjects, especially when investigating a similar large number of viewing conditions.

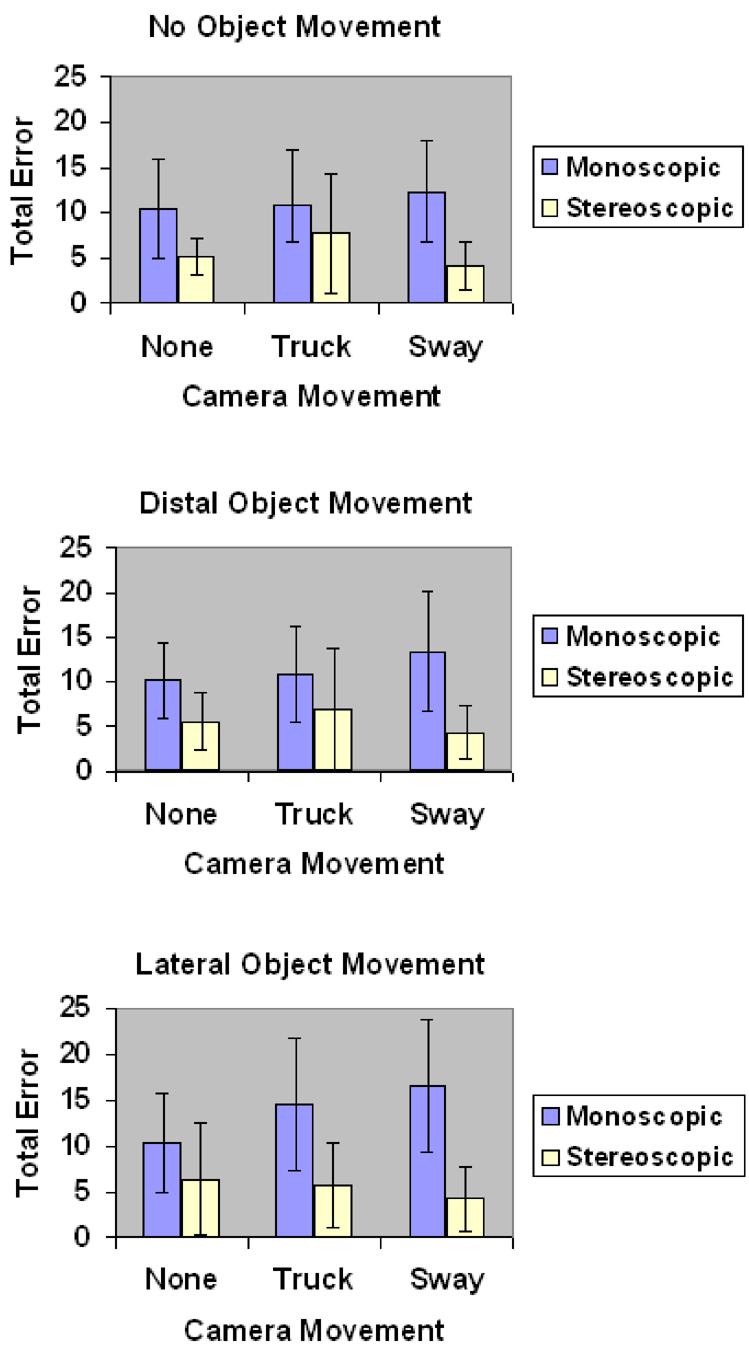


Figure 3. Comparison of the three camera configurations over all 9 camera-target movement combinations

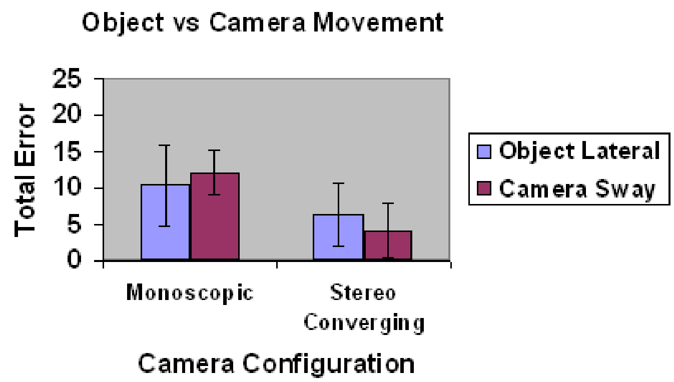


Figure 4. Comparison of scenarios with lateral camera and target movement

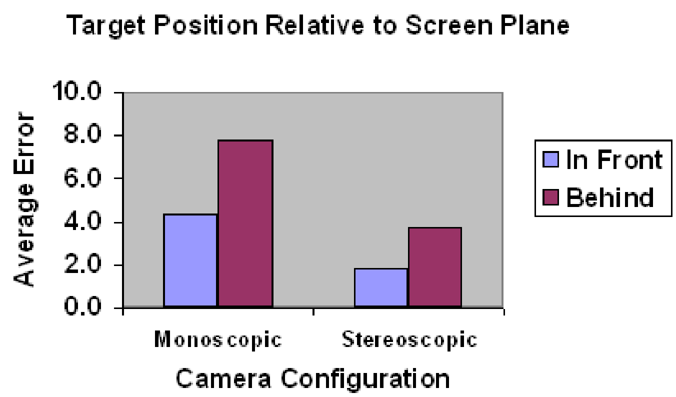


Figure 5. Comparison of error based on target's position relative to the screen plane

We plan to investigate standard methods of guiding the viewers' eyes on the screen and the way that those methods can be enhanced or degraded in a stereoscopic theater environment. We are also interested in investigating standard techniques for layout and scene composition as well as shot selection methods. Furthermore, the viewer in a theater may be seated at a variety of locations relative to the screen, and the effects of this seating position on depth perception deserve empirical study as well.

## 5. CONCLUSION

Our findings seem to suggest that filmmakers can take advantage of the significant gain in depth information offered by swaying camera motion when establishing a scene. These swaying camera movements can be very slight, only a degree or two of rotation about a point is necessary to provide the parallax which makes this technique effective. However further study is needed with a larger subject group in order to establish conclusive statistical significance. Having laid the groundwork for further research in this area, we hope to gain other helpful insights into the use of stereoscopy in cinematographic synthetic imagery in future studies.

## ACKNOWLEDGMENTS

We are grateful to Alan Price for allowing us to use the stereoscopic display equipment at ACCAD and for his numerous helpful suggestions. We would also like to thank Carter McClung for his helpful advice pertaining to our statistical analysis.

## REFERENCES


[1] Woods, A., Docherty, T., and Koch, R., "Image distortions in stereoscopic video systems," *Proc. SPIE* **1915**, 36–48 (February 1993).

[2] Harris, J., "Evaluating accuracy and precision in a stereoscopic display: perception of 3d object motion," *Stereoscopic Displays and Virtual Reality Systems X* **5006**(1), 292–303, SPIE (2003).

[3] Kleiber, M. and Winkelholz, C., "Distortion of depth perception in virtual environments using stereoscopic displays: quantitative assessment and corrective measures," *Stereoscopic Displays and Applications XIX* **6803**(1), SPIE (2008).

[4] Jin, E. W., Miller, M. E., Endrikhovski, S., and Cerosaletti, C. D., "Creating a comfortable stereoscopic viewing experience: effects of viewing distance and field of view on fusional range," *Stereoscopic Displays and Virtual Reality Systems XII* **5664**(1), 10–21, SPIE (2005).

[5] Speranza, F., Tam, W. J., Renaud, R., and Hur, N., "Effect of disparity and motion on visual comfort of stereoscopic images," *Stereoscopic Displays and Virtual Reality Systems XIII* **6055**(1), 94–103, SPIE (2006).

[6] Lambooij, M. T. M., Ijsselsteijn, W. A., and Heynderickx, I., "Visual discomfort in stereoscopic displays: a review," *Stereoscopic Displays and Virtual Reality Systems XIV* **6490**(1), SPIE (2007).

[7] Rupkalvis, J. A., "Human considerations in stereoscopic displays," *Stereoscopic Displays and Virtual Reality Systems VIII* **4297**(1), 268–275, SPIE (2001).

[8] Kelly, J. W., Beall, A. C., Loomis, J. M., Smith, R. S., and Macuga, K. L., "Simultaneous measurement of steering performance and perceived heading on a curving path," *ACM Trans. Appl. Percept.* **3**(2), 83–94 (l 06).

[9] Frenz, H., Lappe, M., Kolesnik, M., and Bührmann, T., "Estimation of travel distance from visual motion in virtual environments," *ACM Trans. Appl. Percept.* **4**(1), 3+ (2007).

[10] Hubona, G. S., Wheeler, P. N., Shirah, G. W., and Brandt, M., "The relative contributions of stereo, lighting, and background scenes in promoting 3d depth visualization," *ACM Trans. Comput.-Hum. Interact.* **6**(3), 214–242 (1999).

[11] Bayas, L. P., "Human factors involved in perception and action in a natural stereoscopic world: an up-to-date review with guidelines for stereoscopic displays and stereoscopic virtual reality (vr)," *Stereoscopic Displays and Virtual Reality Systems VIII* **4297**(1), 251–267, SPIE (2001).

[12] Swanston, M. T. and Gogel, W. C., "Perceived size and motion in depth from optical expansion," *Perception & psychophysics* **39**, 309–326 (1986).

[13] Lipton, L., [*Foundations of the Stereoscopic Cinema: A Study in Depth*], Van Nostrand Reinhold Company, New York (1982).

[14] Froner, B. and Holliman, N. S., "Implementing an improved stereoscopic camera model," *Eurographics Theory and Practice of Computer Graphics* (June 2005).