

Implementation of pseudo-convergence angle in eye tracking for CG-based lenticular naked-eye 3D displays

Yoshihiro Sato, Mirika Yamashiro; Tokyo City University; Yokohama, Kanagawa/Japan

Abstract

Autostereoscopic 3D displays often suffer from inadequate depth perception and visual discomfort due to fixed convergence angles and the vergence–accommodation conflict. This study proposes a novel lenticular-based naked-eye 3D display system that enhances depth perception by dynamically adjusting parallax images using the observer’s real-time gaze data. By generating a pseudo-convergence angle aligned with the viewer’s point of interest, the system aims to provide more accurate and comfortable 3D visualization. An experiment was conducted to evaluate the effectiveness of the system by comparing it with a conventional (fixed-parallax) 3D display and a standard 2D display. Participants judged the depth order of eight spheres in a computer-generated scene. Results showed that the proposed method significantly improved depth perception accuracy, reducing the mean absolute error by approximately 50% compared with the 2D display and by 20% compared with the conventional 3D display. These findings demonstrate that gaze-contingent parallax adjustment is a promising approach for mitigating the inherent limitations of autostereoscopic displays and improving the user experience.

Introduction

The advancement of display technology has significantly contributed to the practical application of 3D displays, which provide users with an enhanced sense of realism and immersion. Among these, autostereoscopic (glasses-free) 3D displays, which enable stereoscopic vision by presenting different images to the left and right eyes without special eyewear, have attracted considerable research and commercial interest [1]. These displays eliminate the need for cumbersome glasses and allow multiple viewers to experience 3D content simultaneously, making them particularly suitable for public installations and digital signage applications. However, the optical principles underlying these displays, such as those based on lenticular lenses or parallax barriers, require the use of low-parallax images to prevent crosstalk during image transitions and to ensure natural stereoscopic vision [2]. Crosstalk refers to the leakage of the left-eye image into the right eye’s view and vice versa, causing ghosting artifacts and degrading the perceived quality of the 3D image. Although stereoscopic vision with high-parallax images can be achieved through parallel or cross-eyed viewing from a fixed position, this approach causes significant discomfort during natural viewing. Consequently, the achievable depth in autostereoscopic displays is fundamentally constrained by the requirement for low parallax. Depth perception in these displays largely depends on **motion parallax** induced by viewer head and eye movements, which cause shifts in the presented parallax images [2, 3].

In recent years, autostereoscopic 3D displays have been widely deployed in digital signage, with content predominantly consisting of parallax images generated in computer graphics (CG) environments [4]. A critical issue in this context is **physiological diplopia**: when viewers observe regions of the image located in front of or behind the binocular fusion point, the parallax images

appear doubled. This phenomenon occurs because the human visual system expects the vergence angle—the inward rotation of both eyes to fixate on a point—to correspond to the accommodative distance at which the eyes are focused. In natural vision, these two responses are tightly coupled through the vergence–accommodation reflex. Under natural viewing conditions, stereoscopic vision is achieved within the fixated region, whereas the peripheral field is perceived without full stereoscopic integration. When gaze shifts from one region to another, the eyes or head must move to adjust the convergence angle to the new fixation point [5]. In autostereoscopic displays, this function is replicated by the virtual camera used to generate the parallax images. However, because the virtual camera is typically fixed, it cannot adjust the convergence angle in response to the viewer’s gaze, resulting in a conflict between focal distance and convergence angle that causes physiological diplopia.

To address these limitations, this study aims to develop an autostereoscopic 3D display system that enhances 3D perception by aligning the parallax images with the viewer’s gaze and the depth of the object of interest. By leveraging real-time eye tracking, the proposed system dynamically generates a pseudo-convergence angle corresponding to the observer’s current fixation point, thereby reducing the mismatch between the displayed parallax and the viewer’s natural vergence response. This approach is expected to improve both depth perception accuracy and visual comfort.

Proposed Method

A 3D display system was constructed that switches the displayed parallax images according to the region or object on which the observer is currently fixating. In conventional 3D display systems, the projected parallax images are generated using a virtual camera with a fixed focal length (e.g., 3.5 m). This setting is typically optimized for a specific viewing position and distance to minimize occlusion artifacts in autostereoscopic displays. However, this constraint limits the depth representation capability of the display.

In the proposed method, the focal length parameter of the virtual camera is adjusted in real time to match the depth of the object on which the observer is fixating, as determined by an eye tracker. Eye tracking is used to identify the observer’s gaze point on the 3D display, and this information is fed back into the Unity rendering pipeline to update the virtual camera parameters on a per-frame basis. This enables the observer to perceive multiple depth layers simply by shifting gaze, without requiring any special operations. Because the object at the center of the observer’s gaze is always rendered with overlapping parallax, visibility is improved even in the presence of crosstalk—a common limitation of autostereoscopic displays. Although the depth information of the fixated object itself does not change because its parallax is minimized, the spatial relationships among surrounding objects are expected to become more clearly perceptible. Figure 1 illustrates the conceptual difference in dynamic convergence angles between the conventional and proposed methods.

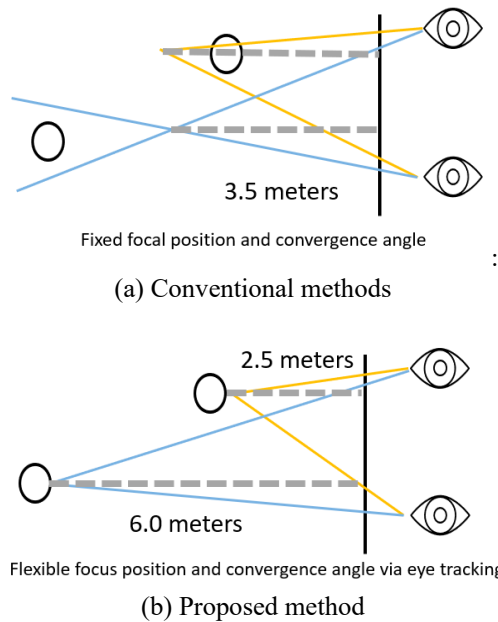


Figure 1. Conceptual diagram illustrating the difference in convergence angles between (a) the conventional method (fixed focal length) and (b) the proposed gaze-contingent method (variable focal length)

Experiments

System Configuration

To construct the proposed 3D display system, the following equipment was used: a 65-inch 4K (3840 × 2160, 30 fps) lenticular 3D display, a Tobii Pro Spark eye tracker, and Unity for generating and rendering nine-view parallax images. Nine virtual cameras were arranged along a single axis with a spacing of 0.014 m, covering a horizontal parallax range suitable for comfortable stereoscopic viewing at the target observation distances. The eye tracker was mounted below the display and calibrated for each subject before the experiment. Gaze data were streamed to the Unity application at 60 Hz, and the virtual camera focal length was updated accordingly for each rendered frame.

- 65-inch 4K (3840 × 2160, 30 fps) lenticular 3D display
- Eye tracker: Tobii Pro Spark
- CG software: Unity (used to generate and render parallax images)

Figure 2 illustrates the constructed display system.

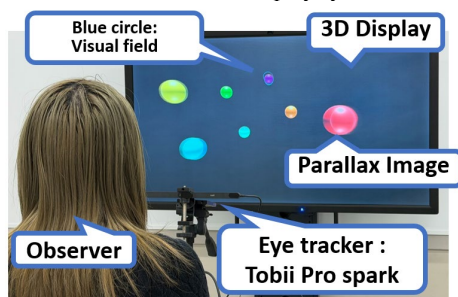


Figure 2. Implementation of the proposed system. Eye tracking shows the observer's gaze as a blue circle on the display. When the observer fixates on an object, the focal length is adjusted to that object, and the parallax image updates in real time

Experimental Design

To evaluate the proposed 3D display system, a comparative observation study was conducted with eight subjects in their 20s and 30s with normal or corrected-to-normal vision. The order of display conditions was counterbalanced across subjects using a Latin square design to prevent order effects and learning biases.

The comparison included three display systems:

1. 2D Display: No depth information.
2. Standard 3D Display: A conventional system capable of conveying depth.
3. Tracking 3D Display: The proposed method.

Nine parallax images were generated in Unity using nine virtual cameras arranged along a single axis with a spacing of 0.014 m. The Unity scene consisted of eight colored spheres placed at eight distances (1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 5.0, and 6.0 m) to allow observation of their spatial relationships. To evaluate not only stereoscopic depth but also apparent depth cues, two types of scenes were created. In the **Random Scene**, sphere radii at each distance were randomized, making it difficult to perceive depth based solely on sphere size. In the 2D condition, this was expected to result in depth judgments based primarily on size. In the **Size Scene**, sphere radii were adjusted according to distance so that all spheres appeared visually identical in size. In this scene, depth perception based on sphere size is difficult, and in the 2D condition, no spatial relationship information is available. These scenes are referred to as the Random Experiment and the Size Experiment, respectively. For each scene type, three variations were created by changing the color arrangement and the left-right and up-down positions of the spheres, resulting in 18 trials per subject (3 conditions × 2 scene types × 3 variations). Figure 3 shows the generated experimental parallax images.

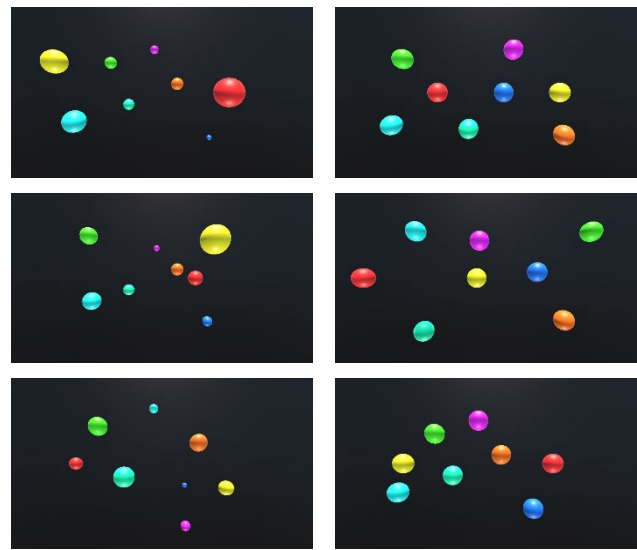


Figure 3. Parallax images used in the experiment. (a) Random Experiment: spheres with random depth positions and sizes. (b) Size Experiment: sphere sizes adjusted according to depth position

To assess each subject's stereopsis and prevent reversed depth perception, a parallax image displaying only two spheres of fixed size was first presented, making their spatial relationship easy to

discern. During this step, subjects also checked for the presence of crosstalk. Next, subjects observed the previously generated scenes and were asked to report the spatial relationships of the displayed spheres using numbers. Each sphere was numbered from 1 to 8, corresponding to depth from back to front, allowing numerical analysis of spatial perception by calculating the absolute difference between the correct number and the subject's response. The evaluation metric for this experiment was the **mean absolute error (MAE)** for each display method, aggregated across all subjects. To prevent order effects, the sequence of display systems was randomized for each subject.

Results

Each subject was presented with a total of 18 displays (3 display methods \times 2 scene types \times 3 variations), and responses regarding the positional relationships of the spheres were collected. The MAE was calculated as the average of the absolute differences between the correct sphere order and the order reported by each subject. The MAE for each display method is shown in Figure 4.

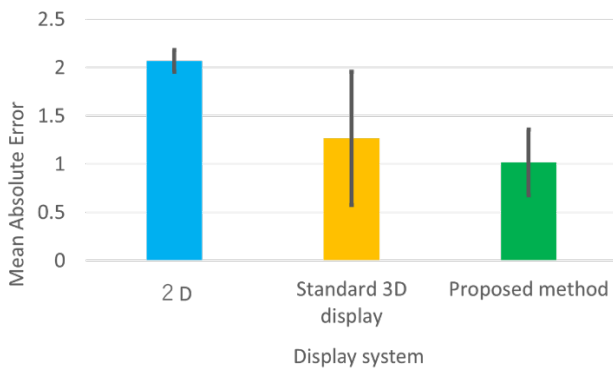


Figure 4. Overall mean absolute error (MAE \pm SD) for each display condition

From Figure 4, the 2D display exhibited the highest error (MAE = 2.06, SD = 1.73). Interestingly, its standard deviation was the lowest among the three conditions, indicating that while it was difficult to select the correct order in the 2D condition, subjects' responses were consistently similar. This suggests minimal inter-subject variation in depth perception derived from cues other than stereopsis and sphere size, such as sphere color or specular highlights from CG lighting. The proposed Tracking 3D method produced the lowest mean absolute error (MAE = 1.02, SD = 1.20), representing a **49% reduction** compared with the 2D display and a **20% reduction** relative to the Standard 3D display (MAE = 1.27, SD = 1.64). The improved depth information enabled by eye tracking is likely the primary factor contributing to this enhancement. The Standard 3D display showed lower error than the 2D display, but its standard deviation was the largest, indicating variability in stereopsis among subjects.

To further evaluate the improvement in depth perception due to fixed versus variable focal lengths during parallax image generation, the absolute error between the correct and reported distances was calculated for each sphere position. Figure 5 shows the MAE and standard deviation for each display method as a function of placement distance.

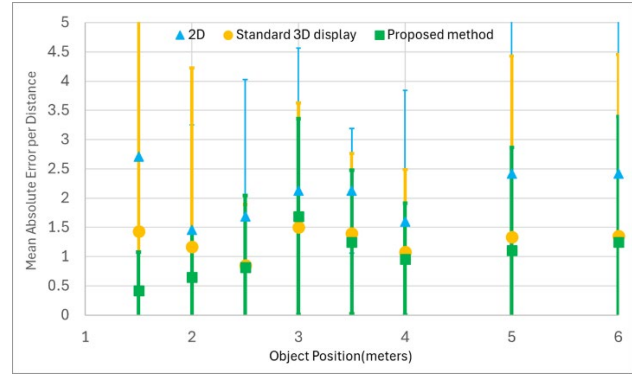


Figure 5. Mean absolute error (MAE) by placement distance for each display condition

Both the Standard 3D display and the proposed display system achieved high accuracy in depth perception at far distances (≥ 4.0 m), with absolute errors of approximately 1. The proposed method also significantly reduced depth errors at near distances (1.5 m: MAE = 0.42 compared with 1.44 for the Standard 3D and 2.70 for the 2D display). Furthermore, the standard deviation of the depth error at near distances was clearly smaller for the proposed method than for the conventional method, indicating more consistent depth perception at close range. These results suggest that dynamic focal length adjustment is particularly effective at near distances, where the difference in convergence angle between the fixated object and the fixed virtual camera is most pronounced.

To evaluate depth perception accuracy based on scene type (Random Scene vs. Size Scene), results were analyzed by scene type and placement distance. Figure 6 shows the MAE for each display method as a function of scene type and placement distance. In the 2D display, errors were substantially higher than in the 3D displays, with the largest errors occurring in the Size Scene. This is likely because, in the Size Scene, all spheres appeared visually identical, making it difficult to distinguish distances and resulting in the highest error.

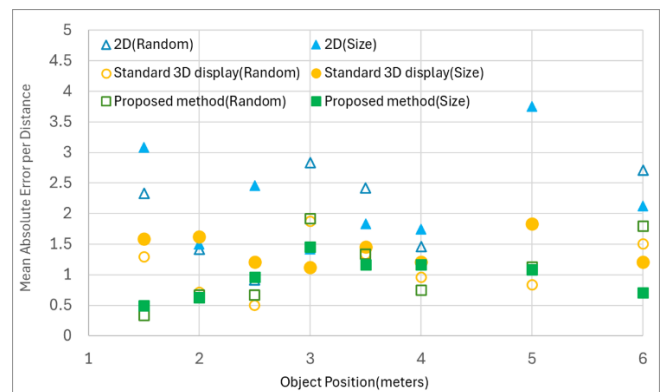


Figure 6. Mean absolute error (MAE) by placement distance, separated by scene type (Random vs. Size)

No significant difference in accuracy was observed between scene types for either the Standard or Tracking 3D displays. This is likely because depth perception in these systems relied on stereoscopic cues rather than sphere size or color.

Discussion

The experimental results demonstrate that dynamically adjusting parallax based on the observer's gaze significantly improves depth perception accuracy in the proposed display system. The 49% reduction in error compared with the 2D display highlights the substantial benefit of stereoscopic information, while the additional 20% reduction compared with the conventional Standard 3D display suggests that the proposed system has potential for broader applications. This improvement is particularly pronounced at near distances, where gaze-contingent focal length adjustment most effectively compensates for the limitations of a fixed virtual camera configuration. An interesting observation is that the uniform perception of CG images in the 2D display condition showed little variability, whereas introducing depth information via autostereoscopic 3D displays resulted in greater inter-subject differences in recognition error. This suggests that current autostereoscopic 3D displays used in digital signage may produce variable stereopsis among viewers, potentially making them less suitable for general audiences compared with 2D displays in advertising contexts. Although the proposed method relies on eye tracking, it reduces variability in depth perception compared with the conventional approach, moving closer to a consistent and uniform presentation of depth.

Some limitations of the proposed method should be noted. Although dynamic parallax switching is effective, it may still induce eye strain or cognitive fatigue during prolonged viewing, as has been observed with conventional methods. This experiment did not measure these effects, and further investigation is needed, as dynamic shifts in the line of sight could increase visual fatigue. The current method addresses only the convergence component of the vergence–accommodation conflict (VAC) but cannot adjust the focal distance to the display surface. The VAC arises because the eyes must focus on the fixed display plane [5], which can cause visual fatigue, discomfort, and distorted depth perception. Future research could address the VAC more comprehensively by combining gaze-contingent parallax adjustment with variable-focus display technologies, such as focus-tunable lenses or light field displays.

Conclusion

This study developed and evaluated a gaze-contingent naked-eye 3D display system that dynamically adjusts parallax to align with the observer's point of interest. Experimental results confirmed that this method significantly improves depth perception accuracy compared with both standard 2D and conventional fixed-parallax

3D displays. The ability to perceive depth across multiple planes simply by shifting gaze is a key advantage of this approach. This work represents a promising step toward more practical and comfortable autostereoscopic displays that overcome the limitations of traditional systems.

References

- [1] F. Chen, C. Qiu, and Z. Liu, "Investigation of Autostereoscopic Displays Based on Various Display Technologies," *Nanomaterials*, vol. 12, no. 3, 429, 2022.
- [2] D. Kang, J.-H. Choi, and H. Hwang, "Autostereoscopic 3D Display System for 3D Medical Images," *Appl. Sci.*, vol. 12, no. 9, 4288, 2022.
- [3] T. Huang, B. Han, X. Zhang, and H. Liao, "High-performance autostereoscopic display based on the lenticular tracking method," *Opt. Express*, vol. 27, pp. 20421-20434, 2019.
- [4] J. Hua, W. Qiao, and L. Chen, "Recent advances in planar optics-based glasses-free 3D displays," *Front. Nanotechnol.*, vol. 4, no. 829011, 2022.
- [5] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence-accommodation conflicts hinder visual performance and cause visual fatigue," *Jour. Vision*, vol. 8, no. 3, pp. 1-30, 2008.

Acknowledgments

This study was supported by the Nakajima Foundation's Grant for the Promotion of Independent Japanese Researchers (FY2024).

Author Biography

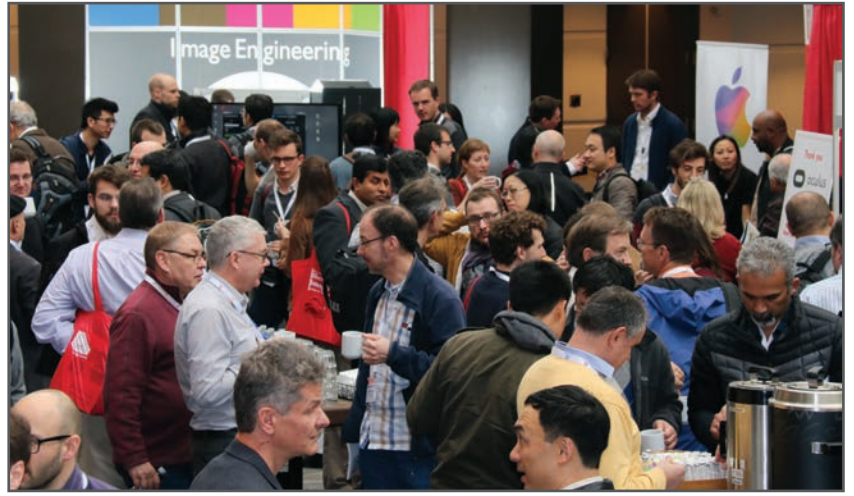
Yoshihiro Sato received his Ph.D. in computer science from Tokyo City University, Tokyo, Japan, in 2022. Since 2023, he has been a lecturer in the Department of Design and Data Science at Tokyo City University. His current research focuses on 3D data processing, 3D vision, and texture structure analysis.

Mirika Yamashiro is an undergraduate student in the Department of Design and Data Science at Tokyo City University. She is currently engaged in the development and evaluation of 3D display systems in the Sato Laboratory. Her research interests include natural stereoscopic vision and immersive experiences.

JOIN US AT THE NEXT EI!

electronic IMAGING

Imaging across applications . . . Where industry and academia meet!



- **SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •**
- **INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •**

www.electronicimaging.org

