Computational Investigation of the Illusory Depth Effect of Luminance Contrast on a Three-Dimensional Architectural Scene with Binocular Vision

Nan-Ching Tai* 1

1 Assistant Professor, Department of Interaction Design, National Taipei University of Technology, Taiwan

Abstract
One of the objectives of architectural design is to enrich the spatial experience within the restrictive confines of the architectural space. Various depth cues such as familiar size and linear perspective have been utilized in architectural design to create illusory spatial depth. Luminance contrast has also been identified as an effective depth cue; its effect in increasing depth perception was verified with perceptual studies using a computer-generated environment. However, previous studies have focused on methods for accurate simulation of the lighting distribution in a scene; they have not addressed visual perception with binocular vision. This study utilizes three-dimensional (3D) stereo display technologies to present a perceptually realistic and computer-generated environment. Psychophysical experiments were conducted to examine the impact of luminance contrast in creating illusory depth effects, with scenes presented on both conventional and 3D stereo displays. The objectives of this study are: first, to investigate the influence of binocular vision on the illusory depth effect of luminance contrast; and second, to verify whether a reliable computer-generated environment can be used to envision possible applications of luminance contrast in architectural design.

Keywords: luminance contrast; binocular vision; space perception; physically based lighting simulation; computer-aided design

1. Introduction
Enrichment of the spatial experience of an architectural space beyond its structural boundary has been one of the objectives of architectural design. Depth cues that contribute to our perception of three-dimensional (3D) spatial layouts have been utilized as design strategies to enhance the spatial dimensions of architectural scenes. The apse of the church of San Satiro in Milan was painted using false perspective to create an illusion of a deeper space (Solso, 1996). In Rome, the sizes and layout of architectural elements in the gallery in Palazzo Spada were distorted to create a forced perspective, which exaggerates the perceived distance of the sculpture located at the end of the gallery (Blunt, 1989).

Lighting was also observed to be an effective element in altering the spatial perception of an architectural scene. The effect of perceived brightness on depth perception has been studied extensively and shown to effectively influence perceptual judgment of the figure-ground spatial relationship (Egusa, 1982). From perceptual studies, O'Shea et al. (1994) have proven that the contrast of lightness and darkness can create illusory depth effects on planar surfaces.

In order to examine the manner in which luminance contrast can create illusory depth effects in 3D spaces, Tai (2012) conducted perceptual studies using a computer-generated pictorial environment that allows precise parametric controls. The causal relationship between the architectural configuration, the lighting distribution in the scene, and the perceived distance of a visual target in the scene was established (Tai, 2013). The computer-generated pictorial environment incorporated physically based lighting simulation and perceptually based tone mapping techniques, which provide the visual realism that closely matches the physical quality of the light distribution in a real scene and reflects the perceptual reality of apparent lighting distribution to the eye. However, the output image of this computational framework is derived from a single-camera viewpoint. Hence, the study can only, at best, account for monocular vision, and it does not address human visual perception from the perspective of binocular vision.

This study incorporates stereo display technologies into the computational framework that uses a computer-generated and perceptually realistic environment. To investigate the effect of luminance contrast on creating illusory depth, psychophysical experiments...
were conducted using this pictorial environment. The objectives were: (1) to examine the influence of binocular vision on the depth effects due to luminance contrast, and (2) to further enhance the visual realism of the alternate pictorial environment, thus enabling it to be used for envisioning possible applications of luminance contrast in architectural design.

2. Representation of Binocular Vision

Binocular vision refers to the visual perception of spatial layout when two retinal images, one from each eye, are combined. Our eyes are separated by about 6 to 6.5 cm, and hence, the retinal image that each eye perceives is slightly different. This difference, known as binocular disparity, allows us to retrieve information about the 3D relationship of an object with a space (Palmer, 1999). Representations of binocular vision have been developed for many years. The stereoscope, discovered as early as the mid-19th century, is a device that deliberately projects a pair of unique and slightly different images of the same scene to the left and right eyes in order to create a stereo viewing experience (Welling, 1978). A concept similar to the stereoscope was developed more recently for digital representation. Stereo display technologies use different methods to deliver the images made for the right eye and those made for the left yet to the corresponding eye. In stereoscopic display technology, the end users must wear special glasses that may feature various optical technologies such as shutters, circular polarization, and/or filtered color in order to distinguish the source images. Autostereoscopic display technology, on the other hand, relies only on the display device to project a specific image to each eye. However, this process initially requires detection of the face and location of the viewer by the device, and then, depends on applied technologies such as lenticular lenses or parallax barriers to project the source image to a particular eye location (Lueder, 2012).

The experimental scenes in this study were simulated with RADIANCE, a physically based lighting simulation program, which creates high dynamic range (HDR) images as output (Ward and Shakespeare, 1998). It has been validated that the RADIANCE-simulated HDR scene accurately encompasses the luminance distribution data of the simulated scene (Mardaljevic, 2001; Ruppertsberg and Bloj, 2006). The photographic tone mapping operator was then used to process the HDR scenes into a JPG image having a dynamic range that can be shown on conventional display devices (Reinhard et al., 2002). Various perceptual studies have verified that the photographic tone mapping operator performs well in many perceptual aspects (Cadik et al., 2008). Thus, this computational framework can generate images that reflect perceptual reality for a single retinal image. In order to incorporate the stereo visual experience in the perceptual realistic representation, in addition to a single camera setting, two additional cameras — each being offset by 3 cm laterally from the original single camera — were set up to simulate the viewpoints of the left and right eye. Three sets of experimental scenes were generated with these three viewpoints. The output set of the single-camera setup was designated as the conventional single-viewpoint experiment set. The pair of images generated by the two additional cameras was processed into a stereo pair and used to generate the experimental scene sets for stereoscopic and autostereoscopic 3D displays.

3. Experimental Setup and Method

Fig. 1. illustrates the experimental setup. A corridor was configured with four modular spaces, each having dimensions of $6 \times 6 \times 4$ m and a $2.5 \times 2.5$ m opening at the center of the ceiling. Each opening could be set up as open, half-closed, or closed to create different interior lighting distributions based on the amount of light permitted into the space. Camera M (the single-camera viewpoint) was placed at one end of the corridor, 1.5 m above the ground, representing the eye level of the viewer. Cameras L and R were located 3 cm to the left and 3 cm to the right, respectively, of Camera M. Each camera was focused on the center of the visual target, a red sphere of diameter 60 cm, floating 1.6 m above the ground. The four openings were configured to be half-closed, open, closed, and half-closed, respectively, to create an interior lighting distribution designated as the "Contrast" condition. The four openings were all configured to be open to create the "Even" condition. In the "Contrast" condition, the luminance contrast of the visual target with the foreground is greater than that of the visual target with its background. In the "Even" condition, the luminance contrast of the visual target with the foreground is equal to that of the visual target with the background.
The experimental scenes were generated using RADIANCE, with all parameters (including location, sky condition, and material properties) maintained constant. Later, the HDR output images were tone mapped with the photographic tone mapping operator and converted into the JPG format. The JPG images produced by Camera M were used to create the single-viewpoint set. For the stereoscopic 3D display, the color filtering method was used. An image-editing program was applied to merge the stereo pair images into an Anaglyph 3D image, which can be viewed stereoscopically with red/cyan glasses.

For the autostereoscopic experimental scene, the stereo pair images were processed and placed adjacent to each other in a single image, and then saved in the JPEG Stereoscopic (JPS) format. When viewed with the autostereoscopic display device, the JPS image will automatically split in half to project the images specific to each eye, thus creating a stereo viewing experience.

In this study, psychophysical experiments were conducted to investigate the quantitative relationships of the binocular disparity and illusory depth effect resulting from the luminance contrast. Psychophysics, a scientific discipline in Psychology, is a research method that was developed to measure the perceptual sensitivity in response to environmental stimulus in the physical domain (Gescheider, 1984).

The experimental setup and method in this study uses a $2 \times 4$ factorial design, testing two independent variables: luminance contrast and viewing conditions. The variable, luminance contrast, has two levels, while viewing conditions has four. Further, the perceived distances of the visual targets in the test scenes are the measured perceptual sensitivity. Fig.2 illustrates the two test scenes rendered under the "Contrast" and "Even" lighting distribution conditions, with the visual target located 15 m from the viewpoint.

Depending upon the intention and the design of the psychophysical experiments, the perceptual sensitivity can be measured using various methods such as free description tasks, rating, and forced choice (Cunningham and Wallraven, 2011). Rating is the most commonly used method in perceptual studies related to depth perception. In this method, the subject must adjust a visual target in the reference scenes in order to match the perceived distance of the same visual target in the test scenes. However, when the reduced-cue condition arises in the presented experiment scene, such as in the case of a two-dimensional (2D) display, the subject may tend to be influenced by the 2D features instead of the 3D nature. Hence, in the case of the reduced-cue condition, the forced choice method has been used. In this method, subjects provide only binary responses, such as left or right, and near or far. The advantage of this method is that it takes significantly less time for the subject to make the perceptual judgment, and thus, a more intuitive measurement can be made.

The constant stimuli method was used to measure the perceived distances of the visual target under different conditions. In this method, the standard stimulus and the comparison stimulus are presented to the subject simultaneously. The controlled values in the standard stimulus are maintained constant, while the controlled values in the comparison stimulus typically have seven, nine, or eleven values, and the difference between consecutive values is the same. The greatest values must be perceived, almost always, to be greater than the standard value, and the smallest values must be perceived, almost always, to be smaller than the standard value. A pair of standard stimulus and comparison stimulus values are presented to the subject in random order. The subjects must specify which stimulus produces a greater magnitude of sensation (Gescheider, 1984).

Based on the principle of the constant stimuli method, seven reference scenes were rendered with the "Even" condition. The visual target was placed at distances of 12 m, 13 m, 14 m, 15 m, 16 m, 17 m, and 18 m from the viewpoint because visual inspection of the targets situated at distances of 12 m and 18 m in the reference scene can almost always be perceived to be nearer and farther away, respectively, than the target situated at a distance of 15 m in the test scene.

Subjects were shown a pair of reference and test scenes, and were instructed to specify the visual target that was perceived to be closer. In order to prevent the test scene from being located only on a single side, the location of test and reference scenes were swapped to create another set, as illustrated in Fig.3. The 14 pairs of test and reference scenes were presented to subjects 5 times in random order. Thus, for each viewing condition, a subject made 10 perceptual judgments for each test scene under "Even" and "Contrast" lighting conditions using one of the seven reference scenes, resulting in 140 trials.

The experimental procedure was repeated for four viewing conditions: Monocular 2D, where subjects use one eye (the dominant eye) to view the experimental

![Fig.2. Test Scenes with Different Display Methods](image-url)
scenes of the single-viewpoint set; Binocular 2D, where subjects use both eyes to view the same single-viewpoint experimental scenes; Anaglyph 3D, where subjects used both eyes and wore red/cyan glasses to view the Anaglyph 3D set; and JPS 3D, where subjects used both eyes to view the autostereoscopic experimental scenes. The three sets of experimental scenes (single-viewpoint, Anaglyph 3D, and autostereoscopic) were displayed on a Toshiba Satellite P850 laptop, a single device capable of displaying these types of experimental scenes.

In a pilot study, experiments were performed by following the sequence of Monocular 2D, Binocular 2D, Anaglyph 3D, and JPS 3D. The results show that luminance contrast affects the judgment of the perceived distance of the visual target under stereo viewing conditions, and that subjects tend to be more reliable under the JPS 3D condition (Tai, 2014). However, the pilot study did not consider the order effects, in which the increased reliability can be attributed to the experience gained from earlier experiments.

As the primary responsibility of the subjects was limited to visual perception and interpretation of the pictorial depth cue, the only factor used for selecting subjects was normal or corrected-to-normal vision. Ten subjects, with an age range of 20–42 years, were invited to participate in the experiment.

Experiments were performed at a campus research lab. The subjects were briefed about the experimental procedure; then, experiments were conducted with the initial sequence of Monocular 2D, Binocular 2D, Anaglyph 3D, and JPS 3D, followed by the reverse sequence of JPS 3D, Anaglyph 3D, Binocular 2D, and Monocular 2D. The subjects took short breaks between the two sequences, and were also encouraged to take breaks whenever they experienced eye fatigue.

4. Results

The constant stimuli method is based on determining the point of subjective equality (PSE) of the standard stimulus within the range of the comparison stimulus. The probit regression model can be used to derive the psychometric functions that represent the cumulative normal probability distribution of the trial data from all subjects and also to derive the PSE (Finney, 1971). Figs.4.(a)–(d) illustrate the probit analyses of the results for the four viewing conditions. In probit analysis, the data (plotted as a curve) indicates the probability of the test target that is located at a distance of 15 m being perceived to be nearer than the reference targets that are located at a distance of 12 m to 18 m. The 0.5 proportion line that intersects the analysis curve is the point of subjective equality (PSE), which represents the point at which the test target is perceived to be at the same distance as the reference target, and is thus considered as the measured perceived distance of the test target. In Fig.4., points A+, B+, C+, and D+ represent the PSEs for the "Even" condition in the initial trial sequence, and points A-, B-, C-, and D- are the PSEs for the "Even" condition in the reverse sequence. Points A'+, B'+, C'+, and D'+ are the PSEs for the "Contrast" condition in the initial sequence, and points A'-, B'-, C'-, and D'- are the PSEs for the "Contrast" condition in the reverse sequence.
Table 1. compares the PSEs of the two sequences under the four viewing conditions. The results indicate that the order of the sequence does not have a significant impact on the results.

5. Discussion

Fig.5. illustrates the probit analysis of the combined results from the two trial sequences. For the "Even" condition, owing to the lighting distribution in the test scene being identical to that in the reference scenes, the perceived distances of the visual target were measured to be close to 15 m, with PSEs of 14.892±0.053 m, 14.975±0.043 m, 14.855±0.049 m, and 14.960±0.040 m for Monocular 2D, Binocular 2D, Anaglyph 3D, and JPS 3D viewing conditions, respectively. However, when luminance contrast was enabled using the "Contrast" condition, the measured perceived distances of the visual target increased to 15.900±0.051 m, 15.865±0.041 m, 16.148±0.049 m, and 15.890±0.041 m for Monocular 2D, Binocular 2D, Anaglyph 3D, and JPS 3D viewing conditions, respectively.

Table 2. further compares the illusory depth effect of luminance contrast under the four viewing conditions. When the architectural configuration was manipulated such that the luminance contrast of the visual target with the foreground was greater than that of the visual target with the background, the perceived distances of

| Condition          | First Sequence | Reversed Sequence |
|--------------------|----------------|-------------------|
| Monocular 2D (Contrast) | 15.913±0.073 | 15.887±0.071      |
| Monocular 2D (Even)   | 14.964±0.077 | 14.813±0.073      |
| Binocular 2D (Contrast) | 15.926±0.057 | 15.804±0.058      |
| Binocular 2D (Even)   | 14.846±0.062 | 15.111±0.059      |
| Anaglyph 3D (Contrast) | 16.291±0.060 | 16.024±0.075      |
| Anaglyph 3D (Even)    | 14.855±0.068 | 14.857±0.071      |
| JPS 3D (Contrast)     | 15.888±0.057 | 15.891±0.057      |
| JPS 3D (Even)         | 14.873±0.057 | 15.049±0.053      |
Humans can perform better visual inspection when using both eyes than when using a single eye. Hence, the D-threshold of Binocular 2D is less than that of Monocular 2D, as seen in Fig.6. However, the D-threshold for JPS 3D is the least among the four viewing conditions for the "Contrast" and "Even" lighting conditions. This indicates that when binocular disparity is incorporated for stereo display, the perceptual judgment on the depth effect is more reliable. Although Anaglyph 3D also provides binocular visual perception, the filtered color may have an impact on perceptual judgment, especially because the visual target is red, and the color perceived may be distorted by red/cyan glasses. Observations during the experiments show that subjects find it slightly difficult to make perceptual judgments when using red/cyan glasses. These observations are also reflected in the results, where the D-threshold for the Anaglyph 3D condition was significantly higher than the D-threshold for the JPS 3D condition. Therefore, this study concludes that the autostereoscopic display, offering the stereo viewing experience without distorting the color representation, is a more reliable pictorial environment for envisioning the perceptual depth effect of luminance contrast.

### Table 2. Comparison of Percentage Increase in Measured Perceived Distance

| Condition          | Measured Perceived Distance | Increase |
|--------------------|----------------------------|----------|
|                    | Even                        | Contrast | %     |
| Monocular 2D       | 14.892±0.053                | 15.900±0.051 | 6.77% |
| Binocular 2D       | 14.975±0.043                | 15.865±0.041 | 5.94% |
| Anaglyph 3D        | 14.855±0.049                | 16.148±0.049 | 8.70% |
| JPS 3D             | 14.960±0.040                | 15.890±0.041 | 6.22% |

6. Conclusion and Application

The objective of this study was twofold: (1) to investigate the illusory depth effect of luminance contrast on a 3D architectural scene with binocular visual perception; (2) to advance the visual realism of the established computational framework that can generate a virtual environment to study the illusory depth effect of luminance contrast, and to envision the design application of utilizing luminance contrast to enrich the spatial experience.

The goal of simulation-based computer-aided design is to utilize the computer simulation to predict complex effects of the proposed design. Thus, the reliability of the simulation-based design depends on the realism provided by the simulation. However, computer simulation often has different degrees of approximation. The pictorial environment generated from the computational framework of physically based lighting simulation and perceptually based tone mapping was used earlier to repeat a classic experiment of investigating the size-distance relationship conducted in a physical setting (Tai, 2012). The results obtained were similar to those obtained from the classic experiment. Thus, it is validated that the computational framework provides the necessary visual realism of the computer simulation and can be an alternative environment for investigating depth perception related to lighting distribution. In this study, the computational framework was extended to incorporate stereo display technologies. The psychophysical experiments used both conventional and stereo display methods to investigate the illusory depth effect of luminance contrast. The results demonstrate that the computational framework (which incorporates physically based lighting simulation, perceptually based tone mapping, and autostereoscopic display technologies) can provide a simulated environment that allows for a more reliable response when judging the perceptual depth effect of luminance contrast. In this particular pictorial environment,
luminance contrast can increase the perceived distance of a visual target in a scene. Thus, this study concludes that luminance contrast is an effective visual cue that can create the illusory depth effect with binocular vision. Further, it indicates that the illusory depth effect can be effectively envisioned with a computer-simulated environment incorporating autostereoscopic display technology, and that luminance contrast can be applied as a design strategy for enriching spatial experiences.

Figs. 7 and 8 illustrate a possible design application. Fig. 7(a) shows a museum exhibition room illuminated by the central skylight and electrical lighting. In Fig. 7(b), partitions were installed in the same exhibition room to create a more dynamic interior illumination with a bright day lit hallway and dark side corridors. Figs. 8(a), 8(b), and 8(c) illustrate the views when approaching the displayed sculpture through the day lit hallway. Figs. 8(d), 8(e), and 8(f) illustrate the views when approaching the displayed sculpture through one of the side corridors. The lightness perception of a visual target by a human depends upon the surrounding luminance distribution owing to sophisticated visual processing that includes lateral inhibition and chromatic light adaptation (Palmer, 1999). Hence, the same sculptures illuminated with the same spotlights were perceived differently when viewed from the foreground with different luminance distributions. The adjacent illustrations in Fig. 8 show that the sculpture was perceived to be farther away in Figs. 8(a) and 8(b) than in Figs. 8(d) and 8(e) owing to the lower luminance contrast of the sculpture against its background than against its foreground. However, this sense of the exaggerated depth perception is eliminated when viewers are finally standing in front of the sculpture without the dynamic luminance contrast, as seen in Figs. 8(c) and 8(f). This example demonstrates that although the depth effect of luminance contrast on the absolute perceived distance of a visual target is limited, it can be applied to sequential views to enrich the spatial experience. In addition, the perceived lightness contrast is a complex visual process that involves adaptation and local interaction (Palmer, 1999), and hence, this example also demonstrates the importance of a reliable virtual environment to envision lighting distribution in space and depth perception when utilizing luminance contrast as the design strategy.

Acknowledgements
This research was funded by the Ministry of Science and Technology, Taiwan, under the grant No: NSC 101-2221-E-027-146. The author would like to express his appreciation to the Ministry of Science and Technology for their support. The author also wishes to thank the people who participated in the experiments.

References
1) Blunt, A. (1989) Borromini. Belknap Press of Harvard University Press, Cambridge, Massachusetts.
2) Cadik, M., Wimmer, M., Neumann, L., and Artusi, A. (2008) Evaluation of HDR Tone Mapping Methods Using Essential Perceptual Attributes. Computer Graphics. 32, pp.330-349.
3) Cunningham, D.W. and Wallraven, C. (2011) Experimental Design: From User Studies to Psychophysics, 1st edition. A K Peters/CRC Press, Boca Raton, Florida.
4) Egusa, H. (1982) Effect of Brightness on Perceived Distance as a Figure–Ground Phenomenon. Perception 11, pp.671-676.
5) Finney, D. (1971) Probit Analysis, 3rd ed. University Press, Cambridge, Massachusetts.
6) Gescheider, G.A. (1984) Psychophysics: Method, Theory, and Application, 2nd ed. Lawrence Erlbaum.
7) Lueder, E. (2012) 3D Displays. Wiley, Hoboken, New Jersey.
8) Mardaljevic, J. (2001) The BRE-IDMP Dataset: A New Benchmark for the Validation of Illuminance Prediction Techniques. Lighting Research and Technology. 33, pp.11-134.
9) O'Shea, R.P., Blackburn, S.G., and Ono, H. (1994) Contrast as a Depth Cue. Vision Research. 34, pp.1595-1604.
10) Palmer, S.E. (1999) Vision Science: Photons to Phenomenology, 1st ed. The MIT Press, Cambridge, Massachusetts.

11) Reinhard, E., Stark, M., Shirley, P., and Ferwerda, J. (2002) Photographic Tone Reproduction for Digital Images, in: Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques. ACM Transactions on Graphics, San Antonio, Texas, pp.267-276.

12) Ruppertsberg, A.I. and Bloj, M. (2006) Rendering Complex Scenes for Psychophysics Using RADIANCE: How Accurate Can You Get? Journal of the Optical Society of America A. 23, pp.759-768.

13) Solso, R.L. (1996) Cognition and the Visual Arts. The MIT Press.

14) Tai, N.-C. (2012) Space Perception in Real and Pictorial Spaces: Investigation of Size-Related and Tone-Related Pictorial Depth Cues through Computer Simulations. Computer-Aided Design and Applications. 9, pp.235-251.

15) Tai, N.-C. (2013) Application of Luminance Contrast in Architectural Design. Computer-Aided Design and Applications. 10, pp.953-963.

16) Tai, N.-C. (2014) Binocular Disparity and Depth Cue of Luminance Contrast: Investigation of Perceptual Influence of Binocular Disparity on Depth Effect of Luminance Contrast through Stereo Display, in: Proceedings of the 19th International Conference on Computer-Aided Architectural Design Research in Asia (CAADRIA 2014). Kyoto, pp.637-646.

17) Ward, G. and Shakespeare, R. (1998) Rendering with Radiance: The Art and Science of Lighting Visualization. Morgan Kaufmann Publishers.

18) Welling, W. (1978) Photography in America: The Formative Years, 1839-1900 - A Documentary History, illustrated edition. Crowell, New York.