Effect of model scale on predicting illusory stereo depth effect of luminance contrast in real and virtual environments

Nan-Ching Tai *

Abstract: Luminance contrast has been identified as an effective depth cue for creating illusory depth effects in scenes and has been validated as a design strategy for enriching spatial experience. To evaluate the effect of complex lighting in a design proposal, computer simulations and physical models are commonly used in the design process. However, computer simulations must provide perceptual reality, and physical models are often constructed at a reduced scale. Thus, validating the perceptual realism of a computer simulation and the effective range of the depth cue provided by luminance contrast is crucial. In this study, psychophysical experiments are conducted to compare the visual realism offered by computer simulations and physical models according to differences in scale. The results demonstrate that computer simulations can provide the necessary visual realism to create an illusory depth effect, and that the influence of model scale (whether full or reduced) on the depth effect is insignificant.

Subjects: Computer Aided Design (CAD); Computer Graphics & Visualization; Design; Lighting; Simulation & Modeling

Keywords: luminance contrast; illusory spatial perception; visual realism; computer-aided design; high dynamic range imaging; binocular disparity

ABOUT THE AUTHOR
Nan-Ching Tai received his Master of Architecture and PhD. in the Built Environment from the University of Washington. He is currently a faculty member at Department of Interaction Design of National Taipei University of Technology. His teaching and research interests are in the area of visual language. He believes that sketching freehand analytically is a means by which to think and create in the field of design and thus can be considered as a native visual language. To communicate and realize a design idea, he believes computer graphics are an advanced common visual language. He notes that the development of high-dynamic-range-imaging-related technologies continue to advance visual realism. To this end, his research interests focus on developing a computational framework for creating a pictorial environment that can reflect perceptual reality and using this alternative environment to investigate the various issues related to space perception in the field of design.

PUBLIC INTEREST STATEMENT
Many depth cues that help to perceive depth can be design strategies to create a sense of depth in architectural space. Lighting in a scene has also been determined to be an effective depth cue for creating a sense of depth and is therefore may be used as a possible design strategy to enrich spatial experience. This study is part of a series of continuing studies that investigate the manner in which we can create a reliable computer environment to establish a relationship between lighting and its effect on the depth perception of a scene. Moreover, we examine the manner in which we can use the same computer environment to assist in the creation of illusory depth. This is a comparative study that examines closely different model scales and their influence on predicting stereo illusory depth effect resulting from luminance contrast in interior lighting.
1. Introduction
Design is a process in which a designer proposes, analyzes, and refines a solution to achieve specific design goals (Kalay, 2004). In this repetitive process, creativity is essential for developing a design solution. However, evaluating and predicting the effects of a design solution are also important. Different design considerations require different evaluation methods. For conceptual considerations such as structural stability, numerical data may be used for assessment; for perceptual quality, various methods of visualization can assist in the design process.

A mathematical drawing method that constructs a perspective view allows a designer to visualize the spatial experience of a proposed conceptual configuration. Thus, a forced perspective is often used to create the illusion of spatial depth (Pérez-Gómez & Pelletier, 1997; Solso, 2003). The light and lighting distribution in an interior scene can also create a sense of depth. However, the cause-and-effect relationship of scene-based luminance contrast and its impact on depth perception could only be determined when alternative experimental environments were established by recent developments in digital representation.

The effect of depth cues on depth perception is often established through perceptual studies that employ psychophysical experiments. Both physical and virtual environments have been used to investigate different depth cues in various studies. For example, to investigate the influence of texture gradient on correct distance perception, Sinai et al. asked participants to judge the perceived distances of visual targets placed at various distances with and without a gap to the ground (Sinai, Ooi, & He, 1998). Although experiment settings using real environments can best reflect the actual perception of depth, restrictions remain because of the difficulty of isolating the investigated depth cue from the real environment. In contrast, physical settings in a lab environment provide more control and therefore can reveal the more specific effect of a particular depth cue. Blessing et al. constructed distorted tunnel models of different lengths but the same apparent distance to investigate the illusory depth effect of false perspective (Blessing, Landauer, & Coltheart, 1967). Huang designed a sophisticated installation that allowed participants to use a rotary knob to control the locations of visual targets for a visual matching task and researchers to manipulate the illumination conditions. Using this installation, specific issues regarding the effect of colored light on correct distance perception for binocular and monocular viewing were revealed (Huang, 2007).

Although a lab-controlled environment allows more precise parametric control of the experiment setting, significant effort is often required to construct the environment, and the spatial scale is often limited to that of interior spaces. This disadvantage can limit the investigation of depth perception to a short range, and the environment used to perform the perceptual study may be difficult to reproduce in order to evaluate the design application of the research finding. To this end, digital technology provides an alternative environment for investigating depth effects that otherwise cannot be investigated in a real setting. For instance, Meng and Sedgwick used a computer environment in which the presence of shadow in nested surface contact relations could be controlled to investigate its effect on distance perception (Meng & Sedgwick, 2001). However, although the digital environment can practically create any scale and condition of a three-dimensional space, whether a depth effect that is revealed in a virtual environment can be applied to actual perception in a real environment is dependent upon the visual realism provided by the computer environment.

Visual perception of the lighting distribution of a scene is derived from complex interactions between light, the built environment, and the visual system. To generate an alternative environment in which the scene-based lighting distribution can be parametrically manipulated to determine its effect on depth perception, Tai and Inanici developed a computational framework incorporating a physically based lighting simulation software called RADIANCE and perceptually based tone-mapping to generate a perceptually realistic computer-generated environment (Tai & Inanici, 2012). This environment was later used to investigate the interrelationships of light, architectural configuration, and the resulting luminance distribution, which revealed that luminance contrast can be used as an effective design strategy to create illusory spatial depth in an architectural scene (Tai, 2013). To
validate whether luminance contrast is effective in both real and virtual environments, a comparative study conducted identical experiments using radiance maps generated from captured images of a physical model and RADIANCE simulations. Results revealed that the RADIANCE simulations create physically accurate luminance distribution data with physically based parameters. The matching experimental results from both settings also demonstrate that the computational framework can generate a perceptually realistic static representation for monocular vision (Tai, 2014). The effect of binocular disparity on both visual realism offered by the alternative experimental environment and the illusory depth effect from luminance contrast was subsequently investigated using three display modes: standard single-image view, anaglyph 3D, and autostereoscopic display (Tai, 2015). The study's results demonstrated that, even with binocular vision, the luminance contrast created an effective illusory depth effect. In addition, observers were more confident in their perceptual judgments when viewing images using an autostereoscopic display.

These previous studies investigated two factors that affect the illusory depth effect of luminance contrast: (1) the visual realism of images produced by a computer-generated pictorial environment and their realism with respect to the visual realism of real images of a physical scene, and (2) the effect generated by stereoscopic displays. However, the combined effect of visual realism and the stereoscopic display of a physical scene has remained unexplored. The difference in scale at which physical and digital models are constructed was also considered as a possible confounding variable.

Binocular disparity refers to the differences of the retinal images perceived by each eye. Although the assumption that some differences in the perceived images might not considerably affect the perception of a general scene may be common, this assumption may not apply when viewing a visual target at different distance scales. This means that viewing a reduced-scale physical model might be different from viewing a full-scale environment, at least when perceiving the depth of objects in space. In other words, if a visualization tool uses a reduced-scale environment to envision the depth effect of luminance contrast, knowing whether the luminance contrast depth cue has the same effect at both short and long distances is important.

This study reviews previous studies that investigated the manner in which luminance contrast can be used as a depth cue, as well as perceptual studies conducted to validate the visual realism of visualizations produced using high dynamic range (HDR) imaging technologies. Two HDR acquisition techniques were employed to generate visual stimuli for psychophysical experiments: a radiance map (in which multiple low-dynamic range images of a physical-scaled model were combined to create a single HDR image) and physically based rendering. The current study has two objectives: to examine the visual realism of HDR acquisition methods and to determine the influence of full and reduced model scales on depth perception when images are viewed binocularly.

2. Experimental design
Depth cues can be categorized based on different criteria such as static or dynamic, pictorial or kinetic, monocular or binocular, and absolute or relative cues. In general, depth cues based on physiological feedback such as the convergence of two eyes focusing can provide absolute distance perception, but only at a short range. In contrast, static pictorial cues enable the perception of relative distances at greater ranges. However, some pictorial cues such as atmospheric perspective can only help to establish a relatively near or far distance for very distant scenes (Palmer, 1999).

Previous studies have established that perceptually realistic computer simulations provide pictorial cues that can match those of real scenes. However, the perspective view presented on common display devices does not address the distance scale issue. Figure 1(a) presents a scene in which a sculpture is viewed in a gallery space. Figure 1(b) shows the same setting with everything at a reduced scale of 1:10. The two images present no visual difference. However, when synthesizing the two perceived images for each eye to simulate stereo visual perception, the reduced scale of the setting increases the angle between the two eyes and target of focus. Figure 2 shows that the disparity of the overlapped images for each eye is considerable in the reduced-scale setting. Therefore, when using a
scale model to envision the depth effect, whether directly or through captured images, this effect may be influenced by the reduced scale resulting from binocular disparity. (Figures 1(a), (b) and 2).

In our study, a scale model of the architectural environment was constructed at a 1:10 scale. Computer simulations at this configuration were conducted with a physically based lighting simulation program and presented to observers using stereoscopic display technology. To account for binocular disparity, images of the scale model were captured from two viewpoints, as in previous studies. The two viewpoints were set apart by 6 and 60 mm to simulate the typical lateral distance between the left and right eyes, which correspond to the full- and reduced-scale environments, respectively.

This experimental design had two objectives: (1) to explore the binocular disparity between the reduced- and full-scale environments and determine the manner in which this influences depth perception and (2) to verify the visual realism produced by the computational framework with respect to images from the physical model.

3. Experiment setup
The experiment tested three independent variables: luminance contrast, HDR acquisition method, and model scale. The luminance contrast consisted of two different conditions. The amount of light entering the scene was controlled through the architectural configuration of the skylights. As shown
in Figure 3, the model used for this experiment was adopted from a previous study and consisted of four modular boxes measuring 6 × 6 × 4 m with a 2.5 × 2.5 m opening at the center of their ceilings. A red sphere 40 cm in diameter floated 1.5 m above the ground and was 15 m from the point of observation. To vary the interior lighting condition, the skylights were set half-open, fully open, closed, and half-open such that the luminance contrast of the visual target to the foreground was greater than the contrast of the target to the background. In the even lighting condition, all skylights were fully open so that the luminance contrast of the visual target to the foreground was equal to the luminance contrast of the target to the background (Figure 3).

Two HDR acquisition methods were used to create visual stimuli for the experiments. The first created radiance maps from low-dynamic range images of the physical model, and the second created computer simulations using RADIANCE. The output images from both methods were then tone-mapped so that they could be displayed on a standard device.

For the physical model, a 1:10 scaled model of the architectural configuration was built using plywood, and its interior surfaces were lined with grey cardboard. A 4 cm diameter ball was spray-painted red and hung by a string 15 cm above the ground at a distance of 1.5 m from the camera viewpoint. Multiple images of the experimental scene were captured with different exposures using a Canon 5D Mark III fitted with a 50 mm prime lens. The sequential images were assembled into a single HDR image using Photospheres software, which is able to calibrate the luminance of the output HDR and provides validated accuracy (Inanici, 2006). For the computer simulation, the digital experimental scene was generated at a 1:1 scale using a three-dimensional modeling program and exported to RADIANCE, which provides physical accuracy (Ruppertsberg & Bloj, 2006; Ward, 1994). The physical colorimetric and reflectance values of the objects in the physical model were measured with a spectrophotometer. These values were then input as arguments to the material modifier in RADIANCE. The digital model was illuminated with an HDR light probe of the office in which the physical model was located when images of the scene were captured (Reinhard et al., 2010).
Figures 4(a) and (b) show the perspective views of the computer-simulated scenes of the architectural configuration at 1:1 full-scale and 1:10 reduced-scale. Both were captured from a single-camera viewpoint. In this two-dimensional perceptual representation, the two images do not present any distinctive visual difference. For the stereo perspective views of the full- and reduced-scale configuration, each stereo view was created using two images captured from two viewpoints and presented to the corresponding eye. Figure 5 shows the superimposed left and right views of the full- and reduced-scale images (Figures 5(a) and (b), respectively) to illustrate the difference in binocular discrepancy. They do not show the actual three-dimensional stereo representation, but a comparison of the two-dimensional misalignment of features in the images can reveal the difference in the binocular discrepancy presented by an autostereoscopic display.

As illustrated in Figure 5(a) and (b), the discrepancy suggested by the misaligned features increases from the 1:1 to 1:10 scale settings. Thus, the visual differences in the stereo representation could be attributed to the distance between the two viewpoints and target of focus. Therefore, to reduce the disparity between the two viewpoints to one-tenth, we can simulate the stereo-viewing mode of the full-scale environment from the 1:10 scale setting (Figures 4 and 5).

To compare the effect of binocular disparity of the experimental scene generated from the physical-scale model to that of the full-scale environment, the experimental scenes were presented in three forms: “standard single-image view,” “reduced stereo view,” and “full stereo view.” To prepare the images for stereoscopic display, the pair of tone-mapped images were placed side-by-side in an image-editing program and saved in JPEG stereoscopic format (JPS). The JPS image was split into halves and each half projected to the corresponding eye when shown on an autostereoscopic display device. For the “full stereo view” condition, two cameras were set 6 mm apart in the physical-scale model to simulate the visual target’s location in a full-scale architectural configuration. For the “reduced stereo view,” the two cameras were placed 60 mm apart to simulate the viewing conditions of the actual distance in the scaled physical model. A third set of experimental scenes was created with a single-camera setting and shown in a standard, single-image mode. This set of images was used as the baseline for comparing the two sets of stereo images.

Figure 6 presents the experimental scenes for the different conditions. The 12 conditions are labelled according to the luminance contrast condition (“E” for “Even” and “C” for “Contrast”), HDR
Figure 6. Experimental scenes for testing different conditions.
Figure 7. Experimental reference scenes for the "monocular viewing condition."
acquisition method ("P" for "Physical" and "S" for RADIANCE simulation), and camera viewpoint distance ("00" for single-camera setting, "06" for 6 mm, and "60" for 60 mm) (Figure 6).

Many methods have been developed to measure perceived distance along with a statistical model to measure perceived distance from collected data. Visual matching is the most commonly used method. It requires participants to view a test target and adjust a comparison target to match the perceived distance of the test target. However, when performing visual matching using a two-dimensional representation of a three-dimensional environment, participants may be distracted by the two-dimensional cues instead of judging the three-dimensional depth. In contrast, a method that forces the participant to make a binary choice allows them to respond quickly and therefore records a more intuitive and reliable perceptual judgment (Cunningham & Wallraven, 2011).

The constant stimuli method was used to measure the perceived distance by requiring a binary response such as indicating whether a test or reference target is perceived to be nearer (Gescheider, 1984). In this method, the location of the test target remains constant, whereas the reference target’s location ranges from the farthest distance needed to perceive the reference target as actually farther than the test target to the nearest distance needed to perceive the reference target as actually nearer than the test target. For the method of constant stimuli, only an odd number of reference scenes is required. Observers need only produce a simple judgment on each pair of test and reference targets, thus generating a more intuitive response. With this method, observers compare the relative distances of the test and reference targets several times, and the collected data are processed using Ogive or Probit analysis to derive a psychometric function. The psychometric function reveals the approximate value of the independent variable when the test target is perceived to be at the same distance as the reference target. Figure 7 shows the seven reference scenes (as a monocular view) of the physical model and computer simulations. Each reference scene uses the same lighting condition as that of the even condition test scene; however, the location of the visual target was placed at 12, 13, 14, 15, 16, 17, and 18 m away from the viewpoint for comparison with the fixed location of 15 m in the test scenes under different conditions (Figure 7).

4. Experiment results and discussion
The constant stimuli method does not require large number of participants; however, it does require each participant to make a large number of perceptual judgments so that the measured response may be derived from the collected binary data. The advantage of this method is that the measured response can be more intuitive and reliable; however, the disadvantage is that the collected data may not be suitable for studying individual differences because of the small number of participants. As a result, the analysis of the experiment results is limited to the visual realism for a normal viewing task of the investigated display methods. Therefore, experiment participants must have normal or corrected-to-normal vision capable of viewing digital images on a computer display. Ten qualified participants volunteered for the experiments. All of them were adults (ages 20–42), with normal or corrected-to-normal vision, and none experienced discomfort when viewing stereoscopic images. Experiments were performed in an artificially lit research lab, thereby ensuring a stable lighting environment for the duration of the experiments. The experimental scenes were displayed on a Toshiba Satellite P850 laptop computer capable of displaying autostereoscopic images. Each test scene was paired with one of the seven reference scenes, and each pair of test and reference scenes was presented to each participant 10 times. The location of the test and reference scenes (whether on the left or right side of the paired images) and the sequence of images were randomized. For each combination of test and reference scenes, participants generated 10 perceptual judgments. Presented with 12 conditions for the test scenes along with seven reference scenes 10 times, each participant produced 840 perceptual judgments. Each combination of paired images was then judged 100 times.

To measure the perceived distances of the visual targets under different conditions, collected data were statistically analyzed using Probit analysis (Finney, 1971). Figures 8 and 9 show the Probit regression curves for each condition using the HDR acquisition methods. The x-axis represents the location of the reference scene targets, whereas the y-axis represents the probability that the test scene
Figure 8. Probit analysis of experimental results based on different HDR acquisition methods: Radiance maps.

Figure 9. Probit analysis of experimental results based on different HDR acquisition methods: Computer simulations.
The intersection point between each regression curve and the dashed line at 0.5 (which represents 50% probability) is the point of subjective equality (PSE) for each condition. The PSE is considered to be the measured perceived distance of the visual target under that specific condition. Table 1 summarizes the measured perceived distances of the test scene targets between the luminance contrast conditions of “Even” and “Contrast,” and lists the increase as a percentage for each condition. The PSEs generated using a radiance map to create the visual stimuli deviate slightly compared to those generated using computer simulations (which all seem to coincide). The same is true of the regression curves. Ignoring these marked differences, we can see that the regression curves of the radiance maps and computer simulations are extremely similar in terms of positive shifts of the PSEs from the “Even” to “Contrast” conditions. These results concur with previous studies, which found that luminance contrast is an effective cue for creating illusory depth effects regardless of the HDR acquisition method employed (Figures 8 and 9, Table 1).

Figure 10 compares the increase of the three viewing modes. In both the “Physical” and “Simulated” conditions, the stereo-viewing mode in the full-scale environment showed the lowest increase of the illusory depth effect. In addition, the increase for images generated by computer simulation for HDR acquisition were much more similar than when radiance maps were used. A possible explanation for this is that the parametric control for generating experimental scenes of the computer simulation method is more precise. The output computer simulations are thus more consistent in appearance, and this yields more consistent perceptual judgments (Figure 10).

![Figure 10](image.png)

**Table 1. Percentage increase of perceived distances of test scene targets**

<table>
<thead>
<tr>
<th></th>
<th>E (Even)</th>
<th>C (Contrast)</th>
<th>% of increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>P00-</td>
<td>14.917 ± 0.065</td>
<td>16.116 ± 0.066</td>
<td>8.038</td>
</tr>
<tr>
<td>S00-</td>
<td>15.001 ± 0.052</td>
<td>15.877 ± 0.049</td>
<td>5.840</td>
</tr>
<tr>
<td>P06-</td>
<td>15.045 ± 0.069</td>
<td>15.952 ± 0.064</td>
<td>6.029</td>
</tr>
<tr>
<td>S06-</td>
<td>15.030 ± 0.055</td>
<td>15.882 ± 0.051</td>
<td>5.669</td>
</tr>
<tr>
<td>P60-</td>
<td>15.090 ± 0.059</td>
<td>16.211 ± 0.057</td>
<td>7.429</td>
</tr>
<tr>
<td>S60-</td>
<td>15.000 ± 0.054</td>
<td>15.901 ± 0.051</td>
<td>6.007</td>
</tr>
</tbody>
</table>
5. Conclusion

Two particular experiments were conducted in this study. The first compared the perceptual equality between the experimental scenes generated using radiance maps and computer simulations created using RADIANCE. The second investigated the effect of binocular disparity on depth perception in reduced- and full-scale environments. The results of both experiments confirm that luminance contrast can be an effective visual cue to increase the perceived distance of a visual target in an architectural scene, regardless of the HDR acquisition method employed. The experimental results also indicate that the influence of the reduced- and full-scale environments on the illusionary depth effect from luminance contrast is not considerable for binocular views. However, analyses of the results reveal that representations generated by computer simulations are more reliable for investigating the effect of luminance contrast depth cues. This study concludes that computer simulations, when employing HDR imaging and autostereoscopic display methods, can provide the visual realism necessary to investigate the effect of luminance contrast on depth, and that the effect of binocular disparity in reduced- and full-scale environments of the scene is insignificant. However, to design and visualize possible applications of luminance contrast for enriching spatial perception, computer simulations that use HDR-related technologies can be a practical and reliable option.

Funding
This work was supported by the Ministry of Science and Technology in Taiwan under [grant number NSC 102-2221-E-027-086-].

Acknowledgement
The author would like to express his appreciation for this support. The author also wishes to thank the people who participated in the experiments.

Author details
Nan-Ching Tai1
E-mail: nctai@ntut.edu.tw
1 Department of Interaction Design, National Taipei University of Technology, No. 1, Sec. 3, Chung-Hsiao E. Rd., Taipei, Taiwan.

Citation information
Cite this article as: Effect of model scale on predicting illusory stereo depth effect of luminance contrast in real and virtual environments, Nan-Ching Tai,Cogent Engineering (2016), 3: 1243007.

References