Disambiguating Stereoscopic Transparency using a Thaumatrope Approach

Yan-Jen Su and Yung-Yu Chuang, *Member, IEEE*

Abstract—Volume rendering is a popular visualization technique for scientific computing and medical imaging. By assigning proper transparency, it allows us to see more information inside the volume. However, because volume rendering projects complex 3D structures into the 2D domain, the resultant visualization often suffers from ambiguity and its spatial relationship could be difficult to recognize correctly, especially when the scene or setting is highly transparent. Stereoscopic displays are not the rescue to the problem even though they add an additional dimension which seems helpful for resolving the ambiguity. This paper proposes a thaumatrope method to enhance 3D understanding with stereoscopic transparency for volume rendering. Our method first generates an additional cue with less spatial ambiguity by using a high opacity setting. To avoid cluttering the actual content, we only select its prominent feature for displaying. By alternating the actual content and the selected feature quickly, the viewer only perceives a whole volume while its spatial understanding has been enhanced. A user study was performed to compare the proposed method with the original stereoscopic volume rendering and the static combination of the actual content and the selected feature using a 3D display. Results show that the proposed thaumatrope approach provides better spatial understanding than compared approaches.

Index Terms—Stereoscopic perception, stereoscopic display, volume rendering.

✦

1 INTRODUCTION

Volume rendering is a widely used visualization technique for scientific computing and medical imaging. A volume renderer usually takes as input a user-defined *transfer function*, and many 2D slices of image data, each cut into a voxel grid to present a 3D volume. Every voxel is assigned a corresponding color and opacity value. For each ray cast from the viewing plane, the colors and opacity values of the voxels along the ray are integrated into the pixel value of the image plane. By modifying opacity values in the transfer function, users can achieve rendering results that resemble X-ray images. With a high opacity setting, the integration is composed of only the front-most layers, and surfaces in the final image look more opaque. On the other hand, a low opacity setting allows rays to pass through more layers and brings up more structure hidden by front materials. It allows the viewer to see more of the interior structure. However, because this kind of volume rendering projects complex 3D structures into a 2D image, increased transparency also increases ambiguity. Although many improvements have been proposed, researchers argued that static volume rendering is still highly ambiguous even in simple cases [1].

A stereoscopic display can help remedy the problem, as it conveys depth perception to the human visual system. Stereoscopic displays are made based on the principle of binocular vision. In binocular vision, different views received by two eyes allow the brain to create more depth cues and resolves many of the ambiguities in single image projections. However, a recent study showed that, for MRI and CT data, stereoscopic displays have not clearly demonstrated their effectiveness on helping understand spatial relationship among objects [2]. It suggested that in translucent volume rendering, both the underlying and overlying structures are visible and fewer depth cues are available, making it more difficult to perceive 3D structure and shape. Similar observations on stereoscopic transparency have been made outside the medical domain. Akerstrom and Todd reported that (1) the 3D structure of overlapping transparent surfaces are significantly more difficult to perceive than opaque surfaces and (2) the perception of transparency is impaired by increased depth differences between overlaid surfaces or by increased density of elements [3]. Tsirlin *et al.* [4] further evaluated the degree of ambiguity. They concluded that the increase in the number of overlapping transparent layers has a detrimental effect on 3D perception and observers are only capable of separating up to six overlapping surfaces.

For stereoscopic displays, Saye and Frisby [5] showed that monocularly conspicuous features are useful for viewers to fuse large disparity images. Moreover, Tyler and Kontsevich [6] attempted to enhance the perception of stereoscopic transparency based on the stereoattention theory. They demonstrated that the ambiguity of stereoscopic transparency can be alleviated by presenting an attentional cue before showing the actual content to prime the viewing focus. Inspired by their study, for improving spatial understanding, we generate an additional cue which removes the underlying structures by rendering the volume data with a high opacity setting. To avoid cluttering the actual content, we only select its prominent feature as the selected feature. The selected feature

- *Y.-J. Su is with the Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei, Taiwan, 106. E-mail: ked@cmlab.csie.ntu.edu.tw*

- *Y.-Y. Chuang is with the Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan, 106. E-mail: cyy@csie.ntu.edu.tw*

(a) Original visualization   (b) Additional cue   (c) Selected feature   (d) Thaumatrope approach
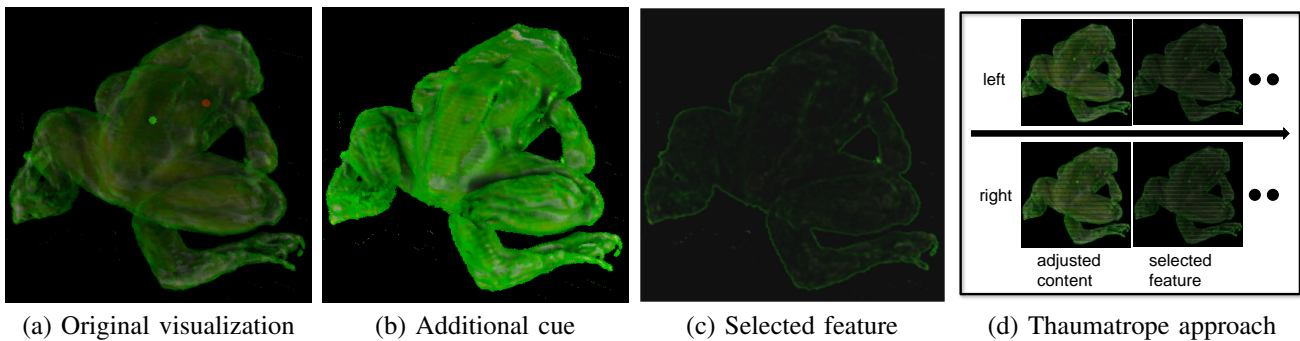
Fig. 1. The proposed thaumatrope approach. (a) The visualization of the *Frog* dataset produced by volume rendering. Because 3D information is projected onto a 2D image, the real structures are often ambiguous and easily misinterpreted. (Only the left view is shown here.) (b) With a high opacity setting, outer surfaces are enhanced and most underlying structures are removed. Viewers often can better understand spatial relations of such a scene. We can use it as an additional cue to supply spatial information. (c) To avoid cluttering the actual content, we only select prominent feature of the additional cue. We call it the selected feature. (d) Our algorithm quickly flips between the (adjusted) actual content and the selected feature. When displayed on a stereoscopic display, the proposed method enhances spatial understanding with stereoscopic transparency.

is then temporally interleaved with the actual stereoscopic content (volume rendering with the normal opacity setting). We call this the *thaumatrope* approach because it shares the same principle with the thaumatrope. A thaumatrope is a toy which flips between two pictures quickly so that they appear to combine into a single image. A famous example, *The Bird in a Cage*, has a bird drawn on one side of a smaller circular disk of paper suspended between pieces of string, and the cage on the other side. When twirling the paper disk, the viewer sees a bird in a cage. This mechanism is used to temporally fuse the actual content and the selected feature. While the selected feature does not overly occlude the actual content, views are primed by the extra information and viewers could have better spatial understanding of the original volume data. Since the two different images are flipped at a sufficiently high frequency, the viewer is not consciously aware of the additional selected feature but her 3D perception of the content is enhanced.

Figure 1 demonstrates the idea of the proposed algorithm. Figure 1(a) shows the original visualization, the actual content we would like viewers to see. Figure 1(b) shows the additional cue generated by a high opacity setting. It exhibits less information within the volume but better spatial perception. To avoid cluttering the actual content, we generate the selected feature by only emphasizing the surface features (areas with high intensity gradients) as shown in Figure 1(c). Although only partially visible, volume completion helps build sufficient scene perception. Figure 1(d) demonstrates the proposed thaumatrope approach which creates interlaced images of the actual content and the selected feature and quickly flips them for both the left and right views. Primed by the spatial cues supplied by the selected feature, viewers have better spatial understanding of the actual content. Our method has the following advantages.

- It enhances depth understanding of stereoscopic volume rendering with transparency. A user study shows that our method improves the ratio of correct depth recognition.

- It is independent of visualization methods. The additional cue is generated by scaling the opacity setting of transfer functions as a simple change of parameters. Thus, the proposed method can be used with any direct volume rendering method.

- It works on normal stereoscopic 3D displays and no additional hardware is required. Although the proposed thaumatrope approach needs to flip images quickly, a display with a refresh rate of 60Hz is sufficient to create flicker-free results.

The rest of this paper is organized as follows. In Section 2, we discuss the related work on improving the quality of volume visualization. In Section 3, we describe our algorithm in details. Section 4 presents the comparisons and user study of the proposed method. Finally, in Section 5, we conclude the paper and describe future work.

2 RELATED WORK

The enhancement methods for volume visualization can be divided into two categories: software methods and hardware methods. The software methods either utilize better visualization algorithms or require extra user effort for obtaining better classification. On the other hand, hardware methods improve results at the expense of requiring additional or specialized devices. We discuss both types in the following.

2.1 Software methods

In the field of physically based rendering, Rushmeier *et al.* [7] used the zone finite element method to calculate isotropic scattering in participating media for more realistic rendering. Based on photon mapping, Jensen *et al.* [8] presented a more efficient method that emits photons in the first pass and gathers the final result by adaptive ray marching in the second pass. This method however requires more memory to store the photons with additional light and transport information. To address this issue, Qiu *et al.* [9] divided a volume into lattices and

constrained the movement of photons. Their method alleviates the memory requirement, but at the same time degrades the quality of visualization. Our method is independent of these methods and can be used along with them.

Our rendering method is derived from simplified luminance models based on visual perception. Metelli [10] summarized important characteristics for perceiving the shape of a transparent surface. Fleming *et al.* [11] showed that lighting and shadows are very important factors for transparency perception. Halle *et al.* [12] demonstrated the influence of lighting on visualization and suggested a multiple light approach. Lee *et al.* [13] proposed a local lighting method that enhances the auto-partition results by different lights. On the other hand, the shadow effect was applied to improve comprehensibility of volumetric data by Stewart [14]. For efficiently generating accurate shadows, a global shading model has been proposed by Schott *et al.* [15] and later improved by Solteszova *et al.* [16]. There are some other perception-based transparency visualization approaches, such as shape-from-texture [17] and shape-from-motion [18]. Chan *et al.* [19] considered the effects of different visual cues and proposed an algorithm to utilize them by finding an optimal combination. Finally, based on quantitative perception models, Zheng *et al.* [20] proposed an automatic approach to enhance the depth-ordering perception of volume by optimizing rendering parameters.

The choice of transfer functions is independent of underlying rendering methods and has significant impact on visual perception. Kindlmann *et al.* [21] described a multi-dimensional approach to refine the classification results and to enhance the boundary structure between different materials. Using multi-dimensional transfer functions, Kniss *et al.* [22] further built an interactive system that allows users to probe in the data space and assign suitable transfer functions. Many features, such as curvatures [23], sizes [24] and visibility [25], have been proposed for better classification. Our method generates additional cues by employing transfer functions made easy and efficient by these methods.

Although not aiming at improving volume visualization, Peterson *et al.* [26] has also worked on reducing ambiguity for stereoscopic displays. Their paper presented a technique to manage visual clutter caused by overlapping labels in complex information displays. They showed that careful selection of disparity differences can be useful for declutter of several visually overlapping planes with texts using stereoscopic displays. Both their paper and ours aim at resolving ambiguity for stereoscopic displays. However, we deal with transparency in volume rendering while they worked on declutter for information displays.

2.2 Hardware methods

Volumetric and stereoscopic 3D displays can improve volumetric rendering results, as both types provide more depth cues than commonplace 2D displays. For volumetric displays, Maeda *et al.* [27] proposed an all-around viewing system that spins a flat display and shows the corresponding view depending on the viewer's position. Agocs *et al.* [28] utilized a holographic screen to modulate the signals generated from

a specially arranged array of projectors to construct a large scale display. Cossairt *et al.* [29] built a commercial volumetric display by applying a swept-volume method to rapidly show 2D slices of volume data and used a horizontal-limited projection plane to preserve the occlusion effect. Jones *et al.* [30] developed a novel projection algorithm collaborated with an anisotropic holographic mirror to achieve better viewing quality for volumetric displays. Barnum *et al.* [31] used water drops as wide-angle lens to reflect the signal from controlled projectors and Wetzstein *et al.* [32] provided a tomographic approach implemented by attenuation layers to recreate a 4D light field. In general, volumetric displays provide solutions to accommodation mismatch and also support wide fields of view [33]. However, they are more expensive and not as popular as flat stereoscopic displays.

This paper proposes a method for enhancing 3D perception of volume rendering on stereoscopic displays. Stereoscopic 3D displays have become popular commodity hardware, due to the success of commercial 3D movies. They can convey very good 3D perception, but also tend to suffer from the problem of convergence and accommodation mismatch. Akeley *et al.* [34] discussed the incorrect focus problem of typical stereoscopic displays and created a prototype to support multiple focal distances for a fixed viewpoint. Liu *et al.* [35] implemented a multi-focus system on a head-mounted display to alleviate the viewpoint limitations. Improving rendering quality is another popular research direction. Damera-Venkata *et al.* [36] combined multiple projectors to enhance the display resolution. Berthouzoz *et al.* [37] recently proposed a method with the similar goal by vibrating displays. As reported by Berthouzoz *et al.*, the brightness and refresh rates are main issues for the fusion of multiple images. Our solution also considers these issues carefully as discussed later.

3 THAUMATROPE APPROACH

Figure 2 demonstrates the flowchart of the proposed thaumatrope approach. This algorithm consists of three main components. First, the additional cue is generated by volume rendering with a high opacity setting. Next, the visual adjustment procedure adjusts for both the actual content and the additional cue such that (1) the selected additional cue does not have too much content so that it will not clutter the actual content and (2) the brightness of the blended image is close to the original one. Finally, an interlace fusion process is used to combine these images for flicker-free results on a 60Hz stereoscopic display. The details are described in the following sections.

3.1 The additional cue

In their study of stereoattention [6], Tyler and Kontsevich suggested that one effective way to enhance the discriminability of two random-dot planes with sinusoidal corrugation is to present a cue plane before showing the actual images. Their experiments showed that, when the cue is close to the test image, the discriminability is dramatically improved. Inspired by their study, we seek to add an additional cue for enhancing spatial understanding. We consider two properties for the
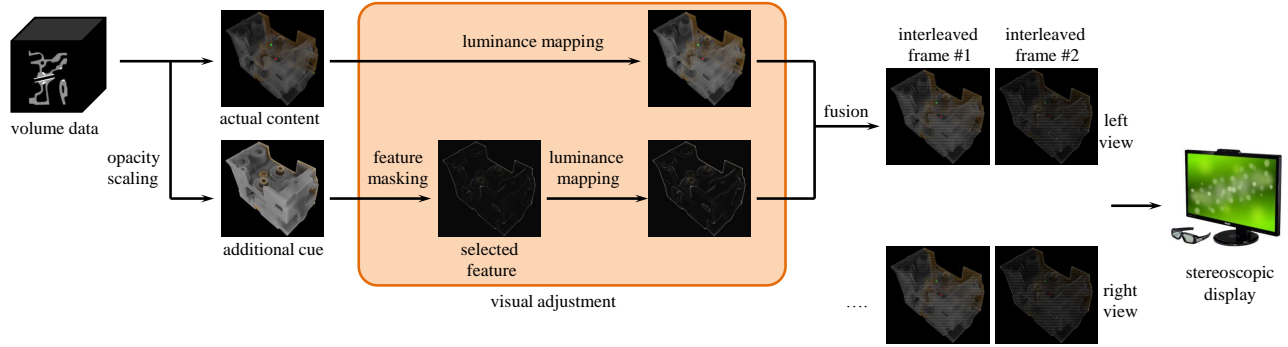
Fig. 2. The flowchart of the proposed thaumatrope approach. The actual content and the additional cue are produced from volume rendering. First, the additional cue is tuned to suitable visibility by masking only features and dimming the image. We refer the tuned additional cue as the selected feature. Next, both the actual content and the selected feature are adjusted by luminance mapping to compensate for the luminance differences caused by the dimmed selected feature. Finally, an interlace fusion process is used to combine these images for flicker-free results.

additional cue: (1) it should be consistent with the scene; and (2) it is better to contain less stereoscopic ambiguity.

There are many different ways of generating an additional cue. One could, for example, only render a partial scene, after chopping off everything at a certain distance from the image plane, or truncate integration of volume rendering like in the peeling approach proposed by Chirstof *et al.* [38]. Our experiments found that opacity scaling gives the best results. We conjecture that it is because it maintains the most coherence in the scene, thus keeping viewing comfort for stereoscopy, while improving perception.

We employ ray casting with emission and absorption for volume rendering. The following numerical integration along a ray is calculated for each pixel.

$$I = \sum_{i=0}^{\lfloor D/\Delta d \rfloor} \left(c_i \prod_{j=0}^{i-1} (1 - \alpha_j) \right), \tag{1}$$

where Δd is a very small ray marching distance that approximates the calculation as the Riemann sum of all segments, i.e., samples are taken uniformly along the ray with the sample interval Δd. D is the distance of the furthest intersection of the ray and the volume; C_i is the color of i-th sample and α_j represents the opacity values of the j-th sample. Equation 1 accumulates contributions of all samples along the ray. Each sample's contribution C_i needs to account for the attenuation of all voxels in front of it which is estimated by multiplying their transparency values $1 - \alpha_j$. Equation 1 can be efficiently calculated by performing linear interpolation on values retrieved from a pre-integrated table.

As mentioned, we decided to generate the additional cue with opacity scaling so that the 3D structure is easier to recognize. This can be achieved by scaling the opacity values with a scale factor β. Figure 3 shows the effects of scaling opacity values of the transfer function for the *Engine* dataset. Figure 3(a) is the volume rendering with the original opacity ($\beta = 1$). A higher opacity setting hides more of the interior structure, reducing information clutter, and thereby alleviating stereoscopic transparency ambiguity. Figure 3(b) shows the

scene with $2\times$ opacity ($\beta = 2$).

The factor β can be specified by users. However, since scenes have very different opacity properties, it requires trial and error to find a proper β. To simplify the task, we make the parameter selection more intuitive. Instead of specifying β directly, we ask users to specify a parameter that is related to the ray attenuation factor. We call it penetration η. For a scene with penetration η, we want that, on average, a ray travels the distance $\eta \times w$ within the volume before being completely attenuated, where w is the width of the volume. Voxels further away than that are almost entirely occluded by those in front. For example, if $\eta = 10\%$, we can only see layers whose distance to the outmost surface is smaller than $0.1 \times w$. We relate the penetration η to the scaling factor β in the following manner. Assuming that all voxels have the same average opacity α, we want that after travelling ηw within the volume, the ray's remaining transmittance is a very small number ε. Using Monte Carlo method for estimating accumulated transmittance, we have

$$(1 - \alpha)^n \approx \varepsilon, \tag{2}$$

where $n = \lfloor \frac{\eta w}{\Delta d} \rfloor$ is the number of samples within the given penetration depth. Thus, the target opacity value $\hat{\alpha}$ would be $1 - \varepsilon^{1/n}$. The scaling factor is set as $\beta = \hat{\alpha}/\bar{\alpha}$ where $\bar{\alpha}$ is the average opacity value of the volume. In our current implementation, we let $\varepsilon = 10^{-6}$. Figure 3(c) and (d) show the scene with $\eta = 10\%$ and $\eta = 2\%$ respectively. Our users found that adjusting penetration η is more intuitive than the scaling factor. In all examples, we used $\eta = 2\%$ to generate the additional cues. Note that, even with different overall transparency settings, opacity scaling preserves good coherence with the actual content.

3.2 Visual adjustment

The process of visual adjustment aims to achieve two goals. First, the strength of the additional cue should convey enough information to enhance spatial perception, while at the same time, it should not clutter or otherwise distort the original
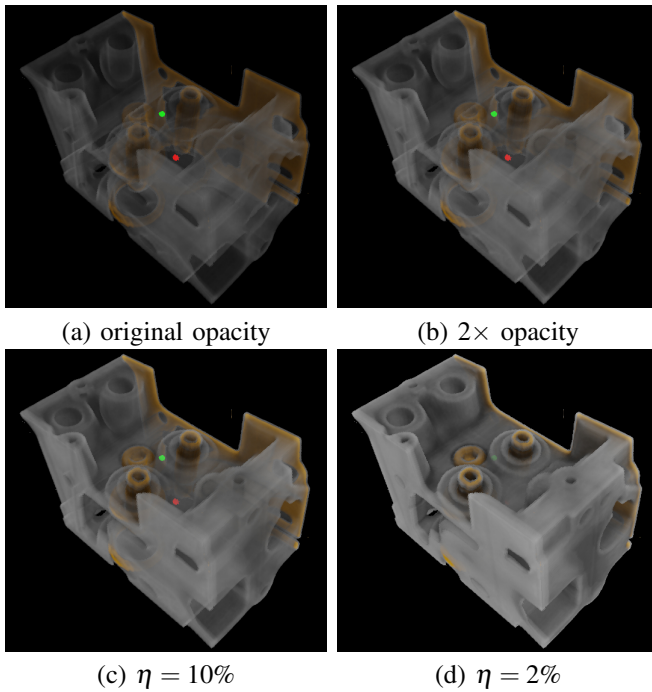
(a) original opacity

(b) 2× opacity

(c) $\eta = 10\%$

(d) $\eta = 2\%$

Fig. 3. Rendering results with different opacity scaling factors and penetration ratios. (a) shows the scene with the original opacity ($\beta = 1$). (b) shows the results with 2× opacity ($\beta = 2$). (c) and (d) show the rendering with 10% and 2% penetration respectively. As the opacity increases (either with a higher scaling factor β or smaller η), the front surfaces are clearer and ambiguity is reduced. Note that the coherence provided by opacity scaling makes it suitable for generating additional cues. We used $\eta = 2\%$ throughout the paper.

(a) Additional cue

(b) Gradient magnitude
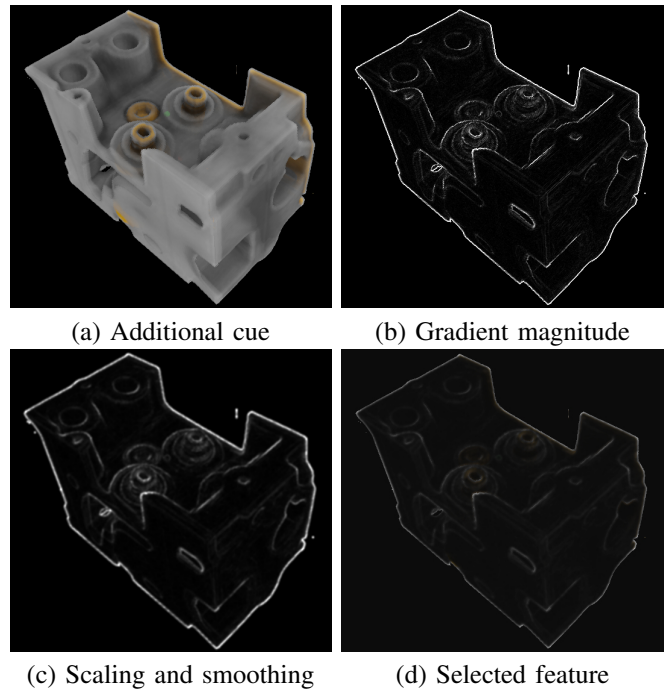
(c) Scaling and smoothing

(d) Selected feature

Fig. 4. Feature masking. Given the additional cue (a) from Section 3.1, we obtain the gradients and calculate their magnitudes (b). The magnitudes are enhanced and a 7×7 Gaussian filter with $\sigma = 1.5$ is applied to smooth out the result to obtain a scale mask (c). We multiply the additional cue with the mask to emphasize only the features of the surface in the additional cue. Finally, the product image is dimmed to match the intensity of the actual content to generate the selected feature (d). Experiments showed that the selected feature alone is sufficient for viewers to build up a baseline shape while not cluttering the actual content.

image. Thus, we have to select only important features and tweak the additional cue's intensity in order to enhance spatial perception without cluttering or overshadowing the original image. Second, since the selected feature is relatively dark, and occupies half of the temporal domain, the resulting interleaved image is also perceived as darker. We employ luminance mapping to compensate for the overall loss of intensity.

3.2.1 Feature masking

We found that an additional cue generated with opacity scaling is too complex and causes visual clutter. We adjust the additional cue by feature masking based on Tse's volume completion theory [39]. It states that the human visual system is able to complete entire 3D enclosures, which are not explicitly visible within a scene. Evidence also shows that the completion is not due to cognitive inference, but rather the automatic outputs of perceptual processing. Based on the theory, we found that it is sufficient to strip everything but prominent features from the image that represents the additional cue. Figure 4(d) shows an example of the selected feature. Our experiments show that seeing a stereoscopic pair of the selected feature alone does allow users to perceive the rendered volume. The results show that the volume completion theory can be applied to stereoscopy.

We explain the detailed steps of feature masking using the example in Figure 4. Given the original additional cue (Figure 4(a)), we first obtain the gradient magnitude for each pixel (Figure 4(b)) as a measure of its prominence. Pixels with high gradient magnitudes are those that we are interested in, and those with small gradient magnitudes are removed. Although simple, the measure proves effective. To reduce the impact of noise and small features, we apply a Gaussian filter to smooth out the gradient magnitude map. Also, the map is scaled up by a scale factor S_{en} to form a scale mask (Figure 4(c)). We multiply the additional cue with the mask and scale the intensity to match that of the actual content. The final product is the selected feature (Figure 4(d)). In all examples, we used $\sigma = 1.5$ for the Gaussian filter and $S_{en} = 10$.

3.2.2 Luminance mapping

The proposed thaumatrope approach displays the selected feature in half of the displayed frames. Since all removed pixels in the selected feature have been replaced with black pixels, we need to make the actual content brighter to compensate for the luminance difference. Note that we can make both the actual content and the selected feature brighter at this stage. As

long as the adjustments are consistent, the actual image will not be cluttered by the selected feature. A naive way would be to scale up the intensity by a constant larger than 1. This however could oversaturate some areas if they are already very bright. The problem is even more disturbing for stereoscopic images as it could lead to intensity mismatch between two views.

To avoid oversaturation, before scaling the intensity, we must apply tone mapping. That compresses the intensity range while retaining the visual appearance of the image. After tone mapping, we can safely apply intensity scaling without introducing oversaturation. For computational efficiency, we modified the global operator of Reinhard *et al.* [40] for tone mapping. It is applied to both, the actual image and the selected feature separately, using the following computations.

$$I' = \frac{S_{key}}{\bar{I}_{log}}I, \qquad (3)$$

$$\hat{I} = C_B \frac{I' + \frac{I^2}{C_{white}^2}}{I' + 1}. \qquad (4)$$

Equation 3 is the initial mapping. The log-space intensity I is normalized by the log-average luminance $\bar{I}_{log}$ for the actual content and then scaled by a user-specified scaling factor S_{key}. S_{key} determines the overall brightness of the results. The resultant intensity is then compressed by Equation 4, in which we let $C_{white} = 2.0$ be the corresponding intensity of pure white and $C_B = 1.5$ for the scale-up factor. Because the ratio of I/C_{white} is smaller than 1, this setting ensures the additional scaling will not lead to oversaturation.

Figure 5 (a) and (b) show results of luminance mapping with a low key ($S_{key} = 0.1$) and a high key ($S_{key} = 0.6$) respectively. We used the medial value $S_{key} = 0.25$ for generating all results in this paper as it generates a result closest to the original image. Figure 5(c) shows the corresponding result.

3.3 Fusion

Once we have obtained the adjusted actual content and adjusted selected feature, the next step is to display them temporally interleaved on a stereoscopic display. Didyk *et al.* [41] reported that 40Hz is the safe frequency to avoid perceived flicker in general cases. Since we need to display both the selected feature and the actual content for one frame, we need a refresh rate of 80Hz. However, most stereoscopic displays only have at most 60Hz[1]. Thus, viewers could perceive flicker if we alternate between the selected feature and the actual content on stereoscopic displays. Fortunately, in addition to increasing the refresh rate, there are other ways to reduce flicker. Makela *et al.* [42] reported that the stimulus size is another important factor to flicker sensitivity and flicker is more perceivable in large uniform regions.

Based on the observation, we decompose the stereo images for the actual content and selected feature by scanlines and

---

1. For some shutter-glass-based displays, although the displays can refresh at 120Hz, its 3D refresh rate is still 60Hz. This is because it needs to display both left and right views to complete a stereoscopic frame.

(a) Low key $S_{key} = 0.1$

(b) High key $S_{key} = 0.6$
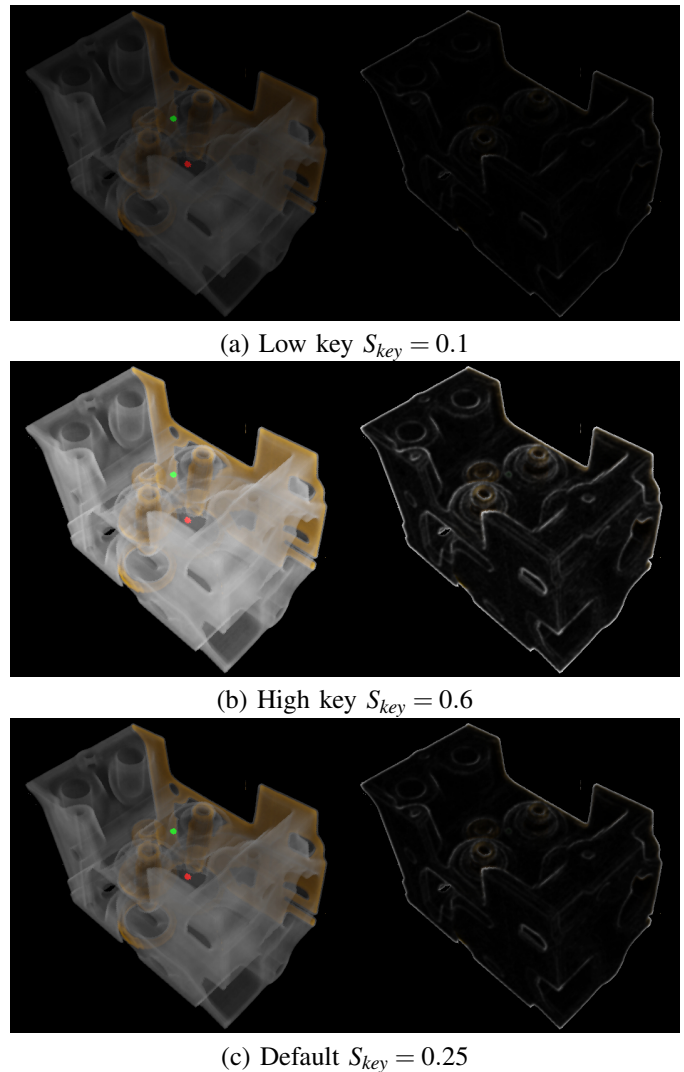

(c) Default $S_{key} = 0.25$

Fig. 5. Examples of luminance mapping. Each subfigure contains two images: The one on the left is the actual content and the one on the right is the selected feature. (a) and (b) show results of luminance mapping with a low key and a high key respectively. (c) depicts the result when rendered with the key that we used for all further results in this paper.

spatially intersperse them to produce two images to be displayed temporally interleaved. To be more specific, for the left view, we take 2/3 of scanlines from the left view of the actual content and 1/3 of scanlines from the left view of the selected feature as the first frame. The second frame, which is the complement of the first frame, is obtained by taking the selected feature's remaining 2/3 scanlines and the rendered image's unseen 1/3 scanlines. The perceived left image is the result of continuously temporally interleaving these two frames in the left view. The right view is processed in the same way. Figure 6 shows an example for one view. This method disrupts the uniform regions of each image, thereby avoiding flicker. Note that, since the method always selects the same corresponding scanlines on both sides, it preserves the disparity information for stereo matching.

TABLE 1
Descriptions of the tested datasets, their resolutions and the rendering and processing
time for the thaumatrope approach.

Dataset	Distribution	Scene (A/B)	Scan	Resolution	Rendering(s) (A/B)	Processing(s) (A/B)
Frog	Whole frog project	Figure 1/Figure 7(a)	MRI	$256 \times 256 \times 44$	0.54/0.52	1.78/1.58
Engine	General Electric	Figure 5(c)/Figure 7(b)	CT	$256 \times 256 \times 128$	0.68/0.72	1.80/1.70
Head	The Chapel Hill test dataset	Figure 7(c)/Figure 7(d)	CT	$256 \times 256 \times 113$	1.00/0.96	3.16/2.54
Brain	The Chapel Hill test dataset	Figure 7(e)/Figure 7(f)	MRI	$256 \times 256 \times 109$	0.76/0.76	2.02/2.85
Bonsai	Stefan Roettger	Figure 7(g)/Figure 7(h)	CT	$256 \times 256 \times 256$	0.40/0.46	1.90/2.52
Chest	Dept. of Radiology, Univ. of Iowa	Figure 7(i)/Figure 7(j)	CT	$384 \times 384 \times 240$	0.88/1.00	2.55/2.79
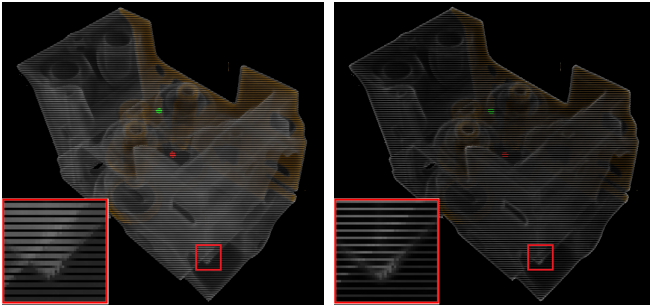
(a) The first frame   (b) The second frame

Fig. 6. An example of interlace fusion. Given the adjusted actual content and the adjusted selected feature, we take $2/3$ scanlines from the left view of the actual content and $1/3$ scanlines from the left view of the selected feature as the first frame (a). The second frame (b) is obtained from the selected feature's remaining $2/3$ scanlines and the actual content's unseen $1/3$ scanlines.

4 USER STUDY

We performed a user study to verify the effectiveness of the proposed approach on improving understanding with stereoscopic transparency. The stereoscopic volume rendering system was implemented using OpenGL shading language and accelerated by GPUs. For the user study, it was deployed on an Asus G60JX laptop, equipped with a 1.6GHz Intel Core(TM)i7 Quad CPU, an NVIDIA GeForce GTS 360M GPU, a 120Hz display of 1024×768 resolution, and a shutter-glass-based stereoscopic display with NVIDIA 3D vision solution. Our system used a fixed interpupillary distance (IPD) for all users. We did not adjust it because it is known that stereo perception is robust to the deviation of the IPD.

Six sets of volume data were used in the experiments. Each was used to generate two renderings with different settings, making a total of twelve test scenes. We used sub-labels A and B to denote the two settings. Table 1 summarizes the datasets with their descriptions, resolutions and processing time. Figure 1 and Figure 5 show the scenes *Frog_A* and *Engine_A* respectively. Figure 7 shows the volume renderings and selected feature of the other ten scenes. Note that CT and MRI data were used in the user study. As reported by Lundstrom *et al.* [43], even with sophisticate filtering schemes, the gradient estimation for MRI data remains an error-prone process. To deal with the issue, we adopted a gradient-free lighting model [44] in our system.

The evaluation of visualization is very subjective and often requires domain knowledge. In addition, the goal of the proposed method is to enhance depth understanding of the scene rather than the volume rendering method. Thus, we designed two tests that are only related to stereoscopic quality and spatial understanding but not visualization quality. We compared the proposed thaumatrope method with two other methods, the *traditional method* which shows only the actual content on a 3D display and the *static combination* of the actual content and the selected feature.

4.1 Comparison with the traditional method

Thirty-three subjects were invited to participate in the user study. For each given scene, users were asked to rate its stereoscopic quality and perform a depth judgment task. Note that it could be biased to show the visualization results by both methods to the same participant since he can often refine his cognition of the 3D structure from the rendering shown first and perform the task better for the one shown later. Thus, the display order of results could introduce bias. Hence, for each scene, a participant only saw one visualization, by either the traditional method or the thaumatrope one, but not both. Each participant saw half of scenes with the traditional visualization method and the other half with the thaumatrope approach. The results are tabulated and reported by datasets but not scenes because the differences between datasets are more essential than scene settings.

Stereoscopic quality. In this study, we asked participants to score visualization results in terms of their stereoscopic quality with a grade within the range [0, 5] in which 0 was the worst and 5 is the best. Note that this is a subjective test and there are many factors on the stereoscopic perception, such as brightness, contour enhancement, visual fusion and so on. The results are depicted in Figure 8. We used the two-tailed t test to analyze users' stereoscopic quality scores. The null hypothesis was $\mu_{tha} = \mu_{trd}$ while the alternative hypothesis was $\mu_{tha} \neq \mu_{trd}$, where μ_{tha} and μ_{trd} were the average grades of the proposed thaumatrope approach and the traditional method. Table 2 shows the statistics within 95% confidence level. It shows that there was no significant difference between these two methods and the grades heavily depended on the datasets. The *Bonsai* dataset had a relatively lower probability for the null hypothesis because of the classification errors caused by the air. Although the proposed thaumatrope approach does not show improvement in the subjective stereoscopic quality study,

(a) *Frog_B*

(b) *Engine_B*

(c) *Head_A*

(d) *Head_B*

(e) *Brain_A*

(f) *Brain_B*

(g) *Bonsai_A*
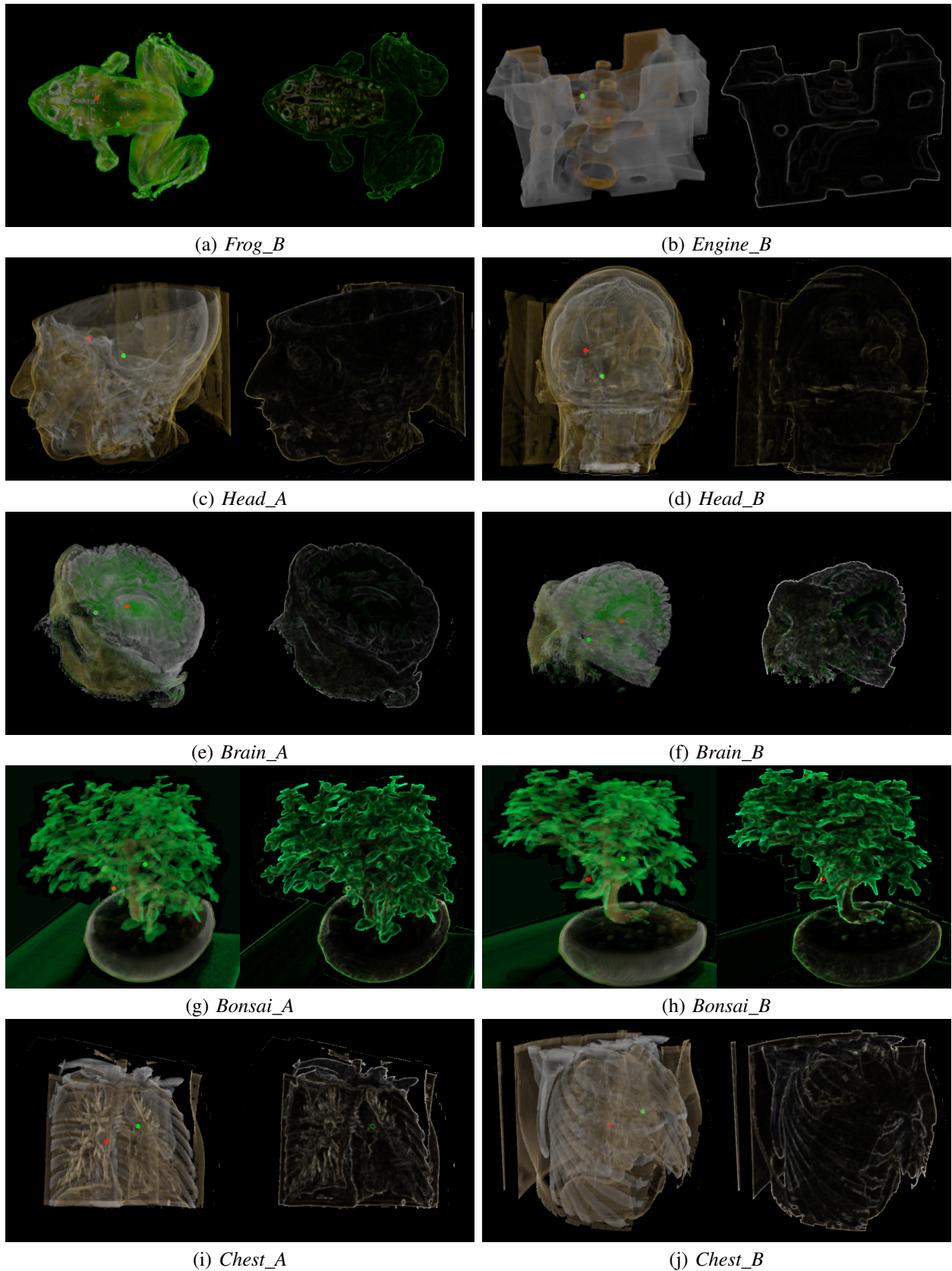
(h) *Bonsai_B*

(i) *Chest_A*

(j) *Chest_B*

Fig. 7. The datasets used in the user study (10 of 12 tested scenes are shown here). We only show the left views of stereoscopic images. For each visualization result, the actual content is shown on the left and the (adjusted) selected feature is shown on the right. All the results were created using the settings $\eta = 2\%$, $S_{en} = 10$ and $S_{key} = 0.25$. The green point and red point were used for the depth judgment study. Their distance in 3D space is smaller than 5% of the width of the volume.
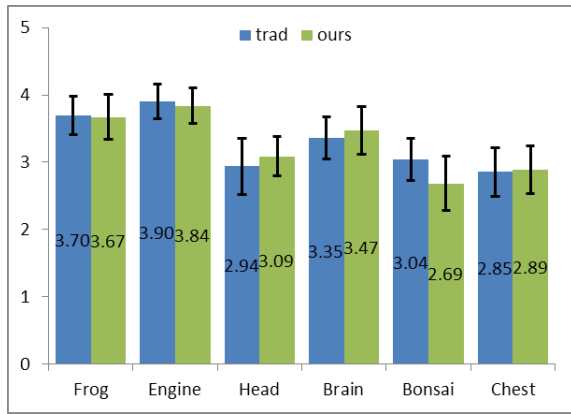
Fig. 8. Comparison of stereoscopic quality with the traditional method. In this subjective study, participants were asked to score visualization results in terms of their stereoscopic quality. For both methods, the mean score for each scene is plotted.
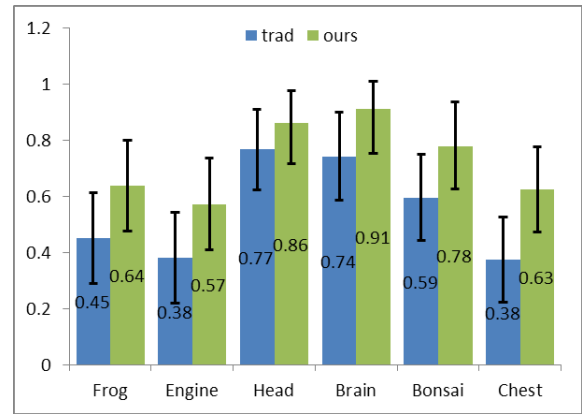

Fig. 9. Comparison of depth judgment with the traditional method. In this objective study, a red point and a green point were embedded into the tested volumes. Participants were asked to point out which point is closer to the viewer. The figure reports the correct rates for both methods.

TABLE 2
The results of the two-tailed t test on users' stereoscopic quality scores for the traditional method and our method.

	Frog	Engine	Head	Brain	Bonsai	Chest
trad	3.70	3.90	2.94	3.35	3.04	2.85
ours	3.67	3.84	3.09	3.47	2.69	2.89
t-value (T(64))	−0.12	−0.35	0.57	0.48	−1.36	0.15
p-value	0.90	0.73	0.57	0.63	0.18	0.88

TABLE 3
The results of the one-tailed t test on the depth judgment test for the traditional method and our method.

	Frog	Engine	Head	Brain	Bonsai	Chest
t-value	1.81	1.71	0.99	1.83	1.71	2.21
p-value	0.04	0.05	0.16	0.04	0.05	0.02

we show that it did improve user's depth understanding of the scene in the next objective study.

Depth judgment. In the depth judgment test, for each dataset, we embedded a red point and a green point into the volume. The positions of these points were arbitrarily chosen by hand without aligning to specific features of the volume data, but we had their distance smaller than 5% of the width of the volume to make the test more challenging. The pixels representing the two points were fully opaque and their colors were set pure red and green without shading so that they can be identified easily.

Given a scene rendered with the thaumatrope approach or the traditional method, a participant was asked to identify which one is closer. We recorded users' answers and measured correctness as a ratio between the amount of correct answers over total answers for each method on each dataset. The results are depicted in Figure 9. The proposed thaumatrope approach outperforms the traditional method in each dataset with up to 25%. The degree of improvement varies between the datasets. We used the one-tailed t test to analyze the results. The null hypothesis was $\mu_{tha} \leq \mu_{trd}$ while the alternative hypothesis was $\mu_{tha} > \mu_{trd}$. Table 3 reports the results within 95% confidence level. In five of six datasets, the alternative hypothesis was accepted, meaning that with the proposed approach, viewers better understand the spatial structure of the volume data. The *Head* dataset is the only dataset that our method does not show significant improvement. The explanation is that, in *Head_A*

(Figure 7(c)), the applied the plane cutting technique [45] already provides very good spatial information. Thus, the gain that our method can provide is limited. Note that the setting of the cutting plane depends on how the user would like to visualize the data and has nothing to do with the use of stereoscopic displays and algorithms. For other datasets applied with the plane cutting technique, *Forg_B*, *Brain_A*, *Brain_B*, and *Chest_A*, the proposed method can still help. Another thing to note is that, for some scenes, the correct rates of the traditional method were even lower than 50%, worse than random guesses. It means that users were significantly misled in these datasets. We can conclude that, in the objective depth judgment test, our method successfully enhances user's understanding of the scene geometry and spatial relationships.

4.2 Comparison with the static combination

In this section, we compare the proposed method with the static combination obtained by adding the content and selected feature. This comparison is designed to verify the effectiveness of the proposed fusion approach which performs the spatially and temporally interleaved rendering (Section 3.3). The same tests and procedure described in Section 4.1 were used in this comparison. Thirty subjects were invited to participate in the user study.

Stereoscopic quality. Figure 10 shows the scores of both methods on each scene in terms of stereoscopic quality. In this subjective study, the static combination provides slightly better perceptual scores than the thaumatrope approach in general.
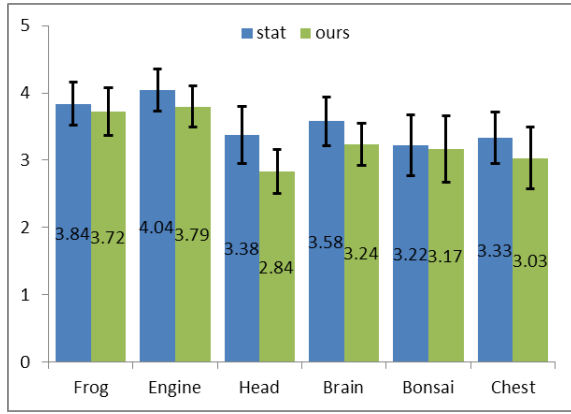
Fig. 10. Comparison of stereoscopic quality with the static combination method.

Fig. 11. Comparison of depth judgment with the static combination.

TABLE 4
The one-tailed t test on the stereoscopic quality for the static combination and our method.

	Frog	Engine	Head	Brain	Bonsai	Chest
t-value	0.47	1.09	1.98	1.39	0.16	0.99
p-value	0.32	0.14	0.03	0.09	0.44	0.16

TABLE 5
The one-tailed t test on the depth judgment for the static combination and our method.

	Frog	Engine	Head	Brain	Bonsai	Chest
t-value	0.46	0.17	1.32	3.47	1.83	1.87
p-value	0.32	0.43	0.09	0.0006	0.04	0.03

We used the one-tailed t test to analyze users' subjective scores on stereoscopic quality. The null hypothesis was $\mu_{tha} \geq \mu_{sta}$ while the alternative hypothesis was $\mu_{tha} < \mu_{sta}$, where μ_{sta} represents the average score of the static combination. Table 4 shows the statistics within 95% confidence level. In one of six datasets (*Head*), the alternative hypothesis was accepted. The superposed selected feature enhances surface features and provides better contrast around contours. Thus, participants generally perceived more vivid images and improved visual quality. However, similar to the observation in Section 4.1, although the static combination seems to improve the subjective scores slightly on stereoscopic quality, our method still outperforms it in terms of user's depth understanding as shown in the next study.

Depth judgment. In the objective study of depth judgment, as shown in Figure 11, the proposed thaumatrope approach outperforms the static combination method in each dataset with up to 40%. We used the one-tailed t test to analyze the results. The null hypothesis was $\mu_{tha} \leq \mu_{sta}$ while the alternative hypothesis was $\mu_{tha} > \mu_{sta}$. Table 5 reports the results within 95% confidence level. In three of six datasets (*Brain*, *Bonsai* and *Chest*), the alternative hypothesis was accepted. It shows that our method is significantly better in half of the scenes. Although the superposed selected feature enhances contours and surface features, it could also occlude the underlying structures. The occlusion becomes more severe when the spatial structure of the scene is complex and the selected feature is rather dense as in *Brain*, *Bonsai* and *Chest*. Thus, in some scenes, static combination could be even worse than the traditional method. The experiment shows that, although the proposed selected feature could improve stereoscopic quality, it is necessary to adopt the proposed interleaved rendering method for improving spatial understanding.
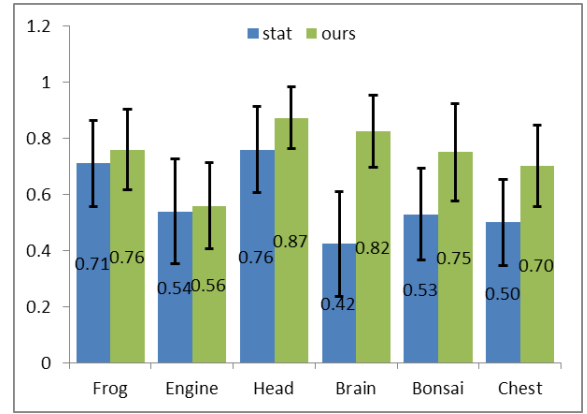
5 CONCLUSIONS

This paper proposes a novel thaumatrope approach to improve understanding with stereoscopic transparency for volume rendering. Inspired by the studies of visual perception, we propose to temporally interleave a selected feature with the actual image to produce a flicker-free visualization. As a result, the proposed method provides viewers with better spatial understanding. Our method is independent of the volume rendering method and can be realized on popular stereoscopic displays.

In the future, we plan to investigate more sophisticated methods to adjust the brightness of the selected feature and the actual content so that the perceived brightness is closer to the original one. In addition, we would like to investigate other stereoattention theories for enhancing depth understanding.

ACKNOWLEDGMENTS

REFERENCES

[1] C. Boucheny, G.-P. Bonneau, J. Droulez, G. Thibault, and S. Ploix, "A perceptive evaluation of volume rendering techniques," *ACM Transactions on Applied Perception*, vol. 5, no. 4, pp. 23:1–23:24, January 2009.
[2] M. H. Beurden, W. A. Ijsselsteijn, and J. F. Juola, "Effectiveness of stereoscopic displays in medicine: A review," *3D Research*, vol. 3, no. 1, pp. 54:1–54:13, March 2012.

[3] R. Akerstrom and J. Todd, "The perception of stereoscopic transparency," *Attention, Perception, & Psychophysics*, vol. 44, no. 5, pp. 421–432, November 1988.

[4] I. Tsirlin, R. S. Allison, and L. M. Wilcox, "Stereoscopic transparency: Constraints on the perception of multiplesurfaces," *Journal of Vision*, vol. 8, no. 5, pp. 1–10, May 2008.

[5] A. Saye and J. P. Frisby, "The role of monocularly conspicuous features in facilitating stereopsis from random-dot stereograms," *Perception*, vol. 4, no. 2, pp. 159–171, February 1975.

[6] C. W. Tyler and L. L. Kontsevich, "Mechanisms of stereoscopic processing: Stereoattention and surface perception in depth reconstruction," *Perception*, vol. 24, February 1995.

[7] H. E. Rushmeier and K. E. Torrance, "The zonal method for calculating light intensities in the presence of a participating medium," in *Proceedings of ACM SIGGRAPH 1987*, August 1987, pp. 293–302.

[8] H. W. Jensen and P. H. Christensen, "Efficient simulation of light transport in scenes with participating media using photon maps," in *Proceedings of ACM SIGGRAPH 1998*, July 1998, pp. 311–320.

[9] F. Qiu, F. Xu, Z. Fan, N. Neophytos, A. Kaufman, and K. Mueller, "Lattice-based volumetric global illumination," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, pp. 1576–1583, November 2007.

[10] F. Metelli, "The perception of transparency." *Scientific American*, vol. 230, no. 4, pp. 90–98, April 1974.

[11] R. W. Fleming and H. H. Bülthoff, "Low-level image cues in the perception of translucent materials," *ACM Transactions on Applied Perception*, vol. 2, no. 3, pp. 346–382, July 2005.

[12] M. Halle and J. Meng, "Lightkit: A lighting system for effective visualization," in *Proceedings of IEEE Visualization (Vis 2003)*, October 2003, pp. 363–370.

[13] C. H. Lee, X. Hao, and A. Varshney, "Light collages: Lighting design for effective visualization," in *Proceedings of IEEE Visualization (Vis 2004)*, October 2004, pp. 281–288.

[14] A. J. Stewart, "Vicinity shading for enhanced perception of volumetric data," in *Proceedings of IEEE Visualization (Vis 2003)*, October 2003, pp. 355–362.

[15] M. Schott, V. Pegoraro, C. D. Hansen, K. Boulanger, and K. Bouatouch, "A directional occlusion shading model for interactive direct volume rendering," *Computer Graphics Forum*, vol. 28, no. 3, pp. 855–862, June 2009.

[16] V. Solteszova, D. Patel, S. Bruckner, and I. Viola, "A multidirectional occlusion shading model for direct volume rendering," *Computer Graphics Forum*, vol. 29, no. 3, pp. 883–891, June 2010.

[17] V. Interrante, H. Fuchs, and S. M. Pizer, "Conveying the 3D shape of smoothly curving transparent surfaces via texture," *IEEE Transactions on Visualization and Computer Graphics*, vol. 3, no. 2, pp. 98–117, April 1997.

[18] E. B. Lum, A. Stompel, and K.-L. Ma, "Kinetic visualization: A technique for illustrating 3D shape and structure," in *Proceedings of IEEE Visualization (Vis 2002)*, October 2002, pp. 435–442.

[19] M.-Y. Chan, Y. Wu, W.-H. Mak, W. Chen, and H. Qu, "Perception-based transparency optimization for direct volume rendering," *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 6, pp. 1283–1290, November 2009.

[20] L. Zheng, Y. Wu, and K.-L. Ma, "Perceptually-based depth-ordering enhancement for direct volume rendering," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 3, pp. 446–459, March 2013.

[21] G. Kindlmann and J. W. Durkin, "Semi-automatic generation of transfer functions for direct volume rendering," in *Proceedings of the 1998 IEEE symposium on Volume visualization*, October 1998, pp. 79–86.

[22] J. Kniss, G. Kindlmann, and C. Hansen, "Multidimensional transfer functions for interactive volume rendering," *IEEE Transactions on Visualization and Computer Graphics*, vol. 8, no. 3, pp. 270–285, July 2002.

[23] G. Kindlmann, R. Whitaker, T. Tasdizen, and T. Moller, "Curvature-based transfer functions for direct volume rendering: Methods and applications," in *Proceedings of IEEE Visualization 2003 (Vis 2003)*, October 2003, pp. 67–74.

[24] C. Correa and K.-L. Ma, "Size-based transfer functions: A new volume exploration technique," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1380–1387, November 2008.

[25] C. D. Correa and K.-L. Ma, "Visibility histograms and visibility-driven transfer functions," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 2, pp. 192–204, February 2011.

[26] S. Peterson, M. Axholt, and S. Ellis, "Managing visual clutter: A generalized technique for label segregation using stereoscopic disparity," in *Proceedings of IEEE Virtual Reality Conference (VR 2008)*, March 2008, pp. 169–176.

[27] H. Maeda, K. Hirose, J. Yamashita, K. Hirota, and M. Hirose, "All-around display for video avatar in real world," in *Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, October 2003, pp. 288–289.

[28] T. Agocs, T. Balogh, T. Forgacs, F. Bettio, E. Gobbetti, G. Zanetti, and E. Bouvier, "A large scale interactive holographic display," in *Proceedings of the IEEE conference on Virtual Reality (VR 2006)*, March 2006, p. 311.

[29] O. Cossairt, J. Napoli, S. Hill, R. Dorval, and G. Favalora, "Occlusion-capable multiview volumetric three-dimensional display," *Applied Optics*, vol. 46, no. 8, pp. 1244–1250, March 2007.

[30] A. Jones, I. McDowall, H. Yamada, M. Bolas, and P. Debevec, "Rendering for an interactive 360 light field display," *ACM Transactions on Graphics*, vol. 26, no. 3, pp. 40:1–40:10, July 2007.

[31] P. C. Barnum, S. G. Narasimhan, and T. Kanade, "A multi-layered display with water drops," *ACM Transactions on Graphics*, vol. 29, no. 4, pp. 76:1–76:7, July 2010.

[32] G. Wetzstein, D. Lanman, W. Heidrich, and R. Raskar, "Layered 3D: Tomographic image synthesis for attenuation-based light field and high dynamic range displays," *ACM Transactions on Graphics*, vol. 30, no. 4, pp. 95:1–95:12, July 2011.

[33] G. E. Favalora, "Volumetric 3D displays and application infrastructure," *IEEE Computer*, vol. 38, no. 8, pp. 37–44, August 2005.

[34] K. Akeley, S. J. Watt, A. R. Girshick, and M. S. Banks, "A stereo display prototype with multiple focal distances," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 804–813, August 2004.

[35] S. Liu, H. Hua, and D. Cheng, "A novel prototype for an optical see-through head-mounted display with addressable focus cues," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 3, pp. 381–393, May 2010.

[36] N. Damera-Venkata and N. Chang, "Realizing super-resolution with superimposed projection," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2007)*, June 2007, pp. 1–8.

[37] F. Berthouzoz and R. Fattal, "Resolution enhancement by vibrating displays," *ACM Transactions on Graphics*, vol. 31, no. 2, pp. 15:1–15:14, April 2012.

[38] C. Rezk-Salama and A. Kolb, "Opacity peeling for direct volume rendering," *Computer Graphics Forum*, vol. 25, no. 3, pp. 597–606, September 2006.

[39] P. U. Tse, "Volume completion," *Cognitive Psychology*, vol. 39, no. 1, pp. 37–68, August 1999.

[40] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Transactions on Graphics*, vol. 21, no. 3, pp. 267–276, July 2002.

[41] P. Didyk, E. Eisemann, T. Ritschel, K. Myszkowski, and H.-P. Seidel, "Apparent display resolution enhancement for moving images," *ACM Transactions on Graphics*, vol. 29, no. 4, pp. 113:1–113:8, July 2010.

[42] P. Makela, J. Rovamo, and D. Whitaker, "Effects of luminance and external temporal noise on flicker sensitivity as a function of stimulus size at various eccentricities," *Vision research*, vol. 34, no. 15, pp. 1981–1991, August 1994.

[43] C. Lundstrom, P. Ljung, and A. Ynnerman, "Local histograms for design of transfer functions in direct volume rendering," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, pp. 1570–1579, November 2006.

[44] Y.-J. Su and Y.-Y. Chuang, "Gradient-free visualization with multiple light approximations." *Studies in health technology and informatics*, vol. 184, pp. 417–423, February 2013.

[45] M. Levoy, H. Fuchs, S. Pizer, J. Rosenman, E. Chaney, G. Sherouse, V. Interrante, and J. Kiel, "Volume rendering in radiation treatment planning," in *Proceedings of the First Conference on Visualization in Biomedical Computing*, May 1990, pp. 4–10.

Yan-Jen Su received the BS and MS degrees from National Taiwan Ocean University in 2002 and 2004, respectively, all in computer science and engineering, and is currently working toward the PhD degree at the Graduate Institute of Networking and Multimedia, National Taiwan University. His research interests include visualization and computer graphics.

Yung-Yu Chuang received the BS and MS degrees from National Taiwan University in 1993 and 1995 respectively, and the PhD degree from the University of Washington at Seattle in 2004, all in Computer Science. He is currently a professor with the Department of Computer Science and Information Engineering, National Taiwan University. His research interests include computational photography, rendering and computer vision. He is a member of the IEEE and the IEEE Computer Society.