Perceptually-Based Depth-Ordering Enhancement for Direct Volume Rendering

Lin Zheng, *Student Member*, *IEEE*, Yingcai Wu, *Member*, *IEEE*, and Kwan-Liu Ma, *Fellow*, *IEEE*

Abstract—Visualizing complex volume data usually renders selected parts of the volume semitransparently to see inner structures of the volume or provide a context. This presents a challenge for volume rendering methods to produce images with unambiguous depth-ordering perception. Existing methods use visual cues such as halos and shadows to enhance depth perception. Along with other limitations, these methods introduce redundant information and require additional overhead. This paper presents a new approach to enhancing depth-ordering perception of volume rendered images without using additional visual cues. We set up an energy function based on quantitative perception models to measure the quality of the images in terms of the effectiveness of depth-ordering and transparency perception as well as the faithfulness of the information revealed. Guided by the function, we use a conjugate gradient method to iteratively and judiciously enhance the results. Our method can complement existing systems for enhancing volume rendering results. The experimental results demonstrate the usefulness and effectiveness of our approach.

Index Terms—Volume rendering, depth ordering, depth perception, transparency, visualization

1 Introduction

Humans perceive the world in three dimensions, and can determine correct spatial relationships between objects based on several key depth cues [38]. 3D graphics researchers consider these depth cues in designing rendering methods. To visualize complex volume data composed of many surface layers and fine features, however, direct rendering methods sometimes fail to produce images that can unambiguously present the depth information of partially occluded structures that are rendered semitransparently. In this work, we show that by optimizing rendering based on a set of perceptual models, it is possible to effectively enhance the interpretation of depth ordering of volumetric and surface features in the resulting 2D images.

Correct depth perception helps users understand spatial relations among structures. For example, in medical visualization for presurgery planning, the correct inference of doctors on spatial relations between structures, such as the relation between important vessels and tumors, is critically important to avoid committing mistakes during surgery. Correct depth perception is also important in volume rendering, given that structures are usually rendered semitransparent to expose additional information. Without appropriate enhancement, semitransparent rendering may lead to misinterpretation of spatial relations between structures and may cause serious problems [5]. Although this problem could be alleviated to some extent by the interactions of users such as rotation, it ultimately relies on the ability or experience of users to perceive the depth relations correctly during their interaction with objects. Furthermore, in some typical scenarios, such as when a static image or poster is shown, user interaction is not available.

Researchers have introduced several methods to ease the difficulty of depth perception without user interaction. Perspective projection is the most common method [24] used to convey depth information in computer graphics and visualization. However, it is insufficient in revealing the correct depth ordering in volume rendering. Fig. 2b shows semitransparent layers presenting ambiguity. Researchers have proposed other methods to provide depth cues in volume rendering. For example, halos or shadows can be used in volume rendering to highlight foreground structures [6], [9], [14]. They work well only when the information of the front structures is required. However, as halos occlude part of the context, depth information is conveyed at the cost of background details.

Recent studies in visual psychology reveal that the perception of depth ordering of semitransparent structures are highly dependent on transparency (see Fig. 2). The *Adelson-Anandan-Anderson's X-junction model* (simply called X-junction model) [2] and its complement, *Transmittance anchoring principle* (TAP) [4], are two of the most important findings in the field. The models can quantitatively evaluate the quality of depth ordering perception of spatial structures based on the structure transparency and luminance. This inspires us to develop a new transparency optimization method with the models to enhance the depth ordering perception of spatial structures without introducing extra illustrative cues. To accomplish this, we set up a novel energy function to quantify the effectiveness of perceived depth ordering perception using the X-junction and TAP models, the effectiveness of transparency perception using

- *L. Zheng and K.-L. Ma are with the Department of Computer Science, University of California, 2063 Kemper Hall, One Shields Avenue, Davis, CA 95616-8562. E-mail: lzheng@ucdavis.edu, ma@cs.ucdavis.edu.*
- *Y. Wu is with Microsoft Research Asia, T2-13131, No. 5 Danling Street, Haidian District, Beijing, P.R. China, 100080. E-mail: ycwu@microsoft.com.*

Manuscript received 12 Aug. 2011; revised 10 Jan. 2012; accepted 30 May 2012; published online 12 June 2012.
Recommended for acceptance by G. Drettakis.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org, and reference IEEECS Log Number TVCG-2011-08-0184.
Digital Object Identifier no. 10.1109/TVCG.2012.144.

Metelli's episcotister model, as well as the faithfulness of the representation using a conditional information entropy measure. Our approach creates a desirable result by optimizing the energy function using a conjugate gradient optimization scheme. We believe that a volume rendered image with minimum ambiguity is necessary before any illustrative effect can be applied. The paper contributions are as follows:

- Investigation of a new problem of how to perceptually enhance the depth ordering of semitransparent structures by adjusting only transparency and luminance.
- Introduction of perception models for quantitative measurement of perceived depth ordering, and a new measure to assess the image quality with the models.
- Design of an optimization framework to enhance the depth ordering perception with inherent visual cues while preserving the original transparency and information.

2 Related Work

Researchers in computer graphics have employed various visual cues to enhance the depth perception of 3D models. However, some methods are not suitable for volume rendering, given that volume rendered images usually contain sophisticated features that are often rendered semitransparently. In this section, we only discuss previous work on depth ordering perception and enhancement of semitransparent features. Interested readers can refer to [11] for additional details on general depth enhancement in computer graphics.

Transparency perception has been studied in psychology for decades. Metelli's episcotister model [23] is the first quantitative model to evaluate visual perception of transparency. Recently, Singh and Anderson introduced the *transmittance anchoring principle* [33] to measure the transparency of layers in images by image contrast. Chan et al. [8] developed a system to optimize transparency automatically in volume rendering based on Metelli's episcotister model and TAP, thus improving the perceptual quality of transparent structures. The depth perception between the layers might be better or worse after the adjustment of transparency. Although our method also deals with transparency perception, we emphasize the adjustment of transparency to enhance the depth perception of semitransparent structures.

Transparency is an important visual cue in the perception of depth ordering for semitransparent structures. The X-junction model [2], [4] and its extensions [28], [32] utilize contour junctions in overlapping semitransparent structures to measure depth ordering perception. These models can be used to depict a special situation in which a user is likely to perceive the correct depth ordering. Several other situations with different configurations are considered to be ambiguous situations by the model. TAP [33] can also be used for depth perception. It can mainly handle one of the ambiguous situations. Based on this theory, the highest contrast region along a continuous contour appears to be the background and the regions with lower contrast are decomposed into multiple layers. X-junction model has also been discussed in volume rendering [5]. In our work, we combine the X-junction model and TAP to guide the optimization of transparency settings. In the previous experiment, when dynamic cues were applied, the ambiguity of depth perception was significantly reduced using motion parallax and perspective projections. In contrast, our work can reduce depth ordering ambiguity by adjusting the transparency and luminance of features.

Researchers have proposed the use of stereoscopic rendering to facilitate volumetric data exploration [7]. However, the usefulness of the stereoscopic technique is still controversial [17]. In monocular systems, shadows and halos have been widely used for depth enhancement in volume rendering. Yagel et al. [40] used recursive ray tracing to add shadows in volume visualization. Kniss et al. [20] proposed an interactive shading model based on volumetric light attenuation effects to incorporate volumetric shadows in volume rendering. Šoltészová et al. [36] proposed a new method called chromatic shadowing based on a shadow transfer function to handle the over-darkening problem, thus allowing for better perception of details in the shadowed areas. Adding halos is another useful technique for conveying depth information [6], [14], [27]. Bruckner and Gröller [6] introduced an approach that enables users to generate halos interactively for volume data using the halo transfer function. Everts et al. [14] presented a technique to create depth-dependent halos for 3D line data. Some researchers advocated the use of color to enhance depth perception [10], [37] in volume rendering. Chuang et al. [10] suggested reducing color saturation in occluded objects only in the overlap region while keeping its lightness. Wang et al. [37] used cold colors such as green, blue, or cyan to encode the foreground, and warm colors such as red, yellow or magenta to encode background. Texture has also been proposed to convey the 3D shape of semitransparent surfaces [18].

Shadows and halos are effective means to convey depth information for opaque surfaces, but come at the expense of occluding other structures. In comparison, our approach aims to enhance depth perception of semitransparent objects by adjusting transparency and luminance of features without additional visual cues. Our method is based on a quantitative depth perception model specifically designed for semitransparent surfaces. Thus, our approach can enhance the perception of depth ordering automatically and quantitatively.

Transfer functions are used to classify data in volume rendering based on data dimensions such as intensity, gradient magnitude [21], curvature [19] and/or local structures [31]. Researchers have also developed various approaches such as visibility histograms [12] and visibility driven visualization [35] to automatically highlight important structures by increasing their visibility in the resulting images. Wu and Qu [39] developed a method for optimizing transfer functions based on a user-specified objective function to enable users to directly edit features on volume rendered images. Takahashi et al. [34] proposed a multidimensional transfer function to emphasize nested inner structures in volume rendered images based on the relative geometric positions as well as the scalar field

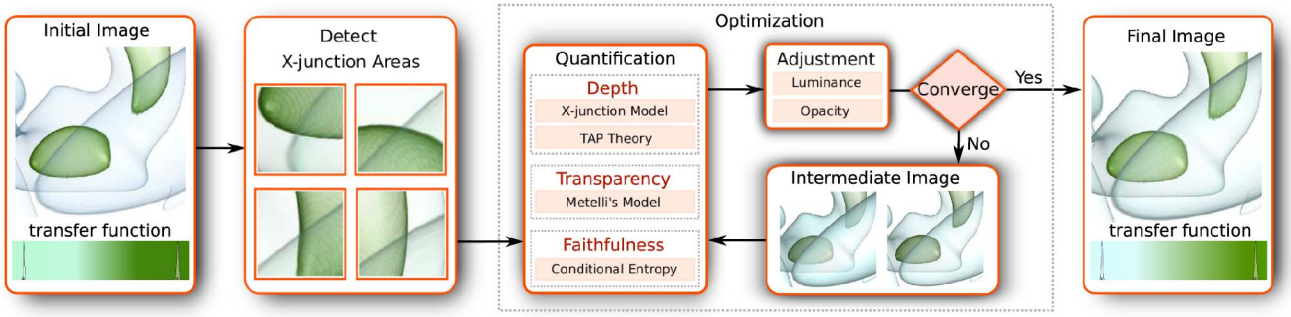

Fig. 1. System overview. Given an image produced with an input transfer function, we preprocess the image to detect X-junction areas. Next, the framework automatically optimizes the image quality through iteratively adjusting the transfer function based on three measures (depth, transparency, and faithfulness) until the optimization converges.

values of the structures. Our work can be viewed as an important complement to transfer function specification to improve depth ordering perception. To the best of our knowledge, we are the first to study the quantitative evaluation and enhancement of depth ordering perception in volume rendering.

3 System Overview

We design a fully automatic framework to enhance depth perception and to correct misleading quantitative orderings. It can measure the quality of semitransparent volume rendered images based on the effectiveness of perceived depth ordering and transparency, as well as the faithfulness of the information revealed. The complete pipeline of our system is presented in Fig. 1. Given an input data set with its initial transfer function, we first preprocess the data by rendering an initial image with current rendering parameters. Furthermore, we detect and record the junction areas in the current rendered image, which are used for later evaluation.

The framework uses a conjugate gradient method to search for an optimal result. We set up an energy function to quantify the effectiveness of candidate solutions during the optimization. The effectiveness quantification consists of three quantitative measures: depth, transparency, and faithfulness measures. The depth measure is for quantifying the difference between the depth ordering perceived by users and the actual depth ordering in the data. By minimizing this difference, we ensure that the correct depth ordering is perceived in a 2D image. The transparency measure evaluates the change in perceived transparency during the optimization based on Metelli's episcotister model. By minimizing this change, we can preserve the transparency perceived by users. The faithfulness measure is used to estimate the faithfulness of the information revealed in the newly generated candidate images. Guided by the energy function, we use a conjugate optimization scheme to optimize the transfer function to iteratively approximate an optimal result until we reach a specified termination condition.

4 Quantitative Perception Models

Despite recent advances in volume rendering, only a few methods exist for the enhancement of depth perception. These include halos, occlusions, and shadows, and each has its own limitations. For instance, occlusion is an intuitive and natural monocular depth cue, but is deemed inappropriate in volume rendering in which structures are rendered semitransparently. Halos and shadows [6], [9], [14] are helpful in some cases, but they could likely occlude contextual structures. To alleviate this problem, halos are usually combined with filtering.

Research in perception focuses on the physiological capabilities of the human visual system. In recent years, quantitative models have been introduced. In particular, several theories have been proposed with respect to the perception of transparency layers and depth ordering, which can be applied towards our goal of depth perception enhancement. Two theories (X-Junction model and TAP) suggest that enhancement of perceived depth ordering can be achieved simply by adjusting the opacity and luminance of the volume. In contrast to previous work, the addition of redundant parameters to make foreground objects stand out becomes unnecessary.

4.1 X-Junctions

Previous studies in visual psychology indicate that transparency and luminance provides an important cue for perceiving the depth ordering of multiple semitransparent layers. The X-junction model [2] reveals that the order of increasing or decreasing luminance value in an X-junction is crucial for depth ordering perception. The red point highlighted in Fig. 2a shows a typical X-junction surrounded by four areas p, q, r, and s. The luminance values of the areas are denoted by l_p, l_q, l_r, and l_s, respectively. The model determines whether a user can perceive the correct depth ordering of layers based on the contrast polarity arrangement of the areas. Reversing the contrast polarity arrangement represents the change of the luminance order in the regions. For example, in Fig. 2a, reversing means that $l_p > l_q$ and $l_r < l_s$.

The X-junction model defines three types of reversing junctions: *nonreversing, single-reversing, and double-reversing junctions*. We do not discuss double-reversing junctions in this research, since they do not allow for transparency and thus are unsuitable for volume rendering. Nonreversing junctions support the perception of transparency but leave the depth ordering ambiguous (see Fig. 2b). In these cases, $l_p < l_q$ and $l_r < l_s$, which means the vertical edge (green edge in Fig. 2a) retains the same sign in both halves of the X

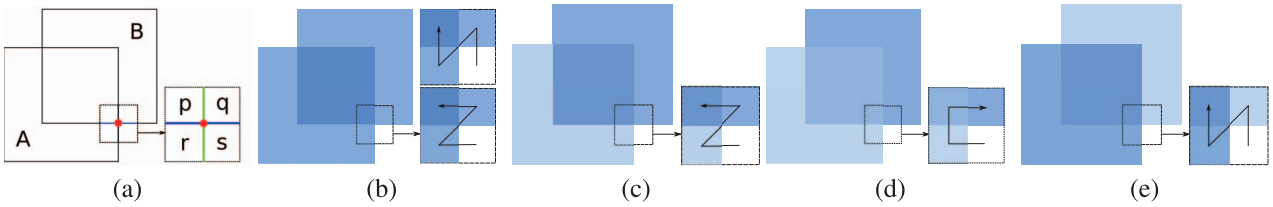

Fig. 2. (a) Illustration of an X-Junction marked by a red point. (b) A configuration. (c) Correct Z configuration. (d) C configuration. In this case, the depth ordering is unambiguous. (e) Wrong Z configuration.

junction; $l_p < l_r$ and $l_q < l_s$, which indicates the horizontal edge (blue edge in Fig. 2a) also retains the same sign in both halves of the X junction. As both edges retain their signs, we call this a "nonreversing" junction. The descending order of the luminance values composes a *Z configuration* (Fig. 2c). If $l_r = l_q$, then it changes to *A configuration* (Fig. 2b). As the depth orderings for these cases are ambiguous, they are also called *Z ambiguity* and *A ambiguity*.

In contrast, single-reversing junctions fully support the perception of depth order and can clearly reveal the correct depth ordering (Fig. 2d). In this case, $l_p > l_q$ and $l_r < l_s$. For the pairs of areas along the vertical edge, the descending order of the luminance changes across the X junction. In addition, we have $l_p < l_r$ and $l_q < l_s$, and the horizontal edge retains its sign. We call this a "single-reversing" junction. The descending order of luminance constructs a *C configuration* (see Fig. 2d). The X-junction model relies only on the relative luminance and transparency values around a junction. Thus, we can simplify our procedure by adjusting only the transparency and luminance of the features.

If only the luminance values of the four areas are known, the depth order of the layers is still left unknown. The X-junction model summarizes four possible local hypotheses about the situation of the frontal layer using these decision rules.

1. If $0 < (l_p - l_q)/(l_r - l_s) \leq 1$, it lies above the middle horizontal line in the X-junction.
2. If $0 < (l_r - l_s)/(l_p - l_q) \leq 1$, it lies below the horizontal line in the X-junction.
3. If $0 < (l_p - l_r)/(l_q - l_s) \leq 1$, it lies to the left of the middle vertical line in the X-junction.
4. If $0 < (l_q - l_s)/(l_p - l_r) \leq 1$, it lies to the right of the vertical line in the X-junction.

Conditions (1) and (2), as well as Conditions (3) and (4), are mutually exclusive unless the ratio is equivalent to one. Thus, we can first calculate the luminance of the areas around the junction for comparison, and then find the composition of the transparent layers using the four local hypotheses. Finally, we can derive the perceived depth ordering of the layers.

One of our goals is to simplify the process for users to perceive depth ordering information. Compared with existing methods that introduce additional information, a better solution is to rely on the transfer functions of volume rendering. During volume rendering, different features are assigned different opacities, colors, and luminances by the transfer functions. When we interact with the rendered volumes, features are usually shown as overlapping semi-transparent layers in 2D images, creating a set of X-junctions. Since opacity and luminance are the basic parameters of volume rendering, they can be utilized for depth enhancement without additional costs. Thus, the X-junction model can be used to estimate how a user perceives the depth ordering of semi-transparent structures.

4.2 Transmittance Anchoring Principle

When the X-junction conforms to the single-reversing situation, it is trivial to determine the depth ordering based on the C configuration. However, in the nonreversing situation (see Fig. 2c), the model fails to predict the perceived depth ordering. It has been reported that in this situation the frontal layer could still likely be distinguished by the human visual system in the Z configuration [33]. To overcome the limitations of the X-junction model, researchers have proposed to use the Transmittance Anchoring Principle model to handle Z ambiguity [33]. Thus, TAP can be regarded as a complement to the X-junction model. According to TAP, the human visual system perceives the regions of highest contrast along their internal or external contours as *transmittance anchors*. In other words, regions with the highest contrast are perceived to be in the background. Regions with lower contrast are perceived to be in the front. For the situation called A ambiguity in Fig. 2b, both areas r and s, and q and s have almost the same contrast. In this case, each layer has the same probability of being perceived as the frontal layer and existing psychological models cannot predict the perceived depth ordering.

4.3 Combination of X-Junctions and TAP

In volume rendering, ideal depth perception corresponds to the situation in which each X-junction can conform to the C configuration, providing clear visual depth cues. Additionally, the contrast between the areas must be high enough for human beings to notice. However, it is usually difficult to let all junction areas conform to the C configuration in practice. When some junction areas do not conform to the C configuration and exhibit the Z ambiguity, we should ensure that they at least conform to the correct Z configuration to resolve the ambiguity. As the TAP model can complement the X-junction model to deal with this situation, we judiciously combine X-junction and TAP models to quantitatively measure the quality of depth ordering perception.

We first use the X-junction model to estimate the perceived depth ordering. If this model fails, which means that depth ambiguity occurs, we then use the TAP model to deal with the Z ambiguity situation. Since TAP is sensitive to the contrast between p, q, r, s areas, it can help us decide the perceived depth ordering when Z ambiguity presents. For

example, in Fig. 2b, the panel in the lower left corner is in the front, representing A ambiguity. If the luminance value of region r is lower than region q (region r is darker), we will perceive an incorrect depth ordering based on the TAP theory, which indicates that the panel in the lower left corner is in the back. We can increase the luminance value of region r and decrease the luminance value of region q. After the Z configuration, which can be resolved by the TAP theory, has been achieved, we proceed to adjust the luminance and opacity and further improve it to C configuration.

5 Quality Optimization

Without enhancement, typical volume rendered images could often mislead users because the perceived depth ordering of features does not always clearly present the actual spatial relationship. For example, in Fig. 2b, the depth ordering between the semitransparent structures is ambiguous. It is difficult to tell which layer is in the front. Additional visual cues such as halos may help, but come at the cost of discarding information. To address this problem, we design an optimization system for automatically and quantitatively enhancing depth perception with a minimal change to the perceived transparency and visual display.

The system requires an initial transfer function to create a preliminary image result. Junction points for each pair of overlapping semitransparent structures are identified to derive the perceived depth relations using the perception models described in Section 4. We set up an energy function for measuring the quality of volume rendered images according to the quality of depth ordering perception and transparency perception as well as the faithfulness of information revealed. Guided by this function, our system can effectively search for optimal rendering parameters to produce a desired result.

5.1 Quality of Depth Perception

This section introduces a quantitative approach based on the X-junction model and TAP for measuring the quality of the perceived depth ordering in a volume rendered image.

5.1.1 X-Junction Detection

Junction points are required by the X-junction model to determine how a user perceives the relative spatial depth between two semitransparent structures. We employ a well-established algorithm based on color signatures [30] to identify X-junction points on an image because of its simplicity and decent performance. Color signatures can represent distributions of features and better characterize the appearance of an image [29]. A color signature is a set of ordered pairs $\{(x_1, v_1), (x_2, v_2), \ldots, (x_n, v_n)\}$, where v_i indicates a vector in a color space and x_i is the weight of v_i. Compared with histograms with fixed bins, color signatures adapt to the data and they do not arbitrarily partition the color space. One color signature can represent a set of feature clusters in one area. In the following, we briefly describe the X-junction detection method. See [30] for a more detailed description.

The basic idea of the algorithm is as follows: given a preliminary image, the algorithm examines every pixel of the image. It repeatedly splits the neighboring region of a certain pixel (see the red pixel in Fig. 4d) at different orientations. For every split, each half of the region is represented by one color signature. If the region is split by the green dash line (see Fig. 4d), then each color signature contains only one feature cluster. If the region is split by the blue dash line (see Fig. 4d), there are two feature clusters for one color signature. The algorithm employs the measurement based on Earth Mover's Distance (EMD) to estimate the perceptual color distance between neighboring color signatures. The farther the two color signatures are, the stronger strength (EMD value) a certain direction has. The orientation with the maximum EMD value is most likely to be a real edge.

The abnormality of EMD values can help to identify junction points. Abnormality represents the minimum EMD value at a certain pixel. The two color signatures are identical (EMD is equal to zero) at an orientation normal to an ideal edge (see the horizontal blue line across the central red point in Fig. 4d). Due to inhomogeneous spatial distribution of colors, the minimum EMD (the abnormality) at an edge pixel is never exactly zero in practice. Ruzon and Tomasi [30] suggest that when the abnormality at an edge pixel is abnormally high compared with its neighborhoods (within its $5X5$ grid) the pixel should be a junction.

Although this method can effectively detect junction points on an image, it has to compute color signatures for every pixel in the image, which is time consuming. To remedy this problem, we first extract the contours with Canny Edge Detector, and then search along the contours for junction points with high abnormality. When we identify a junction, the surrounding regions of the junction can be easily obtained by a region growing method since we have already known the edges (boundaries of the regions) near the junction. To illustrate, in Fig. 2b, there is one region for feature A, one for feature B, one for empty space, and one for the overlapping region of A and B. Thus, if the neighboring pixels of any given pixel can be classified into the four regions, it can be regarded as an X-junction point. Additionally, we can determine the depth relation between features A and B of the junction point by casting a ray into the overlapping region.

5.1.2 Depth Ordering

We formulate an energy term called depth ordering to measure the perceptibility with respect to the depth ordering of semitransparent structures. This formulation requires an accurate identification of junction configuration types in accordance with the depth perception models described in Section 4. We start with an X-junction and examine its four surrounding regions to determine a corresponding junction configuration. Different energy values are assigned to different junction configurations. The minimum energy indicates a scenario in which depth information can be correctly conveyed, while the maximum energy indicates an ambiguous scenario. Assume that object A is in the front of object B in Fig. 2a. Areas p and r belong to object A, and areas q and p belong to object B. We define five luminance situations based on the luminance configuration of p, q, r, and s.

1. **Wrong *C* configuration.** The luminance descending order is $l(s) > l(q) > l(p) > l(r)$. In this situation, the object B in the back is brighter than the object A in the front, such that B appears to be in the front, thus misleading a viewer. The overlapping area p is brighter than r, even though it belongs to the same object as r does. When the opacity and luminance values of A and B are changed, $l(r)$ and $l(q)$ will be changed accordingly. The overlapping area p will be affected by both the changes in A and B. During the ray-casting process, the opacity and luminance values of this area are accumulated in each sample point along the rays that are going through object A and B. If the front object is darker than or equal to the back object, the overlapping area p will always be darker than area r. In other words, area p is always the darkest area in the X-junction given that B is brighter than or equal to A. Thus, in semi-transparent volume rendering with raycasting, situation (1) will never happen.
2. **Wrong *Z* configuration.** The luminance descending order is $l(s) > l(q) > l(r) > l(p)$. Fig. 2e shows a wrong Z configuration where B is brighter than A, thus leading to Z ambiguity that cannot be handled by the TAP theory. To resolve this ambiguity, the luminance of object A should be increased ($l(r) \uparrow$) and/or the luminance of object B should be decreased ($l(q) \downarrow$).
3. ***A* configuration.** The luminance descending order is $l(s) > l(q) = l(r) > l(p)$. By changing the luminance value of object A and/or B, once $l(q) = l(r)$, we could be trapped in A ambiguity (see Fig. 2b). In this case, we must continue to adjust the luminance values by increasing $l(r)$ and decreasing $l(q)$ to leave the A configuration.
4. **Correct *Z* configuration.** The luminance descending order is $l(s) > l(r) > l(q) > l(p)$. By further increasing the luminance of object A and/or decreasing the luminance of object B, we obtain correct Z configuration when object A is brighter than object B. As shown in Fig. 2c, TAP proves that human beings can figure out the correct depth ordering in this situation. In this stage, we need to inspect the luminance of area p. If $l(p)$ is no longer the darkest area, the descending order of the four areas in the X-junction changes to that of the correct C configuration.
5. **Correct *C* configuration.** The luminance descending order is $l(s) > l(r) > l(p) > l(q)$. This configuration is consistent with the X-junction theory and viewers can correctly figure out the real depth ordering. However, if the contrast between the areas is too weak for human beings to notice, it is a weak C configuration. A strong C configuration with high contrast is preferred.

In situation (2), (3), and (4), $l(p)$ is not needed for calculating the energy value of the depth term. We only compare the luminance value between $l(r)$ and $l(q)$. A wrong Z configuration can convey incorrect depth ordering information. Thus, it is worse than an A configuration. The energy value of the wrong Z configuration should be higher than that of the A configuration. A correct Z configuration, on the other hand, is better than the A configuration. Thus, the correct Z configuration has a lower energy value. In situation (5), we rely on the descending order of the local luminance to derive the depth order of the objects. In a correct C configuration, which conveys correct depth ordering, the luminance differences, i.e., $|l(s) - l(r)|$, $|l(r) - l(p)|$, and $|l(p) - l(q)|$, should be high enough for human beings to perceive readily. Higher contrasts of luminance between areas are preferable. Thus, a strong and correct C configuration has the lowest energy value.

Based on the observation, only $l(r)$ and $l(q)$ are involved for measuring the energy in situation (2), (3), (4) , while all $l(p)$, $l(q)$, $l(r)$, $l(s)$ are involved in situation (5). Therefore, we can divide the energy function into two parts, which can be expressed by a piecewise-defined function.

- In the first part, by adjusting the luminance of object A and B, the situation is changed between (2), (3), and (4). We inspect the luminance of p to ensure that it is the darkest area. Note that $l(p)$ is not involved.
- In the correct Z configuration, once the luminance of p becomes higher than that of q, the energy function switches to the second part for the correct C configuration. Note that $l(p)$ is involved.

The energy value for the whole function ranges from 0 to 1 and can be divided into two parts subsequently, i,e., into two ranges $(0.5, 1)$ and $(0, 0.5)$.

For the first part, only the contrast d_1 between r and q, i.e., $d_1 = l(r) - l(q), d_1 \in (-1, 1)$ affects the energy value. We use the hyperbolic tangent function: $y = \tanh(x) = (e^x - e^{-x})/(e^x + e^{-x})$ to estimate the first part energy for situations (2), (3), and (4), as illustrated in Fig. 4a. With this function, we can easily model the rapid changes of the energy from situation (2) to (4) via (3). The first part of the energy function ranging from 0.5 to 1 can be formulated as

$$e_{d_1} = \frac{3}{4} + \frac{1}{4} tanh(-d_1) = \frac{3}{4} + \frac{1}{4} tanh(l(q) - l(r)). \quad (1)$$

When $l(r) = l(q)$, we have $y = 3/4$ as a constant anchor point for the A configuration. When $l(r) < l(q)$, we have $y \in (3/4, 1)$ for the wrong Z configuration. When $l(r) > l(q)$, we have $y \in (1/2, 3/4)$ for the correct Z configuration.

As for the second part (the correct C configuration), the junction showing weak contrast should have a higher energy value than the junction showing strong contrast. In other words, the higher the contrast, the lower the energy value. The contrast of an X-junction can be estimated as the average luminance difference between $l(s)$, $l(r)$, $l(p)$, and $l(q)$. The average luminance difference can be determined by

$$d_2 = \sqrt{\frac{\|l(s) - l(r)\|^2 + \|l(r) - l(p)\|^2 + \|l(p) - l(q)\|^2}{3}}.$$

The second part of the energy function can be calculated by

$$e_{d_2} = \frac{1}{2} \cdot tanh\left(\frac{1}{d_2} - 1\right). \quad (2)$$

Fig. 4b illustrates this function. In this function, the energy value decrease smoothly with d_2 when the contrast d_2 is low, and decrease dramatically with d_2 when the contrast d_2

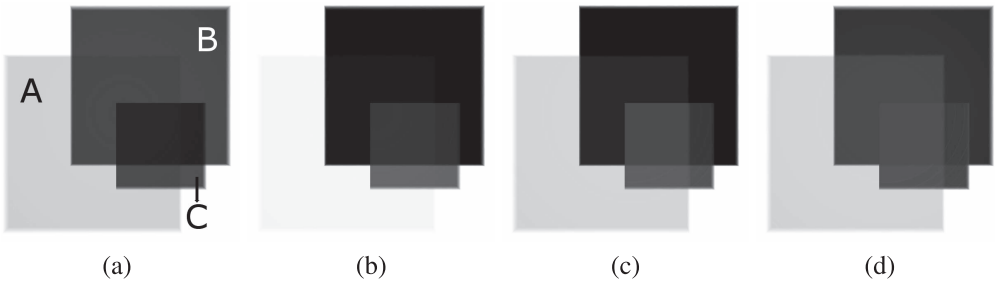

Fig. 3. Illustration of the optimization process. (a) A preliminary volume rendered image showing three panels (A, B, and C) whose depth order is ambiguous (the real front-to-back order is A, C, and B). (b) An intermediate result of depth ordering enhancement without preserving the transparency and considering image faithfulness. (c) An intermediate result of depth ordering enhancement without considering image faithfulness. (d) A final result where the image quality is optimized in terms of depth ordering and transparency perception as well as the information faithfulness.

is high. When $d_2 = 0$, $e_{d_2} = 0.5$. When $d_2 = 1$, $e_{d_2} = 0$. Thus, the function is more sensitive to the changes of higher contrast, which allows the optimization to terminate earlier when the junction has higher contrast.

With (1) and (2), the energy function for measuring the quality of depth ordering perception of an X-junction i can be defined as

$$e_d(i) = \begin{cases} e_{d_1}, & \text{if } l(q) \geq l(p), \\ e_{d_2}, & \text{otherwise.} \end{cases} \tag{3}$$

Let n_j be the number of X-junctions in a volume rendered image. The image quality in terms of depth ordering perception can be formulated as

$$E_d(\Omega) = \frac{1}{n_j} * \sum_{i=1}^{n_j} e_d(i), \tag{4}$$

where Ω indicates the parameter space.

5.2 Transparency

Transparency is regarded as an important factor in the optimization process. It plays an important role in the perception of depth ordering of semitransparent structures. In our optimization, we enhance the perception of depth ordering by adjusting feature opacity in the transfer function. Hence, we change the transparency perceived by users to achieve our goal. For instance, Fig. 3a is a volume rendered image in which three semitransparent structures are present with misleading depth information. If we optimize the image by adjusting the transparency without any constraint, the likely result is shown in Fig. 3b, in which the bottom-left structure becomes almost transparent and unrecognizable. To address this problem, we formulate an energy term to measure the difference between the perceived transparency of the image before and after each opacity adjustment.

Using the general raycasting technique of volume rendering, we add up the opacity value until it is completely opaque based on this formula: $T(s) = exp(-\int_0^s \tau(t)dt)$. This formula can represent the physical properties of the volume. However, it cannot reveal how a user's visual system perceives transparency [22]. Thus, using only a physical model is insufficient for measuring the perceptual quality of transparency. To address this problem, we also use a perceptual model called Metelli's episcotister model [23] to evaluate the perceived transparency. In Fig. 4c, A and B form a bicolored background with a semi-transparent disk T in the front. This disk is perceptually separated into regions P and Q by the background. This perceptual model is defined as follows:

$$\begin{aligned} p &= \alpha a + (1-\alpha)t, \\ q &= \alpha b + (1-\alpha)t, \\ \alpha &= \frac{p-q}{a-b}, \end{aligned} \tag{5}$$

where a, b, p, and q are the reflectances of the respective regions, and t is the reflectance of the transparent layer T. We use luminance instead of reflectance, because luminance is more intuitive for the human visual system, and many physical situations can be combined to form a single mapping from reflectance to luminance [1]. The perceived opacity is interpreted as $1 - \alpha$. We measure the variation of the feature opacity using the following equation:

$$E_t(\Omega) = \frac{1}{n} \cdot \sum_{i=0}^{n-1} \|\alpha_i' - \alpha_i\|, \tag{6}$$

here n is the number of features, and α_i and α_i' indicate the original perceived opacity and the adjusted opacity of the ith feature in the volume, respectively. Using this energy term, we can largely maintain the feature transparency perceived by users. Figs. 3c and 3d show two optimized results where both depth and transparency information are preserved.

5.3 Image Faithfulness

Image faithfulness is another energy term that we establish for optimizing image quality. Merely preserving visual properties of structures in a volumes is insufficient because information can be lost during optimization. Information on an optimized image should be consistent with that of the original image. Fig. 3c shows an example in which information is discarded, although the perception of depth is enhanced with only a minimal change in transparency perception. We use an information-theoretic measure to quantify the image faithfulness (or image similarity). Specifically, we employ the conditional entropy method to measure how faithful the optimized image is to the original version.

Conditional entropy has been widely used in image processing, computer vision, and pattern recognition for image registration and comparison [13]. Let X be a discrete random variable with support $\mathcal{X}$ and Y be a discrete

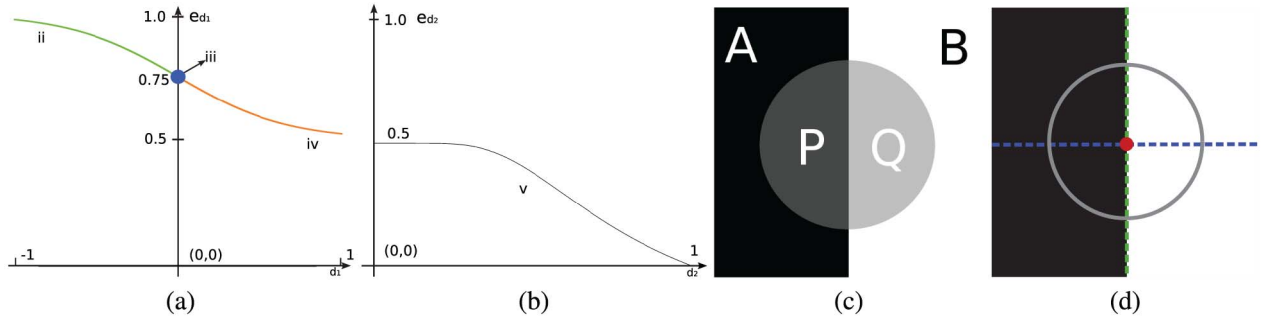

Fig. 4. (a) Illustration of e_{d_1}. The green and orange curves represents the wrong Z configuration (ii) and the correct Z configuration (iv), respectively. The blue point connecting the green curve to the orange curve indicates the A configuration (iii). (b) Illustration of e_{d_2}. The black curve denotes the correct C configuration (v). When $d_2 \to 0$, which means $\frac{1}{d_2} \to \infty$, the e_{d_2} approaches 0.5. (c) Illustration of the Metelli's model. (d) Illustration of the junction detection.

random variable Y with support $\mathcal{Y}$. In information theory, the conditional entropy $H(Y|X)$ for Y given X can be measured by

$$H(Y|X) = \sum_{x \in \mathcal{X}} p(x) H(Y|X = x) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x,y) log \frac{p(x)}{p(x,y)}.$$

where $p(x)$ is the probability mass function of outcome x and $p(x,y)$ is the joint probability of outcomes x and y. We can assume that $\mathcal{X}$ represents the original image and $\mathcal{Y}$ denotes the adjusted image. Thus, $p(x)$ denotes the probability distribution of gray value $x \in [0, 255]$ in image $\mathcal{X}$ [26], and $p(x,y)$ represents the joint distribution of gray values x and y. To estimate $p(x)$, we convert $\mathcal{X}$ to grayscale and then build a histogram of the grayscale image to obtain $p(x)$. The joint distribution $p(x,y)$ can be obtained by converting $\mathcal{X}$ and $\mathcal{Y}$ to grayscale and then build a two dimensional joint histogram for the two grayscale images. Lower $H(Y|X)$ indicates that X shares more similarity with Y. Thus, to preserve information during optimization, we minimize the conditional entropy $H(Y|X)$. We normalize the conditional entropy and obtain the energy term for image faithfulness, $E_e(\Omega)$. Using this image faithfulness term, we can preserve information during the optimization process. Fig. 3d shows a result generated while taking image faithfulness into account.

5.4 Optimization

We can derive the total energy for a volume rendered image using a weighted sum of depth, transparency, and image faithfulness energy as follows:

$$E(\Omega) = \sum w_d \cdot E_d(\Omega) + w_t \cdot E_t(\Omega) + w_e \cdot E_e(\Omega). \quad (7)$$

where w_d, w_t, and w_e represent the weights for each energy term. These weights can be flexibly adjusted by a user to meet his specific requirement. The higher a weight a user assigns, the more important role the associated energy plays during optimization. In our experiments, we found that $w_d = 3$, $w_t = 1$, and $w_e = 1$ produce desirable results that can enhance the depth ordering perception while the transparency and image faithfulness levels are largely preserved. With the total $E(\Omega)$, we can transform the image quality enhancement problem into an optimization problem by minimizing the total energy.

Finding a global optimum of the energy function is very difficult because an analytical form of the function cannot be easily derived. One solution is to perform stochastic search rather than exhaustive search in the parameter space. However, stochastic search algorithms are usually expensive and do not always converge to an optimal solution. We employ a nonlinear conjugate gradient method to solve the problem [12]. The transfer function can be first defined as a mixture of Gaussians

$$T(x) = \sum_i \alpha_i G_{\mu_i, \sigma_i}(x),$$

where G_{μ_i,σ_i} is a Gaussian function with μ and σ denoting the mean and the standard deviation, respectively. This Gaussian mixture model allows for creating good transfer functions without the appearance of aliasing artifacts using only a few parameters. Thus, the parameter space Ω is defined as $\{\alpha_i, \mu_i, \sigma_i, l_i : i \in [1 \cdots N]\}$ for N Gaussian functions where l_i represents the lightness associated with Gaussian i.

Our method searches for the optimal solution iteratively in the parameter space following a greedy mechanism:

$$\Omega_{n+1} = \Omega_n + \gamma_n \Lambda_n,$$

where Λ_n represents the search direction in the space that can reduce the energy. If we let $\Lambda_n = -\nabla E$, the method becomes a steepest descent method that iteratively searches for the optimal solution in the direction opposite to the gradient of the energy. However, this method converges slowly and can get trapped in a local minimum. Our method can address these problems by searching in the direction conjugate to the ones previously explored and resetting the search direction to the steepest direction when only little energy is reduced.

6 Experiments

We have conducted experiments on a PC with a 2.40 GHz Intel Core(TM)2 Quad CPU, 3 GB of memory, and an NVIDIA GTX 280 graphics card with 1 GB of memory. We tested four types of data sets to demonstrate the usefulness of our perception-based approach to improving volume visualization. Ray-casting volume rendering was used to generate image results. We fixed the rendering resolution at 512×512 in our study. Depending on the sizes of the volume data, the time needed to create the following results ranged from 5 to 20 seconds.

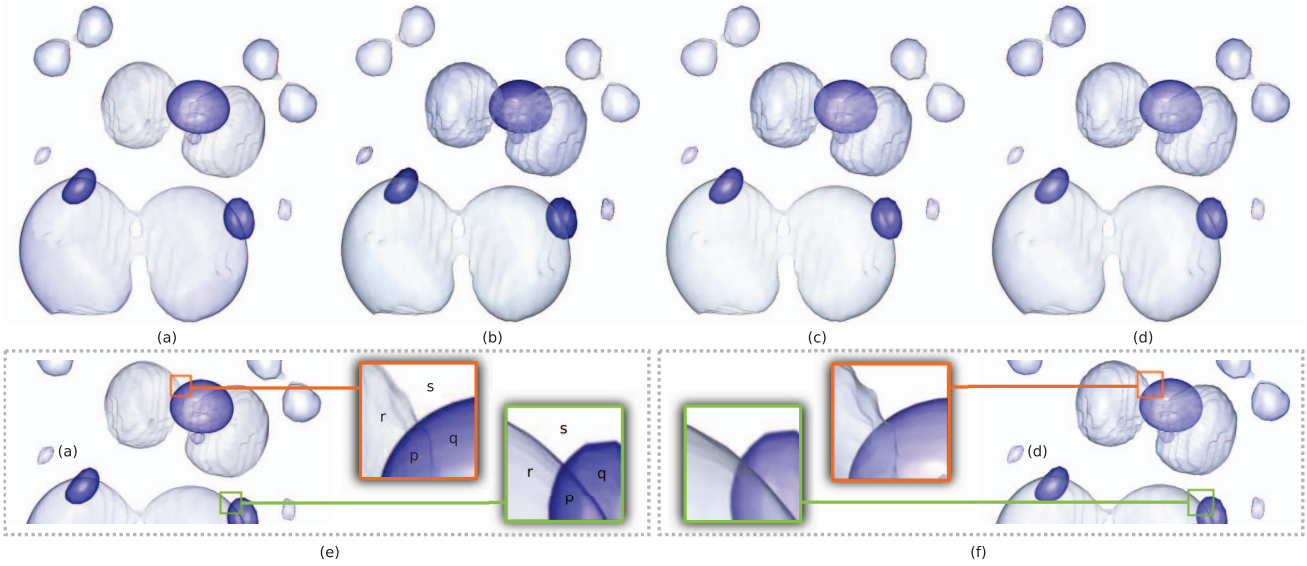

Fig. 5. Experiment on an Neghip volume. (a) An original image in which the depth relations between structures are misleading. (b) An image enhanced by our optimization approach based on merely the depth ordering energy. (c) An image enhanced by our optimization approach based on both the depth ordering and transparency energy. (d) An image enhanced by our optimization approach based on the depth ordering, transparency, and image faithfulness energy. (e) Illustration of two X-Junctions in (a). The X-junction in the orange rectangle shows a weak C configuration. The other in the green rectangle reveals an Z configuration. (f) Illustration of two selected X-Junctions in (d).

In our first experiment, we tested our approach with a simulated protein molecule data set (Neghip). We started by rendering a semitransparent isosurface with the initial luminance and opacity level. As shown in Fig. 5a, the spatial relations between objects are rather confusing. Clearly, it is difficult for a user to discern the correct depth ordering in the bottom two pairs of objects. Regarding the top objects, users may incorrectly perceive that the smaller sphere surrounded by two big spheres is in the front. According to the X-junction model, all these ambiguous and misleading cases can be automatically identified. Fig. 5e illustrates two representative X-junctions. The one in the orange rectangle conforms to the C configuration. However, the contrast between region q and region p is too low to be clearly observed by the human visual system. To better convey the depth ordering, the approach increased the luminance value of the front feature and decrease the luminance value of the back feature to adjust region p, resulting in an enhanced C configuration (see Fig. 5f). Let l_p, l_q, l_r, and l_s denote the luminance values of region p, q, r, and s, respectively. The X-junction in the green rectangle conforms to the A configuration in Fig. 5e. We can see that l_p is the lowest one, while l_r is higher than l_q. Since region r is brighter than region q, observers can still perceive the correct depth order based on the TAP theory. To achieve better result, the approach increased l_r and decreased l_q. We also increased l_p, such that $l_p > l_q$. This results in arranging the luminance of the four regions in an order as follows $l_s > l_r > l_p > l_q$, thus conforming to the C configuration, as shown in Fig. 5f.

The original depth energy value in Fig. 5a computed by our measure is 0.71, indicating that the perceived depth order in this volume rendered image was regarded as being inconsistent with the real depth ordering. Our quantitative optimization framework remedied this problem by iteratively searching for better rendering parameters to reduce the energy value. Fig. 5b shows an optimized result with a lower energy value 0.33. From this result, we can see that the depth ordering of the objects is much clearer. However, the perceived transparency was greatly changed from Fig. 5a, and the computed transparency energy for Fig. 5b is 0.73, which indicates that the transparency level was greatly altered from the original one. Thus, we added one more constraint for transparency to the optimization to ensure that the change in transparency could be minimized. Fig. 5c presents the new result, in which the depth order is clear and the transparency is similar to Fig. 5a. However, Fig. 5c is still suboptimal. Due to the changes in luminance and opacity, the information displayed in Fig. 5c is less than that of Fig. 5a. We can see that the detailed shape information revealed by shading in Fig. 5a is missing in Fig. 5c. After adding the information constraint to the optimization process, we obtained a satisfactory result, shown in Fig. 5d. In this figure, we can see that the depth ordering of spatial objects is clear, while the transparency level and the amount of information are maintained. The computed energy values are 0.40, 0.07, and 0.08 for the depth, transparency, and information energies, respectively. From this experiment, we see that our method can effectively optimize the quality of volume rendered images. Table 1 shows the measured energy at each step (Figs. 5a, 5b, 5c, and 5d) of the optimization for the depth, transparency, and image faithfulness, respectively.

TABLE 1
Energy Optimization of Neghip Data Set

Image	Depth	Transparency	Image faithfulness
Figure 5(a)	0.7104	Ø	Ø
Figure 5(b)	0.3312	0.735	0.1328
Figure 5(c)	0.3978	0.039	0.1177
Figure 5(d)	0.4035	0.073	0.0819

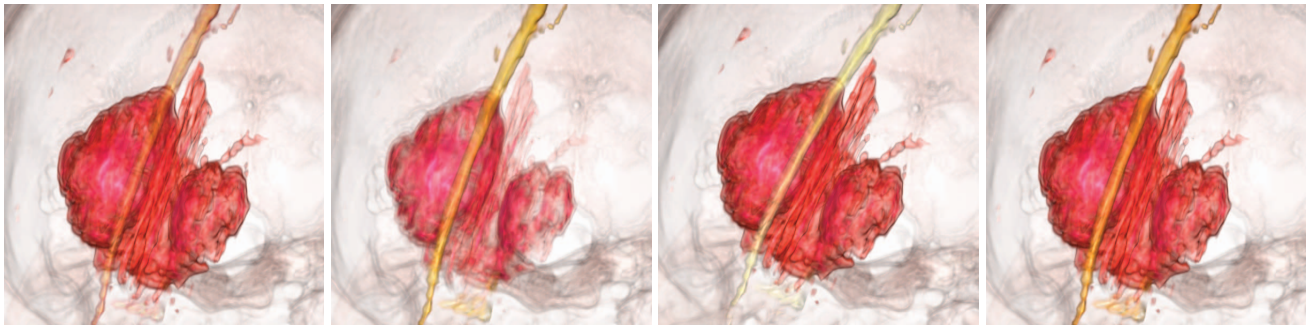

Fig. 6. Experiment on a brain tumor volume. (a) An original volume rendered image without any enhancement. (b), (c), and (d) Images enhanced using only the depth energy, both the depth and transparency energy, and all three depth, transparency, and information energy, respectively.

The second experiment demonstrates how our technique may be applied to medical volume visualization. Direct volume rendering has been widely used in medical diagnosis and presurgery planning. However, presenting multiple semitransparent structures in an image to a doctor without enhancement may increase the risk of making an incorrect decision caused by the ambiguity in depth ordering. Fig. 6 shows a medical volume (a brain tumor data set with size $256 \times 256 \times 60$). Fig. 6a shows a typical scenario using a transfer function, in which the depth relation between the blood vessel and the tumor is unclear. It would be difficult for a doctor to identify the correct depth order of these tissues. This problem can be addressed using our method. Figs. 6b, 6c, and 6d present three enhanced results using only the depth energy, both the depth and transparency energy, and all three depth, transparency, and information energy, respectively. Clearly, the depth order in these results are unambiguously revealed (the blood vessel is in the front). We can see that the image quality of Figs. 6c and 6d are better than in Fig. 6b. Overall, Fig. 6d has the best image quality because it displays the same information as Fig. 6a, using a similar transparency level, while conveying a clear depth ordering.

The third experiment shows the usefulness of our approach using flow visualization in a vortex data set ($128 \times 128 \times 128$). In flow visualization, it is often the case that multiple semi-transparent, overlapping structures are shown simultaneously. Incorrectly perceived depth orderings may lead to totally different analysis results. Fig. 7a shows an original volume rendered image. It has nested structures. Structures with higher intensity values are wrapped by structures with lower intensity values. Fig. 7e illustrates two X-junctions selected from Fig. 7a. Although these two junctions have the correct C configuration, the average luminance differences of p, q, r, and s in these two junctions are very small, resulting in two weak C configuration. Therefore, depth relations appear quite ambiguous in this image; users may incorrectly perceive that the larger structure (in light blue) in the center is behind its neighboring green layers. Fig. 7b presents a result enhanced only using depth energy. After we increased the opacity and luminance of the light blue structure, the depth order

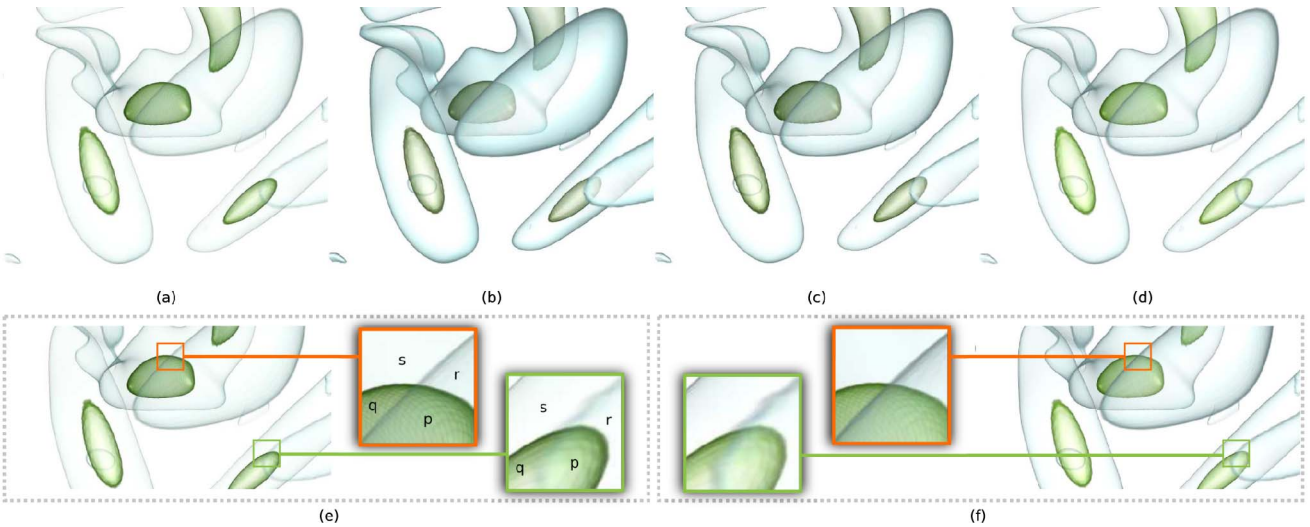

Fig. 7. Experiment on a Vortex volume. (a) An original volume rendered image in which spatial relations between structures are unclear. (b) An image enhanced by making the structures in the front more opaque, which is undesired. (c) An image enhanced by adjusting the structures in the front to be less opaque, considering both the depth ordering and the transparency perception. (d) An image enhanced by the optimization approach taking into account all depth ordering, transparency, and image faithfulness energy in the optimization process, thus creating a desired result. (e) Illustration of two selected X-junctions in (a) showing the weak C configuration which should be strengthened. (f) Illustration of two selected X-junctions in (d) showing the strong C configuration.

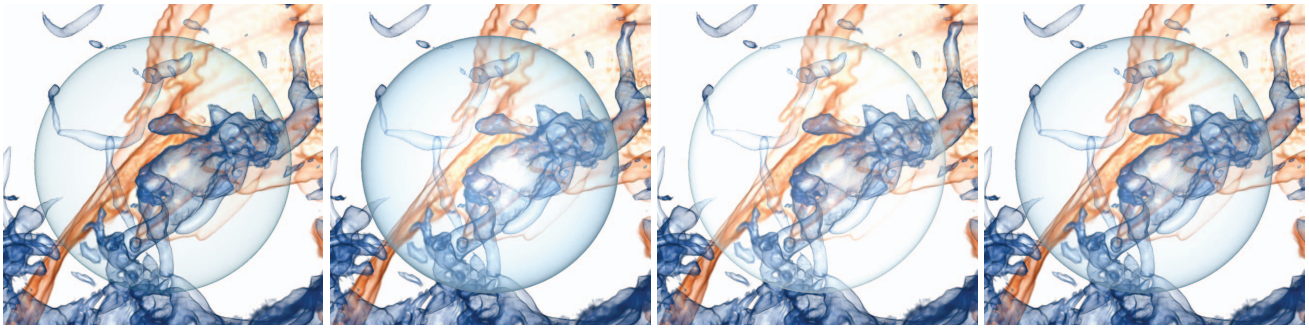

Fig. 8. Experiment on an astrophysical simulated data set. (a) An original volume rendered image without any enhancement. (b), (c), and (d) Images enhanced using only the depth energy, both the depth and transparency energy, and all three depth, transparency, and information energy, respectively.

was clearly revealed but came at the cost of changes in transparency and information. We remedied the transparency issue by taking into account the quantitative transparency change in the optimization process, thus creating Fig. 7c, which has a similar transparency level to Fig. 7a. We further enhanced Fig. 7c by preserving information faithfulness, such that the amount of information revealed in Fig. 7d is similar to that of Fig. 7a. Fig. 7f illustrates the X-junctions selected from Fig. 7d that have higher average luminance differences of p, q, r, and s. Therefore, after the enhancement the X-junctions conform to the strong C configuration. This experiment demonstrates the use of our technique in flow visualization to enhance perceptual image quality and to prevent misleading information representation.

Finally, we also tested our technique using a simulated astrophysical data set ($600 \times 600 \times 600$). This type of data usually contains sophisticated structures with many semi-transparent structures, making it challenging for users to analyze using traditional transfer functions. Fig. 8a shows a typical volume rendered image from the data. In this image, it is unclear whether the sphere structure is in the front or in the back. Figs. 8b, 8c, and 8d show several intermediate steps of our optimization process. Fig. 8b is the result of improving the depth ordering of the structures. Clearly, the sphere is in front of all of its neighboring structures. However, this result is undesirable because the appearance of the sphere changes. We subsequently optimized for transparency and information faithfulness using our techniques and created the results shown in Figs. 8c and 8d. In the final result (Fig. 8d, we can see that the perception of depth ordering has been improved while transparency and information levels are well preserved. This experiment shows how our system can be applied to enhance the perception of depth ordering in astrophysics data.

7 User Study

With the X-junction model and TAP theory, we have developed a method for enhancing the perception of depth ordering by automatically adjusting the opacity and lightness in volume rendering. We conducted a user study to analyze whether the depth perception is improved with the enhancement.

7.1 Hypotheses

To conduct the user study, we made two hypotheses.

1. Our approach improves the accuracy of depth ordering perception.
2. Our approach allows users to spend less time on perceiving the depth ordering.

7.2 Participants

We recruited 12 participants (seven females, five males; six with glasses, six without glasses) on campus. Two of them are designers and the remaining 10 are students studying information visualization(3), visual analytics(2), economics(2), statistics(2), and biochemistry(1). We avoided using subjects who have worked on volume visualization to guarantee that all subjects have no prior knowledge about the depth ordering.

7.3 Study Design

A quantitative user study was designed to validate the effectiveness of our method on different kinds of volume data sets including medical data (brain tumor) and scientific simulated data (simulated astrophysical data set and neghip data set). These data sets cover representative feature types and structures found in volume visualization, which allowed us to assess the general applicability of the two models. We conducted a between-subjects study by having each subject tested only once per image for either the enhanced or the original result.

Fig. 9 shows a snapshot of the task window. The volume rendered image is shown on the left window. Its individual features are shown on the right, from which, the subject was instructed to select the feature perceived to be in front with mouse click. The subjects received a short tutorial (two minutes in general) before the experiment.

During the experiment, each subject was exposed to five images from different kinds of data sets. The order of the images was randomized. For each image, either the enhanced version or the original version was randomly shown. In that way, we could exclude the effect of learning, since each subject is tested only once either with the enhanced version or the original version for each image. The time subjects spent for each image was recorded. Finally, we had 60 cases (12 subjects $\times 5$ images): 30 enhanced cases and 30 original cases. For each case, we collected the following attributes.

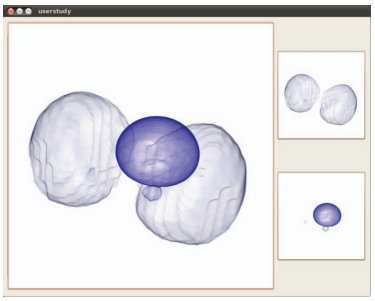

Fig. 9. A snapshot of the task window. The volume rendered image is shown on the left window. Individual features of the volume data are shown on the right, from which, the subject is asked to select the feature perceived to be in front with one click.

- The group it belongs to (enhanced or original group).
- The answer for depth ordering (correct or incorrect).
- The time that a subject spent in the case (in millisecond).

We did the whole study using the same monitor to avoid the color distortion problem on different monitors. All the subjects were asked to make an intuitive decision. Before the user study, the subjects knew nothing about the X-junction model and TAP model. Thus, they could not guess the answer that they thought the researchers preferred. They did not know the real spatial relationship until they finished the whole study.

7.4 Results

The results can be summarized as a contingency table (Table 2) with its two rows corresponding to the group of enhanced images and the group of original images, respectively. As it was a frequency distribution of events, the χ^2 test might be employed for analyzing the result [16]. However, our sample size was small. Since the approximation of χ^2 test is not suitable for small sample size, we applied Fisher' s exact test instead of the χ^2 test. The Fisher's exact test was designed for analysis of contingency table with small sample size [3].

As shown in Table 2, 76.7 percent cases in the enhanced group perceived the real depth ordering, while only 33.3 percent cases in original group got the correct spatial relations. The relative performance of the original group compared to the enhanced group was $33.3\%/76.7\% = 0.43$. We were 95 percent certain that an original image had 43 percent the chance of conveying correct depth perception as an enhanced image using the approximation of Katz. The two-sided P value by Fisher' s exact test was 0.0016 which meets $P < 0.01$, considered very significant. Hypothesis 1 could be accepted at significance level 0.01.

The results of the experiment confirm that the depth perception was improved significantly in the enhanced group. We have succeeded in introducing the X-junction model and TAP model to volume visualization. By adjusting the opacity and lightness parameters slightly based on these theories, it was much easier for users to perceive the correct spatial relation .

A contingency table was used to compare the categorical variables (correct and incorrect). For the continuous time variable, we used t-test to find if there was difference between the enhanced group and original one [15]. The mean time user spent on recognizing depth ordering in the enhanced group was 6983.3 milliseconds ($\sigma = 3920.7$). For the original group, the mean time was 7154.5 milliseconds ($\sigma = 5803.3$). The two-tailed $P <= 0.8940$ which did not meet $P < 0.01$. We did not find statistically significant difference between the average times for decision making of each group. We cannot accept Hypothesis 2 based on these results and this sample size.

TABLE 2
The Correctness of Different Groups

Group	Correct	Incorrect	Correctness Proportion
Enhanced	23	7	76.7%
Original	10	20	33.3%

The results of the t test showed that our method did not speed up the process for the users to perceive the depth.

8 Discussion

Our approach enhances a volume-rendered image by optimizing its associated transfer function, which is defined as a mixture of Gaussians. It is possible to optimize transfer functions that are defined by other parametric models such as triangular and linear ramps, as our conjugate gradient search-based general optimization method is not constrained to the Gaussian mixture model. Our approach does not create a new transfer function from scratch and an initial transfer function is required to start the optimization. Although methods for automatically creating transfer functions exist, our approach can be regarded as an important complement to them, enhancing the quality of their results.

We use a well-established image-based method to detect junctions in a volume-rendered image. Alternatively, one could project the individual 3D structures onto an image, and then identify the junctions based on this information. Unfortunately, this process requires that the volume data be first segmented, which may limit the use of our approach. The results of our junction detection method could be affected by the initial transfer function. If the preliminary image created by the initial transfer function does not show any X-junction (for example, all structures are set to be opaque), our approach will not enhance the image, because the image does not have ambiguity to resolve according to the depth perception models.

Our approach uses the luminance of different semi-transparent structures to determine the quality of the image. In the optimization process, it adjusts the luminance and opacity of the structures through the transfer function to improve the image quality because of the following reasons. First, luminance is regarded in visual psychology as one of the primary visual cues for a viewer to perceive the depth ordering and transparency. Most widely used quantitative models of depth ordering and transparency perception are solely based on luminance. Second, although other visual cues could help enhance depth perception, they also usually introduce additional overhead. For instance, illustrative visual cues such as halos would not only occlude the

background objects, but also lead to fuzzy or unclear boundaries of the front objects.

The perception models have some limitations. For instance, the X-junction model and the TAP theory can only deal with two overlapping semitransparent layers at a time. Nevertheless, this limitation does not affect the effectiveness of our approach, because complicated cases involving more than two overlapping semitransparent layers can still be handled by the models such that the layers are analyzed pair by pair. Another limitation is that the models cannot handle layers of other spatial relations such as the enclosing and separate relations. For enclosing structures, we cannot enhance the correct depth ordering perceptually by adjusting the contrast of luminance. We do not know of any quantitative depth perception models that can deal with these relations. Nevertheless, our technique is still useful because overlapping semitransparent layers widely exist in direct volume rendering. Finally, some real-world volumetric structures may have very complex depth relations with intertwined structures where there is no definite "foreground" and "background" object. As our optimization is based on individual junctions, our approach enhances the image locally at each junction, which might lead to a consistency problem. We plan to investigate this issue further in our future work. One feasible solution to the inconsistency problem is to improve the energy function.

Our approach could be used to enhance animation. One straightforward solution is to optimize each frame of the animation using our enhancement approach. Unfortunately, certain successive frames might appear incoherent because they are enhanced independently. Furthermore, enhancing all the frames using our approach would be quite expensive. To overcome these problems, we can first identify the key frames in the animation and then optimize only those using our approach. Other frames can then be enhanced using parameters obtained by interpolating the parameters of the key frames. The interpolation method can usually reduce the incoherency between successive frames. In the future, we want to improve our optimization approach to enhance the key frames simultaneously rather than independently, thereby eliminating the coherency issue.

9 Conclusion and Future Work

This paper introduces three quantitative perception models, namely, X-junctions, TAP, and Metelli's episcotister, from visual psychology to estimate how a viewer perceives the depth ordering and transparency of semitransparent structures in volume data. Guided by these models, rendering parameters can be effectively optimized to produce volume-rendered images complying with viewer' perceptions in which structures are faithfully revealed. These models can also provide good indications of depth and transparency perception of the images. Therefore, the expressiveness of the images can be adaptively enhanced by additional illustrative visual cues. The experimental results have demonstrated the effectiveness and usefulness of our approach. Importantly, these results also show the potential of interdisciplinary research of visual perception theory and visualization.

It has been reported that color and contrast may play a role in the depth perception of spatial structures [25], [37]. Some heuristic guidelines have been suggested for choosing appropriate color and contrast and providing visual depth cues. Therefore, one possible future direction is to use color and contrast as additional visual cues, such that perception quality can be improved for structures that hold not only the "overlapping" relation but other spatial relations as well. The use of the quality enhancement framework is not limited to direct volume rendering. The whole optimization pipeline, as well as the energy function designed for semitransparent structures, can be easily adapted to other computer graphics applications that deal with semitransparent structures, such as CAD. Our current approach does not assume user interaction. However, adapting our approach for interactive visualization using GPU acceleration is certainly possible. Although there might be some frame-to-frame coherency issues, this concern can be resolved by expanding the energy function to include a coherency term. We plan to further study this area.

Acknowledgments

This research has been sponsored in part by the US National Science Foundation (NSF) through grant CCF-0811422 and US Department of Energy (DOE) with award DE-SC0002289.

References

[1] E.H. Adelson, "Lightness Perception and Lightness Illusion," 1999.
[2] E.H. Adelson and P. An, "Ordinal Characteristics of Transparency," *Proc. AAAI Workshop Qualitative Vision,* pp. 77-81, 1990.
[3] A. Agresti, "A Survey of Exact Inference for Contingency Tables," *Statistical Science,* vol. 7, no. 1, pp. 131-153, Feb. 1992.
[4] B.L. Anderson, "A Theory of Illusory Lightness and Transparency in Monocular and Binocular Images: The Role of Contour Junctions," *Perception,* vol. 26, no. 4, pp. 419-453, 1997.
[5] C. Boucheny, G.-P. Bonneau, J. Droulez, G. Thibault, and S. Ploix, "A Perceptive Evaluation of Volume Rendering Techniques," *ACM Trans. Applied Perception,* vol. 5, no. 4, pp. 23:1-23:24, 2009.
[6] S. Bruckner and E. Gröller, "Enhancing Depth-Perception with Flexible Volumetric Halo," *IEEE Trans. Visualization and Computer Graphics,* vol. 13, no. 6, pp. 1344-1351, Nov./Dec. 2007.
[7] P. Calhoun, B. Kuszyk, D. Heath, J. Calrey, and E. Fishman, "Three-Dimensional Volume Rendering of Spiral CT Data: Theory and Method," *RadioGraphics,* vol. 19, no. 3, pp. 745-764, 1999.
[8] M.-Y. Chan, Y. Wu, W.-H. Mak, W. Chen, and H. Qu, "Perception-Based Transparency Optimization for Direct Volume Rendering," *IEEE Trans. Visualization and Computer Graphics,* vol. 15, pp. 1283-1290, Nov./Dec. 2009.
[9] A. Chu, W.-Y. Chan, J. Guo, W.-M. Pang, and P.-A. Heng, "Perception-Aware Depth Cueing for Illustrative Vascular Visualization," *Proc. Int'l Conf. BioMedical Eng. and Informatics,* pp. 341-346, 2008.
[10] J. Chuang, D. Weiskopf, and T. Moller, "Hue-Preserving Color Blending," *IEEE Trans. Visualization and Computer Graphics,* vol. 15, pp. 1275-1282, Nov. 2009.
[11] Z. Cipiloglu, A. Bulbul, and T. Capin, "A Framework for Enhancing Depth Perception in Computer Graphics," *Proc. Symp. Applied Perception in Graphics and Visualization,* pp. 141-148, 2010.
[12] C.D. Correa and K.-L. Ma, "Visibility Histograms and Visibility-Driven Transfer Functions," *IEEE Trans. Visualization and Computer Graphics,* vol. 17, no. 2, pp. 192-204, Feb. 2011.
[13] F. Escolano, P. Suau, and B. Bonev, *Information Theory in Computer Vision and Pattern Recognition,* second ed. Springer, 2009.
[14] M.H. Everts, H. Bekker, J.B. Roerdink, and T. Isenberg, "Depth-Dependent Halos: Illustrative Rendering of Dense Line Data," *IEEE Trans. Visualization and Computer Graphics,* vol. 15, no. 6, pp. 1299-1306, Nov./Dec. 2009.

[15] R.A. Fisher, *Statistical Methods for Research Workers.* Oliver and Boyd, 1925.
[16] P.E. Greenwood and M.S. Nikulin, *A Guide to Chi-squared Testing.* Wiley, 1996.
[17] G.B. Hanna, S.M. Shimi, and A. Cuschieri, "Randomised Study of Influence of Two-Dimensional versus Three-Dimensional Imaging on Performance of Laparoscopic Cholecystectomy," *Lancet,* vol. 351, no. 9098, pp. 248-251, 1998.
[18] V. Interrante, H. Fuchs, and S.M. Pizer, "Conveying the 3D Shape of Smoothly Curving Transparent Surfaces via Texture," *IEEE Trans. Visualization and Computer Graphics,* vol. 3, no. 2, pp. 98-117, Apr/June 1997.
[19] G. Kindlmann, R. Whitaker, T. Tasdizen, and T. Moller, "Curvature-Based Transfer Functions for Direct Volume Rendering: Methods and Applications," *Proc. IEEE Conf. Visualization,* pp. 513-520, 2003.
[20] J. Kniss, S. Premoze, C. Hansen, P. Shirley, and A. McPherson, "A Model for Volume Lighting and Modeling," *IEEE Trans. Visualization and Computer Graphics,* vol. 9, no. 2, pp. 150-162, Apr./June 2003.
[21] E. Lum and K.-L. Ma, "Lighting Transfer Functions Using Gradient Aligned Sampling," *Proc. IEEE Conf. Visualization,* pp. 289-296, 2004.
[22] F. Metelli, "An Algebraic Development of the Theory of Perceptual Transparency," *Ergonomic,* vol. 13, pp. 59-66, 1970.
[23] F. Metelli, O.D. Pos, and A. Cavedon, "Balanced and Unbalanced, Complete and Partial Transparency," *Perception and Psychophysics,* vol. 38, no. 4, pp. 354-366, 1985.
[24] K.L. Novins, F.X. Sillion, and D.P. Greenberg, "An Efficient Method for Volume Rendering Using Perspective Projection," *Proc. Workshop Volume Visualization,* pp. 95-102, 1990.
[25] R.P. O'Shea, S.G. Blackburn, and H. Ono, "Contrast as a Depth Cue," *Vision Research,* vol. 34, no. 12, pp. 1595-1604, 1994.
[26] J.P.W. Pluim, A. Maintz, and M.A. Viergever, "Mutual-Information-Based Registration of Medical Images: A Survey," *IEEE Trans. Medical Imaging,* vol. 22, no. 8, pp. 986-1004, Aug. 2003.
[27] P. Rheingans and D.S. Ebert, "Volume Illustration: Nonphotorealistic Rendering of Volume Models," *IEEE Trans. Visualization and Computer Graphics,* vol. 7, no. 3, pp. 253-264, July-Sept. 2001.
[28] N. Rubin, "The Role of Junctions in Surface Completion and Contour Matching," *Perception,* vol. 30, no. 3, pp. 339-366, 2001.
[29] Y. Rubner, "Perceptual Metrics for Image Database Navigation," PhD dissertation, Stanford Univ., 1999.
[30] M. Ruzon and C. Tomasi, "Edge, Junction, and Corner Detection Using Color Distributions," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 23, no. 11, pp. 1281-1295, Nov. 2001.
[31] Y. Sato, C.-F. Westin, A. Bhalerao, S. Nakajima, N. Shiraga, S. Tamura, and R. Kikinis, "Tissue Classification Based on 3D Local Intensity Structure for Volume Rendering," *IEEE Trans. Visualization and Computer Graphics,* vol. 6, no. 2, pp. 160-180, Apr.-June 2000.
[32] E. Saund, "Perceptual Organization of Occluding Contours of Opaque Surfaces," *Computer Vision and Image Understanding,* vol. 76, no. 1, pp. 70-82, 1999.
[33] M. Singh and B.L. Anderson, "Towards a Perceptual Theory of Transparency," *Psychology Rev.,* vol. 109, no. 3, pp. 492-519, 2002.
[34] S. Takahashi, Y. Takeshima, I. Fujishiro, and G. Nielson, "Emphasizing Isosurface Embeddings in Direct Volume Rendering," *Scientific Visualization: The Visual Extraction of Knowledge from Data,* pp. 185-206, Springer, 2006.
[35] I. Viola, "Importance-Driven Expressive Visualization," PhD dissertation, Vienna Univ. of Technology, 2005.
[36] V. Šoltészová, D. Patel, and I. Viola, "Chromatic Shadows for Improved Perception," *Proc. ACM SIGGRAPH/Eurographics Symp. Non-Photorealistic Animation and Rendering,* pp. 105-116, 2011.
[37] L. Wang, J. Giesen, K.T. McDonnell, P. Zolliker, and K. Mueller, "Color Design for Illustrative Visualization," *IEEE Trans. Visualization and Computer Graphics,* vol. 14, no. 6, pp. 1739-1754, Nov. 2008.
[38] C. Ware, *Information Visualization: Perception for Design,* second ed., ch. 8. Morgan Kaufmann, 2004.
[39] Y. Wu and H. Qu, "Interactive Transfer Function Design Based on Editing Direct Volume Rendered Images," *IEEE Trans. Visualization and Computer Graphics,* vol. 13, no. 5, pp. 1027-1040, Sept./Oct. 2007.
[40] R. Yagel, A. Kaufman, and Q. Zhang, "Realistic Volume Imaging," *Proc. IEEE Conf. Visualization,* pp. 226-231, 1991.

Lin Zheng received the bachelor's degree in computer science and technology from the Shanghai Jiaotong University. Currently, she is working toward the PhD degree in the Visualization and Interface Design Innovation (VIDi) research group at the University of California, Davis. Her research interests include multimodal volume visualization and perceptually based visualization enhancement. She is a student member of the IEEE.

Yingcai Wu received the BEng degree in computer science and technology from the South China University of Technology in 2004 and the PhD degree in computer science from the Hong Kong University of Science and Technology in 2009. Since June 2010, he has been a postdoctoral researcher at the Visualization and interface Design Innovation (VIDi) research group in the University of California, Davis. Prior to his current position, he was a postdoctoral researcher in the Hong Kong University of Science and Technology. His research interests include information visualization, visual analytics, medical visualization, and user interface design. He is a member of the IEEE.

Kwan-Liu Ma received the PhD degree in computer science from the University of Utah in 1993. He is a professor of computer science and the chair of the Graduate Group in Computer Science (GGCS) at the University of California, Davis. He leads the VIDi research group and directs the UC Davis Center for Visualization. His research interests include visualization, high-performance computing, and user interface design. He was a paper chair of the IEEE Visualization Conference in 2008 and 2009, and an associate editor of IEEE TVCG. He cofounded the IEEE Pacific Visualization Symposium and IEEE Symposium on Large Data Analysis and Visualization. Presently, he serves on the editorial boards of the IEEE CG&A, the *Journal of Computational Science and Discoveries*, and the *Journal of Visualization*. He was a recipient of the PECASE award in 2000. He is a fellow of the IEEE.