Perception enhancement using importance-driven hybrid rendering for augmented reality based endoscopic surgical navigation

YAKUI CHU,1 XU LI,1 XILIN YANG,1 DANNI AI,1 YONG HUANG,1 HONG SONG,2 YURONG JIANG,1 YONGTIAN WANG,1 XIAOHONG CHEN,3,4,5 AND JIAN YANG1,4,6

1Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Electronics, Beijing Institute of Technology, Beijing 100081, China
2School of Software, Beijing Institute of Technology, Beijing 100081, China
3Department of Otolaryngology, Head and Neck Surgery, Beijing Tongren Hospital, Beijing 100730, China
4Co-corresponding authors
5trchxh@163.com
6jyang@bit.edu.cn

Abstract: Misleading depth perception may greatly affect the correct identification of complex structures in image-guided surgery. In this study, we propose a novel importance-driven hybrid rendering method to enhance perception for navigated endoscopic surgery. First, the volume structures are enhanced using gradient-based shading to reduce the color information in low-priority regions and improve the distinctions between complicated structures. Second, an importance sorting method based on the order-independent transparency rendering is introduced to intensify the perception of multiple surfaces. Third, volume data are adaptively truncated and emphasized with respect to the perspective orientation and the illustration of critical information for viewing range extension. Various experimental results prove that with the combination of volume and surface rendering, our method can effectively improve the depth distinction of multiple objects both in simulated and clinical scenes. Our importance-driven surface rendering method demonstrates improved average performance and statistical significance as rated by 15 participants (five clinicians and ten non-clinicians) on a five-point Likert scale. Further, the average frame rate of hybrid rendering with thin-layer sectioning reaches 42 fps. Given that the process of the hybrid rendering is fully automatic, it can be utilized in real-time surgical navigation to improve the rendering efficiency and information validity.

© 2018 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

1. Introduction

Recent advancements in the endoscopic image-guided approach for minimizing surgical invasion have drawn increasing attention. This approach can greatly improve the visibility of lesion areas for surgeons, and its ability to identify important anatomical structures is the key to its successful application in clinical practice. In skull base surgery, the nidus is delicately coupled with adjacent cranial nerves and internal carotid artery, thereby complicating the surgery, possibly extending the duration of the surgery and increasing the risk of misinterpreting their spatial relationships [1,2]. Given the narrow view of navigated endoscopic surgery, augmented reality (AR) technology is especially helpful for precise surgery as it can fuse virtual generated structures with real endoscopic images during surgery by using pre-acquired multi-modality data. Therefore, the AR rendering can greatly extend the visual perception of the surgeon and further help with the data interpretation [3].
With decades of studying, misleading depth cues remain to be the challenge of AR system resulting in perceptual biases [4]. It is obvious that the incorrect depth perception may interfere with the understanding of the spatial relationships of objects inside the rendering view. There are various factors that may influence the perceptual issues, such as environment, capturing and display methods [5,6]. In particular, three kinds of AR-based display methods have been used in medical application in recent years, including see-through, projector-camera and video-based display [1,5]. The see-through method utilizes optical transmissive technology to superimpose information onto a semi-transparent mirror to extend surgeon’s knowledge on in-vivo surgery. Bichlmeier et al. proposed an AR-based see-through method to improve 3D medical data perception [7,8]. With the motion of a virtual window, depth cues of occlusion and motion parallax are intensified. However, the key of this method is the rapid motion of surgeon’s perspective, which may lead to additional physical burden. Moreover, the simple overlaying of semi-transparent images is perceptually not optimal and may cause false spatial relationship judgments between the virtual structures and real images. Liao et al. developed an MRI-guided AR system that superimposed integral videography image with surgical area, and it was viewed through a half-silvered mirror [9]. But the main issue of see-through approach is that it needs more motion tracking sensors to align virtual and real views [10,11]. The projector-camera display allows surgeons to observe the internal structures on patient’s skin surface [12,13]. Gavaghan et al. proposed a handheld image overlay device that projected images of 3D models onto the organ surface [14]. Even with the help of position tracking sensors, their method’s projection error of a rugged liver phantom is uncontrollable. Tabrizi et al. proposed an AR neurosurgery system to project image of virtual model onto patient’s skin, but their rendering was so lack of depth perception that the lesion seemed to be out of the skin surface [15]. The drawback of projector-camera systems is that a small offset between surgeon’s gaze and the optical center of the projector may lead to a significant error of perception. The video-based display uses an external camera to capture the video image around the lesion areas. This approach directly superimposes preoperative patient information on the endoscopic image along the surgeon’s gaze, thereby minimizing the locating error of deep structures inside the projection plane [16–19]. However, as the spatial relations between surgical tissue and organs at different positions are compressed according to the perspective projection principle, the depth cues are difficult to be perceived, which will further result in the abruption of the AR rendering. Therefore, augmented visualization of the preoperative 3D medical data has become a key issue in AR-based surgical navigation.

In clinical application, various medical visualization techniques are available, including basic direct volume rendering, surface rendering and the hybrid rendering that combines both surface and volume rendering [20,21]. Direct volume rendering (DVR) is used to obtain a fast overview, however, emphasizing objects or special parts is difficult. Therefore, to help surgeons understand the rendered anatomical structure of clinical data and remove unimportant details, non-photorealistic rendering method is proposed. This proposed method is extensively used in medical visualization. Volumetric non-photorealistic rendering techniques are also called volume illustration [22], which can handle original data without undergoing any segmentation process, and the gradient direction is estimated based on the differences among adjacent voxels. The volume illustration method can improve the visualization of the important structure and achieve various rendering styles for different features of interest. Ebert and Rheingans presented several illustrative techniques that enhanced structures and added depth and orientation cues [22]. In their applications, exponential functions were employed, showing that color blending of volume could be more effective. Bruckner et al. proposed an interactively defined halo transfer function to enhance the depth perception using GPU-based direct volume rendering [23], but, it led to the expense of occluding other structures. Csebfalvi et al. visualized object contours based on the magnitude of local gradients and the angle between the viewing direction and gradient vector through depth-shaded maximum intensity projection [24], which could clearly show the
outline of objects. Shape-from-shadow is another common approach to enhance shape perception. Lee et al. used multiple lights and local illumination to adaptively enhance the shapes of different parts of structures [25]. However, local illumination may destroy the rendering of different structures and cannot provide the correct location information of the targets. Furthermore, Viola et al. introduced an importance-driven volume rendering method to emphasize important structures. By assigning every individual part of the volumetric data with a different importance factor that encoded visibility priority, a cut-away view was generated to suppress less important parts of the scene and reveal more important underlying information [26]. It provided an alternative idea to intensify structures in complex data.

Generally, the surface rendering needs to extract the surface of object from the volume data. As the objects can be rendered separately, it is very flexible to adjust the color and transparency of a specific object so that special emphasis can be applied to important surfaces. However, a simple sum-up of different structures’ transparency often results in incorrect image with false depth perception. Consequently, many researchers have proposed different methods to enhance the transparency and to improve the depth perception of surface rendering. Porter et al. developed a partial coverage model that defined a set of operations for images with coverage relationship [27], which composited the multiple surfaces iteratively. The rendering methods of transparent surfaces used the alpha blending equation to arrange surfaces either front to back or back to front to obtain the correct visualization [28]. Meshkin et al. [29] first introduced the blended order-independent transparency by formulating the weighted sum of different components and produced plausible images for low alpha values. But their approximation was inadequate that might cause significant deviation from correct blending, because it ignored the order-dependent terms. McGuire et al. proposed a depth weighting operator to improve the perception of shading and color [30]. However, their methods still mislead the depth perception of semi-transparent objects.

Hybrid visualization of volume and surface has been extensively used in computer graphics [31–33]; therefore, it is of great potential to be introduced into AR-based endoscopic surgery. For the hybrid rendering technique, DVR is employed to present the anatomical context, such as skeletal structures, soft tissue and body cavities. Surface rendering is utilized to demonstrate the pre-segmented anatomical structure and the surgical tools. To render the hybrid data of transparent surface and volumetric object, Brecheisen et al. [34] proposed an iterative ray casting approach with surface rendering, which was based on depth peeling [35] to get the Z-buffer of polygonal geometry. Although the aforementioned methods have somehow improved the structural perception, exact blending of transparent object and complex structure remains a problem [36]. The correct composite of color and opacity for multiple objects is important for hybrid rendering, and the tradeoff between the quality and speed need to be carefully considered. Whereas, the combination of illustrative visualization and AR is a beneficial solution to improve perception in surgical applications [6,37]. Hansen et al. proposed an AR-based 3D visualization of planning data which extended illustrative rendering [37]. Their distance-encoding method of silhouettes and surfaces generated accurate depth judgement but was still immersive-less with the laparoscopic images. Tietjen et al. presented a useful rendering method that combined hybrid data and silhouettes for surgical training and planning [33]. The problem for their method is that the order of the nodes in the scene is the basis for depth rendering [38], hence, a slight mistake of the order definition may result in unidentifiable spatial relation of different objects. Moreover, their method cannot render semi-transparent surfaces and the volume simultaneously, hence, the spatial perception is weak. Although employing transparency for every object may be helpful, it also increases the complexity of computation. Thus, the rendering appearance and efficiency cannot satisfy the requirement of real-time surgery navigation and rapidly interaction.
In our previous study, we proposed a hybrid rendering method for dynamic endoscopic vision expansion, by which the 2D endoscope and 3D CT images were effectively fused during surgery [39]. Figure 1 shows the screenshot of our AR-based surgery navigation system for cadaver experiments. The circular area in the middle of the screen is endoscopic image, and its surrounding area is the fused rendering of the hybrid data. However, three major problems need to be addressed. First, a large number of aliasing artifacts, named wood-grain effects, as shown in the yellow rectangles, are caused by the large sample step size of volume rendering. Second, identifying the spatial relationship of different objects is difficult because the scene is cluttered with miscellaneous anatomical structures and artifacts. Moreover, the cost of rendering is high, resulting in a much lower frame rate of scene update. Hence, this study is designed to address these problems through importance-driven hybrid rendering.

In this paper, an importance-driven hybrid rendering method is proposed to enhance structure and depth perception. The contributions in this study are threefold. First, we propose a gradient-based shading method to enhance the volume rendering structure. It can suppress the color information in a low-complexity region while improving the recognition of important structures. Second, by optimizing the order-independent transparency rendering with the priority of surfaces, an importance sorting method is proposed to improve the depth perception of multiple surfaces rendering. Third, we develop a real-time data sectioning method to accelerate data exploration and help surgeons concentrate on important information. Our method has been applied to endoscopic surgical navigation systems to improve rendering efficiency and information validity.
2. Methods

The proposed hybrid rendering process has four major steps, namely, gradient-based volume shading, importance sorting-based surface rendering, real-time data sectioning, and 2D-3D image fusion, as the dashed line rectangles indicated in Fig. 2. The rendering method consists of two pipelines: volume rendering and surface rendering. In the volume rendering pipeline, gradient magnitude and direction of each voxel are first calculated, and then the volume data are enhanced through gradient-based shading and edge-and-contour enhancement methods. In addition, the stochastic jitter method is used for eliminating aliasing artifacts. In the surface rendering pipeline, surfaces of important structures are segmented from 3D CT data during the preoperative preparation process. Then, the importance factors of each surface are sorted and divided into priority and normal. The order independent transparency (OIT) method is employed for color synthesis of all the surfaces. Surfaces with different priorities are reinforced by contouring with different rendering options: higher-priority surfaces are enhanced using edge contour line, whereas the data of normal importance undergo the normal transparency processes. After the rendering process, the two branches are mixed through hybrid rendering. Thus, the endoscopic image and 3D sectioned rendering result can be effectively fused for AR-based navigated surgery.

2.1 Gradient-based volume shading

In volume data, there are no explicitly and discretely defined features like surface models that can distinguish regions clearly [22]. The features indicated by volume characteristics are distributed continuously throughout the whole volume data set, which leads to considerable difficulties to separate disparate regions. But, the important information of boundaries between different regions and spatial relationships between different structures are still of great interest to us, and not yet identified. The local gradient is commonly used to indicate the disparate levels of two different regions. Levoy [40] introduced gradient-based shading and opacity enhancement to volume rendering. In his approach, the opacity of each voxel was scaled based on its gradient magnitude to emphasize the boundaries between different
structures (such as tissue and vessel); this method can also contain the density constant of the same organ. Assuming that the volume data contain a set of precomputed sample points, the value in a location \( P \) can be calculated as follows:

\[
v_{i} = f(P) = f(x_i, y_i, z_i).
\]

By calculating the gradient \( \nabla f(P) \) value of that location, the opacity of enhanced boundary on the basis of the gradient magnitude can be computed as

\[
O_b = O_{i} \left( k_{be} + k_{we} \left( \| \nabla f \| \right)^{ek} \right),
\]

where \( O_{i} \) is the original opacity, and \( \nabla f \) is the gradient value of the volume data. \( k_{we} \) and \( k_{be} \) determine the ability of boundary enhancement from no gradient to full gradient. \( k_{we} \) is a power function to adjust the slope of opacity curve. Whereas, the depth perception is limited in translucent volume rendering as there is no correct obscuration cue to show a clear depth ordering. Inspired by shading concepts in graphics and technical illustration, we develop a gradient-based shading method for the enhancement of structure and depth cue in volume rendering.

To highlight the color of structures with a high gradient magnitude, the following equation is introduced for the gradient color calculation:

\[
C_{e} = C_{v} \left( k_{c} + k_{b} e^{ek} \right),
\]

where \( C_{v} \) is the original color that allows users to define the degree of gradient enhancement \( (k_{c} = 1, k_{b} = 0) \). By using the exponent \( k_{c} \) to adjust the compression range of the gradient color shading, Eq. (3) improves the differentiate of structures in the high gradient magnitude portions.

Additionally, the hierarchical filter based on adaptive octree is implemented to reduce the computation burden and to promote rendering quality. The volume data are first transformed into a hierarchical octree that consists of nodes with constant size of block. All the visible nodes are determined as those within the view frustum and clipping planes, which will be detailed in section 2.3. Then, the invisible nodes are skipped by only traversing the octree nodes that contain relevant data in the lookup table. Notably, the aliasing artifacts mainly come from a large sampling step size. The artifact suppression method based on stochastic ray jitter is used to reduce the artifact of volume rendering with minimal influence on the processing speed [41]. Through random adjustment of the starting depth for each ray, the closing rays are not sampled exactly along the same depth, and then the aliasing of the large step sizes can be minimized.

### 2.2 Importance sorting-based OIT rendering

In surgical navigation or therapy planning, exploring some structures that may touch, cover, or contain one another is necessary. For instance, in tumor resection surgery, vascular structures are often distributed around the tumor and may penetrate into it. The optic nerve, vasculature, and tumors are often very close to each other [39]. In such a complex situation, surgeons are often concerned about the vasculature or optic nerve closer to the tumor, so these structures tend to have higher priority than farther ones. Specifically, the properties of important structures need to be accurately recognized, such as the depth and direction of vasculature and nerve, while other non-important structures can be ignored to avoid distraction [42,43]. Although many solutions exist for the coverage of multiple surface data in complex scenes, their results are not visually correct [30,44]. Hence, this study introduces an importance sorting-based order-independent transparency surface rendering method.
In our method, we sort the objects according to their levels of importance, and the structures are highlighted based on their important impact factors. The importance sorting of different targets is commonly completed during preoperative surgical planning. For our calculation, the importance sorting is determined according to the size of the object and also the distance to the tumor. And, the important factor of the object is determined under the supervision of surgeons, which determines the visibility of the structures. Assuming that the important factor of the \(i\)-th surface is \(S_i \in [0,1]\), the color composition of all the surfaces \(C_f\) can be calculated as follows:

\[
C_f = \frac{1}{\sum_{i=1}^{n} \alpha_i} \sum_{i=1}^{n} C_i \left[ S_i + \left( 1 - \prod_{j=i}^{n} (1 - \alpha_j) \right) \right] + C_0 \left( \prod_{j=1}^{n} (1 - \alpha_j) \right)^{k_i},
\]

where \(C_i\) is the color of the \(i\)-th surface, \(\alpha_i\) is the factor of opacity, and \(k_i\) is the power exponent of surface in the cumulative function that will render the color of \(i\)-th surface distinctly when the surface number of a target is more than two, \(i \geq 2\). With the enumeration of \(S_i\), the surfaces that occluded with a more important factor will be displayed more clearly.

Figure 3 shows a schematic of surface peeling along the depth for the proposed method. The scene consists of one ellipse and two lines that indicate different surfaces of objects. Along the rays of sight, the depth of the scene increases from the left to the right and is normalized from 0 to 1, in 3(a). According to the intersection relationship, the surfaces of the objects are assigned with different importance factors and denoted with different line styles, as can be seen in 3(b). The layers touched by the traveling ray are denoted as Layer 0 to \(i\) according to the sequence of touching, as the bold lines represented in 3(c), 3(d), 3(e) and 3(f), respectively. A layer may consist of series of facets that are touched by the ray in the \(i\)-th sequence.

2.3 Data sectioning and depth buffer rendering

To extend the visualization range of important structures and to improve the rendering efficiency, we propose a real-time sectioning method. The viewing range is adjustable in six freedom, which may enable an arbitrary perspective view of the medical data. During surgical navigation, the image registration process aims to estimate the transformation function that maps the coordinate system of the endoscope to that of the CT image. The endoscopic reference target (ERT, a customized instrument from Northern Digital Inc.) is traced using an optical tracking system (OTS) for the localization of fiducial points on patients and in virtually generated scenes. In the world space, the patient fiducial target (PFT) is located with ERT via OTS, and then a \(4 \times 4\) matrix \(CT_{\text{pret}}\) transform between the CT image and PFT can...
be computed. The transformation from the coordinate system of OTS to PFT can be written as $^{\text{PFT}}T_{\text{OTS}}$. The pose represents the OTS to ERT rigid transformation, which can be written as $^{\text{ERT}}T_{\text{OTS}}$. Hence, the transformation from tracking system to ERT can be calculated as follows:

$$^{\text{ERT}}T_{\text{PFT}} = ^{\text{ERT}}T_{\text{OTS}}\left(^{\text{PFT}}T_{\text{OTS}}\right)^{-1}. \tag{5}$$

![Diagram illustrating volume sectioning at arbitrary perspective view.](image)

Fig. 4. Volume sectioning at arbitrary perspective view. (a) illustrates the building of two orthogonal planes UW and VW at the tip of the endoscope using the virtual endoscopic view. (b) is the view of UW plane in (U) direction and the VW plane viewed in (V) direction after 3D sectioning. The section range ($m, n, d$) can be customized as demanded.

After the preoperative registration, the virtual endoscope is located in the same coordinate system as the CT image is. In this case, the OTS can obtain the position and orientation of the endoscope in real time by using Eq. (5). Assuming the endoscope’s tip end is located at position $P_0(x_0, y_0, z_0)$ with a view direction $\mathbf{W}$ and a view-up direction $\mathbf{V}$. And the $\mathbf{U} = \mathbf{V} \times \mathbf{W}$ is the orthogonal direction in a left-handed coordinate system, as shown in Fig. 4(a). All the three directions are 3-dimensional unit vectors. To achieve the sectioning of the volume in an arbitrary direction, we set an $m \times n \times d$ sectioning cube at the top end of the endoscope in the directions of $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{W}$. The final hybrid rendering, fusion and navigation display are realized in this cubic range. Basing on the information of endoscopic tip end, the three orthogonal planes $\tau_i$ at the point $P_0$ can be expressed as $(\mathbf{\pi}_i - P_0) \cdot \tau_i = 0$, $i \in \{1, 2, 3\}$, where $\tau_i$ is the normal vector of plane $\mathbf{\pi}_i$ and they are corresponding to the three directions respectively, as shown in Fig. 4(a). Thus, the six planes of sectioning cube are generated by translating the three planes in two opposite directions for three half sides length of the cube:

$$\begin{align*}
\left\{ \mathbf{\pi}_i - P_0(x_0 \pm \frac{m}{2}, y_0, z_0) \right\} \cdot \mathbf{W} &= 0, \quad m \in \mathbb{R} \\
\left\{ \mathbf{\pi}_j - P_0(x_0, y_0 \pm \frac{n}{2}, z_0) \right\} \cdot \mathbf{V} &= 0, \quad n \in \mathbb{R}, \quad j \in \{1, 2\}, \\
\left\{ \mathbf{\pi}_d - P_0(x_0, y_0, z_0 \pm \frac{d}{2}) \right\} \cdot \mathbf{U} &= 0, \quad d \in \mathbb{R}
\end{align*} \tag{6}$$

where $\{\mathbf{\pi}'_i\}$ are the two parallel planes generated from the same direction.

Figure 4(b) shows a schematic of the cube section, where the UW plane corresponds to the $\mathbf{V}$ direction and the VW plane corresponds to the $\mathbf{U}$ direction. The rectangular region
shows the hybrid rendering of the volume and surface data that is located in the sectioned cube. The depth of viewing \( m \) can be customized along the virtual endoscopic direction \( W \). When \( m \) decreases, the range of viewing depth will narrow down. The influence of the deep and irrelevant structure on the final rendering result can be effectively minimized, hence improving the efficiency of information. With the variation of viewing range \( m, n \) and \( d \), the visualization of complex structures becomes more effective as the irrelevant information and objects are discarded. It can help surgeons to avoid incorrect perceptual and concentrate on important objects through the data sectioning.

After sectioning the scene, the hybrid data can be composited. First, we render the scene and fill the buffer with facets. Then, these facets are sorted through back-to-front compositing to generate the pixel colors. We assume that each pixel in the composited image is accumulated with the opacity \( \alpha \) and color \((r, g, b)\) along the ray through the hybrid data. Then, the pixel value \((r, g, b, \alpha)\) of each sampling voxel at different depth, \( z \), can be recursively blended as follows:

\[
A(z) = (1 - A(z-1)) \alpha(z) + A(z-1), \tag{7}
\]
\[
C_{r,g,b}(z) = (1 - A(z-1)) c_{r,g,b}(z) + C_{r,g,b}(z-1), \tag{8}
\]

where \( \alpha(z) \) is the transparency at the sample point \( z \), \( c_{r,g,b}(z) \) is the color at \( z \), \( A(z) \) is the accumulated value of the transparency along the light, and \( C_{r,g,b}(z) \) is the cumulative value of color. As the opacity of voxel is closer to 1, the contribution of the pixel becomes less important.

The basic process of depth-based buffer rendering is shown in Fig. 5. The cube represents the CT data, and the contours represent two different surfaces. For the rendering, the \((r, g, b, \alpha, z)\) is the buffered value of the volume data. First, the rendering range is determined according to the depth of each clipping plane in the hybrid data. Then, we rasterize the surfaces to extract \((r, g, b, \alpha, z)\) and synthesize them to the image plane. Finally, the composited and endoscopic images are fused to expand the view.

Fig. 5. Compositing process of hybrid rendering. (a) The cube represents CT data. \( \mathbf{0}_{CT} \) is the hybrid data space with two surface data, Surfaces 1 and 2. (b) The data space is sectioned to obtain the region of interest. (c) is the depth buffer synchronization from volume data rendering to surfaces. (d) synthesizes the color and transparency information of the volume data and the surface data at each sampling point in the hybrid space. Finally, the rendering is fused with the endoscopic image in (e).

2.4 Distance-weighted 2D-3D fusion

To fuse the 2D endoscopic image with the 3D scene, tracking-based registration of 2D–3D is accomplished through the OTS. The real-time posture information of endoscope acquired through the OTS is multiplied by the internal and external transformation matrix of the camera to obtain the position of the endoscopic image in the virtual scene. Therefore, the coordinate transformation between the patient and CT should be calculated before surgery. It is achieved by placing the fiducial markers on the head of the patient prior to acquiring the
CT data set. Thereafter, the surgeon has to manually define or segment the position of these fiducial markers in the recorded CT data set. During surgery, the surgeon touches each fiducial marker using the pointing device of the navigation system. A coordinate transformation between the intraoperative scene and the preoperative CT data set can be established. With the infrared reflective markers attached to the endoscope, the navigation system can continuously provide the position and orientation of endoscope relative to the referenced patient.

The transformation matrix from ERT to the tip pose of endoscope \( \text{TPE}_{\text{ERT}} \) and the relative transformation from TPE to PFT is \( \text{TPE}_{\text{ERT}} \text{TPE}_{\text{PFT}} = \text{TPE}_{\text{ERT}} \text{ERT}_{\text{PFT}} \). Thus, the transformation from CT coordinate to TPE can be calculated as \( \text{ERT}_{\text{PFT}} \text{PFT}_{\text{CT}} \). Then, the CT data are transformed to the virtual view \( (V_{\text{VT}}, V_{\text{VT}_{\text{PTE}}}) \) in the endoscopic image plane through geometric transformation: \( V_{\text{VT}} = V_{\text{VT}_{\text{PTE}}} \text{PFT}_{\text{CT}} \). \( P_{\text{Cam}}(u,v,w) \) and \( P_{\text{Endo}}(i,j) \) are corresponding points in virtual camera \( O_{\text{CT}} \) and endoscope image space, respectively. The relationship between \( P_{\text{Cam}} \) and \( P_{\text{Endo}} \) can be calculated as follows:

\[
P_{\text{Endo}}(i,j) = \begin{bmatrix} i \\ j \\ l \end{bmatrix} = \begin{bmatrix} u/w \\ v/w \\ 1 \end{bmatrix} \times l = P_{\text{Cam}}(u,v,w) \times l/w, \tag{9}
\]

where \( l \) is the distance from the virtual camera coordinate origin to the image plane. The result of Eq. (9) is approximated to the integer pixel position in the endoscopic image.

After the endoscope has been projected onto the image plane in virtual view, the AR image fusion is performed. By combining the pixels of the endoscopic image with the hybrid rendering scene, we can highlight the spatial information and remove redundant structures. To obtain accurate information of the display and additional vivid image effects, we propose a distance weighted fusion method. Given that only the central circular region contains the image information, a three-stages regional transparency modulation is performed. Assuming the endoscopic image is an \( M \times N \) pixel matrix, the illustrative point within a semitransparent layer \( P_{\text{GED}}(i,j) \) is set as the distance \( \text{dis} = \sqrt{(i-M)^2 + (j-N)^2} \) to the center of the image, where \( i \in (0, M - 1) \), \( j \in (0, N - 1) \). The transparency level \( \rho \) of point \( P \) can be calculated as a function of \( \text{dis} \). Therefore, we have

\[
\rho = \begin{cases} 
\alpha, & \text{dis} \in [0, t) \\
\alpha \times e^{-\frac{\text{dis}-t}{R-t}}, & \text{dis} \in [t, R) \\
0, & \text{dis} \in [R, \min(M, N)/2) 
\end{cases}, \tag{10}
\]

where \( \alpha \in [0,1] \) is the transparency constant of the physical view in sight, and \( t \) is the radius of the inner circular view. Meanwhile, the positive integer \( R \in [0, \min(M, N)/2] \) is the radius of the outer circular view. We let \( P_{\text{GED}}(i,j) \) represent the intensity of the pixel in the processed endoscopic image value and \( P_{\text{Hybrid}}(i,j) \) the intensity value of one pixel of the fusion image, as shown in Fig. 5(e). Then, the pixel value at the position \( P_{\text{Fusion}}(i,j) \) of the fused image can be calculated using the pixel intensity of the different sources:

\[
P_{\text{Fusion}}(i,j) = \sum_n P_{\text{GED}}(i,j) + P_{\text{Hybrid}}(i,j) + P_n(i,j) \tag{11}
\]
The Eq. (11) fuses the radial faded endoscopic image $P_{GED}(i, j)$ with hybrid rendering $P_{Hybrid}(i, j)$. The average fusion method is used for the blending of image pixels, which extracts the pixel information from each image and averages them to obtain the final result.

3. Experiments

A series of experiments were performed to evaluate the rendering method. Assessment contents include image similarity, visualization effects, and computational efficiency on the simulative and clinical data sets. The evaluation was conducted using an Intel Xeon E5-2620 computer with 24 GB RAM and an NVIDIA Quadro K2000 graphics card. A viewport with the size of $1000 \times 1000$ pixels was used for all the measurements.

3.1 Evaluation of gradient-based volume shading

Fig. 6. Comparison of gradient-enhanced volume rendering. (a) is the lighting settings of global illumination scene, the main light and the fill light are used to shade the volume data set to enhance the depth perception. The comparisons are direct volume rendering (b) and linear shading (c). The proposed volume shading method with different parameters are (d), (e) and (f). Areas marked with red solid lines and blue dotted lines are enlarged on the right. The differences of SSIM between our method, direct volume rendering and linear gradient shading are shown in color map, wherein the smaller the difference is, the lighter the color will be.
For the gradient-based volume shading, a three-point global illumination system is used to enhance volume visualization [45]. The illumination includes main, fill, and back lights. The main light is white, and its azimuth location and altitude are 60° and 20°, respectively, as shown in Fig. 6(a). The fill light is yellow, and its azimuth location and altitude are 90° and −135°. The back light is blue, and its azimuth location and altitude are 180° and 90°. The main and fill lights are used to shade the volume data set with appropriate contrast and depth perception. The back light is used to highlight the rim and relatively thin structures and to illuminate certain overshadowed regions from the back.

Figure 6 shows the rendering results of the comparative experiments. The direct volume rendering result in 6(b) has no gradient information, and its low texture area in red rectangle has subtle color and spatial relation information. This issue becomes more evident in the blue rectangle due to the extremely intricate anatomy and disordered structural texture in the nasal cavity. In most cases, the linear gradient shading method is used for volume rendering. But its results show a drastic change in color and an indistinguishable importance of different structures, as shown in 6(c). This may cause the neglect of requisite information and easily lead to visual fatigue. In our method, the structural perception on complex and thin areas is enhanced using gradient-based volume shading method. The magnitude of the gradient is computed and mapped to the color of the volume structures, which can intensify the structure information without causing the color change of the small gradient areas. Figure 6(d), 6(e), and 6(f) show the rendering results of our method, where the gradient of the forehead of the skull has minimal changes, whereas the color barely changes. Conversely, in the area of the rectangles, the color enhancement effect can be noticeable at the edge of the structure. The most noticeable effect is that the wood-grain artifacts have been removed completely. The rendering results appear smooth and realistic. The rectangle areas in Fig. 6 indicate that areas with complex structures have better surface consistency.

| Table 1. SSIM values under different shade methods. |
|---|---|---|
| SSIM | Ours | |
| | (d) | (e) | (f) |
| DVR (b) | 0.79 | 0.64 | 0.79 |
| Mean | 21 | 0.16 | 0.18 |
| SD | -0.53 | -0.35 | -0.45 |
| Min | 0.74 | 0.84 | 0.87 |
| Linear Shader (c) | 0.19 | 0.10 | 0.09 |
| Mean | -0.04 | 0.09 | 0.20 |
| SD |

To quantitatively evaluate the results, we use structural similarity (SSIM) [46] to pairwise measure the differences. As the SSIM index uses a perception-based model, it is suitable to be used in image quality assessment. The index values are in the range [−1, 1], where −1 indicates that the two images are quite different from each other, and 1 indicates that they are completely the same. All the SSIM indexes in Fig. 6 are the results between each leftmost method in row and the top one in column, where the bigger the differences are, the smaller the SSIMs will be. Their differences are demonstrated with color mapping and histogram statistics. The most noticeable feature is that all the structural uniform regions are colored lightly, but the edges of the structure are shaded heavily in all the SSIM variograms. The result indicates that the proposed method achieves the desired effect to shade the areas with a high gradient magnitude. Our method can effectively suppress the effect of shading in the low-complexity area and simultaneously enable improved structural perception in the larger gradient magnitude area. In addition, we perform a statistical analysis of all the SSIM values in Table 1. The color mapping of SSIM between Fig. 6(b) and 6(f) is shown in 6(bf), which indicates the most similar structural rendering, in another word, the two methods have
with minimal difference in the ability to represent the structure and texture. Figure 6(ce) has the largest structural difference mainly because the linear gradient shading method takes effect on the entire volume data. The appearances of Fig. 6(d) and 6(f) are apparently similar, but their SSIMs are relatively different. The mean SSIMs of (bd) and (bf) are 0.74 and 0.87, respectively, which indicates that the proposed volume shading methods of Fig. 6(d) and 6(f) generate almost the same rendering effects. However, the differences revealed in the color map of (bd) and (bf) are located at the edge of the structure, which proves that our method performs effectively in enhancing structure and depth information.

![Figure 7](image-url)  
Fig. 7. SSIM between our approach (d, e, f) and volume rendering (b), linear gradient shading (c), respectively.

Figure 7 is the boxplot of SSIM distribution. Among the three tests between direct volume rendering [Fig. 7(b)] and our methods [Fig. 7(d), 7(e), and 7(f)], Fig. 7(e) has the smallest mean index of 0.64, and its SD = 0.16 is also the smallest. The result indicates that although the difference between the two figures is significant, the structures cannot be well differentiated. However, the mean SSIM value of 7(bd) is the greatest at 0.79 ± 0.21, hence indicating its ability to effectively highlight the structural information. In comparison with linear shading in 7(c), the SSIM of 7(cd) is the smallest (0.74 ± 0.19), which indicates better structural prominence. All the assessments demonstrate that the gradient-based volume shading method can effectively highlight the depth and shape perception.

### 3.2 Surface and hybrid rendering

In this work, the simulation data are used to estimate the multi-surface rendering. The data consist of several translucent surfaces with different topological relationships, such as intersecting, traversing, and disjointing. A green sphere in the middle of the scene represents the tumor. The red cylinders with different radii simulate blood vessels and they have three kinds of relationships with the tumor: the 1st vessel penetrates the tumor along the Y-axis, the 2nd vessel traverses the tumor from the inside to the outside, and the 3rd vessel is completely inside the tumor. A blue pyramid inside the tumor simulates the target in the tumor. The yellow prism outside the tumor represents the nerve that is disjoint with the tumor. Table 2 lists the parameters for all the surfaces in the simulated scene, such as the location, size, rotation, and transparency, as shown Fig. 8(a). The rotation can be expressed by a quaternion...
$(x, y, z, \omega)$, where the first three terms define a vector starting from the origin of simulated scene $O_{Sim}$ to point $(x, y, z)$, and $\omega$ is an angle in radians. In the implementation, the surface is rotated around the vector for $\omega$ radians.

Table 2. The surface data settings of simulation scene.

| Surface          | Position (mm) | Parameters (mm) | Rotation (mm & rad) | Transparency |
|------------------|---------------|-----------------|---------------------|-------------|
| 1st Vessel / cylinder | (100, 50, 0)   | Height 500, Radius 10 | (0, 0, 1, 0)       | 0.9         |
| 2nd Vessel / cylinder | (250, 50, 0)   | Height 300, Radius 5  | (0, 0, 1, 1)       | 0.9         |
| 3rd Vessel / cylinder | (267, −265, −200) | Height 150, Radius 4 | (6, 6, 1, 2)       | 0.9         |
| Tumor / sphere   | (0, 0, 0)     | Radius 150      | (0, 0, 1, 0)       | 0.3         |
| Nerve / prism    | (350, 100, 0) | Length = (20, 20, 400) | (0, 0, 1, 1)       | 0.6         |
| Target / pyramid | (0, −30, 0)   | Height 100, Radius 50 | (−3, 3, 1, 2)      | 0.4         |

![Simulation Setup](image1.png)  ![Weighted Average](image2.png)

![Blended OIT](image3.png)  ![Depth Peeling](image4.png)

![No sorting 1, $k_i = 0.5$](image5.png)  ![No sorting 2, $k_i = 1.5$](image6.png)

![Sorting 1, $k_i = 0.8$](image7.png)  ![Sorting 2, $k_i = 1.2$](image8.png)

Fig. 8. The importance sorting rendering results. The first row listed the top-view of simulated scene setup (a) and the rendering result of weighted average method (d). The second row are the rendering of the OIT blend (c) and depth peeling (d) methods. The third row, (e) and (f), are the result of our method under different exponential factors. The last rows, (g) and (h), are obtained using the importance sorting method.

The weighted average, blended OIT, and depth peeling methods are compared. The rendering result of each method is viewed along the Y-axis (on the left side) and inverse the Y-axis (on right side), as shown in Fig. 8. The weighted average method can provide the correct color synthesis results, but the spatial relationship and depth of the structures...
disappear completely in Fig. 8(b). Although the color composition of blended OIT method is correct on the yellow nerve and blue target, its composition of vessels is incorrect, and its structure and direction are unidentifiable in 8(c). Another apparent drawback is that its composites have an incorrect color regardless of the view direction. The absence of depth perception is also observed for the blended OIT method. The depth peeling method shows the correct surface spatial order because when the 1st vessel traverses the tumor, the colors of the three segments are all correct in 8(d). And the location of the nerve is noticeable. But, its color synthesis is distorted; thus, the direction of the blue target cannot be perceived through the vascular structure due to the defect of the depth peeling method. Further, our proposed method is evaluated using different cumulative function parameters $k_i$ without sorting. The color composition is effectively improved when $k_i = 0.5$, as shown in 8(e), thereby improving the depth and spatial perception. The relationship between the 1st vessel and tumor can be clearly identified through its distribution, and the relationship between the blue target and 1st vessel can be recognized easily. When $k_i = 1.5$, the color accumulation is reduced and structures inside and behind the tumor can be easily perceived. We also used the silhouettes of the surfaces to depict the edges of the vessels and nerves so that the basic shapes of the structures can be identified.

Furthermore, we sort the surfaces according to the level of importance and $k_i$, as shown in F. 8(g) and (h). The sorting order of (g) is $S_{vessel} > S_{nerve} > S_{target} > S_{tumor}$ with $k_i = 0.8$. Considering that shapes and directions are complex, the vessel is assigned with the highest priority. Meanwhile, the nerve is inferior to the vessels, and tumor is set as the lowest priority. The relationship among the structures is clear: the 1st vessel penetrates the tumor and the 2nd vessel traverses the tumor. Although the blue target is located in front of the vessels, the contours and shapes of the vessels are still identifiable. Another sorting is set as $S_{vessel1} > S_{target} > S_{nerve} > S_{vessel2} > S_{tumor} > S_{vessel3}$, with $k_i = 1.2$, and the vessels are assigned with different levels of importance. In Fig. 8(h), the target and nerve with high importance are more sharpened in comparison with 8(g). And the 3rd vessel with low priority can be ignored; hence, the surgeon can focus on important objects.

Table 3. Likert scale results depth-enhancement surface rendering.

| Target Depth Perception | Score: 1-5 (Strongly Disagree/ Disagree/ Uncertain/ Agree/ Strongly Agree) | Blended OIT Depth Peeling | No sorting | Ours Sorting |
|-------------------------|--------------------------------------------------|-------------------------|------------|-------------|
|                         | Score                                             | c               | d          | e           | f           | g           | h           |
| (i) Differences of depth between various surfaces are discernible. | 1.8 ± 0.9 | 3.7 ± 1.2 | 4.1 ± 0.9 | 3.7 ± 1.0 | **4.5 ± 0.5** | 3.6 ± 1.2 |
| (ii) The 1st vessel penetrates through the tumor. | 1.5 ± 0.7 | 3.4 ± 1.1 | **4.7 ± 0.5** | 3.3 ± 1.0 | 4.6 ± 0.5 | 4.0 ± 0.8 |
| (iii) Easy to recognize the relationship among the tumor, 2nd and 3rd vessels. | 1.4 ± 0.5 | 2.7 ± 1.2 | 4.0 ± 0.9 | 2.9 ± 1.3 | **4.2 ± 0.8** | 4.1 ± 1.0 |
| (iv) The yellow columnar and green tumor are at different positions. | 1.2 ± 0.4 | 3.6 ± 1.1 | 4.2 ± 0.9 | 4.1 ± 0.7 | **4.5 ± 0.5** | 3.9 ± 1.2 |
| (v) The outline of blue surface indicates a clear direction and shape. | 1.8 ± 0.9 | 2.7 ± 1.1 | 3.7 ± 1.0 | 3.6 ± 0.8 | **4.5 ± 0.6** | 4.0 ± 1.3 |
| **Mean ± SD**              | 1.5 ± 0.8 | 3.2 ± 1.2 | 4.1 ± 0.9 | 3.5 ± 1.0 | **4.5 ± 0.6** | 3.9 ± 1.1 |
| **p-value**               | Method x vs c / p < 0.001 | Method x vs d / p < 0.001 | p < 0.001 | p < 0.001 | p < 0.001 | p < 0.001 | p < 0.001 | p < 0.001 | p < 0.0003 |

To quantitatively evaluate the depth and spatial perception of the proposed multi-surfaces importance sorting method, we designed a five-point Likert scale with increasing consent from strongly disagree (1) through uncertain (3) to strongly agree (5), as shown in Table 3. There were five questions included to investigate the differences of depth perception, color
composition and shape recognition among various methods. As the weighted average method provided no depth information, it was excluded from the experiment. Therefore, six different rendering methods were taken into consideration: the blended OIT, depth peeling and four of our proposed importance sorting methods. In total, 15 participants took part in the experiment: 5 clinicians in otolaryngologic surgery and 10 non-clinical volunteers without any prior experience in volume visualization. In the experiment, each participant completed the questionnaire with randomly arranged pairwise images of the six methods, as Fig. 8 shown. As every method was investigated with 5 questions by 15 respondents, totally 75 counts of response were collected for each method; and 450 samples were collected for all the six methods. The evaluation was performed based on the analysis of variance (ANOVA) in Table 3.

![Fig. 9. Results of the questionnaire rated using a five-point Likert scale. (a) is the histogram of respondents’ scores. The vertical axis represents the number counts of different consent. Each set of bars represent for the score decreasing from 5 to 1. And each method was sampled for 75 times. (b) is the boxplot of the Likert scale. The asterisks were the maximums and minimums of each method. In the middle of the boxes, small square marked the mean of the data and lines marked the median.](image)

The histogram of respondents’ scores was shown in Fig. 9(a). Most responses either strongly disagreed (46/75, 61.3%) or disagreed (19/75, 25.3%) that the blended OIT method provided no perception of depth or spatial relations (9 uncertain and 1 agree). And the method’s average score was 1.5 (SD = 0.8), which was the lowest of the six methods. A one-way ANOVA test showed that the blended OIT method was significantly different from the others ($p < 0.001$). Consistently, 18.7% strongly agreed (14/75) and 22.7% agreed (17/75) that the depth peeling method provided spatial relation perceptual enhancement, while 26.7% uncertain (20/75) or disagree (20/75) and 5.3% strongly disagree (4/75) about it. The average score of depth peeling method was 3.2 ± 1.2, which indicated a quite ambivalent performance of spatial judgement. As one-way ANOVA test showed a significant difference ($p < 0.001$), the depth peeling method could basically improve depth perception. For our proposed method, the four examples all showed a significant effect of depth and spatial enhancement. The no sorting group with different cumulative parameters $k_i$ was strongly agreed by 30% (45/150) responses and agreed by 31.3% (47/150); but there were 26% uncertain, 8% disagree and 2% strongly disagreed about its perceptual enhancement ability. On the other hand, the majority of responses were strongly agreed (44.7%, 67/150) or agreed (36.7%, 55/150) with our importance sorting-based perceptual enhancement method, while 12.7% uncertain and only 4.7% disagree or 1.3% strongly disagree. The one-way ANOVA tests of the two groups against blended OIT method indicated that our methods significantly intensified the depth and
spatial perception ($p < 0.001$), as Table 3 showed. These results indicated that our method was subjectively preferred by most of the participants, which fully confirmed the effectiveness of our spatial perception enhancement method. Furthermore, the two group’s one-way ANOVA tests against depth peeling method showed significant difference. In specific, the sorting groups of two different cumulative parameter $k_i$ were all significant different ($p < 0.001$), but the no sorting group only had one significant different methods with $k_i = 0.5$ ($p < 0.001$) and the other ($k_i = 1.5$) had no significant difference ($p = 0.128$) with the depth peeling method. These results indicate that the proposed cumulative parameters and sorting method are very effective for enhancing of depth and spatial perception. As the boxplot of Likert scale showed in Fig. 9(b), our proposed method in Fig. 8(g) was preferred by most participants, which achieved the highest average score and lowest dispersion (4.5 ± 0.6). This proposed importance sorting-based method was strongly agreed by 50.1% responses (38/75) and agreed by 44% (33/75) ones, only 5.3% was uncertain about its effectiveness (0 disagree or strongly disagree). Hence, it can be concluded that our method is very effective and robust for spatial perception of complex structures.

Subsequently, surface and hybrid rendering experiments are implemented based on clinical data. The surface data from clinical volume data are complex, rugged, and difficult to predict. The existence of additional surface layers along the ray of sight makes surgeons difficult to identify the color and opacity. The results of depth peeling method are not correct, as the white rectangles illustrate confusing spatial relations of surfaces in Fig. 10(a). The green tumor should be located behind the blue eye orbit, but the yellow optical should never be located in front of the tumor. Our method can correctly demonstrate the spatial order in 10(d), and the direction and shape of the vessels are also easy to identify. Besides, the hybrid rendering in 10(c) and (d) both present explicit structures by using our gradient-based volume shading method, which prove the perceptual enhancement ability for additional anatomical information.

Fig. 10. Surface and hybrid rendering of clinical data segmentation. The first column, (a), is the result of the Depth Peel method. The second column, (b), shows our method. (c) and (d) are the hybrid rendering results (see supplementary material Visualization 1).

Subsequently, surface and hybrid rendering experiments are implemented based on clinical data. The surface data from clinical volume data are complex, rugged, and difficult to predict. The existence of additional surface layers along the ray of sight makes surgeons difficult to identify the color and opacity. The results of depth peeling method are not correct, as the white rectangles illustrate confusing spatial relations of surfaces in Fig. 10(a). The green tumor should be located behind the blue eye orbit, but the yellow optical should never be located in front of the tumor. Our method can correctly demonstrate the spatial order in 10(d), and the direction and shape of the vessels are also easy to identify. Besides, the hybrid rendering in 10(c) and (d) both present explicit structures by using our gradient-based volume shading method, which prove the perceptual enhancement ability for additional anatomical information.
3.2 Data sectioning and 2D-3D fusion

The functions of real-time sectioning performance, 2D-3D image fusion effects, and hybrid rendering efficiency were evaluated using the surgical navigation system. In real-time sectioning, we tested the rendering ability with variable sizes of volume data, such as cube and thin layer. The CT volume data we used in the experiments consisted of $328 \times 356 \times 443$ slices, and the resolution is $0.54 \times 0.54 \times 0.6$ mm. The range of the cube section is $134 \times 176 \times 123$ slices, and its volume ratio is 5.6%, as shown in Fig. 11(a). The region of the surgery-related structures that surgeons are interested in can be clearly demonstrated in the cubic section. The spatial relationships among the vessels, tumors, and nerves can be clearly identified. Besides, the number of slices in thin-layer sectioning is eight. Our proposed method can provide the stereoscopic expression of the CT image. Subsequently, we fuse the endoscopic image with the results of the hybrid rendering, as shown in Fig. 11(d) and 11(e). The endoscopic image transits smoothly to the scene, and multiple targets can be easily identified. Finally, we compare the fusion of thin-layer section with full-size CT volume rendering. The results of full-size rendering are in Fig. 11(d) and 11(e), and the corresponding results of thin layer from the same perspective are shown in 11(f) and 11(g). Therefore, the
fusion of thin layer and endoscopic image can preserve the important structures in the sight and discard much of the irrelevant information, which is helpful during surgery.

In addition, we recorded the frame rate of all the aforementioned rendering tests. The frame rates are sampled every 30° as the scene rotates from 0° to 360°, which provides 12 sample points for every rotation axis. To uniformly distribute all the sample points over the observing sphere, we define three Cartesian coordinate systems at the center of the hybrid data space: \( \mathbf{O}_{\text{Hybrid}} - XYZ \), \( \mathbf{O}_{\text{Hybrid}} - X_1Y_1Z_1 \) and \( \mathbf{O}_{\text{Hybrid}} - X_2Y_2Z_2 \). They are denoted in red, green and blue in Fig. 8(a), respectively. The relationship between the two of three coordinates can be calculated as follows:

\[
\begin{align*}
X_1Y_1Z_1 &= \begin{bmatrix}
\cos \theta & \sin \theta & 0 \\
-\sin \theta & \cos \theta & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
\cos \theta & 0 & -\sin \theta \\
0 & 1 & 0 \\
\sin \theta & 0 & \cos \theta
\end{bmatrix}
\cdot XYZ, \\
X_2Y_2Z_2 &= \begin{bmatrix}
1 & 0 & 0 \\
0 & \cos \theta & \sin \theta \\
0 & -\sin \theta & \cos \theta
\end{bmatrix}
\begin{bmatrix}
\cos \theta & 0 & -\sin \theta \\
0 & 1 & 0 \\
\sin \theta & 0 & \cos \theta
\end{bmatrix}
\cdot XYZ.
\end{align*}
\] (12)

The \( X_1Y_1Z_1 \) is obtained by rotating the Y and Z axis by the angle of \( \theta \) successively, while the \( X_2Y_2Z_2 \) is obtained by rotating the Y and X axis. In this study, we set \( \theta = 45\degree \) for the rotation of coordinates. All the nine axes of the three coordinate systems act as the rotation axis for frame rate sample, which generate 108 view angles over the observing sphere. Figure 12(a) demonstrates the frame rates with respect to all the rotating angles, and Table 4 shows the rendering frame rates of different types of data, which also includes the maximum and minimum. The average frame rate of multi-surface rendering is approximately 53 fps, whereas the average frame rate of the full-size volume rendering is only 19 fps. Moreover, for the hybrid rendering, the average frame rates of full size, cube, and section rendering are approximately 13, 25, and 43 fps respectively, as shown in Fig. 12(b). It can be seen that only the surfaces and thin-layer methods obtain an average frame rate higher than 30 fps, which can meet the real-time interaction requirement for surgery navigation in clinical practice. To accelerate the hybrid rendering, the most straightforward way is to increase the step size of sampling, however, it may lead to a large amount of aliasing artifacts. Another way is to truncate the volume data and discard the unnecessary data from rendering. In the clinical application, as the redundant structures may disrupt the surgeon’s judgment, the whole data rendering is generally unnecessary. Hence, the hybrid rendering of cube section and thin-layer section can meet the requirements of surgical navigation instead of the whole data rendering.

Table 4. Rendering frame rates for different types of data.

| (fps) | Volume | Surfaces | Hybrid |
|-------|--------|----------|--------|
|       | Mean ± SD | Mean ± SD | Mean ± SD |
| Max   | 26     | 59       | 16     |
| Min   | 15     | 50       | 10     |
|       | 18.67 ± 2.86 | 53.36 ± 2.27 | 12.75 ± 1.29 |
|       | 24.69 ± 2.17 | 42.83 ± 2.64 |
4. Conclusion and discussion

In this study, we proposed an importance-driven hybrid rendering method to enhance the depth perception for AR-based endoscopic surgery navigation. The method was used to reduce the incorrect demonstration of complex internal anatomical structures and to minimize the cost of rendering, which had considerable potential for clinical application. First, important structures in volume data were highlighted using the gradient-based volume shading method. The shading method could eliminate the color shading in low-complexity areas with the exponential function of gradient magnitude. The structural similarity of different rendering results indicated that our proposed method could effectively enhance the structure and shape perception. Second, an importance sorting-based OIT method was introduced to improve the comprehension of multiple structure rendering. Pre-sorting the priority of different surfaces ensures that the shape and relation of the target were clearly identified in the complex scene. As rated by 15 participants (five clinicians and ten non-clinicians) on a five-point Likert scale, the importance sorting method demonstrated improved average performance and statistical significance. Moreover, the proposed 3D real-time sectioning method allow surgeons to concentrate on critical structures during surgery and decrease the cost of hybrid rendering. The frame rates of simulated multi-surface data and clinical data were evaluated, and the average frame rate of hybrid rendering with thin-layer sectioning reached 42 fps, which could be utilized in real-time surgical navigation to effectively improve rendering efficiency and information validity. The proposed method will greatly improve the structure and depth perception of the hybrid rendering in image-guided surgery.

Funding

National Key Research and Development Program of China (2017YFC0107900); National Science Foundation Program of China (61672099, 81627803, 61501030, 61527827).

Disclosures

The authors declare that there are no conflicts of interest related to this article.

References and links

1. S. Nicolau, L. Soler, D. Mutter, and J. Marescaux, “Augmented reality in laparoscopic surgical oncology,” Surg. Oncol. 20(3), 189–201 (2011).
2. M. Sugimoto, “Recent advances in visualization, imaging, and navigation in hepatobiliary and pancreatic sciences,” J. Hepatobiliary Pancreat. Sci. 17(5), 574–576 (2010).
3. M. Tuceryan, D. S. Greer, R. T. Whitaker, D. Breen, C. Crampton, E. Rose, and K. H. Ahlers, “Calibration Requirements and Procedures for Augmented Reality,” IEEE Trans. Vis. Comput. Graph. 3(1), 255–273 (1995).
4. D. Drascic and P. Milgram, “Perceptual issues in augmented reality reality-virtuality continuum,” Virtual Real. 2351, 197321 (1996).
5. E. Kruijff, J. E. Swan, and S. Feiner, “Perceptual issues in augmented reality revisited,” in 9th IEEE International Symposium on Mixed and Augmented Reality 2010: Science and Technology, ISMAR 2010 - Proceedings (2010).
6. K. Lawonn, I. Viola, B. Preim, and T. Isenberg, “A survey of surface-based illustrative rendering for visualization,” Comput. Graph. Forum 37(6), 205 (2018).
7. C. Bichlmeier, F. Wimmer, S. M. Heinig, and N. Navab, “Contextual anatomic mimesis: Hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality,” in 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR (2007).
8. C. Bichlmeier and N. Navab, “Virtual Window for Improved Depth Perception in Medical AR,” International Workshop on Augmented Reality environments for Medical Imaging and Computer-aided Surgery (AMI-ARCS) (2006).
9. H. Liao, T. Inomata, I. Sakuma, and T. Dohi, “3-D augmented reality for MRI-guided surgery using integral videography autostereoscopic image overlay,” IEEE Trans. Biomed. Eng. 57(6), 1476–1486 (2010).
10. D. G. Armstrong, T. M. Rankin, N. A. Giovinco, J. L. Mills, and Y. Matsuoka, “A heads-up display for diabetic limb salvage surgery: a view through the google looking glass,” J. Diabetes Sci. Technol. 8(5), 951–956 (2014).
11. X. Chen, L. Xu, Y. Wang, H. Wang, F. Wang, X. Zeng, Q. Wang, and J. Egger, “Development of a surgical navigation system based on augmented reality using an optical see-through head-mounted display,” J. Biomed. Inform. 55, 124–131 (2015).
12. F. Volonté, F. Pugin, P. Bucher, M. Sugimoto, O. Ratib, and P. Morel, “Augmented reality and image overlay navigation with OsiriX in laparoscopic and robotic surgery: not only a matter of fashion,” J. Hepatobiliary Pancreat. Sci. 18(4), 506–509 (2011).
13. J. Wang, H. Suenaga, K. Hoshi, L. Yang, E. Kobayashi, I. Sakuma, and H. Liao, “Augmented reality navigation with automatic marker-free image registration using 3-D image overlay for dental surgery,” IEEE Trans. Biomed. Eng. 61(4), 1295–1304 (2014).
14. K. A. Gavaghan, M. Peterhans, T. Oliveira-Santos, and S. Weber, “A portable image overlay projection device for computer-aided open liver surgery,” IEEE Trans. Biomed. Eng. 58(6), 1855–1864 (2011).
15. L. Thalmann, D. Thalmann, M. Mahvash, “Augmented reality-guided neurosurgery: accuracy and intraoperative application of an image projection technique,” J. Neurosurg. 123(1), 206–211 (2015).
16. B. J. Dixon, M. J. Daly, H. Chan, A. Vescan, I. J. Witterick, and J. C. Irish, “Augmented real-time navigation with critical structure proximity alerts for endoscopic skull base surgery,” Laryngoscope 124(4), 853–859 (2014).
17. N. McLaughlin, R. L. Carrau, A. B. Kassam, and D. F. Kelly, “Neuronavigation in endonasal pituitary and skull base surgery using an autoregistration mask without head fixation: An assessment of accuracy and practicality,” J. Neurol. Surg. A Cent. Eur. Neurosurg. 73(6), 351–357 (2012).
18. S. A. Nicolau, X. Pennec, L. Soler, X. Buy, A. Gangi, N. Ayache, and J. Marescaux, “An augmented reality system for liver thermal ablation: Design and evaluation on clinical cases,” Med. Image Anal. 13(3), 494–506 (2009).
19. R. Souzaki, S. Ieiri, M. Uemura, K. Ohuchida, M. Tomikawa, Y. Kinoshita, Y. Koga, A. Sumineo, K. Kohashi, Y. Oda, T. Hara, M. Hashizume, and T. Taguchi, “An augmented reality navigation system for pediatric oncologic surgery based on preoperative CT and MRI images,” J. Pediatr. Surg. 48(12), 2479–2483 (2013).
20. B. Preim, A. Baer, D. Cunningham, T. Isenberg, and T. Ropinski, “A Survey of Perceptually Motivated 3D Visualization of Medical Image Data,” Comput. Graph. Forum 35(3), 501–525 (2016).
21. M. Kersten-Oertel, P. Jannin, and D. L. Collins, “A framework for helping developers in the integration of external tools into virtual learning environments,” in Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality - TEEM ’13 (2013), pp. 127–132.
22. I. Viola, A. Kanitsar, and M. E. Gröller, “Importance-driven feature enhancement in volume visualization,” IEEE Trans. Vis. Comput. Graph. 11(4), 408–418 (2005).
23. T. Porter and T. Duff, “Compositing digital images,” Comput. Graph. 18(3), 253–259 (1984).
24. L. Bavoil and K. Myers, “Order Independent Transparency with Dual Depth Peeling,” Image Rochester NY 107(February), 22020–22025 (2008).
25. E. J. Lorenzo, R. Centeno, and M. Rodríguez-Artacho, “A framework for helping developers in the integration of external tools into virtual learning environments,” in Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality - TEEM ’13 (2013), pp. 127–132.
26. B. Csébfalvi, L. Mroz, H. Hauser, A. König, and E. Gröller, “Fast Visualization of Object Contours by Non-Photorealistic Volume Rendering,” (Proceedings of Eurographics 2001, Manchester, UK, September 2001)
27. I. Viola, A. Kanitsar, and M. E. Gröller, “Importance-driven feature enhancement in volume visualization,” 3D Visualization and Rendering, 9th IEEE
28. L. Bavoil and K. Myers, “Order Independent Transparency with Dual Depth Peeling,” Image Rochester NY 107(February), 22020–22025 (2008).
29. H. Moshkin, “Sort-independent alpha blending,” Perpetual Entertainment, GDC Talk (2007).
30. M. Mcguire and L. Bavoil, “Weighted Blended Order-Independent Transparency,” Journal of Computer Graphics Techniques 2(2), 122–141 (2013).
31. K. Mühler, C. Tietjen, F. Ritter, and B. Preim, “The medical exploration toolkit: An efficient support for visual computing in surgical planning and training,” IEEE Trans. Vis. Comput. Graph. 16(1), 133–146 (2010).
32. Y. Hayashi, K. Misawa, M. Oda, D. J. Hawkes, and K. Mori, “Clinical application of a surgical navigation system based on virtual laparoscopy in laparoscopic gastrectomy for gastric cancer,” Int. J. CARS 11(5), 827–836 (2016).
33. C. Tietjen, T. Isenberg, and B. Preim, “Combining Silhouettes, Surface, and Volume Rendering for Surgery Education and Planning,” EUROGRAPHICS - IEEE VGTC Symposium on Visualization (2005).
34. R. Brechelsen, A. Vilanova, B. Platel, and H. Romeny, “Flexible GPU-Based Multi-Volume Ray-Casting,” Proceedings of Vision, Modeling and Visualization (2008), pp. 303–312.
35. S. S. Snibbe, “A Direct Manipulation Interface for 3D Computer Animation,” Comput. Graph. Forum 14(3), 271–283 (1995).
36. M. Maule, J. L. D. Comba, R. P. Torchelsen, and R. Bastos, “A survey of raster-based transparency techniques,” Computers and Graphics (Pergamon) 35(6), 1023–1034 (2011).
37. C. Hansen, J. Wieferich, F. Ritter, C. Rieder, and H. O. Peitgen, “Illustrative visualization of 3D planning models for augmented reality in liver surgery,” Int. J. CARS 8(2), 133–141 (2010).
38. B. Preim and D. Bartz, Visualization in Medicine: Theory, Algorithms, and Applications (Moegen Kauffman, 2007).
39. Y. Chu, J. Yang, S. Ma, D. Ai, W. Li, H. Song, L. Li, D. Chen, L. Chen, and Y. Wang, “Registration and fusion quantification of augmented reality based nasal endoscopic surgery,” Med. Image Anal. 42, 241–256 (2017).
40. M. Levoy, “Efficient Ray Tracing of Volume Data,” ACM Trans. Graph. 9(3), 245–261 (1990).
41. R. Marques, L. P. Santos, and P. Leškovský, “GPU Ray Casting,” in 17º Encontro Português de Computação Gráfica (En Anexo, 2009).
42. V. R. Ramakrishnan, R. R. Orlandi, M. J. Citardi, T. L. Smith, M. P. Fried, and T. T. Kingdom, “The use of image-guided surgery in endoscopic sinus surgery: an evidence-based review with recommendations,” Int. Forum Allergy Rhinol. 3(3), 236–241 (2013).
43. T. Okamoto, S. Onda, K. Yanaga, N. Suzuki, and A. Hattori, “Clinical application of navigation surgery using augmented reality in the abdominal field,” Surg. Today 45(4), 397–406 (2015).
44. T. Gustafsson, “Concepts of Hybrid Data Rendering,” in SIGRAD 2017 (2017).
45. Y. Zhang and K. L. Ma, “Lighting design for globally illuminated volume rendering,” IEEE Trans. Vis. Comput. Graph. 19(12), 2946–2955 (2013).
46. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” IEEE Trans. Image Process. 13(4), 600–612 (2004).