Virtual Anatomical and Endoscopic Exploration Method of Internal Human Body for Training Simulator

Koojoo Kwon 1, Jin Seo Park 2 and Byeong-Seok Shin 3

1Department of Smart Information Technology, Baewha Women's University, Seoul, Korea
2Department of Anatomy, Dongguk University School of Medicine, Gyeongju, Korea
3Department of Computer Engineering, Inha University, Incheon, Korea

ABSTRACT

Background: Virtual environments have brought the use of realistic training closer to many different fields of education. In medical education, several visualization methods for studying inside the human body have been introduced as a way to verify the structure of internal organs. However, these methods are insufficient for realistic training simulators because they do not provide photorealistic scenes or offer an intuitive perception to the user. In addition, they are used in limited environments within a classroom setting.

Methods: We have developed a virtual dissection exploration system that provides realistic three-dimensional images and a virtual endoscopic experience. This system enables the user to manipulate a virtual camera through a human organ, using gesture-sensing technology. We can make a virtual dissection image of the human body using a virtual dissection simulator and then navigate inside an organ using a virtual endoscope. To improve the navigation performance during virtual endoscopy, our system warns the user about any potential collisions that may occur against the organ's wall by taking the virtual control sphere at the virtual camera position into consideration.

Results: Experimental results show that our system efficiently provides high-quality anatomical visualization. We can simulate anatomic training using virtual dissection and endoscopic images.

Conclusion: Our training simulator would be helpful in training medical students because it provides an immersive environment.

Keywords: Virtual Reality; Education; Anatomy; Gestures; Recognition

INTRODUCTION

In medical education, the training of students through simulations is important because diagnostic procedures are difficult for physicians to learn. Virtual reality (VR) provides the possibility for students to experience real-world clinical situations that require diagnostic skill training. 15 Although the techniques required to use VR devices are important, there is a lack of educational environment to describe them. Also, knowing computer technology can help medical scientists develop software packages that are more beneficial to medicine.
For more immersive systems, various sensors have been proposed to assist with user experiences in the real world. These sensors use a certain device for measuring the distance from the user and recognizing user gestures.\textsuperscript{6} Kinect, Wii, Leap Motion, Playstation Move, and Myo are some typical examples.\textsuperscript{7,12} These motion-recognition sensors improve the immersion of user interaction because of their compact structure. However, their accuracy depends on the type of predefined gestures because they cannot recognize and find user actions accurately. We also defined in advance user gestures that represent similar behaviors of anatomists.

Several systems to visualize anatomical images have been introduced in the studies of medical sciences and training of medical students. The Anatomage Table\textsuperscript{13} provides virtual surgical training because it has been designed to have the form of an operating table. However, it does not provide a realistic stereoscopic model and special interface; it only provides a three-dimensional (3D) model and touch-screen interface. Lundstrom\textsuperscript{14} proposed a tabular computer that can visualize medical images for simulating a real operation. This proposed device is suitable for establishing a surgical plan because orthopedic surgeons have been involved in this particular project. However, its 2D results are less realistic than are stereoscopic images.

Most medical data captured during computed tomography, magnetic resonance imaging, and ultrasound are represented as 2D gray-scale images. Color anatomical datasets including Visible Human,\textsuperscript{15,16} Visible Korean (VK),\textsuperscript{17,20} and Visible Chinese\textsuperscript{21,22} represent human anatomical structures with real-color, high-resolution images. We used the VK dataset for our simulator to produce a high-definition (HD) image display. However, the size of these datasets reaches tens or hundreds of gigabytes (GB) even when we scan a tiny region of the human body, because a series of such images is regarded as a 3D texture sampled using a very high resolution. In our simulator, we crop the area of a specific region of the whole body to concentrate on the surgical training of the target organ.

Our virtual dissection simulator is also operated through user’s gestures from motion-recognition sensors. When performing virtual dissection, some parts of the volume should be moved separately. This requires a type of internal representation to show the cross-sections made through anatomical operations. We need a method to produce an exploded view because both sides of an incised or excised part have an important meaning—the reason being an efficient description of the relationship between the incised or excised parts and the remaining parts exists.

Virtual endoscopy is a method for visualizing the pathological structures in pipe-shaped organs such as the colon, bronchus, and esophagus. Unlike optical endoscopy, there is no discomfort or side effect because the examination is noninvasive. We can observe a wide area inside an organ because the method produces perspective images while navigating the organ. To represent accurate movements of the virtual camera during user manipulation, making a reliable path is essential for diagnostic training. Before a virtual endoscopic simulation, we need a method for producing an exploded view to show the relationship between the separated incised or excised parts.

We propose an interactive anatomy simulator using motion-recognition sensors to improve the sense of immersion. This simulator is composed of two parts: 1) virtual dissection and 2) virtual endoscopy. It also provides HD stereoscopic images to increase the reality with
high-resolution anatomical volume data. The stereoscopic images mean that it provides two images separately to the left and right eye for binocular vision. In our method, we made these images by using the GPU-based visualization method because using GPU can make two images fast. For gesture recognition, we use a Kinect device and its software development kit (SDK). To handle the motion of a virtual camera, we define an area—called a virtual control sphere (VCS)—around every camera position along the central navigation path. Because VCSs define the restricted area of the navigation camera, our simulator can move the camera along the central path, without touching the organ’s wall. When the camera is free from user control, it turns back to the central path. Our simulator can be used to train medical students before they use actual endoscopic equipment.

The contributions of this article are as follows: 1) We provide a simulator with a more realistic and interactive interface than ordinary software which uses a normal input device such as a mouse. Using a gesture-based interface can be useful for reducing contagious infections. 2) It can be used to train medical students before they study anatomical structures. It can be used to observe the endoscope images from various angles by manipulating the motion of the virtual endoscope camera using motion recognition interface.

We describe the materials and methods applied in the “Methods” section, summarize some experimental results and discussions in the “Results” section, and finally, provide some concluding remarks regarding our work in the “Discussion” section.

METHODS

We use color images of the VK dataset. The resolution of these images is \(2,468 \times 1,407\), and each pixel size is \(0.2 \text{ mm} \times 0.2 \text{ mm}\). The interval between the images is also \(0.2 \text{ mm}\). To select the target organ, anatomical structures should be segmented. The volume data, which include color and segment information, are reconstructed (as presented in Table 1). The resolution of each dataset is sufficient to be displayed on an HD screen. As graphics accelerator technology is radically advancing, it is possible to visualize 3D objects even when we use a large dataset. Although graphics accelerators are equipped with 6–8 GB of video memory, we crop the specific region for the target organ because the size of the volume is huge for the memory to accommodate at once. In addition, we separate this dataset into four channels (red, green, blue, and segmentation index) to avoid a large amount of memory usage at the same location.

Our system consists of two parts: virtual dissection (Fig. 1) and virtual endoscopy (Fig. 2). It includes a gesture-recognition sensor module and a volume-rendering module. It provides an HD stereoscopic display to ensure an immersive environment. When the user stands in the HD stereoscopic display, high-resolution anatomical volume data can be observed from various angles.

Table 1. The resolution and capacity of datasets for our simulator

| Dataset          | Resolution | Color data volume | Segment data volume |
|------------------|------------|-------------------|---------------------|
| Whole data       | \(2,468 \times 1,407 \times 8,506\) | 88 GB            | 29 GB               |
| Head             | \(1,162 \times 1,072 \times 1,302\) | 4.8 GB           | 1.6 GB              |
| Abdomen          | \(823 \times 469 \times 1,176\)   | 1.3 GB           | 0.45 GB             |
| Respiratory tract| \(1,185 \times 636 \times 1,805\) | 2.4 GB           | 0.81 GB             |
| Stomach          | \(574 \times 611 \times 2,040\)   | 2.1 GB           | 0.71 GB             |
| Large intestines | \(1,241 \times 898 \times 1,520\) | 5.0 GB           | 1.7 GB              |

GB = gigabyte.
front of the system and manipulates the anatomy simulator using several predefined gestures, it provides an immersive environment because user operations are based on a noncontact interface.

**Virtual dissection simulator**

The virtual dissection part of our system is operated through user’s gestures. When performing dissection, the volume data itself should be deformed and moved. We separate the volume data into several sub-volumes, with user-specific parameters under compute unified device architecture (CUDA). These volume data are used as the scale factor of the corresponding proxy geometry during the rendering phase. During the vertex process, this information is used as a parameter to scale the unit blocks. The scaled blocks are used as
proxy geometry of the volume through a scaling operation. Because proxy geometry can be applied to skip an empty space and terminate an early ray in GPU ray casting during the fragment process, we can accelerate the rendering speed.

Large-sized medical content such as the Visible Male, Visible Female, and Visible Head is very useful in volume rendering. We made volume data using these image contents and cropped the 3D area against the region of interest in our system. The volume data, which only include color values, additionally require segmented information because we cannot recognize the shapes of objects on the basis of color values. These segmented data are created manually or semiautomatically. After segmentation, we can smooth out these segmented data with a Gaussian filter because the boundaries of objects achieved through manual segmentation are clumsy.

To begin with, we separate the volume data into several sub-volumes, with user-specific parameters under CUDA. Each sub-volume has its own resolution width, height, and length. These size data are used as the scale factor of the corresponding proxy geometry during the rendering phase.

Fig. 2. Overall procedure of the virtual endoscopy simulator. (A) Along the navigation path has restricted region, (B) visual representation of user manipulation using a virtual camera near the navigation path.
During the vertex process, this scale information is used as a parameter to scale a unit block \((1.0 \times 1.0 \times 1.0)\) into preferred block sizes. The scaled blocks are used as proxy geometry for a volume with a scaling operation. Because proxy geometry can be used to skip an empty space and terminate an early ray of GPU ray casting during the fragment process, we can accelerate the rendering speed. In the ray-sampling step, we render an image in a frame buffer, which refers to the multiple 3D textures of the sub-volumes separated on CUDA and a predefined opacity transfer function.

**Virtual endoscopy simulator**

Our virtual endoscopy simulator is also operated using user’s gestures. To provide an immersive environment, the simulation system includes an HD stereoscopic display. After virtual dissection, we can select an internal organ to perform virtual endoscopy. The navigation path, which is composed of several control points, and the distance map\(^{28,29}\) are prepared in advance because our endoscopy simulator is designed for medical training and not for diagnostic purposes. In the rendering step, we calculate the position of the virtual camera using rendering points by applying control points and distance values.

**Computing navigation path using VCS**

In the virtual endoscopy part of our simulator, we glide through the interiors of pipe-like organs, and therefore, computing the navigation path is important. We defined some control points using the center of gravity within an organ cavity.\(^{30,31}\)

Fig. 2 shows the overall procedure of the virtual endoscopy simulator. Along the navigation path has restricted region (Fig. 2A). The camera movement needs to be restricted because the camera should glide through the wrinkled wall of a tubular organ. If the degree of freedom of the camera movement increases, an unskilled person may lose track of the camera direction while navigating within the interior of a wrinkled wall. Therefore, we define a restricted region—i.e., the VCS (Fig. 2A)—for the camera movement, using a series of spheres along the navigation path. This region helps avoid a wrinkled wall during navigation, and it is defined by a distance \(r\) centered with the current position of the camera. The camera can move a distance \(d\) from the current position through user gestures. This distance of user movement \(d_{\text{user}}\) should be less than distance \(r\) of the VCS. The right side of Fig. 2B shows the process of camera movement using VCSs. During navigation, the camera can rotate in yaw and pitch and move forward and backward. In addition, it rotates at an angle within a field of view of 70° relative to the marching direction because this angle is the field of view of the endoscope camera. The speed of the camera movement is adjustable, depending on user navigation skills.

If we place VCSs uniformly along the navigation path instead of setting their size, when the size of a VCS has been fitted for a large inside area, the camera may clash with a wall in a smaller inside area. We have to avoid such clashes and observe as wide a range as possible. Therefore, we set the radius of VCSs adaptively using the diameter of the tubular organ.

The experiments were conducted on a system using an Intel Core i7-8700 CPU (3.70 GHz) and 16 GB main memory. An nVidia TITAN X (12 GB video memory) was used as the graphics processor. The Windows 10 operating system and a Kinect SDK 1.8 were also used for the gesture-recognition experiment.
RESULTS

We used tubular organs of the VK dataset, such as the respiratory tract (from the trachea to the segmental bronchus) and gastrointestinal tract (the esophagus, stomach, and large intestines), without small intestines (as provided in Table 1). We used these datasets at the resolution of the original dataset and cropped the dataset sufficiently for inclusion in the video memory.

Fig. 1 shows an example of virtual dissection. This helps us understand the internal structure of an organ itself and the interrelationships of the neighboring organs. After virtual dissection, we can select an organ to navigate inside. Yellow color indicates that the organ (in this case, the large intestines) has been selected by user gestures. When we select the respiratory tract, the user can fly through the inside of the respiratory tract, which is from the trachea to the left and right bronchi of each lobar bronchus.

Fig. 3 shows the results of virtual colonoscopy from the anus to the cecum. The camera movement is controlled through user gestures acquired from the Kinect device. During navigation control using such gestures, the camera can rotate in yaw and pitch and move forward and backward. A collision is detected when the user exits the VCS region, which is helpful in virtual endoscopy training.

The virtual colonoscope enters rectilinearly from the anus to rectum (Fig. 3A). In the sigmoid colon, the colonoscopy moves in the sigmoid shape (Fig. 3B). In the descending, transverse, and ascending colons, the colonoscope passes straightly, but it turns in...
their junctions (Fig. 3D and E). In the cecum, the colonoscope does not go to the ileum. Furthermore, the fold by the appendix can be verified.

The right-side column in Fig. 4 shows virtual bronchoscopy images generated with our simulator. These are rendered images obtained by navigating from the entrance of the trachea to the right bronchus. When we enter the trachea and reach the carina, there is the intersection of the left and right bronchi. Going further into the carina, we find that the left bronchus has a gentle slope and the right bronchus has a steep slope. If we go further, we see several branches, which are called the lobar bronchus. We can go to the segmental bronchus finally. We can also verify the bronchial wrinkles clearly through a realistic color and shape. In addition, a navigation map depicting the current position is provided at the bottom-left of each frame. The left-side column of Fig. 4 shows the results of virtual esophagogastroscope. The camera movement is controlled using user gestures from the Kinect device. The user can move the navigation camera freely inside the VCS region. A collision is detected when the user escapes the VCS region, and this collision detection is helpful in virtual endoscopy training.

**DISCUSSION**

There is not enough time to diagnose a disease during real endoscopy, and we cannot practice it as well. We should not practice on living things ethically, so our simulator is a solution for practice. During navigation, it is possible to effectively observe the entire area of the colon because the user can control the camera freely.

---

**Fig. 4.** Virtual esophagogastroscope and virtual bronchoscopy. (A) The virtual esophagogastroscope enters in the stomach through the esophagus and (B) the cardia of the stomach. (C) The scope passes the longitudinal folds inside the stomach. (D, E) The virtual bronchoscope encounters two ways (right and left main bronchi) and a split spot (carina). (F) in right lung with three lobes, the scope encounters two entrances (middle lobar and inferior lobar bronchi).
In most virtual anatomy software packages, a keyboard, a mouse, and joysticks are used. But user gesture, in a manner similar to this study, is more suitable for practicing real diagnosis and surgery. Several touchless interfaces using gesture recognition have been studied in the medical field.\textsuperscript{32,34} Ruppert et al.\textsuperscript{32} implemented a gesture recognition system based on the Kinect device that enables a surgeon to touchless navigate through the image in the intraoperative setting using a computer. Chiang et al.\textsuperscript{33} presented a novel medical imaging observation system using gesture-based technique to build a touchless interactive environment. The gesture interface can help in teaching realistic clinical medicine and anatomy. Fig. 5 shows an operating person using our simulator and a Kinect device. The user can move the virtual camera during virtual endoscopy using user’s gestures and perform virtual dissection.

Our method can observe the endoscope images from various angles by manipulating the motion of the virtual endoscope camera using motion recognition interface.

Several applications to visualize the internal structure of the human body have been introduced. However, these applications do not provide a realistic simulation environment or adequate training opportunities because they simply display a 2D image and use common interface. Our system includes an HD stereoscopic display device. Our method can provide a rendering speed of 70 fps even though a high-resolution stereoscopic device. Our interactive anatomy simulator that generates realistic images and provides gesture-based interactions is useful for educating in the realistic environment. It also includes an immersive display incorporating a motion-recognition sensor. We also provide three virtual endoscopy (colonoscopy, bronchoscopy and esophagogastroscopy) using the colon, bronchus, and stomach model.

In our previous study,\textsuperscript{35} we created 671 endoscopic images of the colon for virtual colonoscopy and labeled them and provided educational colonoscopy tutorial software only using these fixed images. Since images made along the central path of the large intestine have a limited viewing direction, this paper improved the viewing direction interactively near the central path according to the user’s intention (Fig. 2). And the interface with the user was developed to utilize the user’s gestures (Fig. 5).
Our simulator would be helpful in educating medical students because it provides an immersive environment. In addition, the gesture-based interface of our simulator would be useful for reducing contagious infections.

ACKNOWLEDGMENTS

Raw data of the Visible Korean Human were acquired by the assistance from the Korea Institute of Science and Technology Information.

REFERENCES

1. Ereso AQ, Garcia P, Tseng E, Gauger G, Kim H, Dua MM, et al. Live transference of surgical subspecialty skills using telerobotic proctoring to remote general surgeons. *J Am Coll Surg* 2010;211(3):400-11.
2. Garcia P. Telemedicine for the battlefield: present and future technologies. In: Rosen J, Hannaford B, Satava RM, editors. *Surgical Robotics: Systems Applications and Visions*. Berlin: Springer; 2011, 33-68.
3. Gargiulo T, Helgason T, Ingvarsson P, Mayr W, Kern H, Carraro U. Medical image analysis and 3-D modeling to quantify changes and functional restoration in denervated muscle undergoing electrical stimulation treatment. *Hum-Cen Comput Info* 2012;2(10):1-11.
4. Bostanci E, Kanwal N, Clark AF. Augmented reality applications for cultural heritage using Kinect. *Hum-Cen Comput Info* 2015;5(20):1-18.
5. Park Y, Lee M, Kim MH, Lee JW. Analysis of semantic relations between multimodal medical images based on coronary anatomy for acute myocardial infarction. *J Inform Proc Syst* 2012;2(10):1-11.
6. Ren Z, Yuan J, Zhang Z. Robust hand gesture recognition based on finger-earth mover’s distance with a commodity depth camera. In: *Proceedings of the 19th ACM International Conference on Multimedia*. New York, NY: Association for Computing Machinery; 2011, 1093-6.
7. Zhang Z. Microsoft Kinect sensor and its effect. *IEEE Multimed* 2012;19(2):4-10.
8. Anderson F, Annett M, Bischof WF. Lean on Wii: physical rehabilitation with virtual reality Wii peripherals. *Stud Health Technol Inform* 2010;154:229-34.
9. Deutsch J, Robbins D, Morrison J, Bowlby PG. Wii-based compared to standard of care balance and mobility rehabilitation for two individuals post-stroke. In: *Proceedings of the Virtual Rehabilitation International Conference*. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2009, 117-20.
10. Potter LE, Araulio J, Carter L. The leap motion controller: a view on sign language. In: *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration*. New York, NY: Association for Computing Machinery; 2013, 175-8.
11. Sinclair B. Sony reveals what makes PlayStation Move tick. https://www.gamespot.com/articles/sony-reveals-what-makes-playstation-move-tick/1100-6253435/. Updated 2010. Accessed April 7, 2016.
12. Nuwer R. Armband adds a twitch to gesture control. *New Sci* 2013;217(2906):21.
13. Anatomage Inc. https://www.anatomage.com. Accessed April 7, 2016.
14. Lundström C, Rydell T, Forsell C, Persson A, Ynnerman A. Multi-touch table system for medical visualization: application to orthopedic surgery planning. *IEEE Trans Vis Comput Graph* 2011;17(12):1775-84.
15. Spitzer V, Ackerman MJ, Scherzinger AL, Whitlock D. The visible human male: a technical report. *J Am Med Inform Assoc* 1996;3(2):118-30.
16. Ackerman MJ. The Visible Human project. A resource for education. *Acad Med* 1999;74(6):667-70.
17. Park JS, Chung MS, Hwang SB, Lee YS, Har DH, Park HS. Visible Korean human: improved serially sectioned images of the entire body. *IEEE Trans Med Imaging* 2005;24(3):352-60.

18. Chung MS, Kim SY. Three-dimensional image and virtual dissection program of the brain made of Korean cadaver. *Yonsei Med J* 2000;41(3):299-303.

19. Shin DS, Park JS, Park HS, Hwang SB, Chung MS. Outlining of the detailed structures in sectioned images from Visible Korean. *Surg Radiol Anat* 2012;34(3):235-47.

20. Park JS, Jung YW, Lee IW, Shin DS, Chung MS, Riemer M, et al. Generating useful images for medical applications from the Visible Korean Human. *Comput Methods Programs Biomed* 2008;92(3):257-66.

21. Huang YX, Jin LZ, Lowe JA, Wang XY, Xu HZ, Teng YJ, et al. Three-dimensional reconstruction of the superior mediastinum from Chinese Visible Human Female. *Surg Radiol Anat* 2010;32(7):693-8.

22. Zhang SX, Heng PA, Liu ZI, Tan IW, Qiu MG, Li QY, et al. The Chinese Visible Human (CVH) datasets incorporate technical and imaging advances on earlier digital humans. *J Anat* 2004;204(Pt 3):165-73.

23. Lim S, Kwon K, Shin BS. GPU-based interactive visualization framework for ultrasound datasets. *Comput Animat Virt W* 2009;20(1):11-23.

24. Park JS, Chung MS, Hwang SB, Lee YS, Har DH, Park HS. Visible Korean Human: improved serially sectioned images of the entire body. *IEEE Trans Med Imaging* 2005;24(3):352-60.

25. Shin DS, Jang HG, Hwang SB, Har DH, Moon YL, Chung MS. Two-dimensional sectioned images and three-dimensional surface models for learning the anatomy of the female pelvis. *Anat Sci Educ* 2013;6(5):316-23.

26. Schiemann T, Freudenberg J, Pflesser B, Pommert A, Priesmeyer K, Riemer M, et al. Exploring the Visible Human using the VOXEL-MAN framework. *Comput Med Imaging Graph* 2000;24(3):127-32.

27. Park JS, Chung MS, Hwang SB, Lee YS, Har DH, Park HS. Technical report on semiautomatic segmentation using the Adobe Photoshop. *J Digit Imaging* 2005;18(4):333-43.

28. Felzenszwalb PF, Huttenlocher DP. Distance transforms of sampled functions. *Theory Comput* 2012;8:415-28.

29. Maurer CR, Qi R, Raghavan V. A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions. *IEEE T Pattern Anal* 2003;25(2):265-70.

30. Kwon K, Shin BS. An efficient navigation method using progressive curverization in virtual endoscopy. *Int Congr Ser* 2005;1281:121-5.

31. Barry PJ, Goldman RN. A recursive evaluation algorithm for a class of Catmull-Rom splines. In: *ACM SIGGRAPH 88 Computer Graphics: Conference Proceedings*. New York, NY: Association for Computing Machinery; 1988, 199-204

32. Ruppert GC, Reis LO, Amorim PH, de Moraes TF, da Silva JV. Touchless gesture user interface for interactive image visualization in urological surgery. *World J Urol* 2012;30(5):687-91.

33. Chiang PY, Chen CC, Hsia CH. A touchless interaction interface for observing medical imaging. *J Vis Commu Image R* 2019;58:363-73.

34. Klapaï I, Klapaï L, Majhen Z, Duspara A, Zlatko M, Kubaï G, et al. Do we really need a new navigation-nanvasive “on the Fly” gesture-controlled incisionless surgery? *Biomed J Sci Tech Res* 2019;20(5):15394-404.

35. Chung BS, Chung MS, Park HS, Shin BS, Kwon K. Colonoscopy tutorial software made with a cadaver’s sectioned images. *Ann Anat* 2016;208:19-23.