Comparing the Microsoft® Kinect™ to a Traditional Mouse for Adjusting the Viewed Tissue Densities of Three-Dimensional Anatomical Structures

Monica Berron[1], Adriana Philip[2], Jordan Williams[1]
Bethany Juhnke[3], BS, Eric Foo[3], PhD, Eliot Winer[3], PhD
Iowa State University, Virtual Reality Applications Center, Ames, IA, 50011

ABSTRACT

Advancements in medical image visualization in recent years have enabled three-dimensional (3D) medical images to be volume-rendered from magnetic resonance imaging (MRI) and computed tomography (CT) scans. Medical data is crucial for patient diagnosis and medical education, and analyzing these three-dimensional models rather than two-dimensional (2D) slices would enable more efficient analysis by surgeons and physicians, including non-radiologists. An interaction device that is intuitive, robust, and easily learned is necessary to integrate 3D modeling software into the medical world. The keyboard and mouse cannot readily manipulate 3D models because these traditional interfaces are optimal with two degrees of freedom, not the six degrees of freedom present in three dimensions. Using a familiar, commercial-off-the-shelf device as an interaction device would minimize training time and enable maximum usability in 3D. Different techniques could be used to manipulate 3D medical imaging and provide doctors more innovative ways of visualizing patient data; the technique of windowing to adjust the viewed tissue density of 3D medical volumes would enable doctors to observe select tissue types in three dimensions.

A software package, developed in house, called Isis will be used to visualize and interact with the three dimensional representations of medical data. In this paper, we present the methodology of a pilot study that will examine the usability of windowing 3D medical imaging when comparing the commercial-off-the-shelf device Kinect™ to the traditional mouse.

BACKGROUND

Usability

Usability is an essential factor for integrating 3D imaging into medical environments, but a major barrier for implementing the use of 3D medical imaging into clinical practice is the time-consuming training required to learn a new interaction device[1]. Commercial-off-the-shelf (COTS) hardware such as the Nintendo® Wii Remote™, Microsoft® Kinect™, and gamepad are interaction devices with which medical professionals have some familiarity because they are available to the public, unlike software-specific devices. These devices may enable 3D spatial interaction more efficiently than the traditional mouse because they manipulate position and orientation with six degrees of freedom (DOF) [2]. COTS interaction devices would minimize training time and provide cost-efficient alternatives to software-specific devices.

[1]University of Maryland, Baltimore County
[2]The Pennsylvania State University
[3]Iowa State University

Commercial-off-the-Shelf Devices

After the release of the Nintendo® Wii™, developers discovered the potential of the Wii Remote™ as a COTS device for interacting with a 3D virtual environment (Figure 1). In 2007, a system for head tracking using head-mounted infrared lights was developed which gave positional data to track the user's head when detected by the Wii Remote™ cameras [3]. Research groups began to look at the Wii Remote's™ capabilities for pointing and aiming in a head-mounted virtual reality space. One study analyzed two Wii Remote™ based tracking methods specifically for human computer interaction research because the Wii Remote offered previously unheard of tracking capabilities for a device that was inexpensive in comparison to older devices [4]. However, the study concluded that the Wii Remote™ was imprecise, relying on nothing but two infrared points for location and an accelerometer for each axis of rotation [5].

Figure 1: The Nintendo® Wii Remote™ is a COTS device that operates via one-handed navigation.

The Microsoft® Kinect™ is regarded as an effective COTS solution for hands-free manipulation of 3D volumes (Figure 2) [6]. The Kinect™ uses two cameras and an array of infrared points projected onto the subject. The Kinect™ takes point locations from both cameras and uses the disparities to triangulate each point's position in 3D space. There is another RGB camera for overlaying an actual image onto the 3D mesh to create a lifelike partial 3D model of the subject. In order to sense human movement and hand gestures, the Kinect™ has built-in skeletal tracking which approximates the position of the user's limbs [7]. The Kinect™ has high accuracy, and its ease of use facilitates recognition of hand gestures and natural movement.

Figure 2: The Microsoft® Kinect™ is a COTS device that operates via touch-free navigation.

The gamepad is a controller that uses the fingers and thumbs to provide user input, and it has been used as a COTS device for interacting with 3D virtual environments because it is ergonometric and low cost (Figure 3). It has been found that the gamepad enables easy learning as an interaction device, and its various buttons provide numerous possibilities for manipulation input [8]. The gamepad also offers a high degree of precision and control because of the inclusion of dual analog joysticks [8].

Figure 3: The gamepad is a COTS device that operates via two-handed navigation.

Limited research on the Wii Remote™, the Kinect™, and the gamepad as interaction devices has been conducted to determine how users, including both medical and non-medical personnel, prefer to interact with 3D virtual environments. One study compared the Wii Remote™ and Kinect™ as interaction devices for 3D geographical mapping and used yaw, pitch, and roll gestures to navigate. In the study, the Computer System Usability Questionnaire provided results on a scale of 1 (strongly disagree) to 7 (strongly agree). The Kinect™ performed tasks easier and with less variability (μ=5.4, σ=0.82) than the Wii Remote™ (μ=5.17, σ=0.94). The study also found that the Kinect™ was observed to be more efficient (μ=5.39) than the Wii Remote™ (μ=4.41) because users found it less distracting. The study's findings concluded that "the more the interface is natural (in the sense that it disappears behind the gesture) the more the users are involved in the virtual environment and hosted activities" [9]. Another study compared the Wii Remote™, gamepad, and mouse and keyboard as interaction devices for rotating a virtual object in one task and changing the object's path in a second task. Findings of the study included that the average time to perform both tasks was slowest for the Wii Remote™ (29.09 s), second slowest for the gamepad (21.55 s), and fastest for the mouse and keyboard (20.40 s). Participants were unable to complete the rotation task using the Wii Remote™ in 80% of total attempts, and the Wii Remote™ was selected as the least favorite interface by 90% of the participants [8].

Maintaining a sterile environment, in addition to usability, is an important consideration for integrating 3D medical imaging into medical environments. Sterility is extremely critical in medical settings, especially operating rooms, intensive care units, and autopsy suites. While the gamepad and Wii Remote™ provide the 3D motion mapping that the traditional mouse does not, such interaction devices require physical contact with the medical user and could increase contamination by transferring pathogens [10]. The Microsoft® Kinect™ provides touch-free navigation, allowing remote manipulation of 3D medical images through hand gestures [9].

Windowing

Windowing is defined as the method of adjusting the viewed tissue density of a 3D medical image. The technique of windowing can be implemented by specialists who wish to view select tissue types of 3D medical images (Figure 4). Orthopedic specialists, for example, could utilize windowing as another method to track changes in bone density over time for people at risk of developing osteoporosis [11]. Previous research has been mainly focused on rotation, clipping, zooming, and translation of 3D medical imaging; less has been conducted on windowing [12] [13]. In one study, ten medical professionals tested a Kinect™ and Apple Voice recognition system and a traditional mouse and keyboard as interaction devices to recreate medical screenshots using windowing, rotation, and zooming techniques. While the mean times to reproduce the screenshots using the Kinect™ and Apple Voice recognition interaction device (75.1 s) and a traditional mouse and keyboard (52.1 s) were compared, specific data on how the medical professionals interacted with the windowing technique were not discussed [13].

Figure 4: The windowing technique enables medical professionals to adjust the viewed tissue density of 3D medical images. In this figure, the same 3D cardiac image is displayed in four different viewed tissue density ranges.

RESEARCH QUESTION

In comparing the Kinect™ to the traditional mouse, can users accurately and efficiently identify anatomical structures by adjusting viewed tissue density of 3D medical volumes?

METHODS

In this experiment, the ISIS software will be used to depict the medical data in 3D. The ISIS software includes zooming, rotating, coloring, clipping and windowing functionality; however, all features except windowing will be disabled for this experiment.

Undergraduate and graduate students with novice knowledge of gross anatomy will participate in this pilot study. The participants' previous experiences viewing 3D medical imaging and previous experiences using a Microsoft Kinect™ will be recorded.

In this pilot study, each participant will be randomly assigned an interaction device, Kinect™ or mouse, at the beginning of the study. Each participant will be given a pre-survey and trained on how to use the Kinect™ or mouse to window.

Each participant will be shown an anatomical region on the computer screen and instructed that he/she will have up to one minute to locate a specific anatomical structure by only windowing (Figure 5). Once a participant indicates that he/she had adjusted the position of the tissue density range such that the target anatomical structure is displayed most clearly, researchers will take a screenshot and ask the participant to circle the structure (Figure 6). These screenshots will be evaluated to determine whether the participant correctly identified the anatomical structure. If the participant is unable to successfully locate the anatomical structure during the one minute period, the next anatomical region will be displayed and the participant will be asked to indicate the new specified anatomical structure. The procedure of showing the participant an anatomical region and asking him/her to identify an anatomical structure will be repeated for a total of eight anatomical structures in two different anatomical regions.

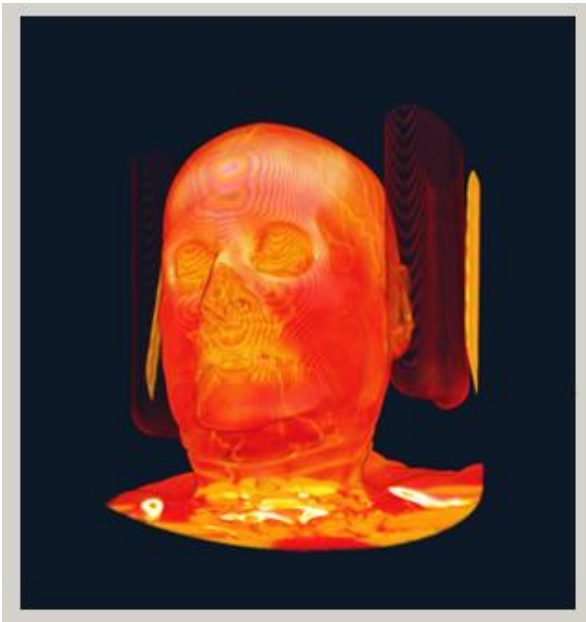

Figure 5: An anatomical region is shown, and the participant will be instructed that he/she has one minute to locate a specific structure by only windowing. For this specific anatomical region, the participant will be instructed to locate the teeth.

Figure 6: When the participant indicates that he/she has windowed to display the target anatomical structure displayed most clearly, researchers will take a screenshot and ask the participant to circle the structure. In this case, the teeth will be the anatomical structure to identify.

Once the first test is completed, the participant will be given training for the other interaction device, mouse or Kinect™, that he/she had not used prior. The participant will be asked to identify the same eight anatomical structures using the same methodology used for the first interaction device. Screenshots will again be captured if the participant indicates that the target anatomical structures are clearly displayed within the one-minute time periods. These screenshots will be later evaluated to determine whether the participant correctly identified the anatomical structure.

FUTURE WORK

In the near future, a user study will be performed with participants from the target demographic of medical students to determine how medical personnel prefer to interact with 3D medical imaging. Medical students are the ideal audience for the user study because they have advanced knowledge of gross anatomy and will use medical imaging in their careers as physicians and surgeons.

The methodology of the user study will be similar to the methodology developed for the pilot study. User study participants will also be asked to identify anatomical structures; however, the anatomical structures to be tested will require an advanced knowledge of anatomy and are not necessarily the same anatomical structures tested in the pilot study.

The primary focus of the pilot study will be on determining the effectiveness of the developed software and design of the experiment. Software development will continue as participants' feedback from the pilot study increases understanding of users' preferences for interacting with 3D medical imaging. The focus of the user study will be specifically on determining if medical personnel can accurately and efficiently identify anatomical structures of 3D medical volumes when comparing the Kinect™ to the traditional mouse.

REFERENCES

[1] Montgomery, K., Stephanides, M., Schendel, S., and Ross, M. User Interface Paradigms for Patient-specific Surgical Planning: Lessons Learned over a Decade of Research, *Computerized Medical Imaging and Graphics 29* (2005), 203–222.

[2] Gallo, L., Minutolo, A., and De Pietro, G. A User Interface for VR-ready 3D Medical Imaging by Off-the-shelf Input Devices. *Computers in Biology and Medicine 40,* (2010), 350-358.

[3] Lee, J. Hacking the Nintendo Wii Remote, *Pervasive Computing, IEEE 7*(2008), 39-45.

[4] Rehg, J.M., and Kanade, T. DigitEyes: Vision-based Hand Tracking for Human-Computer Interaction. *Motion of Non-Rigid and Articulated Objects, 1994, Proceedings of the 1994 IEEE Workshop on* , pp.16-22, 11-12 Nov 1994.

[5] Chow, Yang-Wei. The Wii Remote as an Input Device for 3D Interaction in Immersive Head-Mounted Display Virtual Reality. *Proceedings of the IADIS International Conference on Game and Entertainment Technologies* (2008), 85-92. Amsterdam, Netherlands: IADIS Press.

[6] Rothberg, A. and Bailey, J. Manipulating Medical Images: A Hands-Off Approach. *Southwest Decision Sciences Institute 43rd Annual Meeting* (2012).

[7] Xia, L., Chen, C., and Aggarwal, J.K. Human Detection Using Depth Information by Kinect. *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on* (2011), 5-22, pp.15-22, 20-25 June 2011.

[8] Ardito, C., Buono, P., Costabile, M.F., Lanzilotti, R., and Simeone, A.L. Comparing Low Cost Input Devices for Interacting with 3D Virtual Environments. *Human System Interactions, 2009. HSI '09. 2nd Conference on* (2009), pp. 292-297.

[9] Francese, R., Passero I., and Tortora, G. Wiimote and Kinect: Gestural User Interfaces add a Natural third dimension to HCI. In Proceedings of the International Working Conference on Advanced Visual Interfaces (2012), pp. 116-123.

[10] Hartmann, B., Benson, M., and Junger, A. Computer Keyboard and Mouse as a Reservoir of Pathogens in an Intensive Care Unit. *J Clin Monit Comput. 18* (2004), 7-12.

[11] Kanis, J. A., Melton, L. J., Christiansen, C., Johnston, C. C., and Khaltaev, N. The Diagnosis of Osteoporosis, *J Bone Miner Res 9* (1994), 1137–1141.

[12] Gallo, L., Placitelli, A.P., and Ciampi, M. Controller-Free Exploration of Medical Image Data: Experiencing the Kinect. *Computer-Based Medical Systems (CBMS), 2011 24th International Symposium on*, (2011), 1-6, 27-30.

[13] Ebert, L.C., Hatch, G., Ampanozi, G., Thali, M.J., and Ross, S. (2011) You Can't Touch This: Touch-Free Navigation Through Radiological Images. *Surg Innov*. (2011).