

Multi-Modal VR for Medical Simulation



C Indhumathi¹, Wenyu Chen¹ and Yiyu Cai^{1,2}

¹ School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore

² Bio-informatics Research Center, Nanyang Technological University, Singapore

Abstract—Over the past three decades computer graphics and virtual reality (VR) have played a significant role in adding value to medicine for diagnosis and treatment applications. Medical simulation is increasingly used in medical training and surgical planning. This paper investigates the multi-modal VR interface for medical simulation focusing on motion tracking, stereographic visualization, voice navigation, and interactions. Applications in virtual anatomy learning, surgical training and pre-treatment planning will also be discussed.

Index Terms—Computer Graphics, Medical Imaging, Volume Rendering, Virtual Reality.

I. INTRODUCTION

Computerized X-ray tomography (CT), magnetic resonance imaging (MRI), ultrasound imaging, etc., have revolutionized the medical diagnosis and pre-treatment planning. These scanning technologies have empowered the medical doctors to obtain three-dimensional (3D) recording of the human anatomy or pathology as a stack of two-dimensional (2D) images. In recent years, there has been a rapid increase in the application of VR technology for medical diagnosis, pre-treatment surgical planning and training [1-6].

For efficient medical diagnosis, surgical planning and training, surgeons need to navigate through the 3D anatomical data to gain access to the diseased lesion and analyze them for further treatments. In order to achieve this task, surgeons need a platform to visualize and interact with the medical datasets in order to gain better understanding of the patients and their diseases.

In this paper, we present a high-end human-computer interface for potential medical applications. It is a multi-modality VR environment composed of motion tracking, stereographic visualization, voice navigation, and interactions. Section II discusses the background and motivation for this project. Section III presents an overview of the proposed system. Section IV describes the medical applications in virtual anatomy learning, surgical training and pre-treatment planning. Section VII concludes this research.

II. BACKGROUND AND MOTIVATION

VR is an enabling technology, which offers the possibility to visualize 3D images, with real spatial perception [7]. An

immersive virtual reality (IVR) system allows individual users to immerse themselves fully or partially in the virtual environment and facilitates the efficient exploration of three-dimensional objects [8]. Users can navigate in the 3D environment simply with the aid of visual, audio and other sensorial devices. One of the “selling points” of VR is that it allows users to not only visualize information but also to interact directly and naturally with the data in three dimensions [9]. Interaction can support perception by increasing the efficiency in which users can extract meaningful information from the 3D data. In contrast to desktop-based 3D visualization applications, the uses of stereoscopic displays and spatial interaction devices provide users, a better and more effective 3D experience of datasets.

The interactive nature in VR implies that this advanced human-computer interface plays a prominent role in designing effective visualization applications [10]. With the constant increase in availability and performance of 3D graphics cards, there comes a parallel need to perform operations in 3D environments such as navigation and control intuitively and naturally. In the last 15 years, there is no much changes in the techniques used to interact with computers, and more generally, with digital contents. The window, icons, menus, pointer (WIMP) paradigm using mouse and keyboard is still by far the most common human-computer interface. The research community has been continuously looking for more intuitive and easy to use interfaces [11]. Previous research has shown that manipulation operations are most effective with the ability to operate in six degrees of freedom (DOF) simultaneously [12]. Though keyboard and mouse can be used to navigate in three dimensions, they are not designed for 6 DOF. Thus they lack both intuitive and analogue control. Joysticks are another cheap yet dedicated tool available (can be analogue), offering more than 3 DOF by adding “hat” buttons (mini joysticks) and sliders, which are not again very intuitive [13]. To have intuitive interaction in VR environment, many devices have been developed. Data glove and VR glove devices are designed for native hand-based manipulation. Motion-tracking devices are designed for capturing body or part of body movement. Force feedback devices are designed for haptical operations. Voice navigation devices are designed for audio input.

These multimode interaction technologies offer promising alternatives to existing interfaces because they emulate the natural way in which humans communicate and they provide an useful approach to the effective 3D exploration.

III. SYSTEM OVERVIEW

The system architecture of our multi-modal VR environment is shown in Fig 1. The system basically can be grouped into three major modules namely VR computing platform, VR intuitive devices and VR engine. We have used high-performance graphics workstations, high-speed rendering hardware, voice navigation and 6 DOF devices to deliver an efficient and high quality outputs in real time. Together with a real time rendering, stereovision and 3D interactions, our system provides users a full-fledged 3D interactive VR environment for medical applications.

3.1 VR Computing Platform

3.1.1 Operating System

The multi-modal VR system was developed on the Microsoft Windows platform. The fundamental computer hardware used for the system development is an Intel® Pentium-IV workstation equipped with 2.40GHz CPU, more than 256 MB RAM, accelerated graphics port (AGP) and peripheral component interconnect (PCI) bus slots.

3.1.2 Graphics system

While volume rendering is very popular, the lack of interactive frame rates has limited its widespread use. One of the most promising approaches to achieve real-time frame rates for volume rendering is the use of special-purpose volume rendering hardware. New commodity hardware systems are rapidly emerging for real-time 3D visualization [14]. After a detailed analysis on the algorithms and platforms, the graphics-rendering engine was setup using 2 GB *VolumePro* 1000card by TeraRecon Inc. [15] and nVIDIA GeForce graphics processor running in parallel to render the 3D medical dataset (either volumetric or geometric). Conventional graphics resource such as OpenGL/DirectX is used in the VR system.

3.1.3 Imaging System

A large number of digital imaging modalities are now being used on a routine basis in normal clinical practice. We mainly used CT and MRI medical images for training and surgical planning. Both CT and MRI imaging modalities typically create 2D image of the structures in a thin section of the body by consecutive scanning in x-y plane. These image slices are stacked to form a 3D volume. In addition, our imaging system can also accept volumetric data for cellular images captured with laser fluorescent confocal.

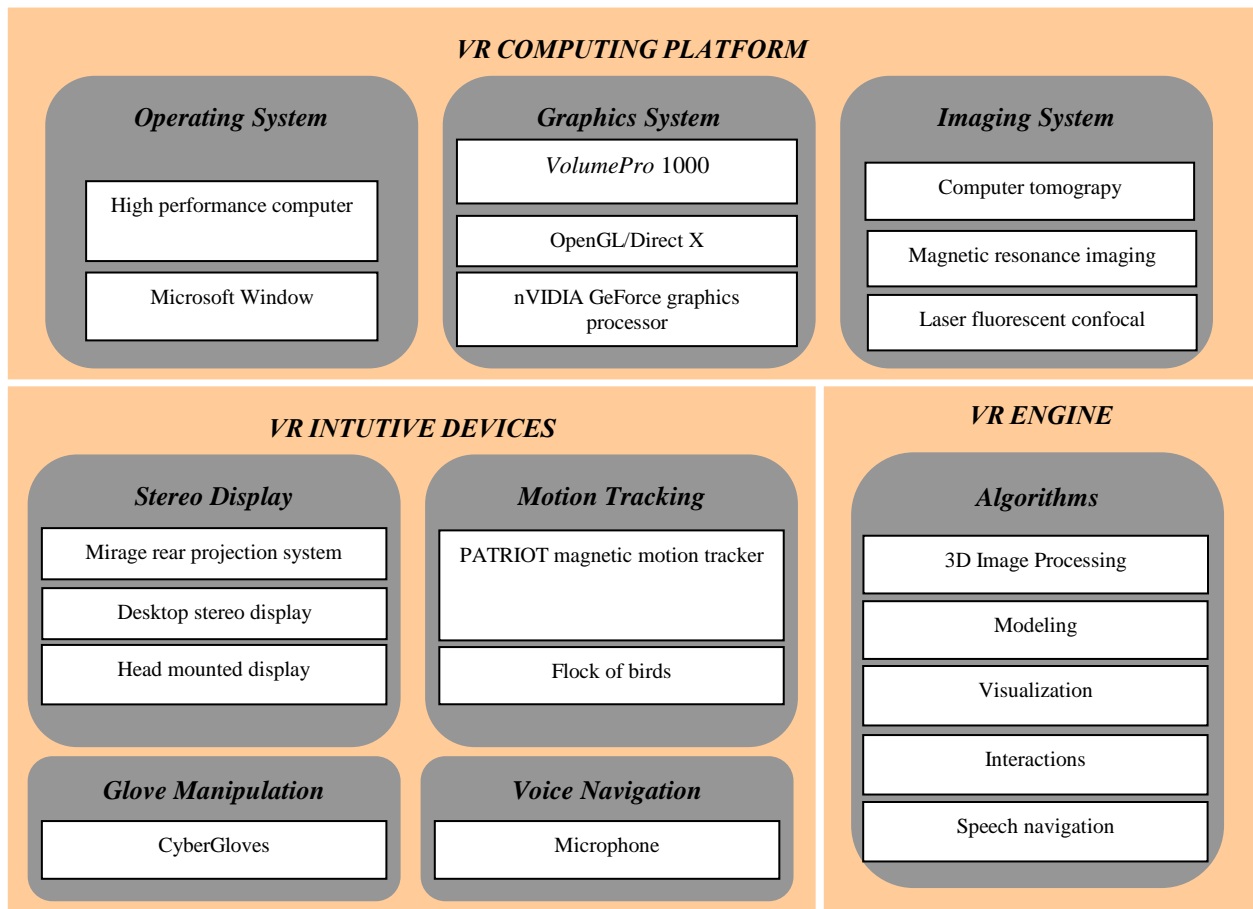


Fig. 1. System architecture of multi modality VR environment

3.2 VR Intuitive Devices

3.2.1 Visual Display Devices

A visual display device is an integral part of a virtual environment (VE) system. Depending on the different degree of immersion, 3D visual display technology can be categorized as desktop displays, head-mounted displays, arm-mounted displays and immersive displays. We have a ProView™ (Kaiser Electro-Optics, Fig 2) HMD with the stereoscopic display device embedded with two small cathode ray tubes (CRT) or Liquefied crystal display (LCD) in front of each eye and an adjustable lens system [16]. It can produce stereoscopic viewing by presenting separate overlapping images to each eye. It tracks the head movement of the user with a position tracker outfitted to HMD, and communicates it to the host computer. We also employ high refresh rate (100 Hz or above) CRT monitors for low cost desk-top based stereo display. To produce high immersive experience, a VR viewing system is built on a Stewart rear projection film-screen together with a Christie mirage 4000 (Fig 2). This high-end stereographic projector [17] is based on high-resolution SXGA 3-chip DLP (Digital Light Processing) technology offering high levels of image brightness and clarity, high bandwidth, frame rates, pixel and clock speeds for both active and passive stereo images.

3.2.2 Motion Tracking Device

Tracking is a critical component of any interactive and immersive environment. Motion tracking is a simple way to track an object's movements or actions as it changes position relative to a fixed point in 3D space. By locating and monitoring position and orientation of peripheral devices, the tracking system can keep update the coordinate information with respect to the virtual world. There are different types of tracking technologies in use today: electromagnetic, mechanical, acoustic, optical and inertial tracking systems. Each has their pros and cons. For local motion tracking, we use a PATRIOT magnetic motion tracking system (Polhemus, Fig 2) to determine the position and orientation of a remote object in 6 DOF [18]. The system essentially consists of a system electronics unit (SEU), source and stylus sensor. SEU, which contains a microcomputer and electronics computes the sensors position and orientation (6 DOF) relative to the source and provide an interface to the host computer. The transmitter allows the sensors to move up to a range of about 152 cm. For whole body tracking, we use Flock Of Birds (FOB) system (Ascension Technology, Fig 2). It is a modular tracker with 6 DOF for simultaneously tracking the position and orientation of one or more receivers (targets) over a specified range of ± 4 feet [19]. Motions are tracked to accuracies of 0.5° and 0.07 inch at rates up to 144Hz. The Flock measures the position and orientation of one or more receiving antenna sensors, typically located on a user's head, hand or body, with respect to a transmitting antenna, which is fixed in space. The microprocessor controls the transmitting and receiving

elements and converts the received signals into position and orientation outputs.

3.2.3 Glove Manipulation Device

Gloves offer far superior data input potential since they provide multiple degrees of freedom for each finger and the hand as a whole. By tracking orientation of fingers and relative position of hand, glove devices can track an enormous variety of gestures, each of which corresponds to a different type of data entry. We use a CyberGlove (Virtual Technologies, Fig 2) to provide data for hand visualization and tactile feedback from micro-vibrators [20]. It is an 18-sensor glove, which can obtain 18 joint angles in hands and flexions of wrist. The sensor resolution is about 0.5 degrees. Many applications require measurement of the position and orientation of the forearm in space. To accomplish this, mounting provisions for Polhemus and Ascension 6 DOF tracking sensors are available for the glove wristband. It has a software programmable switch and LED on the wristband to permit the system software developer to provide the CyberGlove wearer with additional input/output capability.



Fig. 2. VR Interactive Devices

3.2.4 Voice Navigation Device

Due to the noisy environment, a high quality microphone to obtain best results is used for voice navigation system. Close-talking headset microphones are best for passing commands to host PC as it is kept at a consistent distance from the user's mouth. The over head and binaural PC 136 USB

headset (Sennheiser Communications, Fig 2) is used in our VR system for voice recognition. It has a noise-canceling microphone, which comes with a USB sound card adaptor. The Microsoft Speech API (SAPI) [22] is used to provide a high-level interface between an application and the speech engine. SAPI implements all the low-level details needed to control and manage the real-time operations of the speech engine. The two basic types of SAPI engines are text-to-speech (TTS) systems and speech recognizers (SR). TTS systems synthesize text strings and files into spoken audio using synthetic voices. SR engine converts human spoken audio into readable text strings and files. In our system, we use SR engine to retrieve the voice commands from users. Command and control grammar is a type of recognizable utterances by SR engine, which has much higher recognition rate than dictation grammar. A list of voice commands is predefined by users and the SR engine recognizes the commands by matching users' speech with one of those predefined commands during run time of the application. SAPI uses extensible markup language (XML) to create the grammar.

3.3 VR Engine

3.3.1 3D Image Processing

We have been developing 3D medical image processing technique for various purposes [23-24]. Below are some of the typical functions.

- 3D noise-filtering
- 3D volume enhancement
- 3D boundary detection
- 3D dynamic segmentation
- 3D reconstruction, etc

For image stacks, automatic or semi-automatic algorithms are developed using various approaches.

3.3.2 Modeling

Over the years, we have been active in research for various modeling tasks including

- Bifurcation modeling
- Vascular network modeling
- Shape fitting and approximation
- Interpolation
- Tessellation
- Model simplification
- Collision detection
- Finite Element Analysis, etc.

For this, geometric continuity theory, Non-uniform Rational Bezier Surface (NURBS), subdivision and other techniques [25-27] are developed with an emphasis on 3D for the modeling purpose.

3.3.3 Visualization

Real-time volume rendering is a basic requirement for medical training. To achieve this, a special-purpose volume rendering hardware VolumePro 1000 board is employed for

high quality visualization. For conventional 3D, we use several nVidia GeForce graphics cards to enhance the graphics performance incorporating OPENGL/DirectX programming. Furthermore, various visualization functions are implemented

- Graphic optimization
- Level of details
- Stereographic rendering
- Texture mapping
- Animation, etc.

Trade-off between fast speed and high rendering quality is necessary in a VR system. Besides, compatible visualization is also an issue with different display solutions.

3.3.3 Interaction

Interaction is the soul of VR applications. Good VR interaction relies on both software and hardware devices. At the early phase of software design, we carefully investigate all interactive functions for better graphics user interface as well as for best selection of the interactive devices. In 3D VR environment, the emphasis is always placed on the easy use and native use of the graphics menu and interactive device. For medical simulation, typical interactions include

- Volumetric manipulation (e.g., cutting, crop, clipping)
- Object transformation (e.g., zoom, rotation, translation)
- Navigation (e.g., walk-through, fly-over)
- Selection
- Measurement, etc.

Generally, glove device is more suitable for manipulation and transformation, and stylus device is good for selection and measurement. Speech device can be used to execute commands via menu-item selection as well as to conduct more accurate controls positioning.

3.3.4 Integration

The VR environment is implemented using Microsoft Visual C++ programming based on Windows platform. Relevant libraries are integrated into the VR system:

- OpenGL/DirectX
- VolumePro library
- Polhemus Developer Interface (PDI) library
- FOB library
- CyberGlove library
- SAPI
- VLI library

As a high-end and integrated human-machine interface [8], our VR system can provide functions such as real-time rendering (volume or conventional geometry), immersive visualization, motion tracking, voice navigation, etc, for various applications.

IV. MULTI-MODALITY VR FOR MEDICAL SIMULATION

As an enabling technology, VR has found many applications from entertainment to engineering to military. We are interested in the research and development of VR technology for medical simulation. The multi-modality VR system provides new solution for training, diagnosis, and pre-treatment planning.

4.1 Virtual Anatomy Learning

VR is playing an increasingly important role in anatomic education. It can provide medical students comprehensive and effective learning procedures in a more flexible and less costly training environment. It allows the medical students to access to a wide variety of clinical scenarios including rare complications. The Visible Human Dataset is used in this research.

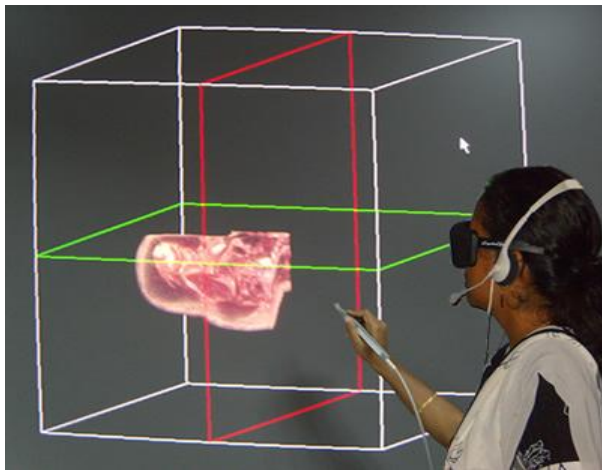


Fig. 3. Human anatomy and interactive training. (See Color Plate 1)

Users can view the human anatomy in a stereographic environment (Fig 3). Cross-sectioning of the volume rendered anatomic structure can be easily achieved by slice tracking. Detailed local structure can be identified with the aid of stylus device. Users can walk-through or fly-over the virtual anatomy via voice navigation.

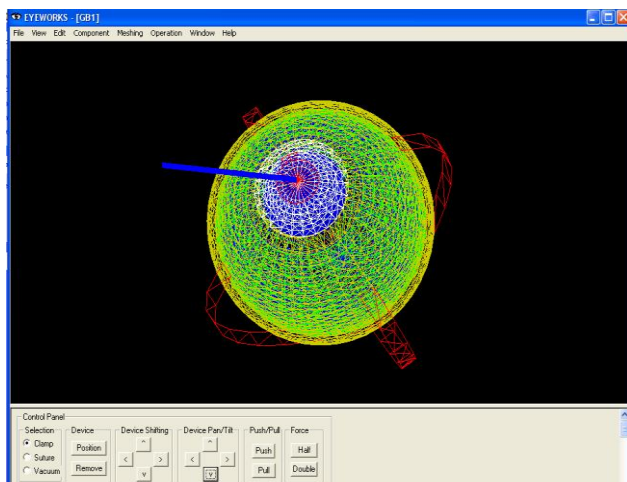


Fig. 4. Eye Model and Interactive Device

4.2 Ophthalmologic Surgery Simulation

Eye diseases are among the most common and rapidly growing public health concerns. Resembling a transparent watch-glass, the cornea is transversely ellipsoid (12 mm in the horizontal meridian and 11 mm in the vertical.) in the anterior aspect, whereas its posterior aspect is circular (about 11.5 mm in diameters). Normally, the cornea forms part of the surface of a sphere. In diseased condition, it is more curved in the vertical than in the horizontal meridian. The radius of curvature of the anterior surface is about 7.8 mm and that of the posterior is 6.5 mm. The sclera forms the posterior opaque approximately five-sixths of the fibrous tunic of the eye. The sclera is thickest behind and gradually becomes thinner when traced forwards. Its anterior portion is visible, and constitutes the “white” of the eye. The choroid is firmly attached to the margin of the optic nerve, and slightly placed at the points where vessels and nerves enter it. A transparent bi-convex body of crystalline appearance placed between the iris and the vitreous, the lens of the eye has an anterior surface and a posterior surface meeting at the equator. The anterior surface is the segment of a sphere whose radius averages 10 mm. The diameter of the lens is 9~10 mm. Its axial diameter varies markedly with accommodation. In this simulation, the cornea, lens, sclera, choroid and retina are 3D modeled.

Ophthalmologic surgery involves very dedicated procedures. The complications of ophthalmologic diseases often pose difficulties for surgeons to make both timely and accurate critical decisions. Subsequently, very intensive training is required for ophthalmic students and young ophthalmologic professionals in order to provide high quality surgical services. VR provides an alternative solution for ophthalmologic training. Interactive device can be used to give trainees simulated experience on ophthalmologic surgery.

In the VR environment, the 3D eye models and virtual device are created for the purpose to simulate operations such as laser surgery, clamping, and vacuuming. Finite element method is applied to model the multi-body deformation using the force-displacement approach under the surgical condition. Suturing operation can also be simulated. Stylus device can provide interactive experience by hands-on VR surgical simulation. Fig. 4 illustrates the interactive simulation of the ophthalmologic surgery.

4.3 Human Middle Ear Pre-treatment Planning

Otosclerosis is a middle ear disease that affects the stapes (the third of the chain of bones in the middle ear) by freezing the stapes into immobility. The immobilized stapes attached to oval window cannot effectively pass vibrations onto the fluids of the inner ear, and hearing is cut off. A surgical operation is usually performed to remove the middle ear ossicles and place a hearing implant.

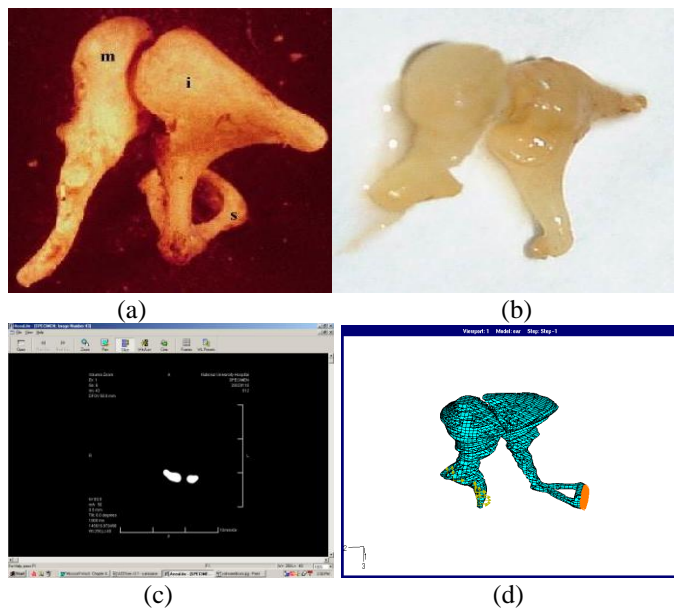


Fig.5. Middle Ear Otosclerosis: (a) Normal middle ear ossicles; (b) Damaged middle ear ossicles; (c) CT scans of middle ear ossicles; and (d) 3D reconstruction of the middle ear ossicles. (See Color Plate 2)

Pre-treatment planning is an important step in the surgery of middle ear otosclerosis. Fig 5(a) shows the normal middle ear ossicles. In Fig 5(b), the incus was still at its original geometry shape while partial damage occurs on the malleus and total damage for the stapes. Fig 5(c) is a CT scan of the damaged middle ear ossicles. The images are in DICOM format. From the CT image stack, a 3D reconstruction is performed to form the 3D shape of the damaged middle ear ossicles. This can be used to study the disease with the middle ear of the patient. Quantitative analysis can be performed to identify the missing parts, and measure the volumes, areas and distances for regions of interest. Interactive simulation can also be applied for the interactive implanting to achieve optimized results.

V. CONCLUSION

In this paper, VR enhanced medical simulation is investigated. We first describe the system components of the multi-modality VR environment emphasizing the hardware and software. It is highly important to select the simple, easy and native solution for interactive tasks. Medical simulation is a booming area. As an enabling technology, VR has made significant contribution to medical simulation. In this paper, we discuss three applications of VR: virtual anatomy learning, ophthalmologic surgery training, and middle ear pre-treatment planning.

ACKNOWLEDGEMENT

The authors would like to thank Ms Wang Shengzhi, Mr Mike Chia Chong Yong, Ms. Wong Kit Mui, and Dr Chui Chee Kong for their helps to this research.

REFERENCES

[1] K.F. Kaltenborn and O. Rienhoff. Virtual reality in medicine, *Methods Inform. Med.* 32, pp.407–417, 1993.

- [2] F.P. Vidal I, F. Bello, K.W. Brodlie, N.W. John, D. Gould, R. Phillips and N.J. Avis. Principles and Applications of Computer Graphics in Medicine, *Computer Graphics forum*, Vol. 25, no. 1, pp.113–137, 2006
- [3] R. F. McCloy and R. Stone. Virtual reality in surgery, *BMJ* 323, pp.912–915, 2001.
- [4] K. Tsuchiya, S. Katase, J. Hachiya and Shiokawa. Volume-Rendered 3D Display Of MR Angiograms in the Diagnosis of Cerebral Arteriovenous Malformations, *Acta Radiologica* 44, 675, 2003.
- [5] Zorana Hadzimahmutovic, David L. Healy, Jim Tsaltas, Anthony S. Lawrence, T. Ian H. Brown and Brendan Flanagan. Gynaecological endoscopy training simulators, *Gynaecological Endoscopy*, Vol. 8, Issue 3, pp. 129-135, June 1999.
- [6] C. Basdogan, M. Sedef, M. Harders and S. Wesarg. VR-Based Simulators for Training in Minimally Invasive Surgery Computer Graphics and Applications, *IEEE Volume 27*, Issue 2, pp. 54 – 66, March-April 2007.
- [7] S.R. Ellis. What are virtual environments?, *IEEE Comp.Graph*, Appl, vol.14, no.1, pp.17–22, 1994.
- [8] Christian Krapichler, Michael Haubner, Rolf Engelbrecht and Karl- Hans Englmeier. VR interaction techniques for medical imaging applications, *Computer Methods and Programs in Biomedicine* 56, 65–74, 1998.
- [9] Philip J. Mercurio, Thomas D. Erickson. Interactive Scientific Visualization: An Assessment Of A Virtual Reality System, Published in *Human-Computer Interaction*, Proceedings of Interact-90, August, 1990.
- [10] M.W. Krueger. Automating virtual reality, *IEEE Comp.Graph. Appl.* 15 (1), 9–11, 1995.
- [11] Emilio Schapira, Rajeev Sharma. Experimental Evaluation of Vision and Speech based Multimodal Interfaces, Proceedings of the 2001 workshop on Perceptive user interfaces, *ACM International Conference Proceeding Series*, Vol. 15, pp 1 – 9, 2001
- [12] C. Ware. Using hand position for virtual object placement, *The Visual Computer: Int. J. Comput. Graph.* 6, 5, pp. 245–253, 1990.
- [13] Eric Woods, Paul Mason, Mark Billinghurst. MagicMouse: an inexpensive 6-degree-of-freedom mouse, Proceedings of the 1st international conference on Computer graphics and interactive techniques in Australasia and South East Asia, 2003
- [14] A. Kaufman, F. Dacheille, B. Chen, I. Bitter, K. Kreeger, N. Zhang and Q. Tang. Real-Time Volume Rendering, *International Journal of Imaging Systems and Technology* Vol. 11, 44–52, 2000.
- [15] TeraRecon, Inc., VolumePro 1000 Principles of Operation, Product Notes, <http://www.terarecon.com/>, 2001
- [16] Head Mounted Display, Kaiser Electro-Optics Inc., A Rockwell Collins Company, Carlsbad, USA.
- [17] Mirage 4000 Brochure, Christie Digital Systems, Inc., Cyrus, CA. <http://www.christiedigital.com/>, 2007
- [18] Polhemus, Patriot Fast and Affordable Digital Tracker Brochure, Polhemus, <http://www.polhemus.com/>, 2004.
- [19] Ascension Technology - Flock of Birds, <http://www.ascension-tech.com/products/flockofbirds/>
- [20] CyberGlove, Virtual Technologies, www.immersion.com/3d/products/, 2001.
- [21] Lehmann, T. M., Gönner, C. & Spitzer, K. Survey: interpolation methods in medical image processing. *IEEE Trans. Med. Imaging* 18(11), 1049–107 5, 1999.
- [22] Microsoft Speech SDK 5.1 Documentation, 2004
- [23] YQ Guan, Y.Y Cai, Y.T. Lee and M Opas (2006), An Automatic Method for Identifying Appropriate Gradient Magnitude for 3D Boundary Detection of Confocal Image Stacks, *Journal of Microscopy*, Blackwell Publishing, Vol. 223, Pt 1 July 2006, pp. 66–72 .
- [24] YQ Guan, YY Cai, XC Zhang, YT Lee and M Opas (2007), Adaptive Correction Technique for 3D Reconstruction of Fluorescence Microscopy Images, *Microscopy Research and Technique*, Wiley, Vol. 71, No. 2, pp. 146-157.
- [25] YY Cai, XZ Ye, CK Chui and J Anderson (2003), Constructive algorithms for GC1 generation of vascular network, *Advances in Engineering Software*, Elsevier, Vol. 34/7, pp. 439 - 450.
- [26] JM Zheng and YY Cai (2006), Interpolation Over Arbitrary Topology Meshes Using a Two-Phase Subdivision Scheme, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 12, No. 3, pp. 301-310.
- [27] YY Cai, CK Chui, YP Wang, ZL Wang, JH Anderson (2001), Parametric eyeball model for interactive simulation of ophthalmologic surgery, *Lecture Notes in Computer Science: Medical Imaging Computing and Computer-assisted Intervention*, Vol. 2208, Springer-Verlag, pp. 465-472.



C Indhumathi is a Ph D student with the School of Mechanical & Aerospace Engineering, Nanyang Technological University, Singapore. Her research interests include Virtual Reality, Imaging and Visualization for Bio-medical Science applications.



Wenyu Chen is a Ph D student with the School of Mechanical & Aerospace Engineering, Nanyang Technological University, Singapore. His research interests include Geometric Modeling, Surface Modeling.



Dr Yiyu CAI is an associate professor with the School of Mechanical & Aerospace Engineering, Nanyang Technological University, Singapore. His research interests include VR, Computational Biomedical Sciences, and Interactive & Digital Media.