

**Bangor University**

**DOCTOR OF PHILOSOPHY**

**A tangible augmented reality anatomy teaching tool**

Thomas, Rhys Gethin

*Award date:*  
2009

*Awarding institution:*  
Bangor University

[Link to publication](#)

#### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 28. Apr. 2024

# A Tangible Augmented Reality Anatomy Teaching Tool

Rhys Gethin Thomas

A Thesis presented for the degree of  
Doctor of Philosophy



Visualization and Medical Graphics Group  
School of Computer Science  
Bangor University  
UK

July 2009



# **A Tangible Augmented Reality Anatomy Teaching Tool**

**Rhys Gethin Thomas**

Submitted for the degree of Doctor of Philosophy

July 2009

## **Abstract**

Augmented Reality (AR) is an emerging technology that allows computer graphics to be superimposed onto real world scenes at interactive frame rates. AR is still in its infancy and the full range of applications that can benefit from this approach is still to be determined. The research described in this thesis has studied AR in the context of the medical applications domain and endeavoured to develop a novel and effective advance in the tools available for anatomy education using off the shelf hardware components.

The use of cadaver dissections as a means of teaching gross anatomy has decreased markedly over recent decades. Several widely used alternatives exist; however dissections are still regarded as being the Gold Standard in anatomy education. The use of Virtual Environments using three dimensional computer graphics models has been reported in the last decade as a method of teaching anatomy. Generally such environments only convey the shape of the anatomy to the student. AR environments have also been investigated for medical education, including several training tools for surgical procedures. Few AR environments are reported that aim to teach gross anatomy. Presented in this thesis is the development of the Bangor Augmented Reality Education Tool for Anatomy (BARETA), a system that combines AR technology, displaying volume and surface renderings of medical datasets, with anatomically correct models produced using Rapid Prototyping (RP) technol-

ogy, to provide the student with stimulation for touch as well as sight. Both the RP model and the user's viewpoint were tracked using separate tracking devices. The use of RP models in this way for anatomical education is unique to this research project.

The principal aims of this work were to provide a more intuitive interface than a mouse and keyboard, and to evaluate such a system as a viable supplement to traditional cadaver based education. A preliminary evaluation was carried out through user studies in which groups of potential end users operated BARETA then filled out questionnaires that asked both about how useful they found the BARETA to be to their education, and how easy BARETA was to use. The versions of BARETA evaluated by students presented them with a rich environment in which to explore and learn the anatomy of the human head.

The results from this work demonstrate that we have developed a compelling tool for anatomy education that augurs well for future innovation in medical AR.



# Acknowledgements

Firstly I would like to acknowledge Professor Nigel John, for without his help and support this thesis would not have been possible.

I would also like to acknowledge Doctor John Delieu, who contributed significantly to the medical aspects of this project, and allowed a user study to be carried out during one of his teaching sessions. Thanks must also be extended to Abigail Malia at Connah's Quay high school and to Doctor Mike Mahon of Keele's School of Medicine for providing students and facilities to perform our user studies.

My fellow PhD students have also been of great help during the last four years, especially James, David and Catrin. Although not named individually here I would also like to acknowledge the support provided by the staff at the Schools of Electronic Engineering and Computer Science (formerly the School of Informatics) for their support and encouragement.

My family have also been supportive throughout the last four years, especially my Dad, and my Mum and Wil. Finally I would like to acknowledge Adam, Greg, Leon and Trig who have provided me with some welcome distractions over the last few years.

This research has been possible thanks to funding from an EPSRC Doctoral Training Account grant.

# Contents

<b>Abstract</b>	<b>ii</b>
<b>Declaration</b>	<b>iv</b>
<b>Acknowledgements</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	2
1.2 Hypothesis . . . . .	5
1.3 Investigation . . . . .	5
1.4 Contribution . . . . .	7
1.5 Publications . . . . .	8
1.5.1 Peer Reviewed Conference Papers . . . . .	8
1.5.2 Peer Reviewed Poster Presentations . . . . .	8
1.5.3 Peer Reviewed Workshop Presentations . . . . .	9
1.5.4 Peer Reviewed Journal Publications . . . . .	9
1.6 Thesis Structure . . . . .	9
<b>2 Literature Review</b>	<b>10</b>
2.1 Introduction . . . . .	11
2.2 Traditional Medical Visualization Techniques . . . . .	11
2.2.1 Surface Rendering . . . . .	13
2.2.2 Segmentation . . . . .	15

2.2.3	Direct Volume Rendering . . . . .	16
2.3	Medical Visualisation in Practice . . . . .	20
2.4	Mixed and Augmented Reality . . . . .	21
2.5	Medical Education Applications . . . . .	24
2.6	Summary . . . . .	33
<b>3</b>	<b>The Components of an Augmented Reality System</b>	<b>34</b>
3.1	Introduction . . . . .	35
3.2	Tracking . . . . .	35
3.2.1	Hardware . . . . .	36
3.2.2	Software . . . . .	37
3.3	Displays . . . . .	37
3.3.1	Hardware . . . . .	38
3.3.2	Software . . . . .	40
3.4	Components of BARETA . . . . .	41
3.4.1	Requirements . . . . .	41
3.4.2	Tracking . . . . .	41
3.4.3	Display . . . . .	44
3.5	System Calibration . . . . .	47
3.5.1	Tracking System Calibration . . . . .	47
3.5.2	Webcam and VisTracker Calibration . . . . .	50
3.6	Graphical Content . . . . .	54
3.7	Summary . . . . .	55
<b>4</b>	<b>Tangible Interfaces</b>	<b>57</b>
4.1	Introduction . . . . .	58
4.2	Tangible Interfaces Overview . . . . .	58
4.3	Generic Anatomy Mannequin Interface . . . . .	59
4.4	Rapid Prototyping Overview . . . . .	61

4.5	Data Extraction . . . . .	64
4.6	Liver Model . . . . .	64
4.7	Ventricles Model . . . . .	66
4.8	Further Props . . . . .	69
4.9	Summary . . . . .	70
<b>5</b>	<b>The BARETA Anatomy Teaching Assistant</b>	<b>71</b>
5.1	Introduction . . . . .	72
5.2	Current Anatomy Lessons . . . . .	72
5.3	ARToolkit-based Teaching Assistant . . . . .	74
5.4	BARETA Liver Teaching Assistant . . . . .	78
5.5	First BARETA Brain Ventricles Teaching Assistant . . . . .	79
5.6	Second BARETA Brain Ventricles Teaching Assistant . . . . .	84
5.7	Summary . . . . .	86
<b>6</b>	<b>User Evaluation</b>	<b>88</b>
6.1	Introduction . . . . .	89
6.2	First User Evaluation . . . . .	89
6.3	Second User Evaluation . . . . .	97
6.4	Third User Evaluation . . . . .	104
6.4.1	Combined Results . . . . .	108
6.4.2	Stand-Alone Results . . . . .	115
6.5	Summary . . . . .	116
<b>7</b>	<b>Conclusions and Future Work</b>	<b>118</b>
7.1	Introduction . . . . .	119
7.2	BARETA . . . . .	119
7.3	Discussion . . . . .	121
7.4	Contribution . . . . .	122
7.5	Future Work . . . . .	123

---

7.6 Conclusion . . . . .	126
--------------------------	-----



# List of Figures

1.1	The Dissection Room at Keele University . . . . .	3
1.2	BARETA in use. . . . .	6
1.3	A typical experimental layout used during the development of BARETA. . . . .	6
2.1	An MPR view of an MRI scan of a human head. . . . .	12
2.2	An example of an iso-surface . . . . .	14
2.3	An image slice of a human brain showing the segmentation of the lateral ventricles. . . . .	16
2.4	CT densities assignments [17]. . . . .	17
2.5	Volume Renderings . . . . .	18
2.6	Milgram's Virtuality Continuum [43]. . . . .	22
2.7	A comparison of a view from a webcam with and without augmentations. . . . .	22
2.8	The Anatomy Lesson of Doctor Nicolaes Tulp . . . . .	25
2.9	A simple plastic model used to teach anatomy to school children. . . . .	27
2.10	The virtusMed project showing an ultrasound simulation of a mannequin [73]. . . . .	28
2.11	The system of Juan <i>et al.</i> [31], showing the intestines of the human body through a slit in the torso. . . . .	29
2.12	A physical representation of a human liver derived from volume data. . . . .	32
3.1	The Ascension miniBIRD 800 . . . . .	42

3.2	The Logitech QuickCam Pro 9000 USB webcam . . . . .	44
3.3	Display devices . . . . .	46
3.4	A comparison of co-ordinate systems . . . . .	48
3.5	The fiducial constellation used by the VisTracker . . . . .	49
3.6	A calibration bracket . . . . .	50
3.7	The application bracket . . . . .	51
3.8	The circular fiducial detection procedure . . . . .	53
3.9	Detecting the centre of fiducial patterns . . . . .	54
3.10	A software architecture diagram showing the relationship between the components of a complete BARETA system. . . . .	55
4.1	The generic anatomy mannequin. . . . .	60
4.2	The Z-Corp ZPrinter 450 at Inition used to produce our models. . . .	63
4.3	Representations of the human liver . . . . .	66
4.4	Drawing of a cast of the ventricular cavities, viewed from above . . .	67
4.5	Using the ITKSnap software to segment the ventricles . . . . .	68
4.6	The ventricles RP model . . . . .	69
4.7	The clipping plane prop . . . . .	70
5.1	Lecture slides used to teach brain anatomy at Bangor's School of Healthcare Sciences . . . . .	73
5.2	A lung of the anatomy mannequin augmented with abdomen CT data.	76
5.3	The known and required transformations to create a clipping plane. .	77
5.4	The ventricles surface model within the volume rendering of the head are exposed using transparency. . . . .	81
5.5	A volume rendering of an MRI scan of the human head with a transfer function applied to produce pseudo colour. . . . .	82
5.6	Steps one and two of the brain ventricles lesson . . . . .	84

5.7	The second ventricles teaching assistant running, showing the location of the third ventricle. . . . .	86
6.1	The experimental setup used in Wrexham (not to scale) . . . . .	90
6.2	A selection of slides used at the Bangor University School of Healthcare Sciences, Wrexham . . . . .	91
6.3	Results from the Wrexham user evaluation . . . . .	93
6.4	Wrexham user evaluation standard deviation. . . . .	94
6.5	Comparing the opinions of two different age groups . . . . .	95
6.6	Comparing the opinions of the two genders. . . . .	96
6.7	The experimental setup used at Connah's Quay High School (not to scale) . . . . .	97
6.8	The experimental setup used at Connah's Quay High School, showing the tracking patterns, the user and the camera. In this instance the user is facing the display device. . . . .	98
6.9	A student using BARETA at Connah's Quay High School . . . . .	99
6.10	Connah's Quay High School results chart . . . . .	100
6.11	Connah's Quay High School male results chart . . . . .	103
6.12	Connah's Quay High School female results chart . . . . .	103
6.13	Connah's Quay High School gender comparison. . . . .	104
6.14	Connah's Quay questionnaire results Standard Deviation . . . . .	105
6.15	A student at Keele University operates the augmented reality system. . . . .	106
6.16	The experimental setup used at Keele University (not to scale) . . . . .	106
6.17	A chart displaying the results from the Keele user evaluation . . . . .	109
6.18	A chart displaying the combined results from the Keele user evaluation . . . . .	110
6.19	A chart displaying the standard deviation for each question asked at Keele University. . . . .	112
6.20	A chart showing the mean responses for each group that took part in the Keele user evaluation. . . . .	113

---

6.21 A chart comparing the difference between groups at Keele . . . . .	113
6.22 A chart comparing the responses of male and female students at Keele	114
6.23 A chart showing the results of stand-alone questions posed at Keele University . . . . .	115
7.1 BARETA running, with the transparency feature enabled. . . . .	121

# List of Tables

3.1	A table comparing tracking hardware . . . . .	36
3.2	A table comparing display output devices . . . . .	46



# Chapter 1

## Introduction

## 1.1 Motivation

Traditional computerised education software uses the window, icon, menu, pointing device (WIMP) interaction style. Although this works well with 2D work spaces, it can prove to be counter-intuitive when working with 3D work spaces as 2D operations have to be used to manipulate 3D objects.

Augmented Reality (AR) allows a user to interact with virtual content in 3D space. AR is a development of Virtual Reality (VR) that allows the user to see both 3D computer generated content and the real world concurrently, and to interact with both elements as though they were both situated in the real world. The use of AR is compelling as it can allow for effective collaboration among students if the environment is shared, as users can see both the virtual objects that they are manipulating and each other without obstruction. AR can also provide the user with effective positional cues as the surrounding real environment is constantly visible. AR provides the user with an interface that requires little learning, as every physical interaction with a real object results in a predictable result since the virtual representation must follow the real object to which it is attached; this allows the user more time to learn the subject matter of the lesson.

Research into AR has been ongoing for more than two decades yet many challenges remain unsolved, including reliable occlusion detection and resolution, the creation of consistent depth cues, and portability of all necessary hardware. However, even with the limitations with the current state of the art, AR presents many exciting possibilities for educational applications as almost any computer graphical content can be integrated with a view of the real world, allowing a vast array of items to be studied without them having to be physically present. Several educational applications lend themselves to having value added through the use of AR, as changes within the field are necessitating changes in the way that education is carried out. Anatomy is one field in which change has been occurring.

Anatomy education has changed dramatically over the last half century, as tra-

ditionally anatomy has been taught through the dissection of cadavers. However, this practice is not as widespread as it once was. This reduction is due to a number of reasons, including financial considerations and ethical issues. In schools small animals such as rats and frogs were often used to teach simple aspects of anatomy, but this too has reduced over recent years because of an increased recognition of animal rights issues. This has led to anatomy being taught in a variety of different ways, however many anatomists believe that cadaver dissection is still the optimal method of anatomy education. Cadaver dissection not only gives the learner a knowledge of the shape and size of the organs, it gives them an appreciation of how they are positioned relative to the rest of the body. Many alternative methods exist for teaching anatomy, however, although such methods do present learners with excellent opportunities for learning about individual organs it is felt that the lack of presentation of spatial relationships can be a disadvantage relative to cadaver dissection.



Figure 1.1: The Dissection Room at Keele University

Medical diagnosis is often carried out using images from medical scanning devices. In the past these images have been viewed in turn by the practitioner who had to interpret information from a number of slices. More recently, scanning work-



stations have become available that can create 3D reconstructions of the data, using volume rendering and other techniques, which can allow for a much more accurate diagnosis. Such techniques have also been exploited for educational purposes.

While 3D approximations and reconstructions of medical images have been used for anatomy education, most systems have used the WIMP interaction style. Using an AR environment can allow a much more natural interaction with the data being represented, especially if some form of tactile interface can be used. Instead of being being restricted by a user interface, as is the case with WIMP-based systems, an AR allows us to view 3D learning material “as 3D entities in our own 3D spaces” [64], and as a result can aid a student’s learning. In schools, small animals such as rats and frogs were often used to teach simple aspects of anatomy. This practice too has decreased over recent years because of an increased recognition of animal rights issues. This has led to anatomy being taught in a variety of different ways, including Prosections, Problem-Based learning Scenarios (PBLs) or, more recently, computer systems derived from the Visible Human Project [74], such as the VOXEL-MAN project [76]. Computer-based systems can also allow the user to perform operations that are not possible on physical specimens, such as undissection [54]. AR systems are now becoming cost effective enough to be deployed in this setting.

Presenting a user with a physical object that they can manipulate and that causes corresponding movements of the image on-screen can help a user to co-ordinate actions in an AR environment, especially if that object has the same shape as the object being viewed on-screen. An exact physical replica can also help the user to find and understand intricate and hidden detail that is not necessarily evident from the visualisation alone [32].

Rapid prototyping provides a method of rapidly creating patient specific models of organs of interest. Models can be generated from a variety of data types, including medical scans. Such a model could on its own be used to provide useful learning material, however if tracked it could be used in conjunction with an AR system

which uses surface and volume rendering techniques, and could be a useful teaching tool.

## 1.2 Hypothesis

Given the above motivation, the hypothesis that is investigated in this thesis is:

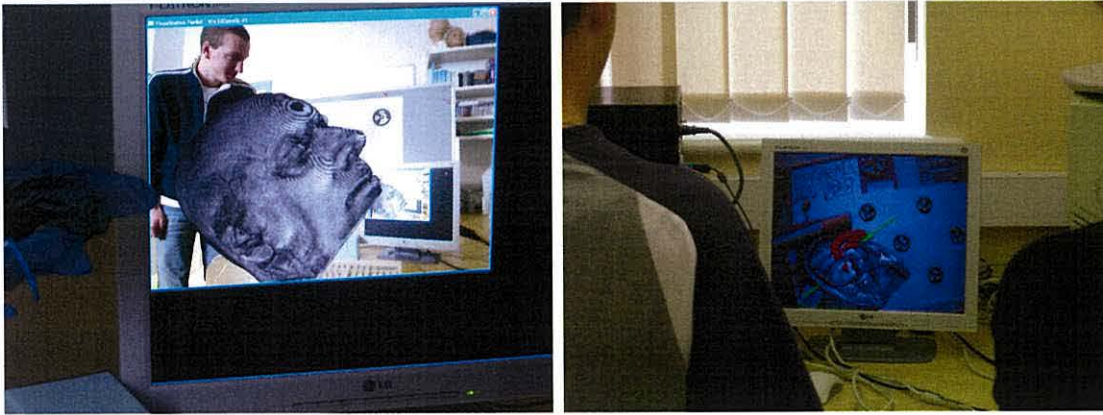
“Is it possible to construct a compelling Augmented Reality system that can offer an alternative learning aid for teaching gross anatomy?”

We believe that by using volume rendering techniques on scan data within an AR environment that we can produce a compelling learning tool for gross anatomy. We also believe that such a tool would be flexible, cost-effective, and an ideal complement to traditional cadaveric education, and more effective than current alternatives to cadaver dissection.

## 1.3 Investigation

To investigate the hypothesis the Bangor Augmented Reality Education Tool for Anatomy (BARETA) was created that used a standard PC in combination with proprietary tracking hardware and a model produced using Rapid Prototyping (RP) technology. The RP model was used to control the pose of the on-screen volume and surface rendering of medical data, tracked by an Ascension miniBIRD magnetic tracking system. The real part of the BARETA environment was provided by a standard USB webcam; attached to the webcam was an InterSense IS-1200 VisTracker, a hybrid optical inertial tracking device that uses fiducial patterns to determine its pose. This allowed the viewpoint of the user to be mobile. In the final version of BARETA described in this thesis renderings of the human head are shown (Figure 1.2(a)). In this case an RP model of the human brain ventricles was produced that was tracked using the miniBIRD. However, in practice, data of any region of anatomy can be used. BARETA allows the data to be interrogated





(a) A user of BARETA viewing a volume rendering of a human head, controlled using an RP feature. (b) Two students using BARETA's clipping model.

Figure 1.2: BARETA in use.

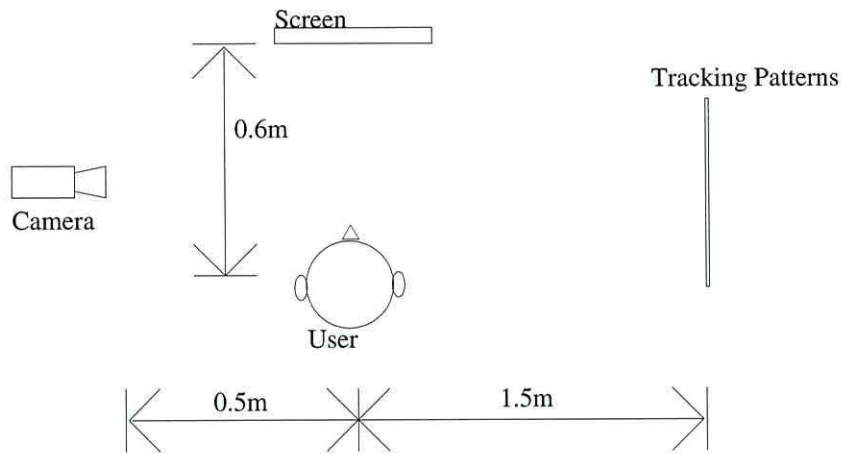


Figure 1.3: A typical experimental layout used during the development of BARETA.

using transparency, clipping (Figure 1.2(b)) and slab rendering, and also highlights regions of interest using arrows, with on screen text providing information about the region pointed to by the arrow.

In the current implementation of BARETA the user views the AR scene on a standard computer monitor. The webcam, although able to move, generally stays in a single location, pointing towards the fiducial patterns used by the VisTracker (Figure 1.3). The user must be careful to sit in a position such that the VisTracker's view of the fiducials is not completely obscured (a small amount of occlusion is acceptable).

To evaluate BARETA we conducted three separate user evaluations at three sites.

The sites chosen were Bangor University's School of Healthcare Sciences, Connah's Quay High School and the School of Medicine at Keele University. These three sites offered the opportunity to see how easy to use and useful the target audience of BARETA found it to be.

## 1.4 Contribution

Several contributions to the state of the art of computerised anatomy education are made in this thesis.

- A novel augmented reality anatomy teaching tool, the Bangor Augmented Reality Education Tool for Anatomy (BARETA), was developed that can be run on a standard desktop PC. BARETA allowed users to interactively control the position and orientation of computer generated anatomy models. A sample lesson based on the human head was developed that focused on the ventricular system within the brain. The head was represented by a volume rendering of MRI data, supplemented by a surface rendering of data segmented from the same MRI data. BARETA is useful for showing spatial relationships between neighbouring objects, especially when combined with features such as transparency, and an interactive clipping plane and slab rendering.
- An anatomically correct plastic model was generated from medical volume data with which a user could interact with BARETA in a natural fashion, avoiding the need for an extended familiarisation period with an interaction style. To generate the model a semi-automatic segmentation was performed on MRI data, from which a surface data set was produced. From this surface data a physical equivalent was fabricated using a 3D printer.
- We have demonstrated through a series of validation studies that students of anatomy are in general willing to use a variety of methods to learn anatomy,

and that augmented environments such as BARETA are a useful supplement to traditional cadaveric education methods. Students agreed that use of BARETA gave them a better understanding of the ventricular system than all other education techniques other than cadaver dissection. Most students also found the system easy to use, and would use it again to learn other anatomical features.

## 1.5 Publications

### 1.5.1 Peer Reviewed Conference Papers

- **Rhys G. Thomas**, Nigel W. John, Ik Soo Lim “A Mixed Reality Anatomy Teaching Tool”, Proceedings of Theory and Practice of Computer Graphics 2006, Middlesbrough, June 2006, pp165–170  
**Winner of the Rob Fletcher Prize for the best student applications paper**
- **Rhys G. Thomas**, Nigel W. John, Ik Soo Lim “Anatomy Education using Rapid Prototyping”, Proceedings of Theory and Practice of Computer Graphics 2007, Bangor, June 2007, pp251–257

### 1.5.2 Peer Reviewed Poster Presentations

- **R G. Thomas**, J M. Delieu, N W. John, M. Mahon “Using Rapid Prototyping to Complement Traditional Anatomical Education”, Summer Meeting of the Anatomical Society of Great Britain and Ireland, Durham University, July 2007
- **Rhys Gethin Thomas**, Nigel W. John “Mixed Reality in Anatomical Education”, Second Peach<sup>1</sup> Summer School, Dubrovnik, July 2008

---

<sup>1</sup>PEACH, and acronym of Presence Research in Action, is a three year FP6 Coordination Action



- Delieu, John, Nigel John, Mike Mahon, Paul Mullins, Hayley Derricott, Ik Soo Lim, **Rhys Thomas** “Magnetic resonance imaging of an embalmed human head”, BACA Summer Meeting, Liverpool, July 2008

### 1.5.3 Peer Reviewed Workshop Presentations

- **Rhys Thomas** and Nigel John “Anatomical Education Using Mixed Reality”, Visual Computing in Wales Workshop 2009, Aberystwyth, April 2009

### 1.5.4 Peer Reviewed Journal Publications

- **Rhys Gethin Thomas** and Nigel William John “Augmented Reality for Anatomical Education”, Journal of Visual Communication in Medicine, To Appear 2010

## 1.6 Thesis Structure

Chapter two of this thesis presents an overview of previous work in the fields of anatomy education, medical imaging and augmented reality. The bulk of this thesis describes the various elements of BARETA including the components used to make the BARETA framework (Chapter 3), the creation of RP models from medical volume data (Chapter 4) and the integration of these components to create a novel teaching assistant (Chapter 5). Chapter 6 describes the evaluation studies that assessed the usefulness of BARETA and the results that were obtained. Finally, Chapter 7 presents some conclusions and ideas for the future development of BARETA.

## Chapter 2

### Literature Review

---



## 2.1 Introduction

This chapter begins with a discussion visualisation techniques that can be used for medical purposes. A brief summary of how these can be used is provided in Section 2.3. The field of Augmented Reality (AR) is then discussed in Section 2.4, then in Section 2.5 general AR medical education tools are discussed, along with current methods of teaching in medicine, focusing on the teaching gross anatomy. A summary of the findings of this chapter is presented in Section 2.6.

## 2.2 Traditional Medical Visualization Techniques

Medical data can be acquired from a variety of scanning devices. For use with computer graphics the most useful scan types are MRI, CT, ultrasound, SPECT and PET. Generally the data is created as a series of slices. These slices are normally stored in the standard DICOM format which can then be readily converted to alternative formats and displayed using computer software (ImageJ [27] is an example of software that can achieve this, there are many others, a list of software can be found at [59]).

Traditionally each slice of data is viewed on a light box, or more recently using the Picture Archiving and Communications System (PACS), in order to gain information. With both of these methods the user has to create a 3D mental image of the content of a number of slices to interpret the data. Another popular viewing technique is the multiplanar reconstruction (MPR) [19]. In MPR a number of dynamic orthogonal slices can be viewed simultaneously (Figure 2.1), allowing features to be tracked through a data set. This too requires the user to create a 3D mental image of the content of a number of slices.

A series of data slices can be combined to create a three-dimensional representation of the section of anatomy that was scanned [17]. In general, 3D scientific data will either be regular, rectilinear, curvilinear, block structured, unstructured, or a

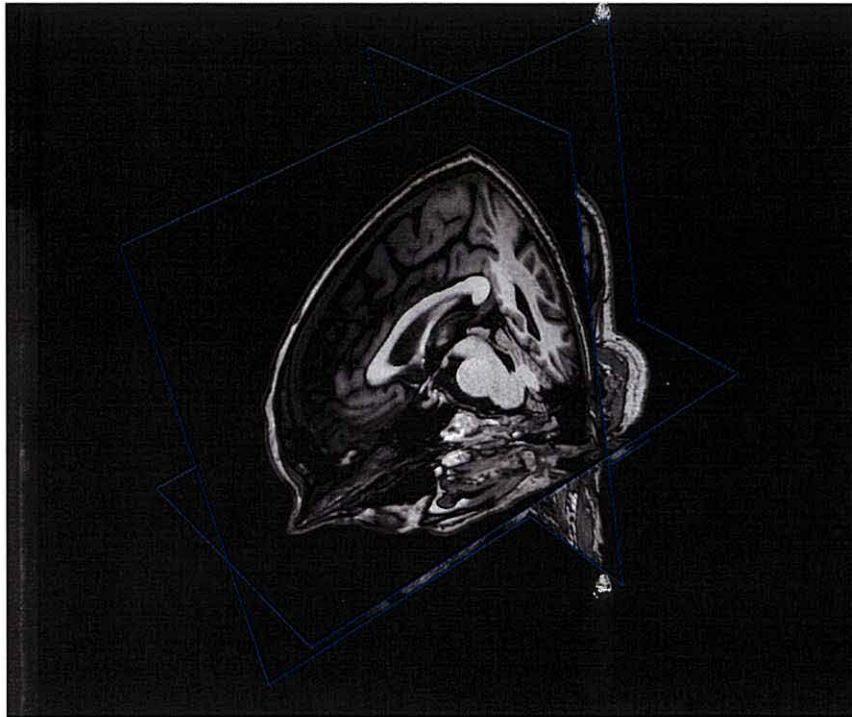


Figure 2.1: An MPR view of an MRI scan of a human head.

hybrid representation. In medical systems the data will most commonly be rectilinear as the resolution in the  $xy$  plane often exceeds that along the  $z$  axis (the interslice distance), however, regular grids are attainable with the latest high resolution medical scanners.

The elements of volume data are referred to as voxels (volume elements, the three-dimensional equivalent to pixels, derived from picture elements in two-dimensional images), which represent the average value over the cell grid. Although there is a single value per voxel, it is recognised that a volume feature may only occupy a part of the voxel; this is called the partial volume effect.

Common data structures used for voxel models are:

- binary: voxel values are either one (object) or zero (no object)
- grey level: each voxel holds an intensity value
- generalised: in addition to an intensity value, each voxel contains attributes describing its membership to various objects, and/or data from other sources.

- intelligent volumes: an extension of the generalised voxel model, properties of objects (such as colour) and their relationships are modelled on a symbolic level. Useful for creating medical atlases.

For this data to be useful it must be converted from a three-dimensional representation into a two-dimensional representation that can then be displayed on a two-dimensional medium, generally a computer screen (stereoscopic representations also require this conversion as two two-dimensional images are produced, one for each eye). Several methods for converting from three to two dimensions exist, several of which are introduced in Sections 2.2.1 and 2.2.3. The method used will depend upon the requirements of the specific application.

The data that is obtained may contain features that we do not wish to visualise, for example a medical data set will likely contain air surrounding the subject. Different rendering methods deal with this in different ways. Two of the most common methods are Surface Rendering (described in Section 2.2.1), and Direct Volume Rendering (described in Section 2.2.3).

### 2.2.1 Surface Rendering

Surface-based rendering methods do not directly render medical data. Instead an intermediate representation is extracted from the volume data set. Usually this intermediate representation consists of a collection of polygonal facets, such as triangles.

A common method of extracting regions of interest from volume data is to create an iso-surface. An iso-surface is a surface that connects voxels of a specific value. The marching cubes algorithm (first described by Lorensen and Cline [38]) is an example of a method that creates an iso-surface. In the marching cubes algorithm each voxel is processed in turn. Each of the eight corners of the voxel are assessed to see if its value is above or below the desired iso-value. Each corner contributes a bit to an eight bit integer. This is then used as an address into a look up table





Figure 2.2: An isosurface of blood vessels in the brain where an aneurysm has occurred. The data was derived from a 256x256x256 Rotational C-arm x-ray scan of the arteries of the right half of a human head. A contrast agent was injected into the blood and an aneurysm is present. Courtesy of Philips Research, Hamburg, Germany.

that stores a description of where the surface should pass through the cube. A total of 256 different polygon configurations exist, however, this can be reduced to 15 configurations by taking advantage of rotation and symmetry. The values of the vertices are then assessed for a second time to interpolate where the surface meets the edge between the vertices. A set of polygons is output that can then be saved as a surface data set, or rendered on-screen using a ray-casting method.

Iso-surfacing is not without problems. Often an iso-surface will contain redundant data. Also an iso-surface will often contain some unwanted features, and will miss some fine detail from the structures that are required (note in Figure 2.2 that the thin arteries appear to be discontinuous).

The simplest method of visualising a surface model is to use a ray-casting technique that shoots rays at the surface model and computes pixel values for each ray. Several variations on this technique exist. Most implementations use a basic local illumination model where shadows and indirect illumination are not computed, which leads to fast rendering rates. Global illumination methods that do take shadows and

indirect illumination into account also exist, however, rendering performance tends to be much slower; methods are in development that aim to reduce some of this deficit [4].

Other methods of surface rendering exist. One such method is nonphotorealistic rendering (NPR) that can use stylisation to enhance certain features of a surface model to give the user information about properties such as curvature [36].

### 2.2.2 Segmentation

It is often useful to partition a data set into its constituent components, a practice that is called segmentation. Several methods of segmentation exist. The method used will depend both on what the user requires as an output, and the data input. Segmented regions are often shown as coloured regions on the original image (Figure 2.3). Segmentation can be achieved

- manually
- using a threshold
- using an edge-based method
- using a region-based method
- using deformable methods
- using statistical analysis.

Automatic segmentation is a desirable method of segmentation, and finding such a method is an area of active research. Further discussion of segmentation is beyond the scope of this thesis. The interested reader may, however, find the survey of Ashton *et al.* [1] useful.



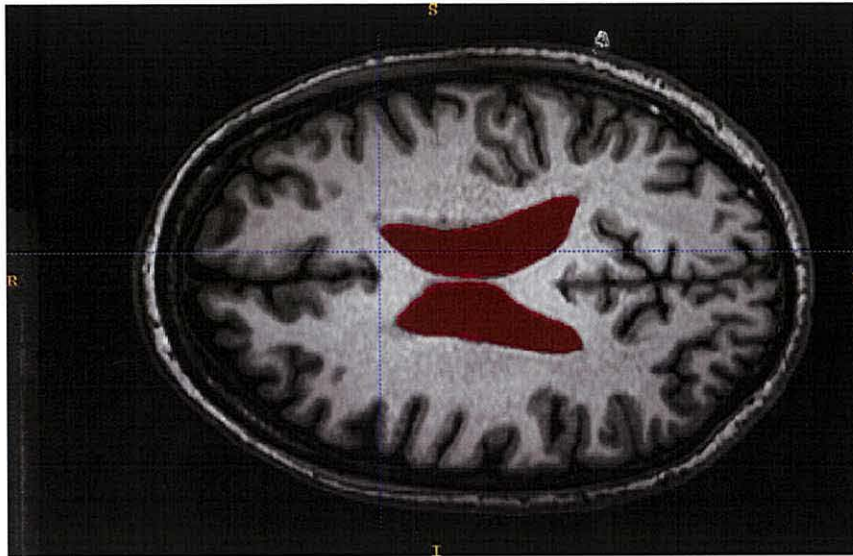


Figure 2.3: An image slice of a human brain showing the segmentation of the lateral ventricles.

### 2.2.3 Direct Volume Rendering

Volume Rendering is the process of creating a two-dimensional representation from a three-dimensional volume data set. In contrast to surface rendering, it does not require an intermediate representation to be derived, therefore retaining all of the information present in the data. This makes it an ideal technique for interactive data exploration. Threshold values and other parameters that are not clear from the beginning can be changed interactively. Several different methods of direct volume rendering exist (some of which are introduced below). The choice of which method to use is often the result of a compromise between the speed of rendering and the rendering quality.

Regions of interest from a volume data set can also be highlighted during the direct volume rendering method using transfer functions that can alter both the colour and opacity of a greyscale volume data set. Transfer functions can depend on voxel value, voxel gradient, curvature or other parameters. As different tissues and structures have different densities (appearing as different values in greyscale images), they can to a certain extent be separated by using transfer functions (Figure 2.5). In CT data sets bone is the densest structure (and has the highest greyscale value),

followed by soft tissue, fat and finally air (which has the lowest greyscale value) [17] (see Figure 2.4). This method of defining transfer functions works well if adjacent structures vary greatly in image intensity, however this is often not the case.

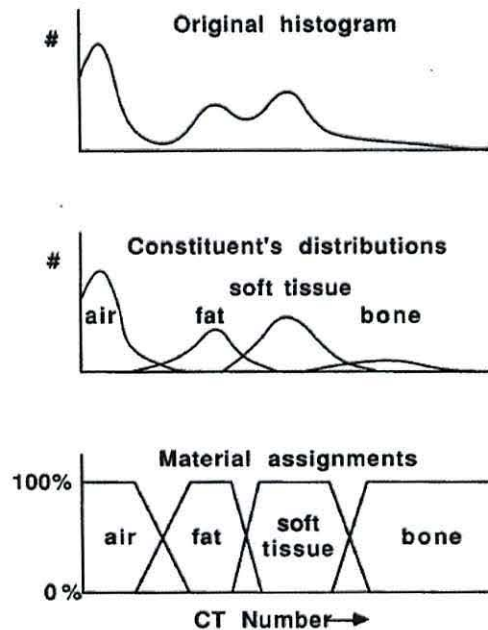
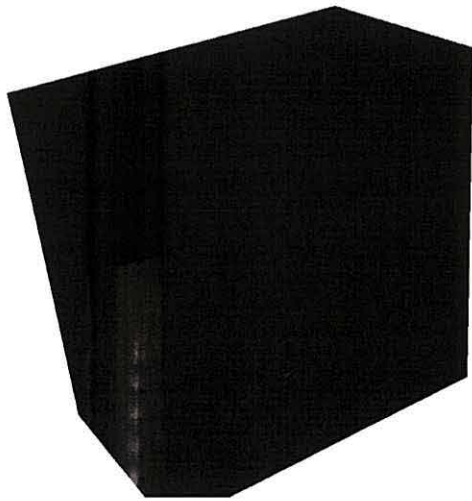


Figure 2.4: CT densities assignments [17].

The volume data is often abstracted as a set of transparent gels that have two properties: colour, itself a set of three values (red, green and blue), and opacity. This model allows for the concept of being able to see all of the data. Without a transfer function changing opacity values, every voxel is rendered with an opacity value of 1 (fully opaque), resulting in only the outer voxels of the data set being visible (see Figure 2.5(a)).

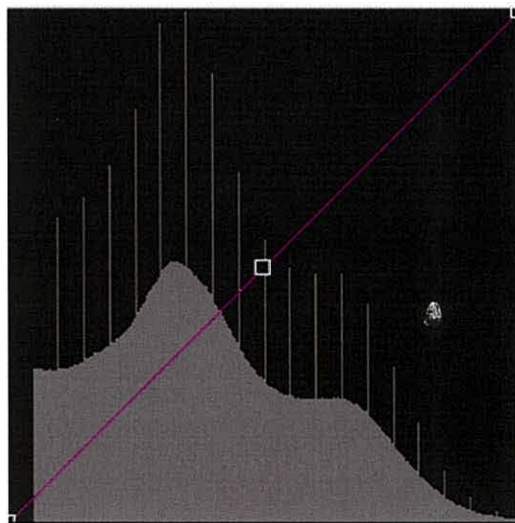
Under certain circumstances it may be advantageous to combine two or more data sets of the same area of interest in a single visualisation. For instance the combination of PET and CT scan data can aid the detection of lesions; the PET scan data can effectively detect lesions and distinguish between tissues, while the CT data helps identify anatomical landmarks and the detection of boundaries between organs and lesions [33].



(a) Data of a human head rendered without a transfer function (all voxels completely opaque.)



(b) Data of a human head rendered with a transfer function.



(c) A linear transfer function shown on a histogram of volume data.

Figure 2.5: Volume Renderings



There are two basic scanning strategies for traversing the volume data:

- Feed Backward Projection or Image Order Traversal. The pixels in the image plane are traversed and imaginary rays are cast through each pixel into the volume. The path of the ray determines the pixel value.
- Feed Forward Projection or Object Order Traversal. The data volume is traversed and each voxel in the volume is projected onto the image plane.

These strategies correspond to the image and object order rasterisation algorithms used in computer graphics.

Several methods for rendering the data exist, including:

- Volume Ray-Casting
- Splatting
- Shear Warp
- Texture Mapping
- Hardware-Based Volume Rendering

Each method has its advantages and disadvantages: generally methods that produce renderings of greater quality do so at the expense of frame rate. The choice of the method to use therefore depends on the relative importance of these two factors to the application. Further discussion of Volume Rendering is beyond the main scope of this thesis. The interested reader should refer to the surveys by Elvins [18] and John [29] for further information.

The Visualization Toolkit (VTK) [62] is a software collection of software libraries that provides many volume rendering options. A large variety of hardware and software platforms are supported due to its use of OpenGL. VTK also supports a large variety of data types, and can open many common data formats. It is therefore ideal for visualising medical data. VTK's pipeline architecture allows for filters to



be placed between the source data and the rendering components, and has its own event processing system. VTK was written in C++, with wrappers available for Java and Tcl.

## 2.3 Medical Visualisation in Practice

The visualisation methods described above are used in a wide variety of medical specialities. Computer generated visualisations can be used to:

- Locate tumours [26], [11]
- Locate aneurysms [39]
- Locate breakages and fractures in bones [52]
- Detect prenatal conditions [35], [42]

Many dedicated medical visualisation systems are currently in use within the medical professions. The systems available are continuously becoming faster and more sophisticated, aiding efficiency of diagnosis [41]. Such systems include:

- Philips Brilliance Workspace
- Siemens Syngo
- General Electric Advantage Workstation VolumeShare 2
- Vital Images systems
- Terarecon Aquarius

Each system can perform a wide variety of tasks and can be specified to suit many medical specialities.

## 2.4 Mixed and Augmented Reality

The field of Mixed Reality (MR) can be thought of as a part of a continuum in which reality lies to one extreme and Virtual Reality (VR) to the other [43] (Figure 2.6). MR itself can be thought of as the central part of this continuum in which the real and the virtual are combined. Within MR are Augmented Reality (AR), in which the majority of the environment is real, and Augmented Virtuality (AV), where most of the environment is virtual. More rigid definitions of AR exist. Azuma [2] defines an AR as an environment that has the following characteristics:

1. Combines the virtual and the real
2. Interactive in real-time
3. Registered in 3D

This definition deliberately steers clear of any discussion regarding display types, of which there are many. In contrast, Milgram and Kishino define AR simply as systems where computer generated images are added to the real environment [43], as in Figure 2.7. He does however offer a discussion on display types, in which they state that see-through Head-Mounted Displays (HMD), where reality is supplemented by projecting images onto half-silvered mirrors (also referred to as optical see-through displays), are the most common method of delivering an AR and suggests that monitor-based “Window on the World” systems are also worthy of the definition AR. Other types of AR displays described by Milgram are fully immersive Head-Mounted Displays in place of a regular monitor, and so-called video see-through systems in which the real world is recorded on a camera mounted on the HMD and delivered to the display.

Registration is perhaps the most challenging of Azuma’s three conditions to meet. It requires that the many components of the system (both hardware and software) are integrated. Misregistration for even a short amount of time can lead to a loss of

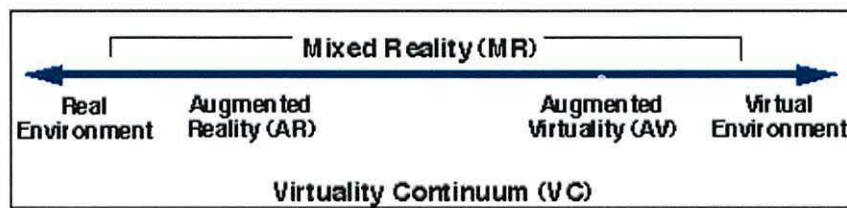
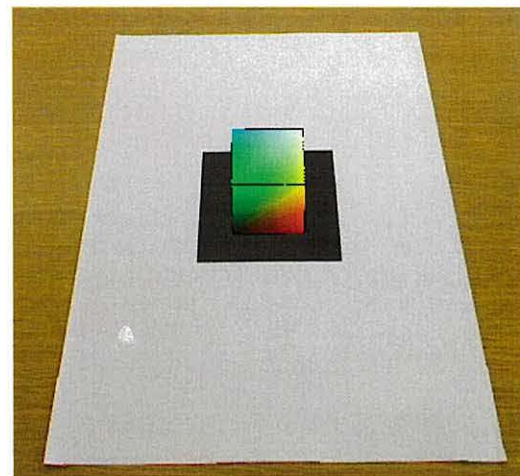


Figure 2.6: Milgram's Virtuality Continuum [43].



(a) A standard video feed.



(b) A standard video feed augmented with a cube.

Figure 2.7: A comparison of a view from a webcam with and without augmentations.

user presence in the system (the impression that the user is co-located with both the real and virtual objects being presented). Misregistration in an AR can be either static or dynamic [2].

- Static misregistration occurs when everything within the system is stationary, yet virtual objects are mis-registered.
- Dynamic misregistration occurs when virtual objects are not located as they should be in the real-world during motion of the viewpoint or of a real object. This is often caused by one component of the system updating at a slower rate than another component, often the computer graphics lagging behind the video stream in a video see-through system. Dynamic misregistration can lead to simulator sickness, which is caused by conflicting sensory information [16] as the user's view of the computer generated imagery lags behind the view of



the real world.

Depth perception is also an important issue, especially if a virtual object is intended to be displayed within a real object. Several factors contribute to our perception of depth. Swan *et al.* [68] provide a list of ten recognized depth cues. These are:

1. binocular disparity
2. binocular convergence
3. accomodative focus
4. atmospheric haze
5. motion parallax
6. linear perspective and foreshortening
7. occlusion
8. height in the visual field
9. shading
10. texture gradient

Landy states that a combination of several of these depth cues may lead to a more accurate perception of an object's distance from the viewer [34]. Drascic and Milgram discuss the relevance of depth cues to AR systems [16]. Some depth cues are easier to simulate in computer graphics systems than others. A number of ways in which depth can be added to an AR environment are investigated and evaluated in [66].



## 2.5 Medical Education Applications

In this thesis we are particularly interested in how medical visualisation techniques and AR technology can be applied to medical education, in particular anatomy education. Anatomy education can be split into several types, including microscopic anatomy, which requires the use of a microscope, and gross anatomy which encompasses anatomy that can be seen with the naked eye.

The amount of time dedicated to the teaching of gross anatomy has decreased significantly over the last fifty years, and a reduced proportion of the available teaching time is spent lecturing and performing dissections, in favour of other topics; this has had a knock on effect on the knowledge acquired [12]. As such, the quality of education must increase to maintain the standard of knowledge of graduates.

The Gold Standard for human anatomy education for medical and Professions Allied to Medicine (PAMS) students has for centuries been cadaver dissection (Figure 2.8). Also in schools, anatomy in biology classes is taught by dissecting small animals and amphibians such as rats and frogs. However the availability of cadavers has reduced over the years due to financial, legal [65] and ethical pressures, and because of availability issues. Dissection allows the student first-hand experience of viewing and manipulating a cadaver, which many deem as being the best way of learning gross anatomy as it not only shows the student the shape of the internal organs, it also gives them an awareness of the location within the body and also knowledge of anatomical variation and anomalies.

Many alternatives to a full dissection exist. For example a student can alternatively view a prosection. A prosection is either a dissection by a professional carried out for students to view, or a pre-dissected specimen that a student can study (for example an arm or a leg). Students are divided on the usefulness of full dissections and many prefer prosections [15] as they allow more time to study the feature of interest, and less time attempting to find it; in some cases a large amount of fat and muscle has to be removed before the feature of interest is found. It is also almost



Figure 2.8: The Anatomy Lesson of Doctor Nicolaes Tulp, Rembrandt van Rijn. 1632

impossible to exactly repeat a dissection due to the nature of the procedure [58]. This is however part of a much larger debate as to the usefulness of teaching gross anatomy [49], and the stance taken by researchers often depends on their specialism. An analysis of this debate is beyond the scope of this thesis.

Gross anatomy can also be taught using Problem Based Learning (PBL) scenarios. PBL presents students with a scenario that is to be solved as a group [78]. The students must then individually seek the required knowledge which is then shared with the other members of the group. This relies on the motivation of the individual students and peer-pressure. This type of scenario not only teaches anatomy, it also encourages teamwork, self-directed learning and presentation skills that will be useful in later stages of a student's career. It also promotes a student's development of differential diagnosis.

The purpose of using PBL is to present the students with knowledge that can be easily applied [60]. It has been suggested that traditional education gives students a lot of knowledge, however it is not always useful because the students are often unable to apply it. It is argued that traditional education relies on prior knowledge



for understanding of new topics.

The use of textbooks is widespread in the education of anatomy and there are several well established books of note. Gray's Anatomy, first published in 1858, is now in its 40<sup>th</sup> edition [67]<sup>1</sup>. Each edition contains illustrations of human anatomy as well as detailed textual accounts of the anatomy being presented. Anatomy Atlases are a common form of anatomy textbook, containing a large number of anatomical illustrations. Grant's Atlas of Anatomy is one example of such a publication. Other popular books include Grant's Dissector [69], which provides students with a detailed set of steps for dissecting a cadaver, and Netters Atlas of Human Anatomy, part of a wide range of anatomy textbooks.

Plastic models can be used to teach students about the shape of a feature of interest, with the advantage that a three-dimensional impression is easier to gain than from a textbook. Models vary in realism, with some being very stylised (Figure 2.9), and most are scaled up or down. Often only prominent features are shown, and only a limited portrayal of location and context of a feature can be made. Plastic models are also costly; a basic brain model costs £165, and a more sophisticated full head and neck model costs around £2300<sup>2</sup>.

Another approach to show three-dimensional anatomy is to have artists create virtual surface models with reference to real medical data. The Primal Pictures 3D human anatomy software [53] uses this approach to create visualizations of human anatomy. The software is available as several packages, dependant on speciality, and allows the user to add or remove layers, and to rotate the anatomy being viewed. The software has been used on the BBC television series "Waking the Dead." Zygote Media Group, Inc. [81] offer three-dimensional models created in a similar fashion, and which are available for use with a number of surface editing packages, such as 3D Studio Max, Maya and Softimage. Zygote's models have been used by the

---

<sup>1</sup>an online version of the 20<sup>th</sup> edition, first published in 1918, can be viewed for free at <http://www.bartleby.com/107/>

<sup>2</sup>prices taken from the January 2009 price list at <http://www.adam-rouilly.co.uk/>

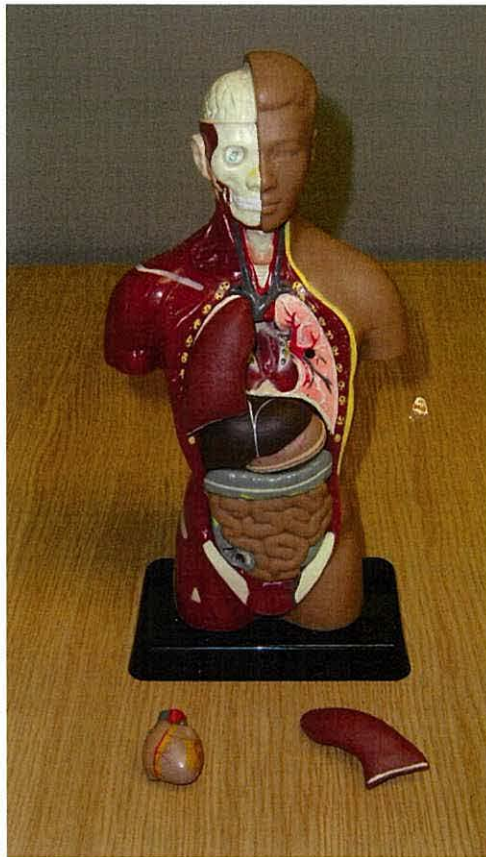


Figure 2.9: A simple plastic model used to teach anatomy to school children.

Coca-Cola Company, The Cartoon Network and Fox News, among others.

Many different software tools have been produced over the last twenty years to attempt to aid the teaching of anatomy. The BodyWorks software by The Learning Company provides three-dimensional models for the user to view, and lectures describing parts of the anatomy<sup>3</sup>. The images presented are of good quality, however it is not considered to be of a sufficient standard to be used as an alternative to anatomy textbooks [46] as it is too simplistic.

The advent of the National Library of Medicine's Visible Human Project [74] has allowed medical data from human subjects to be widely disseminated and used by the medical community for a wide variety of purposes. The entire project consists of data from both a male, sampled at one millimetre intervals, and a female, sampled

---

<sup>3</sup>The BodyWorks software is no longer produced.



at a third of a millimetre intervals. Three types of data were produced for each subject: transverse CT, MR and cryosection images. The Visible Human male data set is approximately 15 gigabytes in size, and the Visible Human female data set is approximately 40 gigabytes. The Chinese Visible Human is a similar project which contains data from Chinese subjects [10], and similarly there is also a Korean Visible Human [75].

Several projects exist that use data from the Visible Human Project. Such projects include:

- MedImage [24]
- Radiologic Anatomy Atlas Viewer [37]
- ITK
- virtusMed project [73] (Figure 2.10)
- The Real Anatomy software [30]

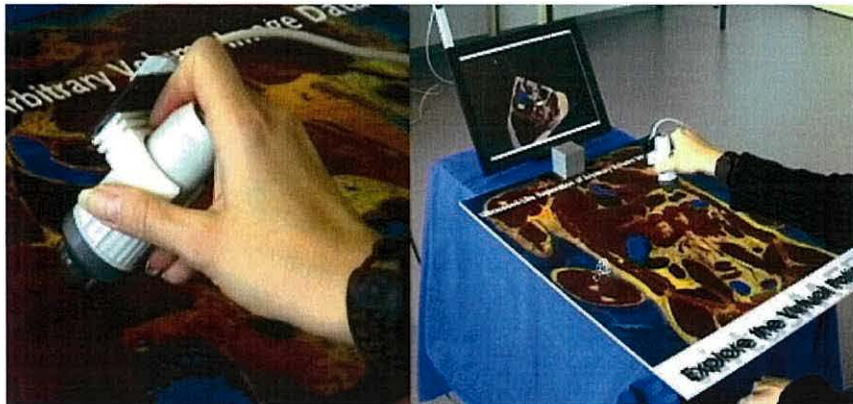


Figure 2.10: The virtusMed project showing an ultrasound simulation of a mannequin [73].

Many other medical training tools are either available or are in development, most of which are aimed at training older students to perform medical procedures, such as the system of Rolland *et al.* for teaching paramedics and other medical staff how to perform Endotracheal Intubation (ETI). In this system [56], [57], [55], [13] a

real (physical) phantom, augmented with an overlay produced using a head-mounted projective display, is used to guide the insertion of the Intubation tool. An optical tracking system is used to track both the phantom and the ETI tool.

The MEDARPA project [61] focuses specifically on assisting needle implantation and the insertion of endoscopy instruments. In this system pre-operative medical data is placed into the surgeons view via a freely positionable semi-transparent LCD. This and other simulations are highly specialised and prior knowledge of gross anatomy is a necessity.

One example of a system aimed at younger people who are not expected to have prior knowledge of gross anatomy is the system by Juan *et al.* [31], which provides a simple AR system for anatomy education that is aimed at young children. The system allows the user to view some of the interior of the human torso by opening up a hole in white fabric that has been stretched over a wooden frame, revealing the virtual content (Figure 2.11). However the augmentations remain in a fixed position in the real world and appear through narrow openings in a fixed model.

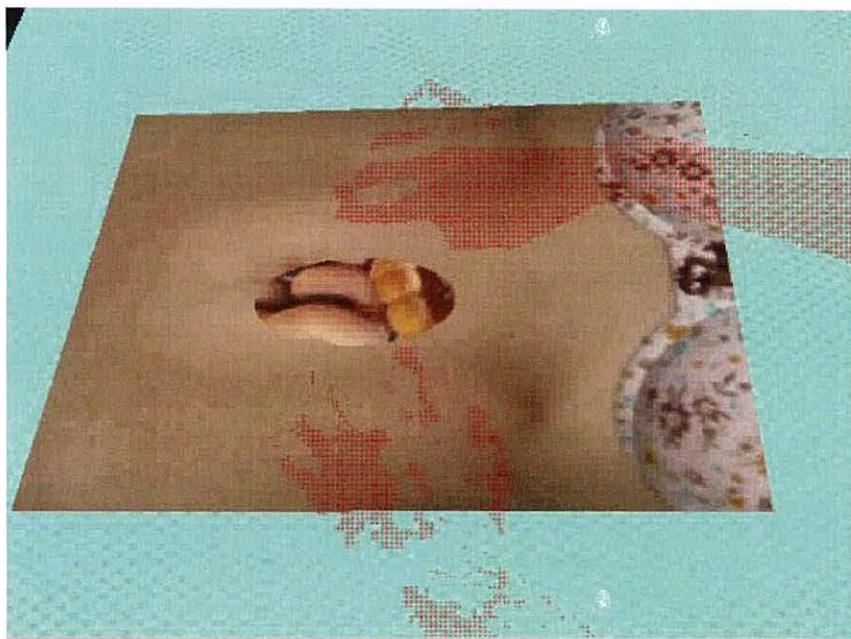


Figure 2.11: The system of Juan *et al.* [31], showing the intestines of the human body through a slit in the torso.

It has been shown that computer simulations can compare favourably with more



traditional anatomy education methods. In a 2004 study by Hariri *et al.* [23], learning the anatomy of the shoulder from a textbook was compared with learning the same anatomy using a computer simulation. Following a ten minute learning period the subjects were asked to identify a subset of anatomical features. The computer-simulator-trained students performed slightly better in this identification task than their textbook trained counterparts, and rated it more highly than those evaluating the textbook with respect to its ability to teach anatomy, its use of modality available, its ease of use and its image realism.

Computer simulations and educational software can provide the user with a wealth of information about the object that is being studied, however sight is but one sense, and a person's perception of an object is created from a combination of all five senses. The perception of an environment,  $P$ , has been described as a function over time ( $t$ ) of task ( $\tau$ ) and preconditioning ( $\rho$ ) [8], [9]:

$$P(\tau, \rho)(t) = \omega_V V(t) + \omega_A A(t) + \omega_S S(t) + \omega_T T(t) + \omega_F F(t) + \omega_\Delta \Delta(t) \quad (2.1)$$

where  $V$  = Visuals,  $A$  = Audio,  $S$  = Smell,  $T$  = Taste and  $F$  = Feel.  $\Delta$  is a measurement of distraction indicating how focussed the user is on the environment. The  $\omega_i$  term is the particular perceptual weighting that each of the senses, and any distraction, has for the perception of that particular moment, with  $\sum \omega_i = 1$ .

The stimulation of multiple senses in medical educational AR has to a certain extent been investigated. Nieder *et al.* suggested that having a physical representation to hold, as well as a virtual environment to view, can aid a user's understanding of the anatomy being viewed in a virtual environment [47]. The use of a physical model as an interface to a virtual environment could also provide benefits to the learner. In his survey of tangible interfaces, Marshall suggests that "three-dimensional forms might be perceived and understood more readily through haptic and proprioceptive perception of tangible representations than through visual representation alone" [40].

Similarly Panchaphongsaphak contends that using a tactile interface (often referred to as Tactile Augmentation or Passive Haptics) in educational systems increases ease of use and the accuracy obtained [50]. In his system, a reproduction of a piece of anatomy is mounted on a 6-DOF force-torque sensor that is then used to manipulate an on-screen equivalent. The system has two modes, the first of which can detect where the user is pressing the model and highlights this point on-screen as well as providing text and audio information. The second mode allows the user to move an on-screen cross section of the anatomy by moving their finger along the model. The model is placed between the user and the display (a standard monitor), possibly restricting the features that can be seen by the user. Passive haptics can also be used to dissuade a user from following an inappropriate path. Insko [28] uses this to guide a user around a maze, and concludes that the use of passive haptics improves cognitive mapping and increases training transfer.

Moody, [44], describes surgical training in which the user is presented with a model of the exterior of a human knee for practicing arthroscopy technique. A set of experiments are described which first determine whether performance is improved in a search task when a plastic model (the purpose of which is to create a degree of passive haptic feedback) of a knee joint is inserted into the knee model compared to without. The knee model remained static in both instances. It was determined that transition from the version of the system with passive haptics to the one without was simpler than the opposite transition, and would therefore be more useful as an education tool for the procedure. Although the task took longer on the system with the passive haptics, it is more realistic than performing the task without. The second experiment assesses the difference between consultants, residents and complete novices in arthroscopy to see how performance varied. A questionnaire was deployed in which the consultants responded that the system was useful for basic training in arthroscopy.

A mixed reality system has also been used to plan cardiac surgery. Seitel *et al.*



have produced a system which allows the user to delineate the mitral annulus, a ring-shaped structure surrounding the mitral valve [63]. This too uses a physical model, in this case a model of the human heart, with which the user can interact. In this case the interaction is carried out using tracked pointers which are used to perform the delineation.

Learners could also benefit from being able to interact physically with the object that they are viewing in an unrestricted fashion so that it can be seen from any angle that the user desires and also provides tactile information. Gillet *et al.* [22] demonstrate a system in which this kind of interaction is possible. The real object that the user is interested in is a simple rapidly prototyped molecule model — tracked using the ARToolkit — onto which various complex computer-generated representations are superimposed. It is argued that the tactile and kinesthetic cues are useful in helping to understand the spatial characteristics of an object.



Figure 2.12: A physical representation of a human liver derived from volume data.

## 2.6 Summary

Digital medical images are now available for many medical specialities, providing an alternative to more traditional medical imaging. These images come in a variety of forms and can vary greatly in the level of detail provided. The image data must be processed in some way before it can be visualised. One technique for visualising medical data is surface rendering, which visualises a small part of the data, often extracted using segmentation or iso-surfacing. Another technique is direct volume rendering, a technique that can show all of the data or smaller parts. The visibility and colour of portions of a volume rendering can be altered using transfer functions. Such techniques are becoming widespread in medical practice.

An Augmented Reality (AR) is a type of virtual environment in which computer generated imagery is combined with reality. In an AR the integration of the real and virtual should be seamless. In recent years AR technology has been used for educational purposes, including use for surgical simulators.

The education of gross anatomy has seen much change over recent decades, as the use of traditional techniques — often involving the dissection of cadavers — are becoming less widespread because of many pressures, including financial and ethical issues. Many alternatives to cadaver dissection exist, however cadaver dissection is still the Gold Standard despite its diminishing usage.

## Chapter 3

# The Components of an Augmented Reality System

## 3.1 Introduction

In this chapter the Bangor Augmented Reality Education Tool for Anatomy (BARETA) framework is introduced. The BARETA framework allows the production of bespoke Augmented Reality (AR) anatomy lesson aids using Rapid Prototyping (RP) technology (discussed in Chapter 4) and volume data (discussed in Chapter 5) in which a user can view volume and surface renderings of anatomical features, controlled using a physical representation of the same anatomical feature.

The BARETA framework contains two main component types: tracking components (discussed in Section 3.2) and display technologies (discussed in Section 3.3). The components used by the BARETA framework are then discussed in Section 3.4. Section 3.5 then describes the integration of the individual components of BARETA: to achieve correct registration a calibration step is required to discover transformations between the co-ordinate systems of the devices used, and the system used to generate the graphical content. Calibration software was created to solve for these transformations. This allows BARETA to be used in more than one location, all that is needed to ensure correct registration a transformation that is obtained by running the calibration software. Section 3.7 then presents a summary of this chapter.

## 3.2 Tracking

In an effective AR system the location and orientation (pose) of objects in the real world must be known. As objects are likely to be in motion, a tracking system is required to ensure correct registration. Several different types of tracking system exist, many of which are suitable for use with AR systems. Tracking systems are sometimes described as being either inside-out or outside-in. Inside-out tracking systems are those where the tracking system is in motion within a fixed environment. Outside-in tracking systems remain static and track the pose of objects that are in motion. Often the choice of tracking system will depend upon the individual



Table 3.1: A table comparing tracking hardware

Technology	Advantages	Disadvantages	Cost From
Mechanical	Accurate	Heavy, Restricted range	£140 (3-DOF)
Magnetic	Accurate under ideal conditions	Susceptible to noise caused by magnetic fields in the operating environment	£1170
Optical	Fast, potentially a wide range of operation	Do not operate well under certain lighting conditions, line of sight to fiducials required, fiducial position must be accurately known	£5435
Inertial	Sourceless	Suffer drift, initial position must be known to find absolute values	£990
GPS and Differential GPS	Large operating range	Only work well outdoors in wide open spaces, provides position only	£35
Hybrid	Can operate under a wide variety of conditions	Expensive	£6840

application. Factors affecting the choice of tracking system include the location at which the AR system will be used (indoor or outdoor), the range over which tracking is required, and budgetary requirements. Ideally for an AR system a 6 Degrees of Freedom (6-DOF, position and orientation) tracking system will be used.

The most common types of tracking system used for AR are outlined below. This section begins with a discussion of hardware-based tracking systems, and concludes with a discussion of software-based tracking systems.

### 3.2.1 Hardware

Several hardware solution for tracking objects in the real world exist. Each has its advantages and disadvantages that must be matched to the environment in which the AR system will run. A summary of hardware tracking technologies is provided in Table 3.1.

### 3.2.2 Software

Several software tool kits exist that aim to act as a tracking component in an AR system, and some provide an interface to underlying display libraries. A typical basic system will include a webcam that will provide images to the tool kit. Image processing techniques will be used to find fiducials in the 2D image and further processing of these fiducials will result in a 3D pose, which can then be used to render graphics that are correctly registered with the original image. Software toolkits include:

- The ARToolkit
- The MXRToolkit
- Magic Symbol
- Metaio Unifeye
- OpenCV

The ARToolkit, MXRToolkit and OpenCV are open source, Magic Symbol and Metaio Unifeye are proprietary.

## 3.3 Displays

As with tracking there are a large number of display solutions that can be used for AR. A complete display solution will include both hardware and software components. Perhaps the biggest factor affecting the choice of display solution is whether the system is optical see-through or video see-through. In optical see-through systems the user has a direct view of the real world to which graphics are augmented. In video see-through systems the view of the real world is provided by a video stream

from a video camera. Alternatives to these two choices exist, however optical see-through and video see-through are the most common.

### 3.3.1 Hardware

In an AR system there are often several hardware components that can be considered as being display items. In video see-through systems the camera that captures the real-world image could be considered to be a part of the display hardware setup, as could a video capture interface connected to this camera, if such a device is used. The use of video capture interfaces was common in the past, however in recent years USB web cameras have increased in resolution and in frame rate, as well as being more readily available and cheaper. A second issue is the computer hardware on which the AR environment is being run. Finally the video output device must be selected.

#### Workstation

The workstation on which an AR system is run can have a significant effect on the performance of the system. AR systems have been run on dedicated high-end graphics workstations in the past, however current PCs are capable of rendering many AR applications at real-time frame rates, especially those with high-end graphics cards. Basic AR is even possible on PDAs and some high-end mobile phones [51]. In general the workstation used will depend upon the requirements of the application. Indoor systems are less likely to require mobility and will be able to take advantage of the computing power of a desktop PC or a graphics workstation. Outdoor applications usually require more mobility and must use portable devices. Notebook PCs have been used in such circumstances [70], however this is very much an experimental solution and not suited to end users. The notebook and peripherals are usually attached to a backpack and can be fairly cumbersome.



### Output Devices

As previously discussed AR displays can be classed as either optical see-through or video see-through. As well as classifying AR displays in this way they can also be classified as head-mounted and non-head-mounted. Head-mounted displays allow the user to easily and intuitively manipulate the viewpoint of the AR system. HMDs can however be cumbersome devices and can cause eyestrain among some users, due to the proximity of the displays to the users eyes, combined with the conflicting depth cues discussed in Section 2.4; this can also induce motion sickness. Such systems also require a connection between the display and the machine producing the graphics, meaning that either the user's movements are restricted by a cable connecting the two, or a mobile wearable PC has to be carried by the user.

Head-mounted displays are available as both optical see-through and video see-through, although the latter is more common. One cheap solution for a video see-through head-mounted display is to use a unit designed for viewing movies, such as the Icuiti DV920 (or its successor the Vuzix iWear series). Optical see-through devices are often much more complex as they utilise some form of optical combiner, often a half-silvered mirror. One example of a Head-Mounted optical see-through display is the Sony Glasstron PLM-S700 that was released in 1997 and has been used in several AR projects. This device has since been discontinued.

There also exists a type of head-mounted display that cannot be classified as either optical or video see-through: the Head-Mounted Projective Display (HMPD) [25]. Such displays project virtual images onto the real objects that are being viewed by the user. This does however rely on there being a real object that can be projected onto.

Non-head-mounted displays are usually standard computer monitors. The AR environment displayed is usually composed of a video stream onto which computer graphics are added. Such systems are often less intuitive than those utilising HMDs, however the lack of cumbersome headgear can mean that the system can be used for



longer periods without experiencing discomfort. Such systems are useful when the user is controlling a remote object. A more intuitive system can be created using modified LCD display that has had its back panel and back light removed, allowing it to be used as an optical see-through display [61].

Although not yet a widely used technology, Organic Light Emitting Diodes (OLEDs) present exciting opportunities for use in AR. OLEDs can be made in a variety of shapes and can even be flexible. OLEDs have several other advantages over traditional LCDs including increased brightness, contrast, viewing angle and a lower power consumption because no backlight is required. Currently the biggest disadvantage of OLEDs is their lack of longevity, however this is improving.

Another emerging technology that could be exploited by an AR system is 3D holography. 3D holography devices typically project images onto a transparent screen, or some other reflective medium. 3D holography can have a large view angle, however direct user interaction is not yet possible.

### 3.3.2 Software

Several software options exist for creating graphical augmentations. Some of the software tracking systems described in Section 3.2.2 provide an API for drawing graphics, whilst others simply supply tracking data that can then be used by the user in whichever environment they see fit. The two most common environments for creating graphics are OpenGL and Microsoft's Direct3D. The Virtual Reality Modelling Language (VRML) can also be used. Another option is to use a separate software library that builds upon one of these APIs. One such library is the Visualization Toolkit (VTK), a cross-platform library that uses OpenGL. VTK provides C++ classes that can read and render a large variety of data types. It has a pipeline architecture that allows filters to be inserted that can manipulate the input data to produce a variety of different results.

## 3.4 Components of BARETA

### 3.4.1 Requirements

A list of requirements for the BARETA anatomy teaching tool platform was drawn up in collaboration with Doctor John Delieu, an anatomy lecturer based at Bangor University. It was decided that in addition to Azuma's requirements for an AR [2], BARETA must:

- track a mobile viewpoint
- track two physical objects (an anatomy model and a tool)
- be capable of rendering high resolution medical volume data in real-time
- be suitable for use in a classroom environment
- be easily moved between locations

These requirements influenced our choices of technology, which are discussed below.

### 3.4.2 Tracking

As BARETA required that both an object and the viewpoint were tracked, our tracking system needed to be able to track over the likely locations that both of these could assume. For tracking the two physical objects a magnetic tracker was deemed suitable because although it provided a limited range its accuracy over this range was good, and did not require a line of sight between the transmitter and the sensor, therefore an object could be attached to the sensor without adversely affecting tracking; this also meant that the position of a user's hands did not affect tracking accuracy. The magnetic tracker used was an Ascension miniBIRD 800 (miniBIRD) (Figure 3.1) with two sensors and electronics units. Each miniBIRD

unit is about the same size as a small digital set-top box, and the transmitter smaller, therefore the entire system is very easy to transport. The miniBIRD's tracking radius of about 36 inches from the transmitter is very small, however the user was not expected to move the physical objects very far and the objects' rotation is much more important in our application. Viewpoint tracking could also have been carried out using magnetic tracking, however this would have limited the range of movement of the viewpoint too much, because it would have to be close to the origin of the magnetic tracking system to track accurately, and therefore to the object.

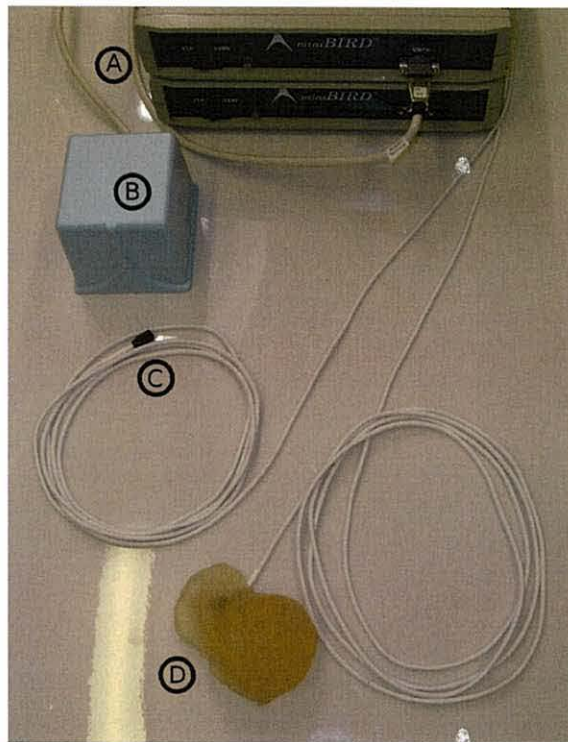


Figure 3.1: The Ascension miniBIRD 800 showing the electronics units (A), the transmitter (B), a sensor (C) and an object attached to a sensor (D)

The system used for viewpoint tracking was the InterSense IS-1200 VisTracker (VisTracker), a light-weight device small enough to fit in the palm of the hand. It allowed a wide range of tracking within an environment prepared with fiducial markers. Markers needed to be precisely placed for maximum tracking accuracy. As the VisTracker is a hybrid system it is not necessary for the fiducial marker to be in view for the entirety of a BARETA session; the inertial component of the tracker can



produce accurate tracking results for a short time when no markers are in view of the optical sensor. The user application did not retrieve data from the VisTracker directly: data was retrieved from a TCP/IP server provided by InterSense that communicates directly with the device through a USB port.

Our two tracking devices had different APIs that offered a variety of functions, many of which were not required by our application, and different ways of representing positions and orientations. The miniBIRD was capable of returning data in a variety of different representations for use in different applications. Best suited to our application was the *positionmatrix* option that returned a  $3 \times 3$  rotation matrix and a position vector for each sensor. The matrix and the vector are, in BARETA, combined to produce a  $4 \times 4$  homogeneous transformation matrix. Such a matrix is commonly used in computer graphics applications to represent transformations.

The VisTracker reported data as two vectors, one of which stored the position of the VisTracker in 3D space, and the second stored the rotation as three Euler angles; rotations are in the sequence z, y, x. The individual rotations and the translation were encoded in separate  $4 \times 4$  homogeneous transformation matrices and multiplied in sequence to produce a matrix representing the transformation of the VisTracker in a similar format to that produced from the miniBIRD tracking data.

To simplify system development the most common functions were included in a new tracking library that performed operations such as initialising and shutting down the trackers, retrieving and pre-processing data, and passing standardised data to the application. This provided a common programming interface to the two tracking systems, and allowed for the creation of an abstract tracker. This also allowed for more tracking systems to be added, or replaced without making a significant alteration to application source code.



### 3.4.3 Display

BARETA was run on a standard Intel Pentium 4 Windows desktop PC with 1GB of RAM and a high-end nVidia GeForce 8800GTX graphics card. This graphics card allowed for the use of fast hardware 3D texture mapping of the volume data that was to be used for BARETA, rendering it in real-time.

An off-the-shelf Logitech Quickcam Pro 9000 USB webcam (Figure 3.2) was used to provide the real part of the BARETA environment. It could capture video at resolutions up to  $1600 \times 1200$  pixels at 30 frames per second. Features of this camera included a capability to automatically adapt to a wide range of lighting conditions, auto focus and face tracking. The final two features were both disabled for use with BARETA as they could be detrimental to the registration of the system.



Figure 3.2: The Logitech QuickCam Pro 9000 USB webcam

The VTK was used as the rendering library for BARETA, allowing the use of a diverse range of data types and real-time rendering options (although frame-rate decreased with large data sets), including video streams from a webcam and a variety of volume rendering methods. VTK also allowed for applying transformations to objects and the virtual camera using a variety of methods, including  $4 \times 4$  homogeneous

transformation matrices, and several useful mathematical functions including functions that returned constants to convert degrees to radians and radians to degrees, and functions to solve a variety of equations.

Several display output devices were evaluated. For initial experiments a standard computer monitor was used. An Icuiti DV920 HMD was also evaluated (Figure 3.3(a)). The DV920 is a lightweight monoscopic device designed primarily for use for watching movies. Such a device would help to create a seamless AR, however, for use with BARETA the VisTracker and the webcam would have to be mounted on it, adding a considerable weight to the device. The light weight of the DV920 meant that attaching these peripherals was not practical as it became heavy on the user's nose and was prone to slipping off as it was much heavier at the front than at the back. Its small resolution of  $640 \times 480$  meant that full advantage of the host PC's graphics capabilities could not be fully exploited. A stereoscopic monitor was also evaluated for use with BARETA. The Zalman Trimon ZM-M220W was chosen, which uses circular polarisation combined with polarising glasses to create a stereoscopic image (Figure 3.3(b)). Although the stereoscopic image was good, the horizontal view angle is only about  $10^\circ$ , making head position critical to correctly view the stereoscopic image. This made the display impractical for use with BARETA as the user was expected to spend some time looking at the physical object and would have to re-adjust their viewing position every time that their attention was switched back to the display. A summary of the advantages and disadvantages of each output device is presented in Table 3.2.

For our user studies (described in Chapter 6) a standard monitor was used to allow several students to view the augmented reality at the same time and to ease the changeover between participants. This also allowed students to focus on the information being presented to them rather than difficulties adapting to the display. The set up of the system must ensure that the user is able to view the screen whilst using BARETA's tracking devices.



(a) The Icuiti DV920 Head-Mounted Display



(b) The Zalman Trimon ZM-M220W stereoscopic monitor

Figure 3.3: Display devices

Table 3.2: A table comparing display output devices

Technology	Advantages	Disadvantages	Cost From
Standard Monitor	Wide view angle, no wires restricting user position	Can cause co-ordination difficulty, monoscopic	£50
Icuiti DV920	Lightweight, if tracked can provide a view manipulated by the user's head movements, can help make an AR seamless	Difficult to mount peripherals onto, monoscopic, low resolution, can cause eye strain if used for an extended time	£260
Zalman ZM-M220W	Stereoscopic, high resolution	Small vertical view angle, special glasses required	£550



## 3.5 System Calibration

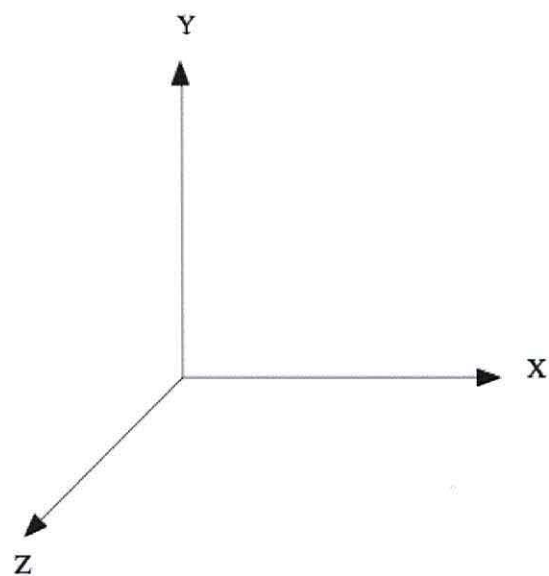
BARETA was composed of several components that were required to interact with each other to produce an AR environment. As each component had its own co-ordinate system the components would not interact to produce the desired result without first determining transformations between co-ordinate systems, and applying them to the tracking results. Provided that each tracking device was within range of its source during calibration and during the operation of BARETA, the relative locations of the tracking devices is not important.

### 3.5.1 Tracking System Calibration

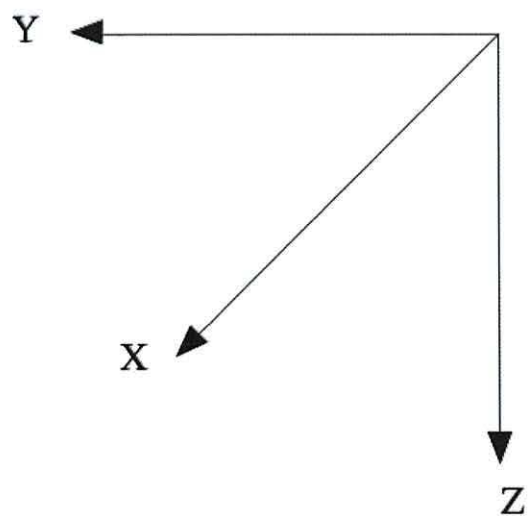
Although both of our tracking systems and VTK used right-handed co-ordinate systems, they were not oriented the same way, and had different origins and scalings; the miniBIRD reported displacement in inches, the VisTracker in metres. This made it necessary to apply transformations to ensure that the two were using the same co-ordinate system. The co-ordinate systems of the two tracking devices were rotated within the tracking library so that they matched the standard graphics orientation used by VTK.

To allow the VisTracker to operate correctly fiducial patterns were placed at known positions on a wall. One pattern was chosen to be at the origin, the rest were positioned at regular intervals from this origin (Figure 3.5). This became the world co-ordinate system, and would be the co-ordinate system used by VTK. This meant that the VisTracker reported its position and orientation within the world co-ordinate system.

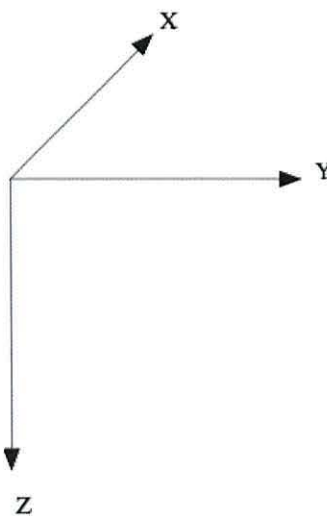
For the object to be correctly registered in the AR environment, the miniBIRD transmitter was positioned at a fixed location and a transformation from the miniBIRD co-ordinate system into the world co-ordinate system was required. Although an estimate of origin offset could have been made manually, and the difference in scale



(a) The standard graphics co-ordinate system



(b) The Ascension graphics co-ordinate system



(c) The VisTracker graphics co-ordinate system

Figure 3.4: A comparison of co-ordinate systems



Figure 3.5: The fiducial constellation used by the VisTracker

was a known constant, rotations were much harder to measure manually, especially when they must be known about three axes; therefore another method was required to compute the transformation.

To find this transformation a miniBIRD sensor and the VisTracker were placed within a unique precision made calibration bracket (Figure 3.6) that was designed and manufactured by the workshop at Bangor University. The bracket was designed to maintain the relative position and orientation of the VisTracker and miniBIRD sensor at known values, whilst separating them to reduce the electromagnetic interference from the VisTracker that was received by the sensor. Readings from both trackers were taken at forty locations within the range of the miniBIRD. At each position an average of one hundred readings from the miniBIRD was taken to further reduce the effect of the interference from the VisTracker. A least squares estimation was then used to compute a transformation from the miniBIRD co-ordinate system to the world co-ordinate system used by the VisTracker.



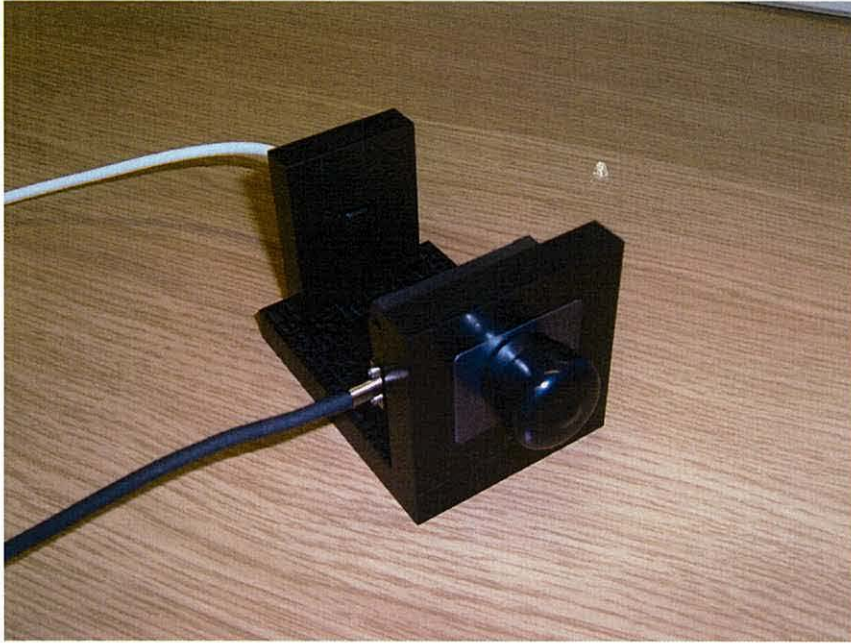


Figure 3.6: A calibration bracket with the VisTracker and miniBIRD sensor in place

### 3.5.2 Webcam and VisTracker Calibration

Calibration was also necessary between the webcam used and the VisTracker because the two had slightly different viewpoints, and the purpose of the VisTracker was to track the viewpoint of the webcam. A second bracket was produced (Figure 3.7) to maintain the webcam and the VisTracker at a constant distance from each other. This distance could either be measured manually, which would have produced only a position offset, or computed, which would have given a position offset and rotations about three axes. This required the determination of the position and orientation of the webcam in the world co-ordinate system.

The VisTracker used the algorithm of Naimark and Foxlin [45] to determine its pose by detecting circular fiducial markers at known positions in its environment. As this was the environment in which the webcam would also be operating, a modified implementation of this algorithm was ideal to determine the positions of fiducials in the image from the webcam. The algorithm proceeds as follows:

1. An image is captured by the webcam (Figure 3.8(a)). In our implementation a reading is also taken from the VisTracker and stored.



Figure 3.7: The application bracket with the VisTracker and webcam in place

2. The original image is converted into a greyscale image (Figure 3.8(b)).
3. The lighting in the greyscale image is made uniform using the following equation taken from [45]

$$p(n, m) = 105.89 * \log(f(n, m) + 1) \quad (3.1)$$

where  $f(n, m)$  is a pixel in the original image, and  $p(n, m)$  is a pixel in the resulting image where  $f(n, m) \in [0, \dots, 255]$  and  $p(n, m) \in [0, \dots, 255]$  (Figure 3.8(c)). This allows reliable detection to occur under most lighting conditions as it ensures that there are at least 40 grey levels between adjacent black and white regions.

4. An edge detection is performed on the uniform image. Our implementation uses a standard Sobel operator to detect edges in the image (Figure 3.8(d)).
5. The image showing fiducial edges is blurred using a simple  $3 \times 3$  mean blur filter (Figure 3.8(e)).



6. A threshold is applied to the image. The threshold value is discovered by first locating the median value in the histogram. From this two peaks can be located; one of these peaks is of a lower value than the median, the black pixels. The other peak has a higher value than the median, the white pixels. The local minimum value between these two peaks is used as the threshold value (Figure 3.8(f)).
7. The binary image has a horizontal erosion applied to it to remove any small artifacts (Figure 3.8(g)).
8. Candidate features in the eroded image are highlighted (Figure 3.8(h)). Bounding boxes are generated around all white features in the binary image. All features that are too large or too small are immediately disregarded. Features that are long and thin or L-shaped are then disregarded. Candidate fiducials are then coloured grey.
9. Candidate fiducials are extracted from the greyscale image using the bounding boxes found in the previous step and saved in separate files (Figure 3.9(a)). Each image is then thresholded as previously (Figure 3.9(b)) then the center of the fiducial is located by finding the centre of gravity of white objects found at the centre of the fiducial (Figure 3.9(c)). It is not necessary to make adjustments to the grey levels at this stage as only small regions are used, which are unlikely to have large variations in greyscale of black or white portions. The centre is then highlighted on the greyscale image as a small grey dot (Figure 3.8(i)).

Following this process, the location of the centres of the fiducials and the Vis-Tracker were output to a text file. The fiducial centres were then manually compared to the original image to obtain correspondences between fiducial centres, and their location in the world co-ordinate system. The pixel co-ordinates of the centres and the world co-ordinates of the fiducials were then written into another text file that



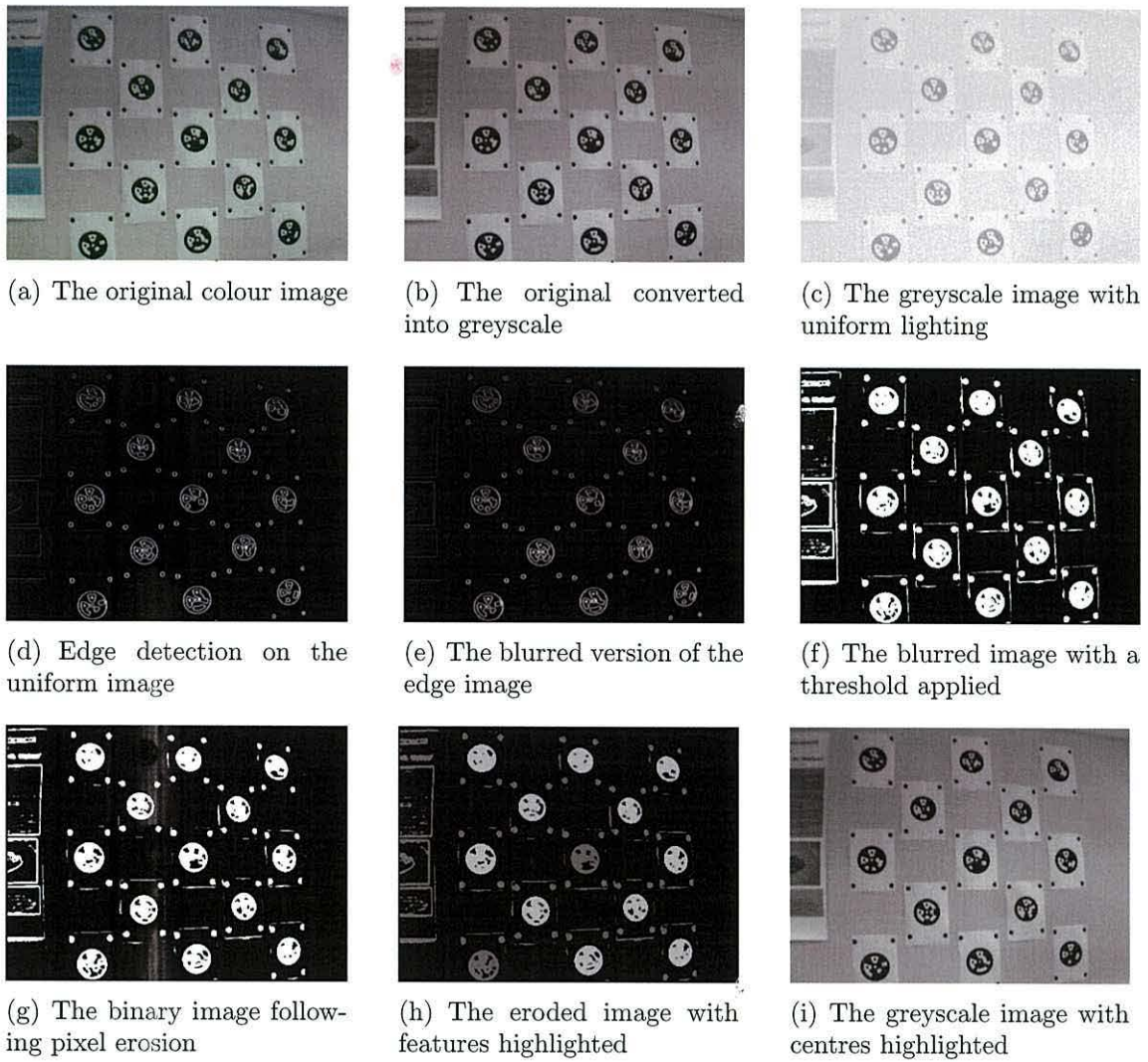


Figure 3.8: The circular fiducial detection procedure

was then input into an implementation of Tsai's calibration algorithm [71] along with the locations of the corresponding fiducial locations in world co-ordinates. The output was an intrinsic parameter matrix describing the camera's internal parameters and an extrinsic matrix describing the camera's location in world co-ordinates. The extrinsic matrix could then be input into a least squares function, along with temporally corresponding readings from the VisTracker to compute the transformation between the two.

The parameters of the virtual camera were also adjusted to ensure correct registration. The view angle was changed so that it matched that of the webcam, and a perspective view transformation was used to ensure that the rendering appeared

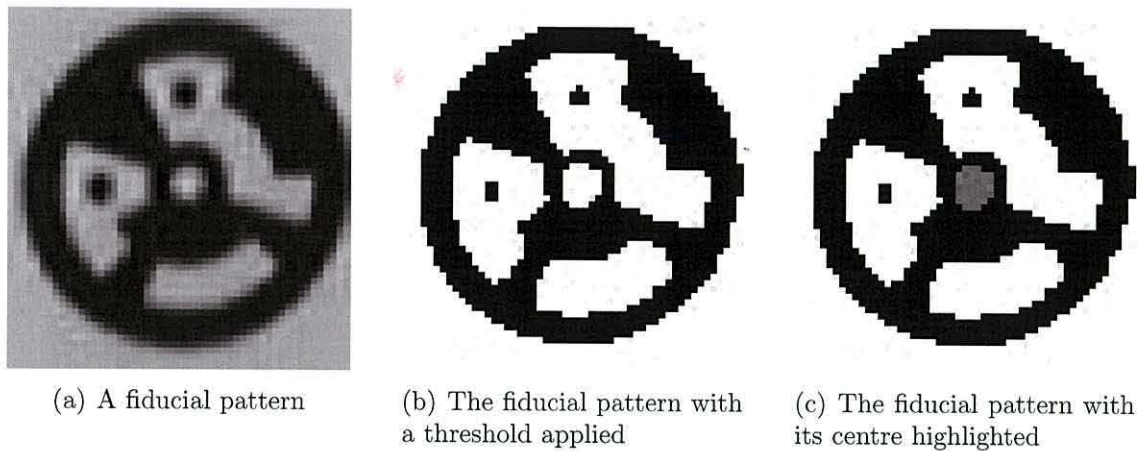


Figure 3.9: Detecting the centre of fiducial patterns

smaller when the object was moved further from the camera.

## 3.6 Graphical Content

The addition of content to the BARETA framework to create a full BARETA is a straightforward process. BARETA allows for the addition of a class that uses VTK to introduce content. Such classes must provide a number of methods that are required by the BARETA framework, the most important of which is a method that pass pointers to the framework, allowing access to instance of a *vtkProp3D* (a class that acts as a container for a 3D object or objects) containing all of the content of the lesson. Other important methods are the method that allows the next step of the content to be shown (described in more detail in Section 5.5), and the method that allows a *vtkRenderWindowInteractor* to be passed to the content. The *vtkRenderWindowInteractor* allows for time dependent events to occur within the content.

The content class will typically load volume and surface data of the anatomy that is to be displayed. This data will then be placed into a VTK pipeline that may edit filter the data in some way. In the case of volume rendering a transfer function may be applied. The class will also include ways of altering some of these



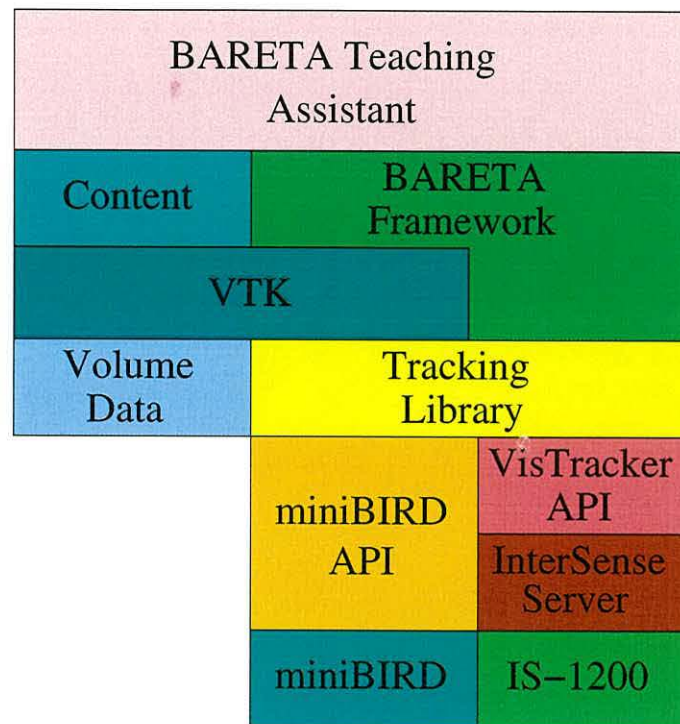


Figure 3.10: A software architecture diagram showing the relationship between the components of a complete BARETA system.

parameters when certain keys are pressed to reveal different aspect of the anatomy that is being shown. The content fits into the BARETA framework as shown in Figure 3.10.

## 3.7 Summary

A large number of technologies exist that support the creation of an AR. The choice of technology often depends on the intended use of the AR system, and how much money can be invested in it. The main challenge in creating an AR environment is in the combination of the different technologies used. In the case of BARETA several different hardware components had to be integrated, including two different tracking systems and a webcam. An InterSense IS-1200 VisTracker was attached to a webcam which allowed us to track the viewpoint, and a two-sensor Ascension miniBIRD system that allowed us to track two separate objects.

To satisfy our requirements we had to ensure that all of the tracking components



worked with a common co-ordinate system so that a change in pose for any of them would not result in a loss of registration. This meant that two calibration steps were required to recover the difference in pose, firstly between the VisTracker and the miniBIRD, and secondly between the VisTracker and the webcam; this first required that the pose of the webcam be determined from an image. This was determined using an implementation of the algorithm used by the VisTracker.

To allow medical volume data to be rendered the VTK was used. VTK allowed the use of a variety of volume rendering techniques. This data could then be easily transformed using the output of the tracking systems. The data was rendered using a standard desktop PC with a high-end graphics card that supported several rendering techniques that could execute in real-time. The use of a standard PC also allows the software to be run at several locations, enhancing the system's portability; only the peripherals need to be moved. We also used a standard computer monitor as our display, which allows the user more freedom of movement than a head-mounted display and allows several students to collaborate with each other in a classroom as they can all see the AR.

## Chapter 4

# Tangible Interfaces

## 4.1 Introduction

In this chapter the use of tangible interfaces as a novel interface to the BARETA environment is discussed. Initially a brief overview of tangible interfaces is presented. The third section describes an AR environment that uses a basic tangible interface that was developed during this project, using a generic anatomy mannequin and the ARToolkit. Rapid Prototyping (RP) is then discussed in Section 4.4; RP technology can be used to create unique physical models from surface data. Section 4.5 discusses how surface data was extracted from medical volume data sets. The following two sections discuss how this extracted data was refined to create two anatomically correct RP models (a liver and the ventricles from within the brain) that would later be used as a novel interface to our AR anatomy lessons. Section 4.8 discusses other tangible interfaces used to enhance the user experience of the AR environment. A summary of this chapter are then presented in Section 4.9.

## 4.2 Tangible Interfaces Overview

Tangible interfaces are widely used in every day life for performing a large variety of tasks. Tangible interfaces are important because they give the user both a target at which to aim their hand, and usually some feedback indicating the reaction that is generated by a user's action. For example when using a computer mouse it is useful to feel that the mouse is within your grasp and to feel the resistance of the buttons as they are pressed. Often users expect tactile sensation to accompany visual stimuli, or as an indicator of success when the object that is being touched is not visible.

In Augmented Reality (AR) environments there is typically no tactile sensation associated with the system. If a user is to interact with a virtual object the only indication that they have of their success or failure in interacting is in the resulting virtual content. It is because of this that the use of tactile interfaces is an interesting research topic relevant to AR. Tangible interfaces can also provide a user with more

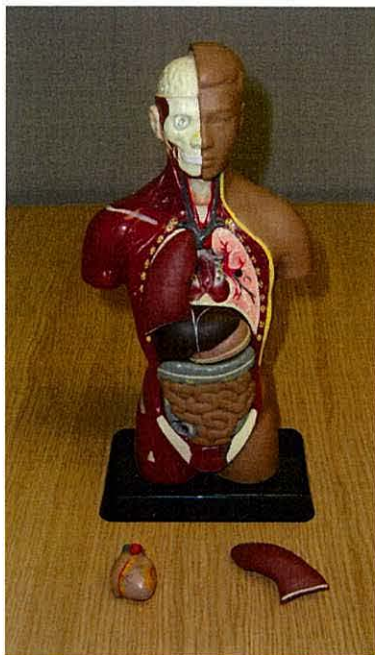


information about the shape of an object than is immediately evident from looking at computer generated graphics, and also information about the surface texture of the object if the tangible model can be produced to replicate this. Tangible interfaces can also help a user to resolve ambiguities in data that may appear onscreen [32]. Another advantage of tangible interfaces is that they reducing the multiplexing of input that a user must be aware of when they interact with a system [20], and also allowing the input of the user to directly correspond with the output. This can help the AR system to seamlessly integrate into the user's physical environment [6].

In many computer systems, such as the medical training system by Vidal *et al.* [72], giving tactile feedback is possible using active haptic joysticks that provide feedback through computer controlled electric motors acting on mechanical joints. Such devices are usually only suited to single user systems and are restrictive in possible displacements and rotations available to users. Passive haptics (also referred to as a tactile interface or tactile augmentation) do not use any electronic components and simply give the user an object of equivalent shape to the on-screen virtual model to hold and manipulate in unrestricted space. However unlike active haptics devices, reactionary forces cannot at present be accurately replicated using passive haptics, therefore information on deformation cannot be included. Passive haptics does, however, allow the user more freedom to move the object of interest in any direction or rotation that they choose, but if the object is to be tracked, restrictions may be imposed on this freedom.

### 4.3 Generic Anatomy Mannequin Interface

Initially the removable parts of a simple anatomy mannequin (Figure 4.1(a)) were used as a tactile interface. Although the parts provided a reasonable approximation to the shape of some of the body's organs it was felt that the shapes were too stylised and lacking in detail to be of benefit to medical students, and the organs were much



(a) A generic anatomy mannequin used to teach anatomy.



(b) A generic anatomy mannequin prepared with AR-Toolkit fiducial markers.

Figure 4.1: The generic anatomy mannequin.

smaller than actual size. Larger life-sized models are available; however, although more detail is present these too are stylised and some of the larger organs can be difficult to handle because of their size. Larger models are also more expensive than their smaller equivalents.

Another issue with using parts of our anatomy mannequin as an interface was that it was difficult to attach an Ascension miniBIRD sensor (see Section 3.4.2) to any of the organs because of their small size. Indeed, instead of using the miniBIRD the software for this interface used the ARToolkit for tracking, with miniature fiducial patterns attached to the organs with sticky tape (Figure 4.1(b)). The ARToolkit uses computer vision techniques to track these patterns in an image from a webcam, therefore the patterns must be visible to the webcam.

Although the above solution worked well within a restricted environment, an interface that facilitates the use of a more accurate tracking system was desirable. We experienced problems with the markers attached to the plastic models: they wore out and slid slightly over time, and issues with tracking were caused by both the



reflectivity of the sticky tape and the small size of the markers. Also the markers were only detectable at small displacements from the camera's position. This problem would be reduced with the use of larger models, however as the ARToolkit uses computer vision techniques the fiducial patterns attached to an object must be visible to the camera at all times for correct registration to be maintained. Both the rotation of the object beyond about  $45^\circ$  and occlusion, often by the user's hands, meant that the fiducial patterns were not always in view. Also models that better matched the shape of real anatomy cases would be advantageous.

## 4.4 Rapid Prototyping Overview

Rapid Prototyping (RP) is the practice of taking 3D virtual models (usually an stl file, a stereolithography file format for CAD software) and creating a physical equivalent. RP machines can use either additive or subtractive processes. Additive RP machines construct models layer by layer, which are then attached to each other by a process such as gluing, or fusion using a laser. Subtractive RP machines start with a block of material that is then cut to shape using a laser or a similar device. The models tend to be made from a plastic material, although some machines may use paper, cardboard or a metal. RP has been referred to as Automated Fabrication [48]. An Automated Fabrication system has been described by Burns [7] as one in which:

1. The process should take in raw material in some shapeless form such as blocks, sheets or a fluid, and produce solid objects with a definite shape.
2. The process must do this without a significant amount of human interaction.
3. The process must produce shapes with some degree of three-dimensional complexity. This criterion eliminates the forming of simple tubes or rods by extrusion and cutting or drilling of simple holes in sheet material.



4. The process must not involve the manufacture of new tools for each different shape to be generated (part specific tooling). This criterion eliminates all types of moulding and casting, EDM (Electrical Discharge Machining), die sinking and copy milling.
5. Each item produced must be a single object, not an assembly of component parts, thus eliminating joining operations such as gluing, welding and riveting.

RP is now becoming more affordable as hardware costs fall and because companies such as Ambler (<http://www.expressprototyping.com>), and Inition (<http://www.thinglab.co.uk>), provide a bureau service so that the purchase of specialised equipment is no longer necessary. RP models can be produced to faithfully reproduce anatomy segmented from CT and other medical data and can even use pumps to circulate fluid, thus mimicking blood flow and permitting contrast media injections, with realistic guidance using through-transmission of light or real fluoroscopy. Webb [77] and Gibson *et al.* [21] provide useful surveys of most current uses of RP to assist medical applications. In particular, they identify oral and maxillofacial surgery, orthopaedic applications, forensics, prosthesis development and tissue engineering. Despite these wide ranging uses of RP models, the use of RP models for general anatomy teaching has not been previously reported.

For our application a plastic RP model was deemed to be the most suitable option. The plastics used are quite robust and were therefore suitable for use in a classroom environment. A bureau service was used to produce our models. The bureau used a Z-Corp ZPrinter 450 (ZPrinter, Figure 4.2) that is capable of using several different materials for constructing models. The ZPrinter lays a fine layer of powder onto a build platform, onto which a binder is precisely applied in the shape of the object that is being constructed, in a manner almost identical to the way in which a 2D inkjet printer operates. The ZPrinter can achieve a resolution of 600dpi on each layer. Binders are available in a variety of colours as well as transparent, allowing the production of coloured models. Successive layers are then constructed

in the same way. Once all layers have been completed the loose powder is removed, leaving just the RP model. Powder that is not a part of the resulting model can be re-used in future constructions.



Figure 4.2: The Z-Corp ZPrinter 450 at Inition used to produce our models.

The use of an RP model has several potential advantages that we investigate in this research, in particular the use of patient specific data to highlight natural anatomical variations among different people, and for the on-screen renderings being shown to exactly match the object that the user is holding, and can allow hidden or ambiguous features to be understood [32]. This also allows the effects of various diseases to be shown. New variations and cases would be very easy to add as no additional tooling needs to be made — only data of the object of interest is needed. This project is the first to apply RP to anatomy education in this way.



## 4.5 Data Extraction

Our volume data in all cases was taken from patient medical scan data, however the scan type was different for each region; different scan types are suited to viewing different features. For example CT scans are ideally suited to detecting bone and calcifications within the body. MRI on the other hand is better suited to applications where a high contrast between soft tissues is required [80], including neurological and oncological imaging.

Regions of interest in the volume data were segmented using the ITKSnap software [79]. ITKSnap allows the use of both manual segmentation on a per slice basis, or a user supervised semi-automatic segmentation that uses a snake evolution algorithm to isolate features in a stack of pre-processed images. Two variations of the algorithm are provided that accept different pre-processing steps. The first variation creates a thresholded binary image from the source data to be supplied to the snake algorithm that then attempts to fill white regions in the image; the second variation uses an edge detection algorithm to create a binary image in which the snake expands towards image edges and no further. The snake parameters can be altered to alter the growth of the snake from user positioned seed points (a manual segmentation can be used as a starting point), allowing it to be adapted to the characteristics of the source image and the shape of the desired region. Once the segmentation is started it can be stopped by the user when it has converged to an acceptable solution. This solution can then be refined using a manual segmentation if necessary. The segmented region can be saved in a number of different formats, including the standard surface file type supported by RP printers.

## 4.6 Liver Model

Our first RP model was of a liver, taken from an anonymised abdominal CT scan that contained 8-bit voxels, scanned at a resolution of  $512 \times 512 \times 246$  voxels. The

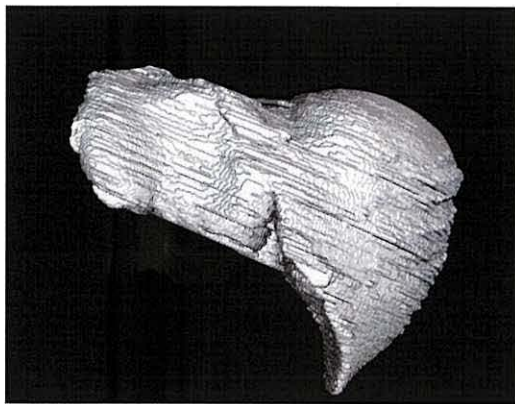


abdomen was segmented using the ITKSnap tool. An automatic segmentation was used to extract a general outline, which was then refined using ITKSnap's manual segmentation tool to compensate for areas where "leaking" had occurred due to neighbouring tissues having similar intensities in the CT data. Initially the file was large because of inconsistent "leaking" amongst neighbouring layers and variations in the manual segmentation operations, which lead to a bumpy surface (Figure 4.3(a)).

To reduce the bumpiness of the surface and the size of the STL file, a VTK-based program was created that performed a Gaussian convolution and then decimated the surface data (an operation that reduces the polygonal complexity of a surface model). Several iterations of each step were performed to reduce the size of the resulting STL file whilst adjusting the available parameters (both of the Gaussian function and the decimation function) to preserve the shape of the features of the liver.

Following this post-processing the surface file was sent to an RP bureau for production. The finished model (Figure 4.3(b)) weighed 192 grams and measured approximately  $100 \times 70 \times 70$ mm, which was appreciably smaller than the liver of a human adult would measure, and cost £202.13 + VAT. The small size and weight of the model meant that the model was easy to handle — a life size model would have been cumbersome and awkward to use during the development of BARETA. The RP model was also different in colour to a human liver, however in an AR environment this discrepancy shouldn't be important because the computer generated model should completely obscure the physical model.

The primary issue with the liver model was that no recesses had been created in the STL file into which a miniBIRD sensor could be located. This was compounded because the potential temporary solution of using sticky tape to locate the sensor failed as it did not sufficiently adhere to the surface of the liver model. This also prevented the use of the ARToolkit as its patterns would have required adhesion to the surface of the model. As such it was difficult to achieve consistent registration



(a) The liver surface model before post-processing



(b) The Liver RP model

Figure 4.3: Representations of the human liver

whilst using the liver model as an interface.

## 4.7 Ventricles Model

The ventricular system (Figure 4.4) is a particularly difficult part of the human anatomy to teach. Students have great difficulty in locating the ventricles whilst dissecting a cadaver; therefore the ventricles were an ideal candidate for our prototype AR anatomy teaching tool and allowed us to explore our hypothesis. The ventricles are located within the brain near the centre of the head and are responsible for the production of Cerebrospinal Fluid (CSF), which is distributed around the interior of the skull, and also to the the central canal of the spinal cord.

As with the liver model the ventricles model was segmented using the ITKSnap software. However in contrast to the liver model, MRI scan data was used. Two separate scans were taken, one of a cadaver head that had been embalmed three years previously with formaldehyde and phenol, and one of the head of a live human volunteer. Both subjects were scanned in Bangor's School of Psychology's Philips Achieva 3 Tesla MRI scanner, the first of its kind to be installed in the UK. The embalmed head was scanned at a resolution of  $448 \times 448 \times 320$  using 16-bit voxels with a voxel spacing of 0.5mm. The live subject was scanned at a resolution of



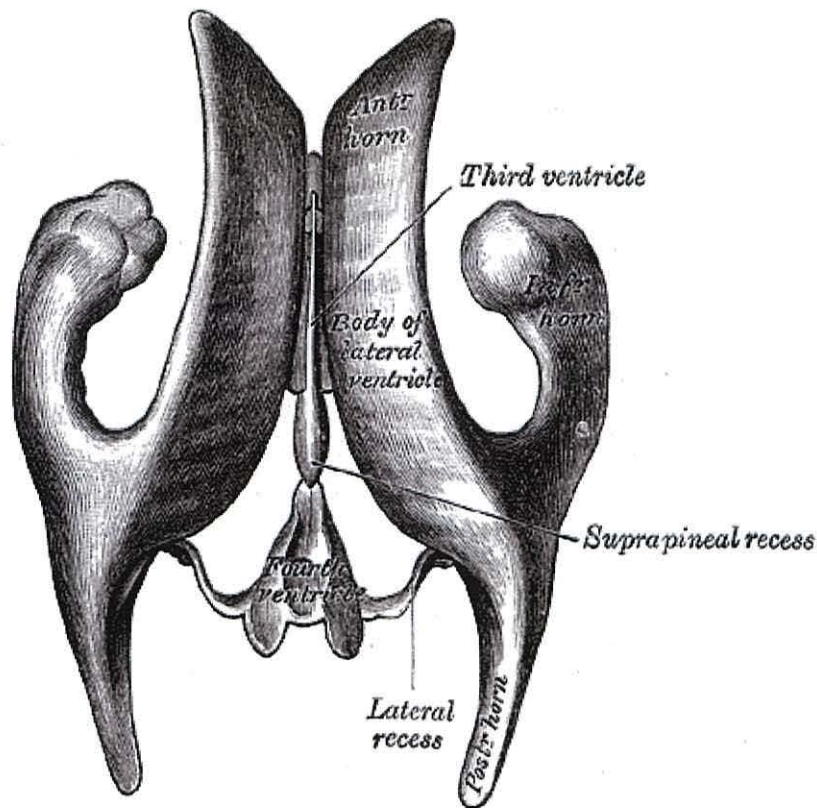


Figure 4.4: Drawing of a cast of the ventricular cavities, viewed from above, from the 20th U.S. edition of Gray's Anatomy of the Human Body, originally published in 1918.

384 × 384 × 220 using 16-bit voxels, each measuring 0.625 × 0.625 × 0.7mm. It was decided that the data of the live subject should be used to create our RP model; although the data of the embalmed head was of a higher resolution and had not suffered a loss of homogeneity, an appreciable contraction of the white and grey matter of the brain had occurred, exaggerating the size of the ventricles and also decreasing the separation between them.

The segmentation of the ventricles model was carried out in conjunction with an expert in anatomy and was therefore much more accurate than the liver model, and required less post-processing because of better delineation between the ventricles and neighbouring tissues in the source data (Figure 4.5). Following the segmentation the data was sent off to a bureau service to create a life-size model for use as an interface to BARETA (Figure 4.6(a)). The fourth ventricle was omitted from the RP model



because the cerebral aqueduct that joins the third ventricle to the fourth was too thin to be produced without a large probability of it breaking.

Once again the colouring of the RP model bore no resemblance to the colour of the anatomical feature, however as the model is present only as a tactile feature, with visual information being presented by computer generated imagery, the colour of the model was not considered to be important.

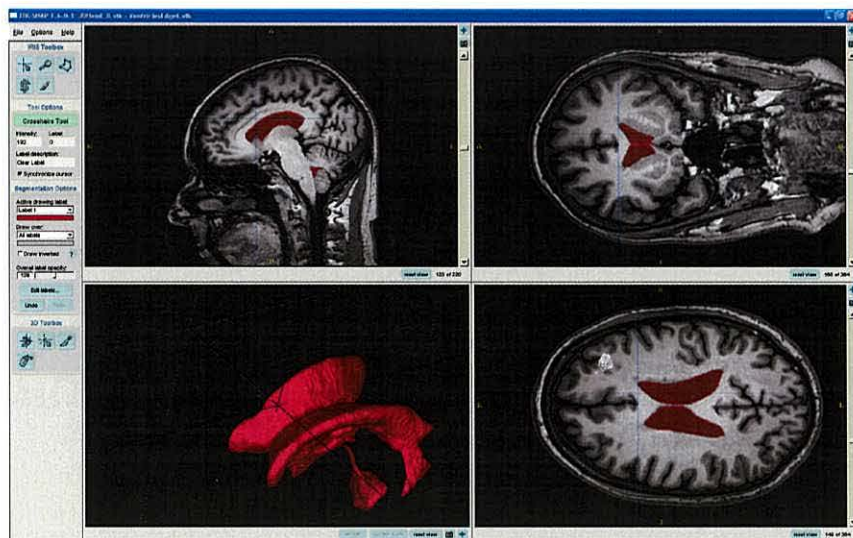


Figure 4.5: Using the ITKSnap software to segment the ventricles from the live subject data

A sensor for the miniBIRD was attached to the ventricles model between the two lateral ventricles (the larger structures) and the third ventricle so that it could be tracked (Figure 4.6(b)). In contrast to the liver model it was possible to attach the miniBIRD sensor despite no recesses having been created. The third ventricle is a thin region which allowed the sensor to be attached rigidly using a cable tie, aided by the narrowness of the gap between the two lateral ventricles. This allowed the position of the model to be continuously tracked, allowing on-screen graphics to follow its movements around the workspace. With the position of the sensor taken into account during calibration correct registration can be achieved.



(a) The ventricles RP model without a miniBIRD sensor



(b) The ventricles RP model with a miniBIRD sensor attached

Figure 4.6: The ventricles RP model

## 4.8 Further Props

An additional prop was used in conjunction with all of our primary tactile interfaces to manipulate a virtual clipping plane to interrogate the computer generated models. The benefit of having a tactile interface for both the clipping plane and the anatomy piece was that the user could intuitively co-ordinate the position of the clipping plane with the position of the anatomy model, so that the resulting clipping plane on-screen matched the location of the interface relative to the anatomy model. To our knowledge, this approach is novel to BARETA.

For the generic mannequin interface a second ARToolkit marker was attached to a long thin piece of cardboard. The centre of this plane was the centre of the ARToolkit marker, and the plane normal pointed straight up and away from the marker. For the versions of the software using RP models and the miniBIRD, a small square piece of paper attached to the end of a miniBIRD sensor was used to represent the clipping plane (Figure 4.7) to give the user an appropriate visual cue. The centre for this plane was the square surface at the end of the miniBIRD sensor, with the plane normal pointing opposite to the direction of the sensor cable.





Figure 4.7: The clipping plane prop

## 4.9 Summary

In this chapter the use of tangible interfaces has been discussed. Our first tangible interface using a generic anatomy mannequin and its associated AR application is described and evaluated. RP technology has also been described and the use of RP models constructed from medical data evaluated as a unique tangible interface to an AR environment. To create RP models from medical volume data first relevant data had to be segmented, then refined and converted into surface data. Such a process can be repeated on further data sets to create a series of models that can be used to highlight natural variations in anatomy, as well as example cases of disease. The use of further tangible interfaces as a novel data interrogation method is also discussed.



## **Chapter 5**

# **The BARETA Anatomy Teaching Assistant**

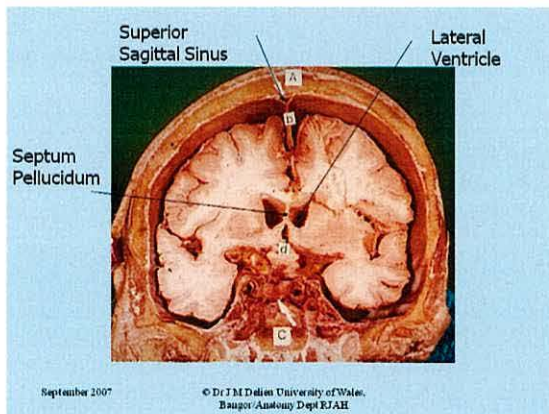
## 5.1 Introduction

In this chapter the development of novel anatomy teaching assistants from the Augmented Reality (AR) platforms discussed in Chapter 3 is considered. Section 5.2 describes one way in which the anatomy of the brain ventricles is currently taught. Section 5.3 discusses our first AR teaching assistant that used the ARToolkit as its tracking system. Section 5.4 describes an early version of BARETA in which an anatomically correct RP model derived from CT data of a liver is used in conjunction with a magnetic tracking system as a novel interface to an AR environment. Following the development of this version of BARETA a clear distinction between the reusable components of BARETA (the framework) and the content was made. The first version of BARETA that used the BARETA framework is described in Section 5.5. This version of BARETA used an anatomically correct brain ventricles RP model derived from an MRI scan of a live subject. This version of BARETA was subsequently improved, and is described in Section 5.6. A summary of this chapter is presented in Section 5.7.

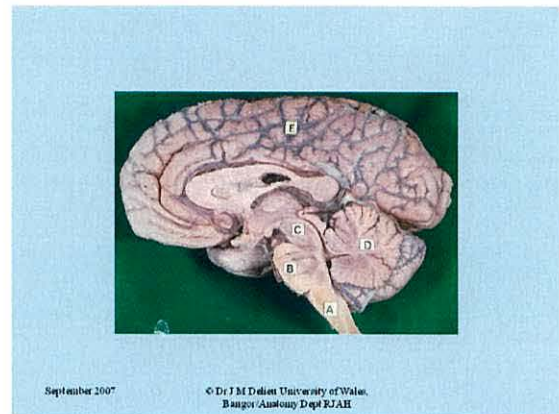
## 5.2 Current Anatomy Lessons

Medical students have to study the whole of the human anatomy during their training. After consultation with Doctor John Delieu, the Anatomy Lecturer with whom we have been collaborating, it was decided to produce an example BARETA lesson based on the human brain. Currently several methods are used to teach students about the ventricles within the human brain. A typical lesson might consist of a lecture using Powerpoint slides (such as those shown in Figure 5.1), followed by exposure to cadavers or prosections. The aim of the lecture is to explain to the students the layout and the function of the ventricular system within the brain prior to performing the subsequent dissection. As a cadaver/prosection (i.e. brain) can only be dissected once, it is important that students have a basic knowledge of the

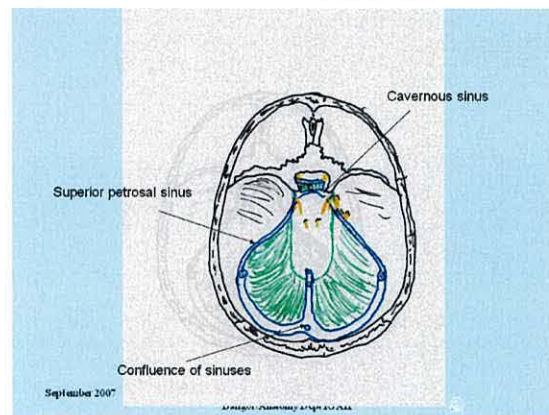
anatomy that they are to dissect.



(a) A slide showing a coronal section of the human head



(b) A slide showing a sagittal section of the human head



(c) A slide containing a diagram of a horizontal section of the human head

Figure 5.1: Lecture slides used to teach brain anatomy at Bangor's School of Health-care Sciences

A difficulty that arises with this type of lesson is that the lecture slides are 2D, whereas the human head is 3D, therefore it can be difficult to explain exactly where the ventricles can be found in preparation for exposure to human material: material is often presented as images of slices through a cadaver head, hand-drawn diagrams of regions of interest, and text descriptions. The ventricular system is a difficult concept for students to envisage, therefore the location of the ventricles within the human head was an important relationship that any potential teaching assistant would have to present to students. Having important features highlighted in a teaching assistant would also be beneficial as knowledge of the locations of



these could be useful if a student is to be guided through a dissection by verbal instruction.

## 5.3 ARToolkit-based Teaching Assistant

The first anatomy teaching assistant implementation was a simple system that used the ARToolkit to perform tracking functions, and the Visualization Toolkit (VTK) to perform the rendering of the graphical augmentations. The use of the ARToolkit allowed for the rapid creation of a proof of concept system that could be used to try out various ideas. Similarly the use of VTK removed the need to create bespoke rendering software components, and provided many different rendering options that could be changed and evaluated rapidly.

As discussed in Section 4.3 this initial implementation of the anatomy teaching assistant used parts of a small anatomy mannequin prepared with ARToolkit markers as an interface to manipulate the graphical augmentations. On top of each ARToolkit pattern visible in the video stream an organ was volume rendered by VTK, using the *vtkVolumeRayCastMapper* class that implements volume ray cast volume rendering.

As this teaching assistant implementation was derived from an ARToolkit sample program the program functioned in a similar way to the sample programs. The OpenGL Utility Toolkit (GLUT) was responsible for window management and event handling functions and the ARToolkit was responsible for capturing images from the webcam and deriving tracking information from them. VTK was required to draw into the window created by GLUT. VTK provides a method for achieving this but note that its capability to erase the background of the window must be disabled to ensure that the webcam image remains visible in the background. VTK is then able to volume render on top of the ARToolkit marker present in the webcam image.

In any AR environment, rendering a virtual object so that it appears at the

correct depth within the scene is critical to a user's perceived immersion within the environment. To achieve realistic depth of virtual objects the ARToolkit provides a camera calibration utility that can determine the intrinsic parameters of a camera (parameters relating to the lens) that provides a transformation matrix that can be used to alter the characteristics of the virtual camera so that it behaves like a physical camera; the ARToolkit also provides a function (the *argInit* function) to pass this matrix to OpenGL; this function was used for the implementation of this teaching assistant. This ensures that the perceived depth of the virtual object is correct, as long as no real object is between the ARToolkit marker and the physical camera. Occlusion is also a problem when real object does not occlude the ARToolkit marker, but ought to occlude the virtual object. The ARToolkit cannot detect this type of occlusion as no knowledge of the environment is stored by the ARToolkit: it only has knowledge of the markers that are expected to be present. Therefore this teaching assistant also cannot detect this type of occlusion.

The volume data rendered by VTK in this implementation was taken from an anonymised CT scan of a human abdomen (converted into a raw file format from the original DICOM data, a standard used by medical scanners), that was segmented using the ITKSnap software [79]. Several organs were visible in the data set, including the heart, the liver and the kidneys, as well as the ribs and a section of the spine. The two organs that were used from this data set were the liver and the heart, although the heart was incomplete. The segmentation data produced by ITKSnap was used to produce new volume data sets containing only the required organ, which were then clipped using the smallest bounding box that entirely contained the organ, using a simple Java application created specifically for this purpose. This not only increased the speed of rendering, it also meant that the centre of the data set would be within the organ, as well as the centre of rotation.

Following this each raw data set was converted into the *.vtk* file format using



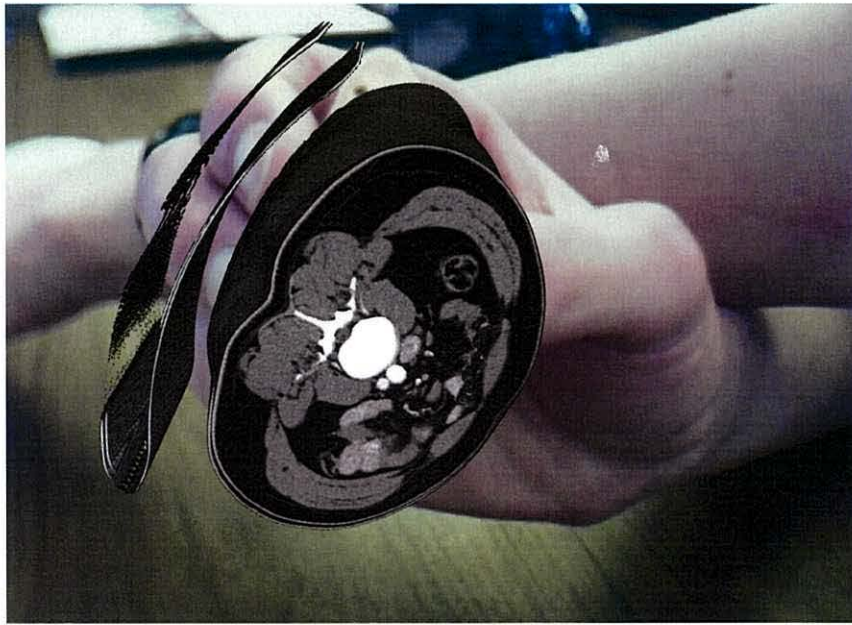


Figure 5.2: A lung of the anatomy mannequin augmented with abdomen CT data.

Erik Vidholm’s conversion utility<sup>1</sup> so that it could be read by VTK. This data was then used to augment the video stream containing a view of the real world. As this was only an experimental system not all of the augmentations used organ data that matched the physical model that could be seen on screen. In some cases where no matching data was available the entirety of the volume data was used (Figure 5.2). In all cases precise registration could not be achieved because the shape of the organ model did not match the shape of the data set being visualised.

To prevent the air in the volume data from being displayed on-screen, and allowing the organ of interest to be viewed, a piecewise transfer function was utilised during volume rendering. All voxels with a value of less than 2 were made fully transparent, whilst the remaining voxels were made completely opaque (see Section 2.2.3 and Figure 2.4), showing the exterior of the volume data.

This version of our system allowed the user to manipulate a clipping plane or a parallel opposing pair of clipping planes (slab rendering) using a second physical object that had also been prepared with an ARToolkit marker. To implement this

---

<sup>1</sup>The source code for Erik Vidholm’s raw to VTK conversion utility can be downloaded from <http://www.cb.uu.se/~erik/vtk/rawToVTK.cpp>



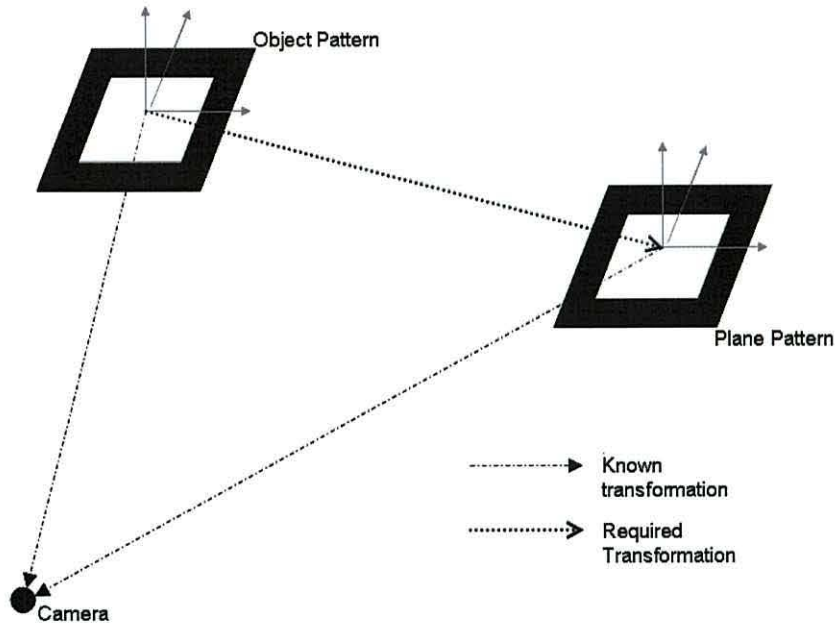


Figure 5.3: The known and required transformations to create a clipping plane.

a transformation between the two visible ARToolkit markers was required, as VTK required that a plane be specified relative to the volume's co-ordinate system, and not the camera or the plane pattern co-ordinate systems. Using our matrix for the plane and the matrix for the current volume transformation, the origin for the plane can be derived using simple vector geometry as shown in Figure 5.3.

To create a plane the normal vector of the plane must also be determined in the object's co-ordinate frame. The normalised normal of an ARToolkit marker points in the direction  $(0, 0, 1)$  in the marker's co-ordinate frame. The direction of this normal in the object's co-ordinate frame can be discovered by rotating the translation vector  $(0, 0, 1)$  by the rotation matrix describing the rotation of the ARToolkit plane pattern in the object pattern's co-ordinate frame. Using  $4 \times 4$  homogeneous transformation matrices this can be expressed as:

$$\begin{bmatrix} a & e & i & m \\ b & f & j & n \\ c & g & k & o \\ d & h & l & p \end{bmatrix} * \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a & e & i & i+m \\ b & f & j & j+n \\ c & g & k & k+o \\ d & h & l & l+p \end{bmatrix}$$

As d, h, l, m, n and o are all zero, and  $p = 1$

$$\begin{bmatrix} a & e & i & i+m \\ b & f & j & j+n \\ c & g & k & k+o \\ d & h & l & l+p \end{bmatrix} = \begin{bmatrix} a & e & i & i \\ b & f & j & j \\ c & g & k & k \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

This result meant that we do not have to multiply any further matrices to discover the normal, the values  $i$ ,  $j$ , and  $k$  of the matrix for the plane pattern in the object pattern's co-ordinate frame could be used to define the normalised normal of the plane.

## 5.4 BARETA Liver Teaching Assistant

The BARETA liver teaching assistant used the same CT data that had been used in the previous ARToolkit-based teaching assistant, however in this case its movements were controlled by the movements of a sensor from an Ascension miniBIRD 800 system (miniBIRD), using the tracking library discussed in Section 3.4.2. The intention of the miniBIRD was to track the position of a physical representation of the same liver model as was being viewed on-screen (Section 4.6). As well as having a more realistic physical representation of the liver, this implementation also allowed a greater freedom to move and rotate the physical model because unlike the ARToolkit-based implementation, a line of sight between the webcam and the tangible interface was not required. The hardware set up was similar to that described in Section 3.4, however, at this stage, the two brackets had not yet been constructed.

Once again volume rendering was performed by VTK, using 3D texture-mapped volume rendering for performance reasons (provided by the *vtkVolumeTextureMapper3D* class), and used the same transfer function as the ARToolkit-based teaching assistant. Unlike the ARToolkit-based implementation the video stream was captured using VTK. VTK also controlled the window management and event handling.

The use of VTK's event handling allowed a much larger range of event options.

The GLUT event handling system used for the ARToolkit-based implementation used a single polling loop and a single function for handling keyboard events. The VTK event handling system allows the application programmer to create callbacks that can be triggered by a variety of events, including timed events (used in this case to capture frames from the webcam at regular intervals), keyboard events that can be used to edit parameters at runtime, and mouse events that by default in all VTK programs controls the size and position of objects onscreen, and also the zooming of the virtual camera.

As VTK was receiving and rendering the video stream from the webcam its integration with the virtual content was implemented differently to how it was in the ARToolkit-based teaching assistant. Rather than having VTK render into a pre-existing window, VTK created its own window into which content would be rendered. Two separate instances of *vtkRenderer* were used: one for the video stream and one for the virtual content. The two instances of *vtkRenderer* were placed in separate layers of a *vtkRenderWindow*, and the erase capability of the *vtkRenderer* containing the virtual content was disabled to ensure that the video stream appeared in the background and that the virtual content appeared to be between the user and the video stream.

## 5.5 First BARETA Brain Ventricles Teaching Assistant

The first BARETA implementation to support teaching of brain anatomy, in particular the ventricles, built upon the software of the BARETA liver teaching assistant. At this stage of development a clear delineation between the BARETA framework (Section 3.4) and content was now evident. This means that the miniBIRD system is again used here, this time with a sensor attached to the RP model of the human ventricular system described in Section 4.7. The RP model was used to manipulate



renderings of a human head. The dataset used for the volume rendering was the same dataset that had been used to produce the RP model, allowing the volume rendering to be registered with the RP model. Rather than rendering only the segmented region as in the BARETA liver teaching assistant, the entire head was volume rendered.

In this case, as well as visualising the volume data of the human head, the surface model of the ventricles that we had used to create the RP model was also visualised. The surface rendering was also performed by VTK. The surface was rendered in such a way that it was registered with the ventricles present within the volume rendering. The volume data and surface data were placed into a *vtkAssembly*<sup>2</sup>, which allowed both representations to be manipulated using a single call to a transformation method. This also ensured that correct registration was retained when transformations changed.

To ensure that registration was correct between the RP model and the on-screen rendering, two transformations were applied to the *vtkAssembly*. Firstly the pose of the centre of rotation of the *vtkAssembly* was changed to reflect the position of the sensor within the RP model. The second transformation scaled the *vtkAssembly* to match the size of the RP model on-screen.

Using the same transfer function as was used in the BARETA liver teaching assistant, the surface model of the ventricles cannot be seen as it is obscured by the outer voxels of the volume rendering of the head. Therefore, in order to see the ventricles, the way in which the volume was viewed had to be changed. This was initially accomplished by using a different transfer function, one that reduced the opacity of the volume rendering to a level at which the ventricles could be seen, yet still allowing the volume rendering to remain visible (Figure 5.4), thus allowing the

---

<sup>2</sup>*vtkAssembly* is a class in VTK that is derived from *vtkProp3D* that can contain several instances of *vtkProp3D*, including additional instances of *vtkAssembly*, and allows them to be manipulated as though they were all a single *vtkProp3D*. Both the volume and surface data are stored within subclasses of *vtkProp3D* allowing them to be stored in a *vtkAssembly*.



Figure 5.4: The ventricles surface model within the volume rendering of the head are exposed using transparency.

position of the ventricles within the head to be seen. A transfer function could have been chosen to render the volume in a variety of colours that could have enhanced certain details (Figure 5.5), however a simple greyscale function was chosen so that the appearance of the volume rendering would match the colouring that would be encountered when viewing the MRI data as slices, as is currently the case in medical diagnosis. This also allowed for a large contrast between the volume rendering and the red colour chosen for the surface rendering of the ventricles.

An alternative approach to viewing the volume in a way that allowed the ventricles to be seen was by using clipping plane and slab rendering features; this allowed the surface rendering of the ventricles to be seen within the volume rendering, demonstrating their location relative to the surrounding tissue. These features were closely related to the clipping plane and slab rendering implemented in the ARToolkit-based teaching assistant, however minor changes were required to adapt



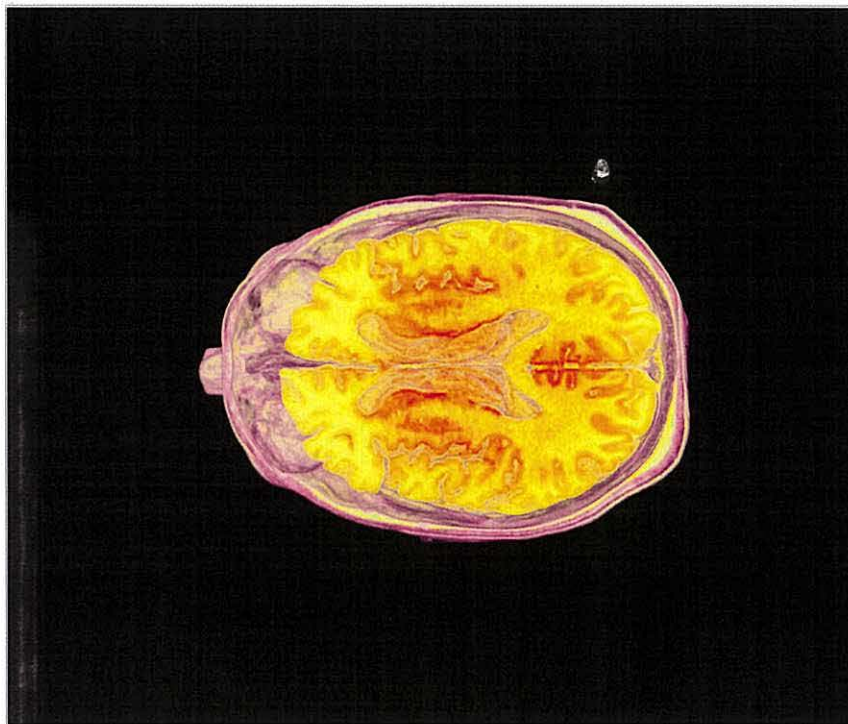


Figure 5.5: A volume rendering of an MRI scan of the human head with a transfer function applied to produce pseudo colour.

them to work with this implementation of BARETA.

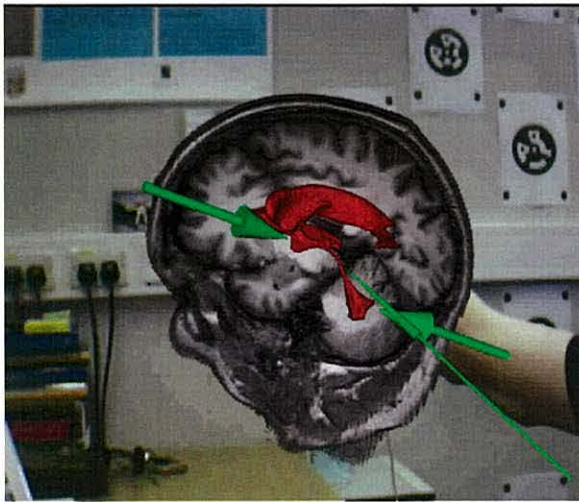
The inclusion of the ventricles surface model within the whole head volume rendering is useful as a reference to the location of the ventricles, however the ventricles have many detail features that are important for students of anatomy to learn, which may not be obvious from this rendering alone. To help students to identify key features of the ventricles arrows were used (created using the *vtkArrowSource* class); these were coloured green to contrast the greyscale of the volume rendering and the red of the ventricles. Attached to each arrow was a 3D text label (created using the *vtkVectorText* class that models the text as a surface) naming the feature to which the arrow was pointing. The arrows and labels remained static relative to the volume rendering as they were also included in the *vtkAssembly* that contained the volume and surface data. In this implementation of BARETA three items were annotated. The three items were: the third ventricle, the fourth ventricle and the cerebral aqueduct.



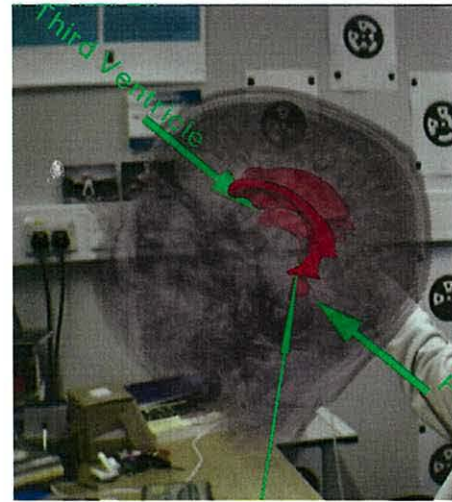
VTK does include a billboarding function that could ensure that the text labels were always facing the camera, the *vtkFollower*. The *vtkFollower* works by observing the transformation of an object and altering the rotation to ensure that it always faces the camera. The possibility of using the *vtkFollower* in BARETA was investigated as a way of making the text labels easier to read, however it became apparent that it would not function as BARETA required, as a *vtkAssembly* was being used. This meant that the final rotation of a label was always the result of at least two rotations (in one case it was the result of three). The *vtkFollower* worked when the interface was rotated about a single axis, however when a rotation about a second axis was also introduced the results became unstable. The implementation of a variation of the *vtkFollower* that could read multiple rotations was investigated but the value added by its inclusion was not great as other factors also influenced how easy the labels were to read, including occlusion by the volume and surface renderings and the labels being moved off-screen by the movements of the RP interface.

From the aforementioned features our anatomy teaching assistant aid was constructed. The anatomy teaching assistant consisted of three different steps, each providing the user with a different way of viewing the ventricular system. During each of the three steps the clipping and slab rendering features could be enabled and disabled by the user by a single keystroke. Transitions to successive steps were activated using a press of the space bar. The steps were as follows:

1. The volume is rendered using a transfer function that makes the entire head opaque, and the surrounding air transparent. Arrows are visible, however the associated labels are not (Figure 5.6(a)).
2. The volume rendering gradually becomes more transparent, allowing the surface rendering of the ventricles to be viewed within. Labels associated with the visible arrows are made visible (Figure 5.6(b)).
3. The volume is made completely transparent, allowing the student to view the



(a) Step 1 of the ventricles teaching assistant showing the clipping plane in use, exposing the ventricles surface model.



(b) Step 2 of the ventricles lesson showing the transparent volume rendering that allows the ventricles surface model to be seen

Figure 5.6: Steps one and two of the brain ventricles lesson

surface rendering, arrows and labels.

This implementation of BARETA allowed the student to explore the head anatomy and the ventricles at their own pace. The only restrictions placed on the user's movements of the RP model were those imposed by the tracking devices. To evaluate this version of BARETA we carried out two user studies, the results of which are presented in Sections 6.2 and 6.3.

## 5.6 Second BARETA Brain Ventricles Teaching Assistant

The second BARETA ventricles teaching assistant was a refinement of the first, using the same volume and surface data, and the same RP model. The volume and surface rendering capabilities were retained. The labelling was however modified to enhance the readability of the text labels, and to reduce visual clutter, especially since eight features were now labelled. The eight features labelled were: the third ventricle, the fourth ventricle, the anterior horn, the inferior horn, the cerebral aqueduct, the



posterior horn, the right lateral ventricle and the collateral trigone.

The inclusion of labels for eight features, if displayed in the same way as in the previous version of the BARETA teaching assistant, would lead to a large amount of visual clutter and a strong possibility of labels occluding each other, and also the features that are being annotated. Any suitable solution would not be trivial, both in the time required to implement and runtime complexity. This problem can only get worse as more arrows and labels are added. Several solutions to the label occlusion problem have been proposed, including that by Azuma and Furmanski [3]; however such methods are not best suited to arbitrarily positioned labels. In addition to ensuring that the labels will not occlude other objects, the labels should be positioned and oriented in such a way that they are always easily visible to the user. In the previous version of BARETA it was possible for labels to appear at any angle, including backwards. This meant that labels were not always easily readable.

Instead of having several arrows and labels present at the same time it was decided that only one arrow and its corresponding label should be displayed at any one time. In this improved version of BARETA the label was permanently positioned at the bottom left-hand corner of the screen, and therefore would always be visible to the student (Figure 5.7).

As only one arrow and one label were visible at any given time only one arrow needed to be placed into the *vtkAssembly* containing the volume and surface data. When the user decided that they wanted a different region to be annotated the transformation of this single arrow would be changed, rather replacing the arrow that was already in the *vtkAssembly* with another, or making the first arrow invisible and the second arrow visible. The change of arrow position and text label was activated by the user pressing a number key in the range 1-8. The transformations to point the arrow at the correct feature were stored in an array. Similarly only one text label was displayed at any one time (with text for each feature stored in an array), therefore only one *vtkTextActor* was included in the VTK pipeline, and all of the



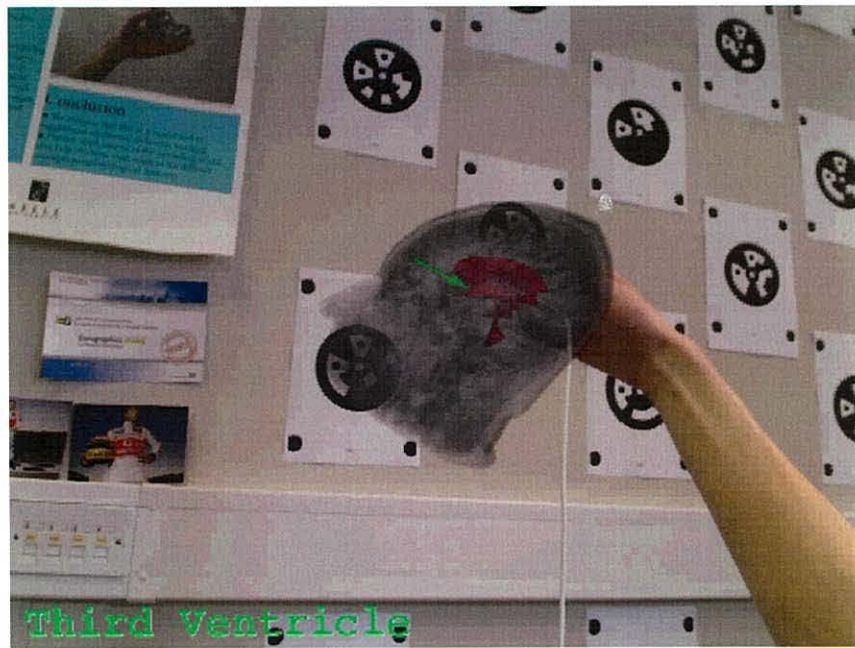


Figure 5.7: The second ventricles teaching assistant running, showing the location of the third ventricle.

necessary text was stored in an array. *vtkTextActor* renders the text as a 2D object in the application window, in contrast to the *vtkVectorText* that produces 3D text represented by a surface that was used in the previous version of BARETA.

As with the previous version of BARETA the transitions between steps in the teaching assistant were activated using the space bar. The steps were the same as in the previous version of BARETA. This version of BARETA was also the subject of a user study, the results from which can be found in Section 6.4.

## 5.7 Summary

In this chapter the creation of novel augmented reality applications for assisting anatomy lessons has been described. The first implementation described used the ARToolkit, creating a simple and inexpensive system that displayed real patient data augmented on a video stream in which organs from a plastic anatomy mannequin were visible. Subsequent sections described how the BARETA teaching assistants were created using the technologies and processes described in Chapters 3 and 4,

in the first case using liver data from a CT scan of a human abdomen, and then in the second case using ventricles data from a high resolution MRI scan of the human brain. This version of BARETA allows users to view the ventricular system from any angle and to have arrows pointing at important features, with related text also appearing on screen. A volume rendering of the human head can be viewed at a variety of opacity values, allowing a surface rendering of the brain ventricles to be seen. The ventricles can also be exposed using the clipping plane and slab rendering tools, controlled using a second tangible interface.

The BARETA teaching assistant successfully demonstrates to students the location of the ventricles within the human head using transparency and an arbitrarily positionable clipping plane. Features of importance within the ventricular system are highlighted using arrows and text annotations. The BARETA teaching assistant also provides an effective platform onto which further learning material can be added, and may in future be an effective replacement for cadaver dissection. One advantage that BARETA has over cadaver dissection is that it can be used repeatedly, and without supervision, allowing it to be used as an exam revision aid as well as an introduction to the ventricular system. Students only have a very limited time in the dissection lab, whereas time available in a computing or resource room has no such restrictions.

# Chapter 6

## User Evaluation



## 6.1 Introduction

In this chapter user studies evaluating the the BARETA teaching assistant are described and the results presented. Three user studies were conducted (with radiology students, sixth form pupils, and medical students learning anatomy), where students that fit into our intended end user group would view and use BARETA, then fill out a questionnaire to assess BARETA's usefulness as a teaching aid together with its ease of use. Following each user evaluation, BARETA was evaluated and improved where necessary.

During each evaluation the layout of BARETA's components was different. This was a result of the space that had been allocated for BARETA to be set up in. BARETA is flexible enough to allow for these differences, as long as the components of the two tracking systems are within operating range of each other. As the positioning of fiducial markers for the VisTracker is a time consuming process a poster was created with several precisely located fiducials attached to allow for rapid deployment of BARETA.

## 6.2 First User Evaluation

We conducted our first user evaluation at Bangor University's School of Healthcare Sciences in Wrexham. Twenty second-year Radiology students took part in the evaluation, five of whom were male, the remaining fifteen were female. The participants covered a large age range, although over half of them were aged between eighteen and twenty-one, as is to be expected for students on an undergraduate course. The user evaluation was conducted during the students' introduction to the ventricular system, therefore allowing a comparison between BARETA and other teaching methods. This group of students was selected because the lecture on the ventricular system during the session was presented by Doctor John Delieu, our collaborator from the School of Healthcare Sciences. His input meant that areas highlighted by

BARETA were relevant to the students. Whilst the main aim of this evaluation was to evaluate how BARETA compared with other methods of teaching anatomy, we also sought opinions on usability issues.

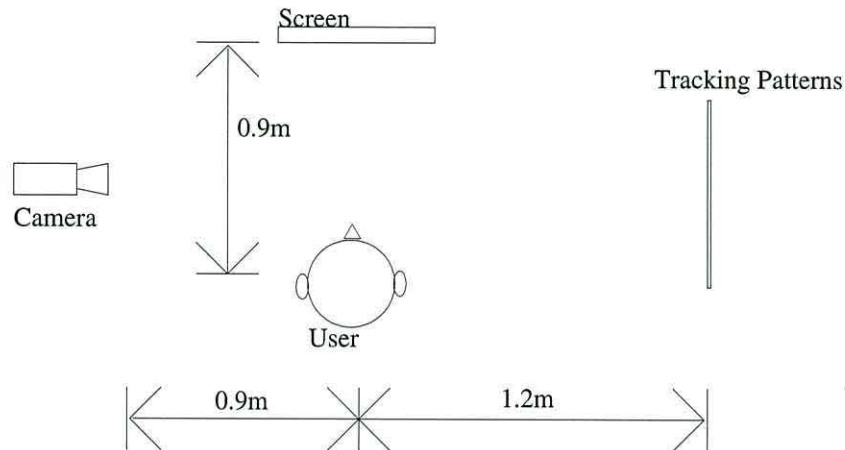


Figure 6.1: The experimental setup used in Wrexham (not to scale)

Each of the students present took part in a classroom lesson that described the ventricular system in which the lecturer used slides and diagrams to illustrate different features (Figure 6.2). Following the lecture, students were provided with various materials to research more information about the ventricles. They had access to textbooks, the Internet, and our BARETA system in the configuration shown in Figure 6.1. The user in this case was seated, with the camera located to their left and in an elevated position so that only the top of their head was visible on the screen. The camera was left at this static position to simplify user interaction. The version of BARETA used was that described in Section 5.5, without the clipping plane and slab rendering features (these features had not been implemented at this time).

At the end of the session, the students were asked to fill in a questionnaire that posed questions both on the BARETA system alone, and in comparison with the teaching media presented to them on the same day. The questionnaire presented used a five point Likert scale. For analysis purposes we subsequently assigned a value of five to the response “strongly agree”, four points to “agree”, three points

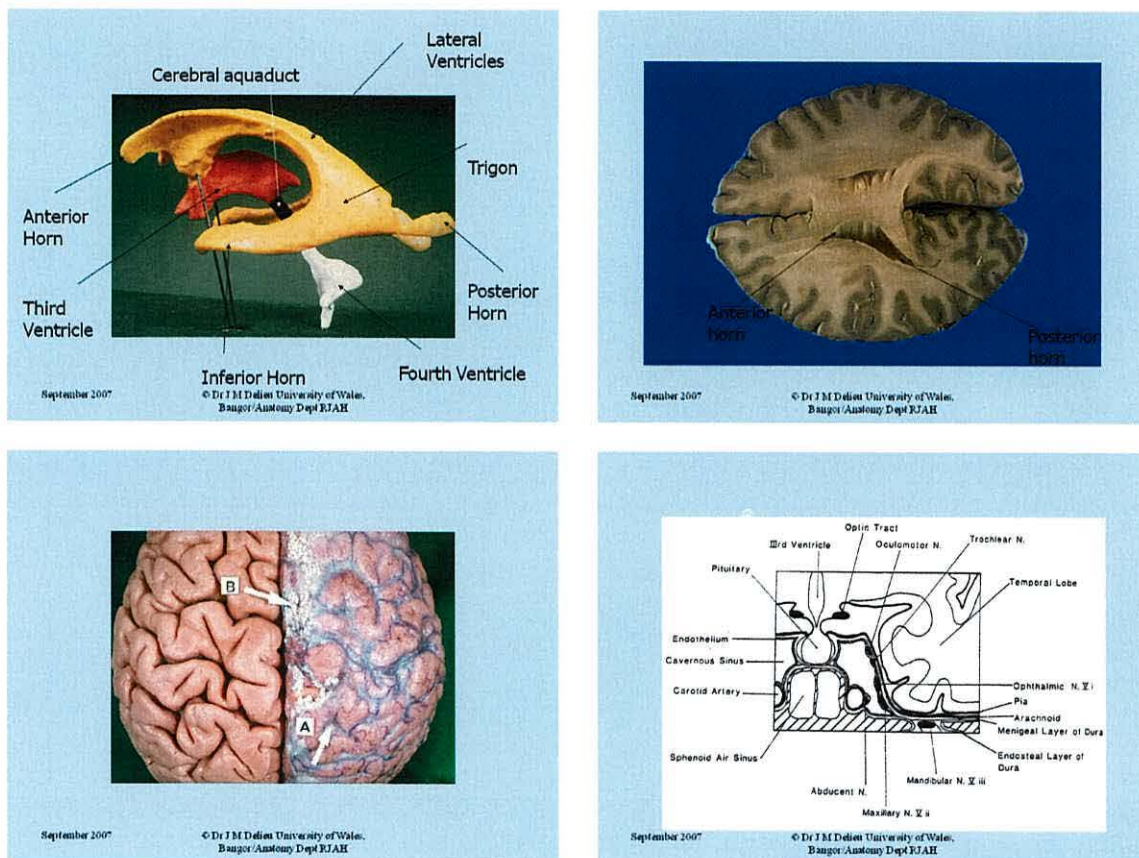


Figure 6.2: A selection of slides used at the Bangor University School of Healthcare Sciences, Wrexham

to “neutral”, two points to “disagree”, and one point to “strongly disagree”. Also included was space for the students to record additional comments. Our results are summarised in Figure 6.3.

The questions were as follows:

1. Have you previously considered that access to virtual material may be beneficial to you?
2. The system was straightforward to use
3. The system conveyed information effectively
4. The transparency feature was useful
5. The arrows and labels were well positioned



6. Using the plastic model as an interface was more intuitive than using a mouse and keyboard to move the onscreen image
7. The plastic model helped me to understand the shape of the ventricular system.
8. The on-screen representation of the ventricles followed the movements of the plastic model
9. The system helped me to understand the shape and location of the ventricular system
10. I would like to use this system to learn other anatomical features
11. The system helped me to understand the ventricular system better than the Powerpoint slides/lecture
12. The system helped me to understand the ventricular system better than the textbook
13. The system helped me to understand the ventricular system better than the plastic models
14. The system helped me to understand the ventricular system better than viewing a cadaver
15. The system helped me to understand the ventricular system better than the anatomy website

One of the most important questions in providing feedback on the research hypothesis of this thesis was question 9, which asked the students whether they agreed that BARETA helped them to understand the shape and the location of the ventricular system, as this relates directly to our hypothesis. We found that all of the students either agreed or strongly agreed that BARETA helped them to understand the shape and location of the ventricular system.

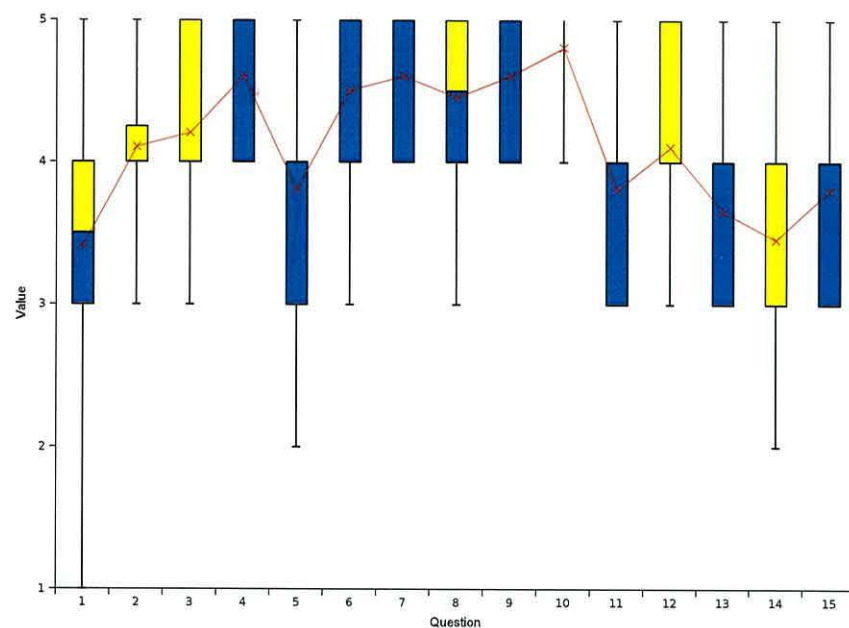


Figure 6.3: Wrexham questionnaire results. Yellow represents the second quartile, and blue the third. The red line represents the mean answer.

For BARETA to be an effective learning aid it must be at least as good as other methods of teaching anatomy. From the chart we can infer that BARETA helps the group of students to understand the ventricular system better than the textbook, the lecture, the standard plastic anatomy models and the anatomy website (Questions 11, 12, 13 and 15). The mean response to each question was greater than 3 (“neutral”) in each case, with no student disagreeing. One student commented that BARETA was “much better than learning from Powerpoint.” Students were less sure about the increase in understanding over the viewing of a cadaver (Question 14). Although the mean was higher than 3 (3.58), it was the lowest mean of all of the comparison questions; several students disagreed that using BARETA helped them to understand the ventricular system better than viewing a cadaver. This question also showed the greatest standard deviation of students opinions (Figure 6.4), suggesting that there was genuine uncertainty as to whether BARETA was more useful than cadaver dissection. Question 14 was the only question for which the standard deviation was greater than 1, suggesting that the majority of students felt the same way about many aspects of BARETA.

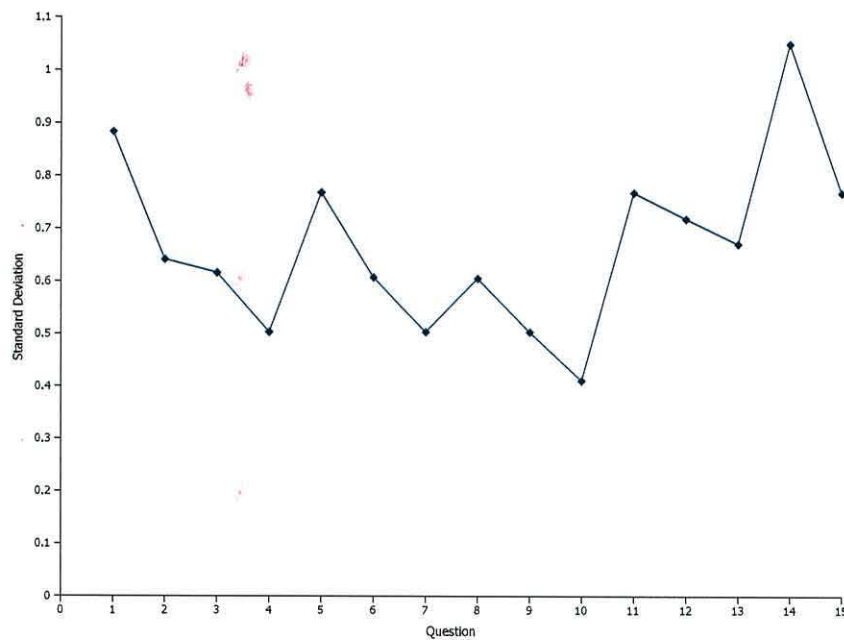


Figure 6.4: Wrexham user evaluation standard deviation.

As well as comparing our system to other educational methods, we also wanted to find out how intuitive to use the students found BARETA to be. Several questions on the questionnaire asked the students about how useful and usable they found various aspects of BARETA. Many of the students found that the system was straightforward to use (Question 2), and more straightforward than using a keyboard and mouse to control the visualisation (Question 6). The students also found that using the RP model as an interface helped them to understand the shape of the ventricular system (Question 7).

One novel feature of BARETA that we wanted to evaluate was the transparency feature. The students found the transparency feature of BARETA useful (Question 4). Every student either agreed or strongly agreed that this was the case.

The results from this user evaluation did reveal a few areas that required improvement. Although many of the students thought that the arrows and labels were well positioned several of the students felt that the labels could be difficult to read and were not well positioned (Question 5). One student commented that the “labels were too large and difficult to see on the screen at the same time.”



The results show that not all of the students had considered that the use of virtual material could be helpful to their education (Question 1); however students would like to use such a system again to learn other features of anatomy (Question 10), with a large majority of students strongly agreeing that this was the case; the remaining students agreed. This result is useful because it demonstrates that the students felt that BARETA could benefit their education.

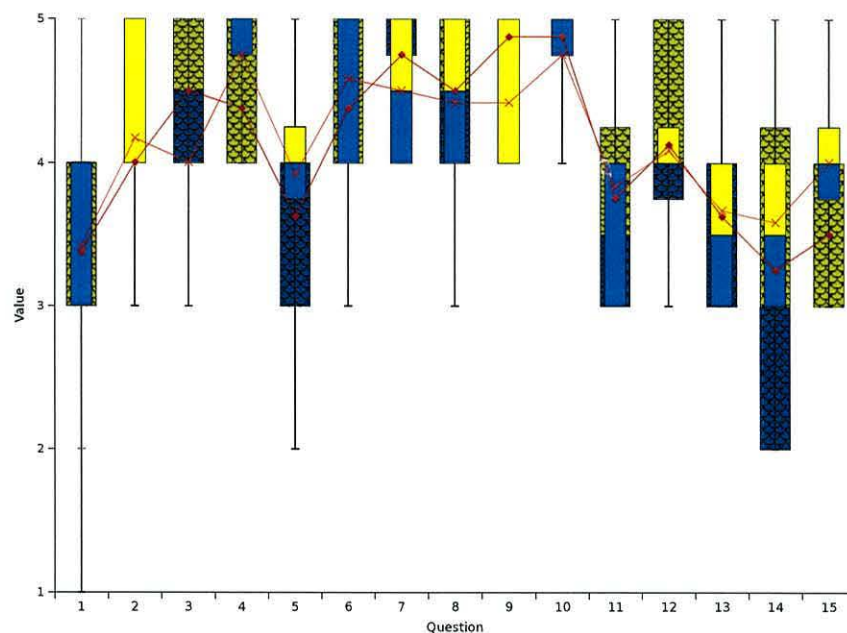


Figure 6.5: Comparing the opinions of two different age groups. The lighter un-patterned areas represent the students aged between 18 and 21, whilst the darker patterned areas represent the students aged 22 or over. The red lines represent the mean answers for each group, with the crosses representing the younger age group, and the diamonds the older. The black error bars relate to the younger age group, and the grey to the older.

We were interested to see if the age of a student had an effect on their perception of BARETA, and its merits relative to other educational methods. To study this we analysed the results of the user evaluation when partitioned into two age groups. The first group comprised all of the students aged between 18 and 21, and the second group comprised those students aged 22 or more. Although the sample sizes of the two groups is small (only 8 students were aged 22 or above) it can be seen that there is little variation between the two groups (Figure 6.5).

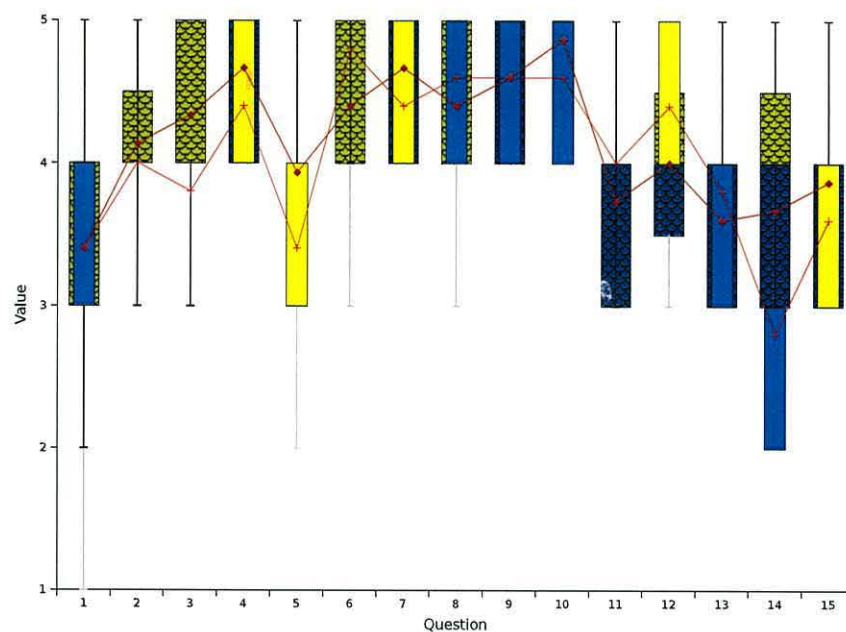


Figure 6.6: Comparing the opinions of the two genders. Males are represented by the lighter areas, and females by the darker patterned areas. The red lines represent the mean answers for each group, with crosses representing the males, and the diamonds the females. The black error bars relate to the male students, and the grey to the female students.

We also wanted to see if there was a difference in opinion between male and female students. Again the sample sizes were too small to make any definitive conclusions (only 5 males took part in this user evaluation), however it can be seen that although there is some variation between the two the general trends are the same (Figure 6.6). The largest variation can be seen in the responses to Question 14 which asked students how BARETA compared with cadaver dissection.

This user evaluation provided us with some useful user feedback, however, more opinions were needed in order to draw any proper conclusions. Also some new features were added to BARETA after this user evaluation, therefore another user evaluation was required both to gather further data on some aspects and new opinions on additional functionality.

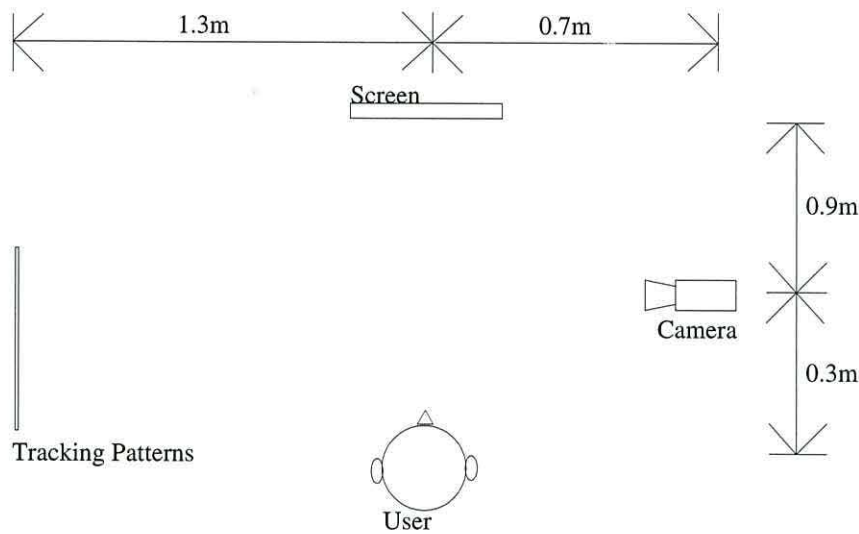


Figure 6.7: The experimental setup used at Connah's Quay High School (not to scale)

### 6.3 Second User Evaluation

The second user evaluation was carried out at Connah's Quay High School. Thirty-eight students took part in the evaluation, and were either in year ten (aged fourteen or fifteen), or in year thirteen (aged seventeen or eighteen). Nineteen of the participants were male, seventeen were female; the remaining two students did not specify a gender. The high-school setting allowed the opportunity to focus our questionnaire on usability issues, allowing us to find any weaknesses in this area prior to evaluating improvements in the educational content.

As in the first user evaluation, each student used BARETA, on this occasion in the configuration shown in Figures 6.7 and 6.8. The users at Connah's Quay High School viewed the BARETA environment whilst standing. The camera was located approximately at chest height and required the user to stand back to avoid obscuring the camera's view of the tracking patterns. As was the case at Wrexham, the camera remained static throughout the evaluation to simplify user interaction. In this case BARETA included the new options of using the clipping plane and slab rendering features described in Section 5.5 (Figure 6.9), then filled in a questionnaire.

The questionnaire presented at Connah's Quay was similar to that used during





Figure 6.8: The experimental setup used at Connah's Quay High School, showing the tracking patterns, the user and the camera. In this instance the user is facing the display device.

the first user evaluation, with some alterations made to remove questions irrelevant to the students and to investigate perceptions of BARETA's clipping plane and slab rendering functionality. As the participants in this evaluation were not studying the ventricular system as part of their curriculum, the questionnaire focused more on the usability of the system. Also included were some negatively worded questions to see if the wording of questions had introduced bias into a student's response.

Once again we assigned a value of five to the response "strongly agree", four points to "agree", three points to "neutral", two points to "disagree", and one point to "strongly disagree". Also included was space for the students to record additional comments. Our results are summarised in Figure 6.10.

The questions were as follows:

1. I have previously considered that access to virtual material may help me to learn new topics
2. The system was straightforward to use



Figure 6.9: A student using BARETA at Connah's Quay High School. The clipping plane facility is being used to view the surface rendering of the ventricles within the volume rendering of the entire head.

3. The system conveyed information effectively
4. The transparency feature was useful
5. The clipping plane was useful
6. The slab rendering feature was useful
7. I found that co-ordinating the two sensors was difficult
8. Using the plastic model as an interface was more intuitive than using a mouse and keyboard to move the onscreen image
9. The on-screen representation of the brain ventricles followed the movements of the plastic model
10. The arrows and labels were well positioned
11. The on-screen rendering was difficult to control using the plastic model

12. The system helped me to understand the shape and location of the ventricular system
13. I would like to use this system to learn other anatomical features
14. The plastic model helped me to understand the shape of the ventricular system
15. I found the sensor cables distracting

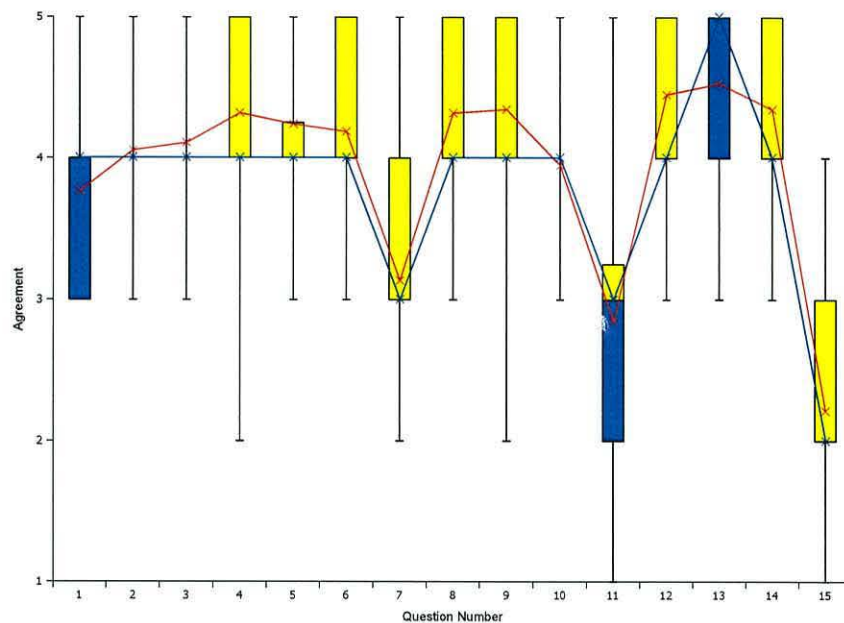


Figure 6.10: A chart showing the results from the questionnaire where the yellow areas represent the second quartile and the blue areas the third quartile. The red line represents the mean for each question, and the blue line the median.

Again the most important question asked the students whether they agreed that BARETA helped them to understand the shape and location of the ventricular system (Question 12). Once again the reaction from the students was positive, with all but one student either agreeing or strongly agreeing. One student commented that they could see how BARETA could “help learn about the anatomy of the Brain.”

As well as showing the students the shape and position of the ventricular system, we also wanted to find out how easy to use the students found BARETA to be. Several questions asked the students about how useful and usable they found various



aspects of BARETA. When asked if they found the system straightforward to use (Question 2) a large majority of students responded with “agree”, with only a few either responding “strongly agree” or “neutral.” One student commented that the system was “easy to use,” however another commented that it was “hard to control.” Most students either responded “strongly agree” or “agree” when asked if they found using the plastic model as an interface was easier than using a keyboard and mouse interface (Question 8). A large majority of students also either agreed or strongly agreed that using the plastic model helped them to understand the shape of the ventricular system (Question 14).

As well as usability issues we also sought opinions on some of the novel viewing features that had been implemented. The first such feature was the transparency feature. Question 4 asked the students whether they agreed that the transparency feature was useful. In general the students either agreed or strongly agreed, three students replied with “neutral” and one with “disagree.” A mean of 4.3 for this question indicates that most of the students found this feature useful.

Another novel viewing feature of BARETA is the clipping plane feature. Question 5 asked the students if they found this feature useful. Again most students either agreed or strongly agreed that this feature was useful, with four students responding with “neutral.” The mean for this question was 4.2. The slab rendering feature is related to the clipping plane feature. We asked the students if they found this feature useful (Question 6). All but two of the students agreed or strongly agreed that this feature was useful; the remaining two students replied with “neutral.” Again the mean for this question was 4.2.

Despite the usefulness of the clipping plane and slab rendering features, Question 7 revealed that not all of the students found co-ordinating the two sensors easy. The median response when asked if they found co-ordinating the two sensors difficult was 3, with a mean of 3.1; this suggests that about a half of the students found this to be a problem. Similarly the responses to Question 11 indicate that some of the students

found BARETA difficult to control with the plastic RP model. The median response when asked whether they found it difficult to control the on-screen rendering with the plastic model was 3, and the mean was 2.8; this has similar implications to the result of the co-ordination question. We also asked the students if they found the miniBIRD sensor cables distracting (Question 15). Only two students agreed that this was the case, with most students either disagreeing or strongly disagreeing.

In contrast with the previous user evaluation the students felt that the arrows and label were quite well positioned. More than half of the respondents agreed that this was the case when asked (Question 10). The remaining students either strongly agreed or were neutral.

None of the students disagreed that they had considered that virtual material could be helpful to their education (Question 1). A mean of 3.7 suggests that several of the students were unsure whether or not they agreed. After using BARETA all but one of the students either agreed or strongly agreed that they would like to use BARETA to learn other anatomical features (Question 13), showing that the students found BARETA useful. One student commented that they would “want to use it in class,” and another that BARETA “could further learning in key areas.”

As in the previous user evaluation we were interested to see if any demographic factors had an influence on a user’s perceptions of BARETA. In this evaluation we did not record any age information as the age difference between students was quite small. We did however record gender information. We analysed both the data from males and females separately. As we saw in the previous user evaluation there were some small differences between the answers of the male students (see Figure 6.11) and the answers of the female students (see Figure 6.12). Again the small size of the samples could explain the differences. When the mean answer of the two gender groups are compared with the values for the entire population (see Figure 6.13) the similarity in opinions is more evident; the difference in mean between male and female is no more than 0.47 for any question, and in some cases is much smaller.

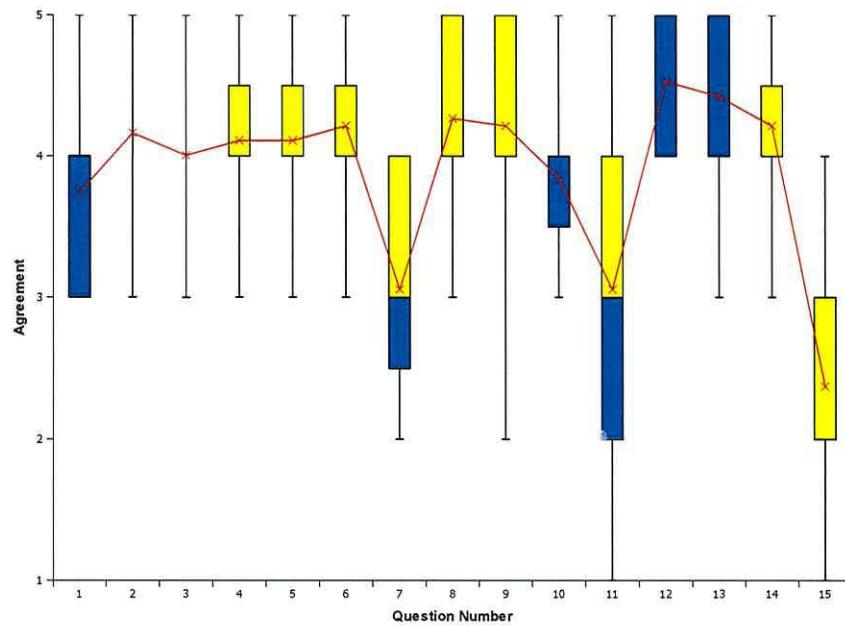


Figure 6.11: A chart showing the results from male respondents of the questionnaire where yellow areas represent the second quartile and the blue areas the third quartile. The red line represents the mean for each question.

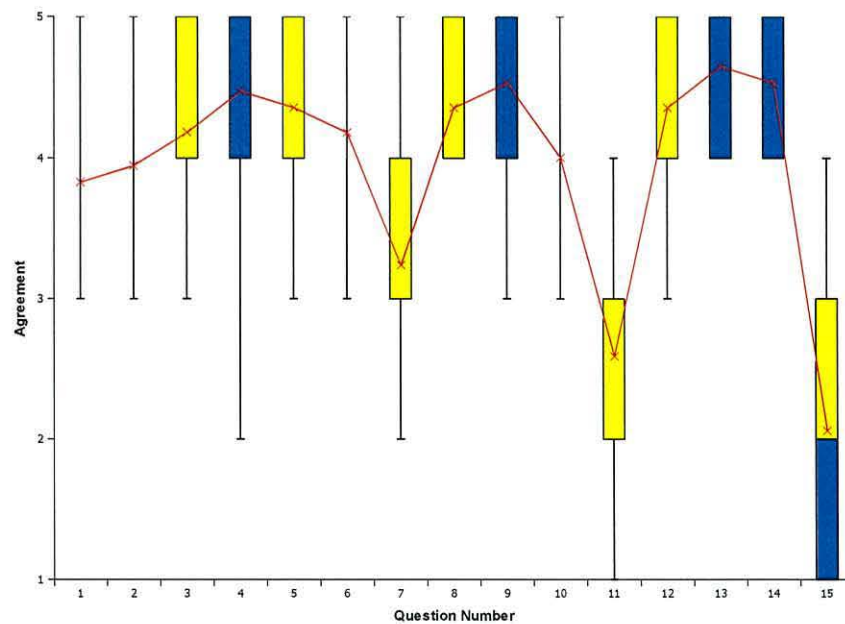


Figure 6.12: A chart showing the results from female respondents of the questionnaire where yellow areas represent the second quartile and the blue areas the third quartile. The red line represents the mean for each question.

As we did with the previous user evaluation we also calculated the standard deviation for each question (see Figure 6.14). As was the case in the previous user evaluation the standard deviation was less than one for all but one question, and



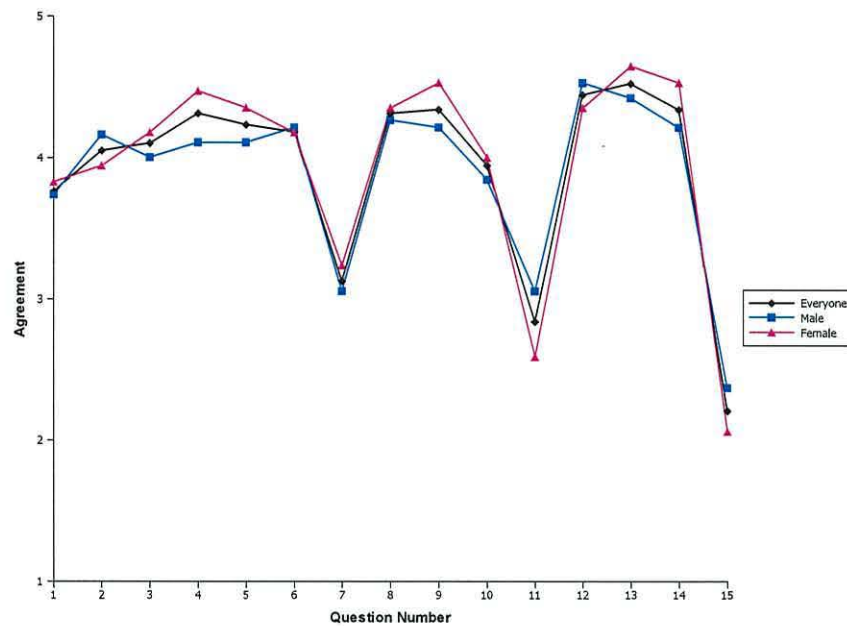


Figure 6.13: A chart comparing the mean answers of the entire population to that of the male and female students.

in several cases it was less than 0.5, suggesting that most of the students shared similar opinions on many aspects of the system. We also compared the population standard deviations with the standard deviations of both the males and females, which revealed small differences between the two; however the differences were not large and varied between questions.

This user evaluation provided us with more useful feedback that aided the development of BARETA; in general the students were impressed with BARETA, however some issues still required resolution, including the positioning of the arrows and labels and the lack of a reset function. Following this evaluation some changes were made to BARETA that reflected what we had learned. To evaluate these changes another user evaluation was required.

## 6.4 Third User Evaluation

Our third user evaluation was carried out at the School of Medicine at Keele University. This user evaluation was conducted in a similar manner to the previous

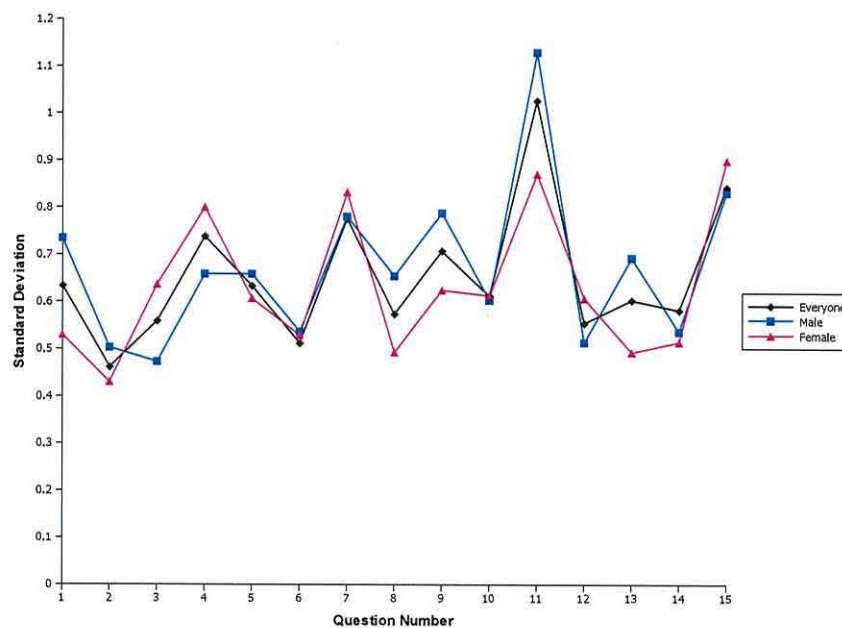


Figure 6.14: Connah's Quay questionnaire results Standard Deviation

two, where the students used BARETA then filled out a questionnaire. Thirty-four first year medical students took part in the evaluation, twelve of whom were male, twenty-one of whom were female, and one who did not specify a gender. This evaluation took place during the students' introductory session to the ventricles of the brain and included a dissection component. Doctor John Delieu was also involved with this group of students, in this case as an assistant during the dissection. As was the case during the first evaluation, this meant that the material presented by BARETA was relevant to the students.

On this occasion the updated version of BARETA described in Section 5.6 was used. Each student used BARETA in the configuration shown in Figure 6.16. Students assumed a seated position to use BARETA. The camera in this case was located in a slightly elevated position behind the user, pointing towards the computer screen. The view provided by the camera showed an image that looked over the student's right shoulder.

Each participant filled in a questionnaire after using BARETA. Again the questionnaire was focused on the usability of BARETA as the user evaluation took place

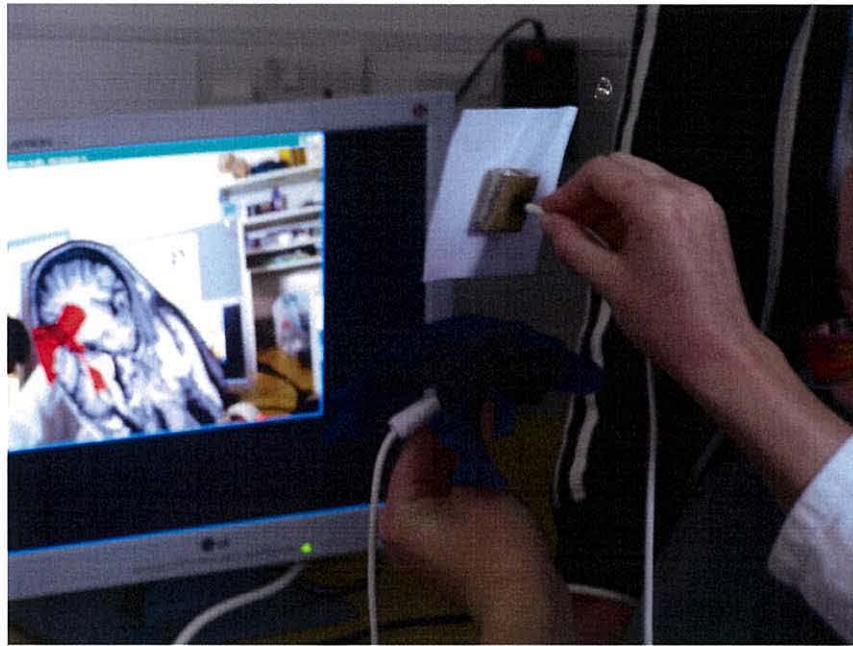


Figure 6.15: A student at Keele University operates the augmented reality system.

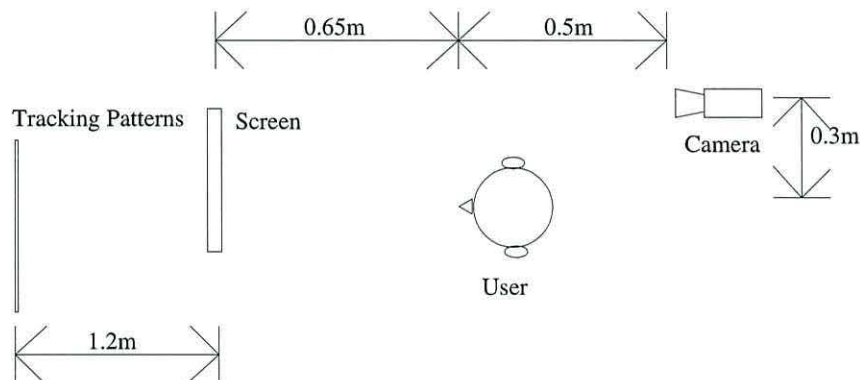


Figure 6.16: The experimental setup used at Keele University (not to scale)

during their introductory session on the ventricular system, therefore no experience of other teaching methods used for teaching the ventricular system could be assumed. The students were divided into groups of four, five or six, and tried BARETA during their dissection session (Figure 6.15). The first group did not participate in the dissection until after they had used BARETA. Each subsequent group had spent more time dissecting before using BARETA. This gave the opportunity to assess whether or not contact with cadaveric material altered the students' perception of BARETA.

The questions were as follows:



1. I have previously considered that access to virtual material may help me to learn new topics
2. The system was straightforward to use
3. Co-ordinating the two sensors was easy
4. The transparency feature was useful
5. The clipping plane was useful
6. The slab rendering feature was useful
7. I found that co-ordinating the two sensors was difficult
8. Using the plastic model as an interface was more intuitive than using a mouse and keyboard to move the onscreen image
9. I found that the system was difficult to use
10. The arrows and labels were well positioned
11. Using a physical representation of the ventricles made the system easier to use
12. I found the sensor cables distracting
13. I did not find the transparency feature useful
14. The frame rate of the system was good
15. I did not find the slab rendering feature useful
16. I thought that the arrows and labels were not well positioned
17. I did not find the system intuitive to use
18. The system did not help me to understand the shape and location of the ventricular system

19. The on-screen rendering was difficult to control using the plastic model
20. The system helped me to understand the shape and location of the ventricular system
21. I did not find the clipping plane useful
22. The update rate of the system was poor
23. I did not think that using a physical representation of the ventricles made the system easier to use
24. The camera could have been positioned better
25. I would like to use this system to learn other anatomical features

Questions were posed both positively and negatively to avoid leading the respondents. The questionnaire also allowed the user to include additional comments on BARETA. The results from this questionnaire are more difficult to interpret from the chart (Figure 6.17) than previously because of the mix of positively and negatively posed questions. Therefore a method was required to make the results easier to interpret.

#### 6.4.1 Combined Results

To make the interpretation of the results easier the results from a positively posed question was combined with the corresponding negatively asked question to produce a third value between one and five where one was the worst possible and five was the best possible. This produced ten different lines of enquiry, labelled A to J on the chart (Figure 6.18) and were as follows:

- A. The system was easy to use
- B. Co-ordinating the two sensors was easy

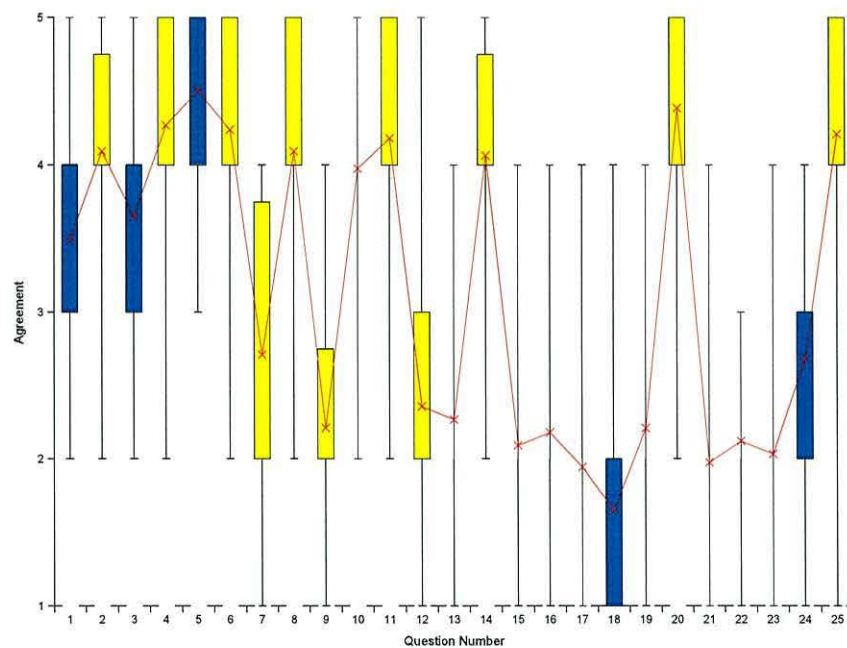


Figure 6.17: A chart displaying the results from the Keele user evaluation where the yellow areas represent the second quartile and the blue areas the third quartile. The red line represents the mean for each question.

- C. The transparency feature was useful
- D. The clipping plane was useful
- E. The slab rendering feature was useful
- F. Using the plastic model as an interface was more intuitive than using a mouse and keyboard to move the onscreen image
- G. The arrows and labels were well positioned
- H. Using a physical representation of the ventricles made the system easier to use
- I. The frame rate of the system was good
- J. The system helped me to understand the shape and location of the ventricular system

The formula used to produce our figures was as follows:



$$Answer = \frac{(Positive + (6 - Negative))}{2}$$

The term  $(6 - Negative)$  is included to convert an answer from one where 1 was the best and 5 the worst into one in which 5 was the best and 1 was the worst ( $6 - 5 = 1$ ,  $6 - 1 = 5$ ). The average of this and the score for a positively posed question is then calculated.

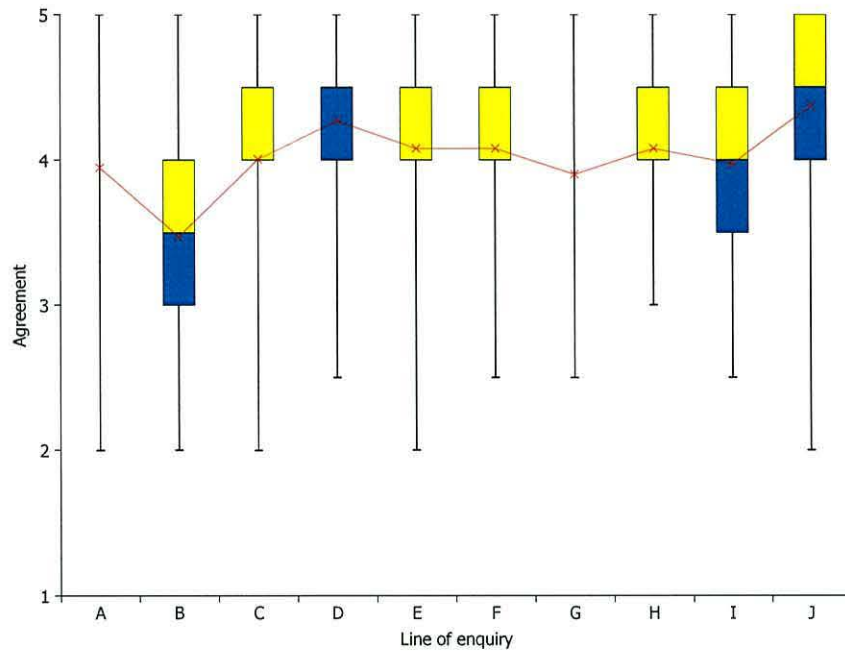


Figure 6.18: A chart displaying the combined results from the Keele user evaluation where the yellow areas represent the second quartile and the blue areas the third quartile. The red line represents the mean for each question.

The most important result for our hypothesis was that the students found that BARETA helped them to understand the shape and the location of the ventricles within the human head (Line J). All but two of the students recorded a score of 4 or greater for this line of enquiry. A mean of 4.3 and a median of 4.5 suggest that the students' agreement is quite strong.

Ease of use of BARETA is another important area of questioning. Several questions dealt with this. Line of enquiry A asks users if they found BARETA easy to use. Although results ranged from 2 to 5 a mean of 3.9 and a median of 4 suggests

that most of the students found that the system was easy to use. Line of enquiry F also relates to ease of use; this pair of questions asks the students if they found that using the plastic model as an interface was more intuitive than using a keyboard and mouse interface. The results here were also positive, with a mean of 4.1 and a median of 4 strongly suggesting that the students found that using the plastic RP model as an interface was more intuitive than using a keyboard and mouse interface. Line of enquiry H was similar, and showed almost identical results.

Another aspect of BARETA for which we sought opinions was the inclusion of the novel viewing features. Line of enquiry C asked the students whether they felt that the transparency feature was useful. The majority of the students agreed that this feature was useful, showing a mean and median of 4. The remaining students disagreed. Another novel viewing feature that was demonstrated to the students was the clipping plane feature. Line of enquiry D asked the students for their opinions on this feature. Most of the students found that this feature was useful, recording a mean of 4.3 and a median of 4.5. The lowest score was 2.5, suggesting that none of the students strongly disagreed. The final novel viewing feature that we demonstrated to the students was the slab rendering feature, which was dealt with by line of enquiry E. Once again opinions were favourable with a mean score 4.1 and a median score of 4 suggesting that most of the students found this feature useful.

As the use of the clipping plane and slab rendering features required the co-ordination of two tracked physical objects we wanted to see how easy the students found co-ordinating the two sensors, which was addressed by line of enquiry B. The results reveal that not all of the students found that the two sensors were easy to co-ordinate although a mean and a median of 3.5 suggests that several of the students found the two sensors easy to co-ordinate. The maximum score was 5 and the minimum 2 suggesting that a few of the students found co-ordination difficult, but not beyond their capability.

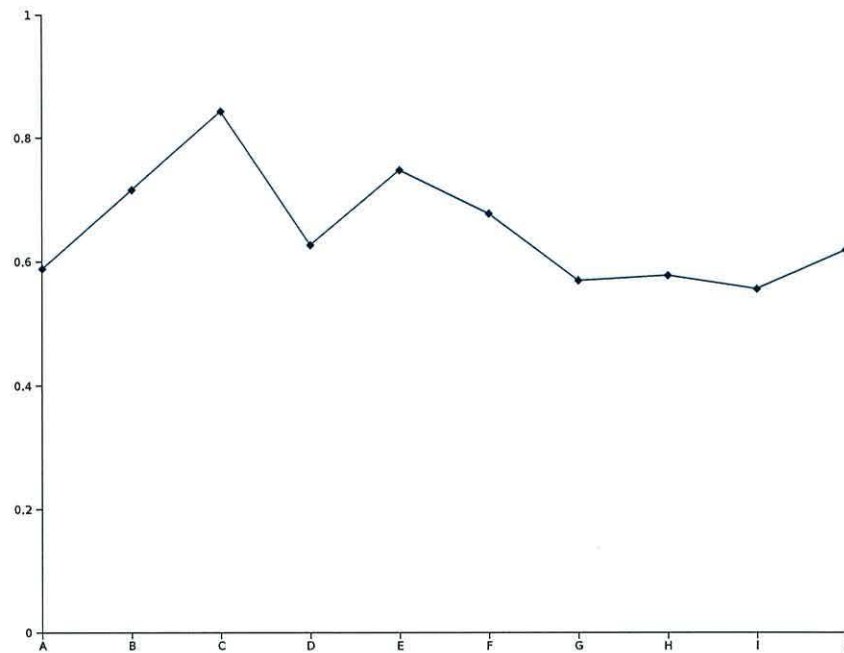


Figure 6.19: A chart displaying the standard deviation for each question asked at Keele University.

The way in which text labels and arrows were displayed had changed since the previous user studies, therefore we wanted to evaluate the new system. Line of enquiry G asked the users for their opinions on this. Most of the students found that the arrows and labels were well positioned, with a median score of 4 and a mean of 3.9

The final line of enquiry (Line I) asked the students how they felt about the frame rate of BARETA. A mean and a median of 4 suggests that the students were satisfied with the frame rate of the system; only two students recorded a score of less than 3.5 for this line of enquiry.

The standard deviation for each line of enquiry was also calculated, and can be seen in Figure 6.19. The minimum standard deviation for a line of enquiry was 0.56 and the maximum 0.83 suggesting that although the students had differing opinions about each aspect of the system, they did not disagree by a large amount.

As students in later groups had some experience with dissecting a cadaver immediately prior to using BARETA, an opportunity arose to see how opinions on



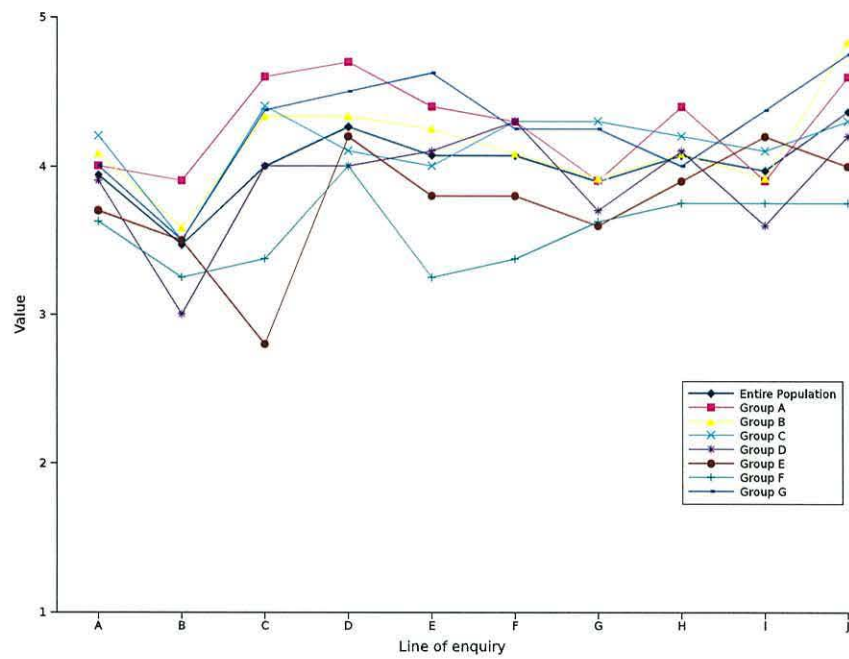


Figure 6.20: A chart showing the mean responses for each group that took part in the Keele user evaluation.

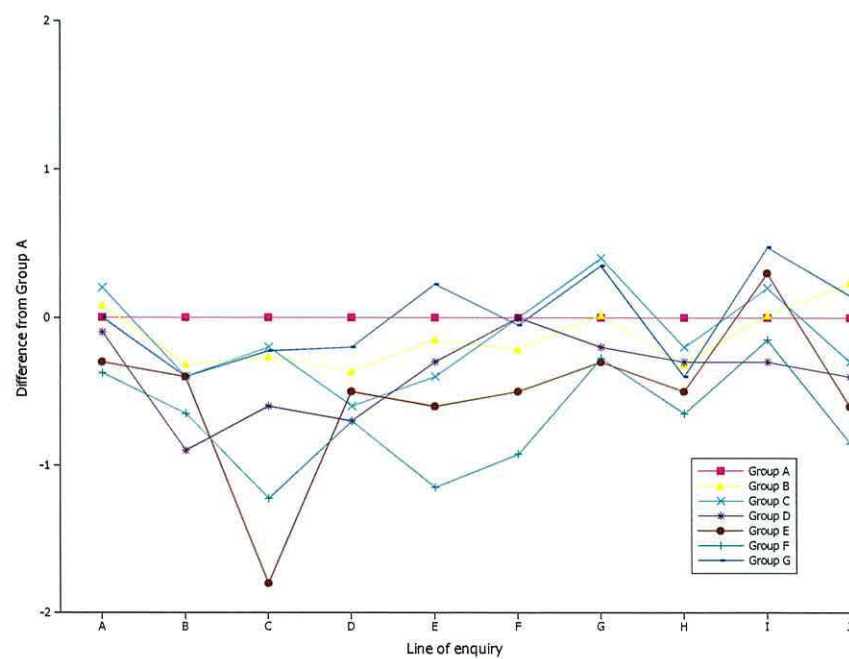


Figure 6.21: A chart showing how the mean response of each group of students differs from that of Group A at Keele University

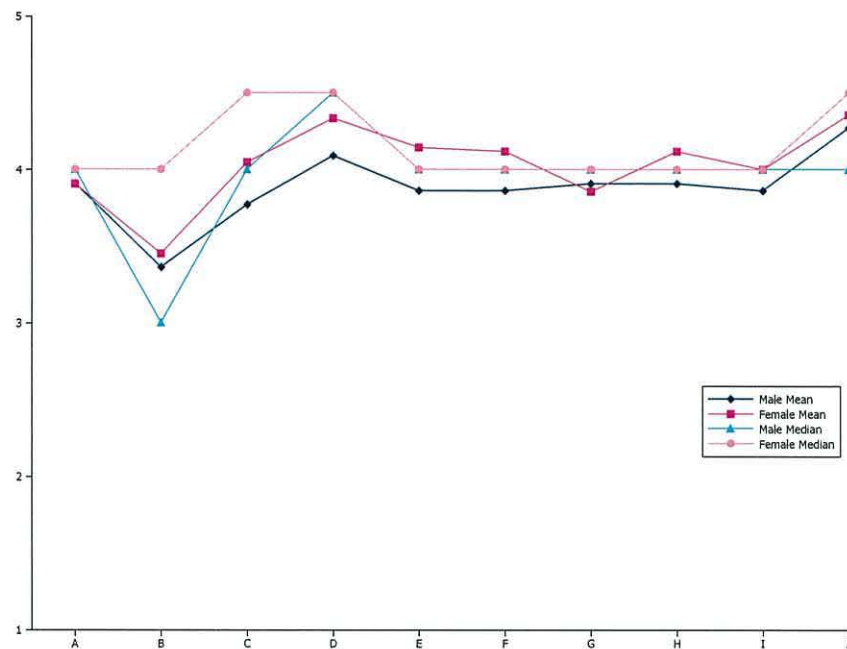


Figure 6.22: A chart comparing the responses of male and female students at Keele the system changed as a student had carried out more dissection of the brain. To investigate this we took a mean value from each group for each line of enquiry which allowed us to make a comparison (Figures 6.20 & 6.21). It would appear that in general that the students opinions of BARETA do change slightly as they carry out a real dissection. With the exception of group G, each successive group generally records a slightly lower mean for each line of enquiry. The line of enquiry featuring the largest difference is line C, asking about the usefulness of the transparency feature.

To investigate the effect of gender on a student's perception of BARETA we calculated the mean and median score of each line of enquiry for the two gender groups, the results of which can be seen in Figure 6.22. The difference between the mean score of the two groups is no greater than 0.28 for any line of enquiry, however the median shows differences as great as 1, although in all but three of the lines of enquiry it is identical. It is interesting to note that the male participants found co-ordinating the two sensors more difficult than the females.

### 6.4.2 Stand-Alone Results

Five of the questions on the questionnaire did not have corresponding questions to produce a combined result. The results of these questions are displayed in Figure 6.23. The questions were re-numbered for inclusion on the chart and are as follows:

1. I have previously considered that access to virtual material may help me to learn new topics
2. I found the sensor cables distracting
3. The on-screen rendering was difficult to control using the plastic model
4. The camera could have been positioned better
5. I would like to use this system to learn other anatomical features

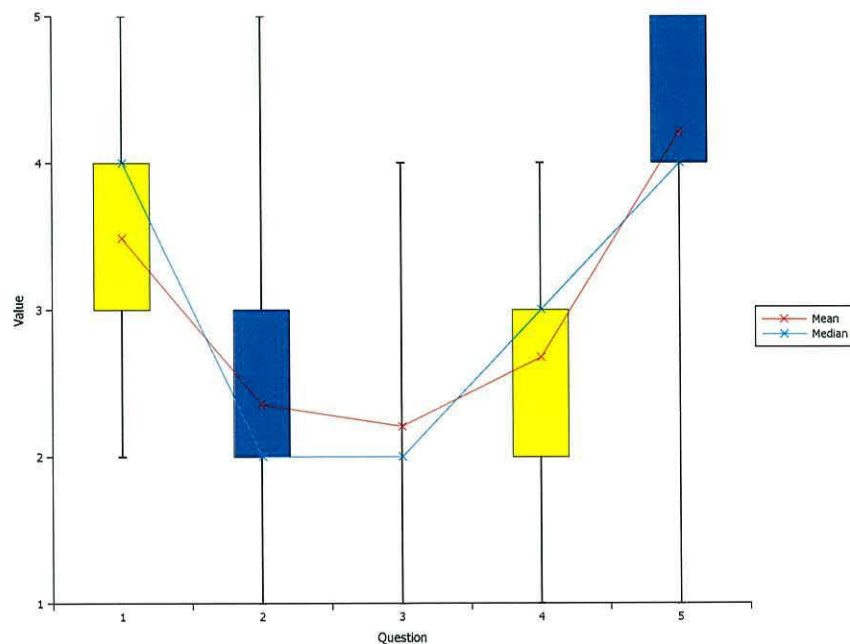


Figure 6.23: A chart showing the results of stand-alone questions posed at Keele University

As in the previous user studies several of the students had considered that virtual material may be beneficial to their education (Question 1). We also asked the



students if they would like to use BARETA to learn other aspects of anatomy (Question 5). Many of the students either agreed or strongly agreed that this was the case, and the mean for this question was higher than for Question 1, as was the case in the two previous user studies. It was interesting to note that two students strongly disagreed that they would like to use BARETA again: we had not encountered this in the previous user studies.

Questions 2, 3 and 4 all related to the usability of BARETA, and were negatively posed. Question 2 asked the students if they found the sensor cables distracting. Few students agreed that this was the case, however a median of 2 suggests that students did not disagree very strongly. Question 3 asked the students if they found BARETA difficult to control. As with Question 2 a small number of students agreed that this was the case, whilst twenty seven students responded with “neutral” or “disagree,” which concurs with our combined results enquiring about ease of use. Question 4 asked the students whether they felt that the camera was well positioned. A mean of 2.7 and a median of 3 suggests that the students weren’t sure whether or not this was the case; however whilst four students strongly agreed that the camera was well positioned, none of them strongly disagreed.

## 6.5 Summary

The results from our user studies were promising. BARETA was targeted at sixth-form and University students who were at an early stage of their education in gross anatomy. To prove our hypothesis the results needed to show that users felt that using BARETA helped their understanding of the anatomy being shown to them. A large majority of the students through all three of the user studies believed that BARETA helped them to learn the shape and position of the ventricular system within the human brain. Students also felt that some of the novel viewing features that had been introduced to them were helpful.

Another thing that has become clear through the user studies is that some students found that BARETA system was difficult to use as a lot of co-ordination was required. In the short time in which the students used the system they could not fully resolve appropriate movements to achieve a desired result. This could be because of the experimental setup used where the camera was positioned in such a way that movements on the screen did not match the actual movements made by the user. Such a situation may put students off further use of this system, or one that is similar. The results did however show that a large majority of the students were interested in using BARETA again to learn other aspects of anatomy, even though many had not previously considered that virtual material could be beneficial to their education before using the system.

All of our user studies presented the ventricles as either a new area of study, or as part of a during scheduled teaching activity that focused on the ventricular system. During the user studies several students commented that if such a system was made available to them in a study room or library then it could also be a valuable revision tool that could help them to improve their examination results.

## Chapter 7

### Conclusions and Future Work



## 7.1 Introduction

In the introduction to this thesis the hypothesis “Is it possible to construct a compelling Augmented Reality system that can offer an alternative learning aid for teaching gross anatomy?” was postulated. In an attempt to prove this hypothesis the Bangor Augmented Reality Education Tool for Anatomy (BARETA) was developed and subsequently evaluated using a series of user studies in which potential end users tried BARETA and gave feedback on a questionnaire.

## 7.2 BARETA

In Section 3.4.1 a number of requirements were set out for BARETA, which were developed in collaboration with an Anatomy lecturer based at Bangor University. It was decided that BARETA must:

- combine the virtual and the real
- be interactive in real-time
- be registered in 3D
- track a mobile viewpoint
- track two physical objects (an anatomy model and a tool)
- be capable of rendering high resolution medical volume data in real-time
- be suitable for use in a classroom environment
- be easily moved between locations

During this thesis the development of BARETA was presented that was created to satisfy these requirements. BARETA was developed to operate on a standard desktop PC and using commercially available tracking hardware: an Ascension

miniBIRD 800 (miniBIRD) was used to track two physical objects (an anatomy model and a tool, though it is possible to add more sensors to track more objects) and an InterSense IS-1200 VisTracker (VisTracker) was used to track the user's viewpoint. Each of these devices is easily carried by a single person and therefore BARETA is portable.

A view of the real world was provided by a Logitech Quickcam Pro 9000 (webcam), creating a video see-through Augmented Reality system that allowed the user to see both the real world and virtual objects. The Visualization Toolkit (VTK) open-source software toolkit was used as a software component for creating virtual output. VTK supports a large number of data types, and can render data in many different ways. In the case of BARETA, VTK was used to generate real-time volume and surface renderings of parts of the human head.

To ensure that BARETA achieved correct registration all of the components of BARETA had to be integrated and calibrated (Chapter 3). To calibrate BARETA a novel bracket was designed and constructed to maintain a constant displacement between a sensor for the miniBIRD and the VisTracker whilst the positions of the two in their own co-ordinate systems were sampled. A similar bracket was also constructed to maintain a constant displacement between the VisTracker and the webcam during the operation of BARETA. This ensured that the registration of BARETA was correct.

To one of the miniBIRD's sensors a tangible interface derived from real medical data and produced using Rapid Prototyping (RP) technology was attached (Chapter 4), with which the user controlled an anatomy lesson based upon the same medical data. The anatomy lesson used both volume and surface rendering and provided the user with a number of options for viewing the ventricles (Chapter 5, Figure 7.1) including transparency, an arbitrarily positionable clipping plane and an arbitrarily positionable slab rendering, controlled using a second tangible interface, that was also tracked using the miniBIRD system. Also provided were annotations



Figure 7.1: BARETA running, with the transparency feature enabled.

of some of the important features of the ventricular system.

In its current state BARETA can be used in a classroom, however supervision is required. The tracking hardware in particular is delicate and could easily be damaged if used incorrectly. The BARETA software itself is easy to use, and if used with physically robust tracking hardware could be used without supervision.

## 7.3 Discussion

Key to proving the hypothesis was determining whether or not students found that BARETA helped them to understand the shape and the location of the ventricular system. To investigate if this was the case three user studies involving potential end users were carried out during the development of BARETA. The results of our user studies suggest that the students did find that BARETA helped them to understand the shape and the location of the ventricular system.

The perception of the usefulness of BARETA's novel viewing features was also important as these contributed to the effectiveness of BARETA as a learning aid.



The novel viewing features were the transparency feature, the clipping plane and the slab rendering. Many of the students felt that these features were useful, however it is interesting to note that the different groups of students preferred different features. The students at Keele showed a slight preference for the clipping plane and slab rendering over the transparency feature whereas the students at Connah's Quay High School showed a slight preference for the transparency feature. However none of the user studies showed that students found that any these features were not useful. This difference in opinion could be explained by the different educational goals of the two groups of students. A university medical student will be a lot more interested in the fine details of the ventricles than the high school students, who will not have performed a cadaver dissection.

The students were also asked how useful they found BARETA in comparison to other teaching methods. Of the students questioned most felt that BARETA was at least as good for teaching the layout of the ventricular system as each of the other teaching methods, except for cadaver dissection.

Our user studies also assessed how easy to use BARETA was, and how useful they found each of the individual features. Some of the users felt that using the system could be a little difficult, however this was potentially caused by the positioning of the camera providing the real world scene, which was not always pointing in the user's view direction.

## 7.4 Contribution

During the undertaking of this project the main contributions made to the state of the art were:

- A novel augmented reality anatomy teaching tool, the Bangor Augmented Reality Education Tool for Anatomy (BARETA), was developed that can be run on a standard desktop PC. BARETA allowed users to interactively control the

position and orientation of computer generated anatomy models. A sample lesson based on the human head was developed that focused on the ventricular system within the brain. The head was represented by a volume rendering of MRI data, supplemented by a surface rendering of data segmented from the same MRI data. BARETA is useful for showing spatial relationships between neighbouring objects, especially when combined with features such as transparency, and an interactive clipping plane and slab rendering.

- An anatomically correct plastic model was generated from medical volume data with which a user could interact with the system in a natural fashion, avoiding the need for an extended familiarisation period with an interaction style. To generate the model a semi-automatic segmentation was performed on MRI data, from which a surface data set was produced. From this surface data a physical equivalent was fabricated using a 3D printer.
- We have demonstrated through a series of validation studies that students of anatomy are in general willing to use a variety of methods to learn anatomy, and that augmented environments are a useful supplement to traditional cadaveric education methods. Students agreed that use of our tool gave them a better understanding of the ventricular system than all other education techniques other than cadaver dissection. Most students also found the system easy to use, and would use it again to learn other anatomical features.

## 7.5 Future Work

The current version of BARETA shows the user features of the human ventricular system within the human brain. Adding surface renderings of other anatomy to the current version of BARETA is straightforward as it only requires an additional segmentation of the region of interest. The creation of lessons about other regions of the human body are also possible given relevant scan data and a segmentation



that produces data suitable for the production of an RP model. BARETA could also be used to visualise more complex anatomy with some adaptation. For instance the function of the elbow joint could be illustrated using separate RP models of the upper arm and the forearm that meet at the elbow. This would require the use of two tracking sensors (one for the upper arm, one for the forearm), and software to calculate the movement of the muscle as the elbow joint is manipulated.

Although the colour of the volume rendering produced by BARETA matches what a user could expect to see from a medical scan, it does not show the real colour of the anatomy being studied. The use of transfer functions can provide pseudo-colour, however it is very difficult to create transfer functions that produce anatomically correct colouring. In computer graphics a technique that can be used for photorealistic rendering is to apply a bidirectional reflectance distribution function (BRDF), which is a 4-dimensional function that defines how light is reflected at an opaque surface. A technique such as that used in [14] to recover BRDF data that can then be fitted to a BRDF model (such as the widely used Lafortune model) could be used. This can then be used to render colour accurate computer-generated images.

As another novel viewing feature alongside the clipping plane a novel viewing glass could be introduced into the system. Using a hollowed out cardboard square as the tangible component, the user could control the position of a tunnel into the data, similar to the AR window used by Bichlmeier and Navab [5]. This could allow a greater perception of the depth of the ventricles (or other highlighted features), and the relationship with the rest of the anatomy.

The introduction of occlusion detection and resolution could also enhance BARETA. It is possible that this could make the clipping plane easier to use. At present the prop for controlling the clipping plane is always occluded by the volume rendering when the two are co-located on the image plane, if the clipping plane was to always occlude the volume rendering when it is active this might make co-ordination easier.



As the prop for the clipping plane is a small square it is not a complex shape to model, and could be modelled as a tunnel into the volume rendering.

The inclusion of a head-mounted display (HMD) would provide the possibility of providing the user with an easy to use stereoscopic environment. It would also make BARETA seamless and more intuitive, as the movement of the user's head could be tracked as well as the viewpoint of the webcam, reducing the effort that user has to make to co-ordinate their manipulations of the props available to them. This would require the mounting of the webcam and the VisTracker on the HMD, which would need consideration when choosing a HMD. The unit should be light, and have good balance with these peripherals attached.

Better exploitation of the capabilities of a programmable GPU could enhance the rendering speed of BARETA, for instance using CUDA with nVidia graphics card or the VTKEdge extensions of VTK could dramatically increase the frame rate of the system. Although this is not an issue with the data sets currently used, it would enable much larger data sets to be used and for more surfaces to be displayed without a loss of performance. The data sets used for BARETA were much smaller than the MRI scanner used was capable of producing, and were of a relatively small area of the human anatomy.

The introduction of the typical smell of a dissection room into a virtual environment could enhance a user's perception of the ventricles, and may improve recall when in contact with real anatomy. As discussed on Page 30 and in [9], a user's perception of an environment is formed from a combination of all five senses, not just sight, often the only sense stimulated by a virtual environment. Similarly the incorporation of active haptics into the BARETA environment could enhance a user's understanding or recall of the anatomy that is being studied. If the tracked RP model represented a bone, the haptics device could be used to provide resistance where muscle or other tissue is present near to this bone.

We have established that potential end users are broadly in favour of using a

system such as BARETA, and believe that it is beneficial to their education. What is not yet known is if the system produces a greater learning effect than other education methods such as cadaver dissection, viewing a prosection or reading a textbook. A user study that has students doing each type of learning followed by a short examination could provide the answers to this question. Such a study requires more planning than the user studies already carried out, suitable test material having to be devised. Also more time would have to be allocated to it at the relevant stage within the students' curriculum, or for students to participate in their own time.

BARETA could also be extended to aid a user's understanding of flows within the body. The ventricular system distributes Cerebrospinal Fluid (CSF) around the brain and into the spinal cord. If this could be shown in a BARETA lesson it could add extra value to the tool. The way in which this can be shown could be through the use of static arrows, moving arrows, moving blobs, or some other dynamic glyph.

Currently RP models are solid and although they can convey the shape of anatomical features, the models do not deform; the tissues that make up the human body however do deform. An RP model, or similar, that can deform in a realistic manner would be an intriguing addition to the Augmented Reality system, especially if the deformation could be measured and represented on the screen somehow.

## 7.6 Conclusion

Cadaveric dissection has long been the Gold Standard for the education of gross anatomy. This thesis has presented the development of BARETA, a system that aims to act as an compelling supplement to cadaveric dissection. Although the system is effective in conveying information about anatomy it is not yet able to fully replace cadaveric dissection. Although both visual and tactile representations of anatomy are presented it cannot at present reproduce all of the sensations of cadaveric dissection. BARETA does however allow students to view samples of

living subjects, a potential advantage over cadaver dissection; cadavers are known to deform over time, and may give a false impression in certain areas. Future technological advancements and the implementation of ideas discussed above will mean that such a system will be possible within the next decade. Despite this we believe that our hypothesis has been proven correct by the evidence presented in this thesis.



# References

- [1] E.A. Ashton, M.J. Berg, K.J. Parker, J. Weisberg, C.W. Chen, and L. Ketonen. Segmentation and feature extraction techniques, with applications to MRI head studies. *Magnetic Resonance in Medicine*, 33(5):670–677, 1995.
- [2] Ronald Azuma. A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385, Aug 1997.
- [3] Ronald Azuma and Chris Furmanski. Evaluating label placement for augmented reality view management. In *ISMAR '03: Proceedings of the The 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality*, page 66, Washington, DC, USA, 2003. IEEE Computer Society.
- [4] D.C. Banks and K. Beason. Fast global illumination for visualizing isosurfaces with a 3D illumination grid. *Computing in Science and Engineering*, 9(1):48–54, 2007.
- [5] Christoph Bichlmeier and Nassir Navab. Virtual window for improved depth perception in medical ar. In *AMIARCS - Virtual Window for Improved Depth Perception in Medical AR*, Copenhagen, Denmark, Oct. 2006. MICCAI Society.
- [6] Mark Billinghurst, Hirokazu Kato, and Ivan Poupyrev. Tangible augmented reality. In *SIGGRAPH Asia '08: ACM SIGGRAPH ASIA 2008 courses*, pages 1–10, New York, NY, USA, 2008. ACM.

- [7] M. Burns. *Automated fabrication: improving productivity in manufacturing*. Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 1993.
- [8] A. Chalmers and A. Ferko. Levels of realism: From virtual reality to real virtuality. In *Proceedings of SCCG'08*, page 2733, 2008.
- [9] Alan Chalmers, David Howard, and Christopher Moir. Real virtuality: A step change from virtual reality. In *Keynote Talk, 25th Spring Conference on Computer Graphics, Slovakia*, April 2009.
- [10] Chinese visible human-tmmu [http://www.chinesevisiblehuman.com/index\\_e.aspx](http://www.chinesevisiblehuman.com/index_e.aspx).
- [11] D.M. Coll, B.R. Herts, W.J. Davros, R.G. Uzzo, and A.C. Novick. Preoperative Use of 3D Volume Rendering to Demonstrate Renal Tumors and Renal Anatomy 1, 2000.
- [12] W.W. Cottam. Adequacy of Medical School Gross Anatomy Education as Perceived by Certain Postgraduate Residency Programs and Anatomy Course Directors. *Clinical Anatomy*, 12:55–65, 1999.
- [13] L. Davis, F.G. Hamza-Lup, J. Daly, Y. Ha, S. Frolich, C. Meyer, G. Martin, J. Norfleet, K.C. Lin, C. Imielinska, et al. Application of augmented reality to visualizing anatomical airways. *Proceedings of SPIE, the International Society for Optical Engineering*, 4711:400–405, 2002.
- [14] P. Debevec, C. Tchou, A. Gardner, T. Hawkins, C. Poullis, J. Stumpfel, A. Jones, N. Yun, P. Einarsson, T. Lundgren, et al. Estimating surface reflectance properties of a complex scene under captured natural illumination. *Conditionally Accepted to ACM Transactions on Graphics*, 2004.

- [15] C.E. Dinsmore, S. Daugherty, and H.J. Zeitz. Teaching and Learning Gross Anatomy: Dissection, Prosection, or “Both of the Above?”. *Clinical Anatomy*, 12:110–114, 1999.
- [16] David Drascic and Paul Milgram. Perceptual issues in augmented reality. *SPIE: Stereoscopic Displays and Virtual Reality Systems III*, 2653:123–134, April 1996.
- [17] Robert A. Drebin, Loren Carpenter, and Pat Hanrahan. Volume rendering. *SIGGRAPH Comput. Graph.*, 22(4):65–74, 1988.
- [18] T.T. Elvins. A survey of algorithms for volume visualization. *ACM Siggraph Computer Graphics*, 26(3):194–201, 1992.
- [19] G.R. Ferretti, I. Bricault, and M. Coulomb. Virtual tools for imaging of the thorax. *European Respiratory Journal*, 18(2):381–392, 2001.
- [20] G.W. Fitzmaurice, H. Ishii, and W.A.S. Buxton. Bricks: laying the foundations for graspable user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 442–449. ACM Press/Addison-Wesley Publishing Co. New York, NY, USA, 1995.
- [21] I. Gibson, L.K. Cheung, S.P. Chow, W.L. Cheung, S.L. Beh, M. Savalani, and S.H. Lee. The use of rapid prototyping to assist medical applications. *Rapid Prototyping Journal*, 12(1):53–58, 2006.
- [22] A. Gillet, M. Sanner, D. Stoffler, D. Goodsell, and A. Olson. Augmented reality with tangible auto-fabricated models for molecular biology applications. *Visualization, 2004. IEEE*, pages 235–241, 2004.
- [23] S. Hariri, C. Rawn, S. Srivastava, P. Youngblood, and A. Ladd. Evaluation of a surgical simulator for learning clinical anatomy. *Medical Education*, 38(8):896–902, 2004.



- [24] O. W. Henson Jr. Med image home page  
[http://cbaweb2.med.unc.edu/henson\\_mrm/VisibleHuman/VHMain.html](http://cbaweb2.med.unc.edu/henson_mrm/VisibleHuman/VHMain.html).
- [25] Hong Hua, Chunyu Gao, Leonard D. Brown, Narendra Ahuja, and Jannick P. Rolland. A testbed for precise registration, natural occlusion and interaction in an augmented environment using a head-mounted projective display (hmpd). In *VR '02: Proceedings of the IEEE Virtual Reality Conference 2002*, page 81, Washington, DC, USA, 2002. IEEE Computer Society.
- [26] N.M. Hylton. Breast Magnetic Resonance Imaging Techniques. *Breast MRI*, 1:7–14, 2005.
- [27] Imagej <http://rsbweb.nih.gov/ij/>.
- [28] B.E. Insko. *Passive Haptics Significantly Enhances Virtual Environments*. PhD thesis, University of North Carolina, 2001.
- [29] N.W. John. Chapter 5 Volume Rendering. *3D Image Processing: Techniques and Clinical Applications*, page 35, 2002.
- [30] John Wiley & Sons, Inc. Real anatomy dvd  
<http://www.wiley.co.uk/alertimages/tortora/realanatomy/index.html>.
- [31] Carmen Juan, Francesca Beatrice, and Juan Cano. An augmented reality system for learning the interior of the human body. *Advanced Learning Technologies, 2008. ICALT '08. Eighth IEEE International Conference on*, pages 186–188, July 2008.
- [32] C.C. Kai, C.S. Meng, L.S. Ching, E.K. Hoe, and L.K. Fah. Rapid prototyping assisted surgery planning. *The International Journal of Advanced Manufacturing Technology*, 14(9):624–630, 1998.

- [33] J. Kim, S. Eberl, and D. Feng. Visualizing Dual-Modality Rendered Volumes Using a Dual-Lookup Table Transfer Function. *Computing in Science & Engineering*, 9(1):20–25, 2007.
- [34] M.S. Landy, L.T. Maloney, E.B. Johnston, and M. Young. Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research*, 35(3):389–412, 1995.
- [35] A. Lee, J. Deutinger, and G. Bernaschek. Three dimensional ultrasound: abnormalities of the fetal face in surface and volume rendering mode. *BJOG: An International Journal of Obstetrics & Gynaecology*, 102(4):302–306, 1995.
- [36] T.Y. Lee and C.R. Yan. Stylized rendering for anatomic visualization. *Computing in Science & Engineering*, 9:13–19, 2007.
- [37] Robert Livingston. Raaviewer homepage <http://web.mac.com/rlivingston/RAAViewer/Downloads.html>.
- [38] William E. Lorensen and Harvey E. Cline. Marching cubes: A high resolution 3d surface construction algorithm. In *SIGGRAPH '87: Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, pages 163–169, New York, NY, USA, 1987. ACM.
- [39] B. Marro, C.A. Valery, A. Bitard, M. Sahel, A. Zouaoui, B. Randoux, C. Oppenheim, and C. Marsault. Intracranial Aneurysm on CTA: Demonstration Using a Transparency Volume-Rendering Technique. *Journal of Computer Assisted Tomography*, 24(1):96, 2000.
- [40] Paul Marshall. Do tangible interfaces enhance learning? In *TEI '07: Proceedings of the 1st international conference on Tangible and embedded interaction*, pages 163–170, New York, NY, USA, 2007. ACM Press.

- [41] A.J. Megibow. Three-D offers workflow gains, new diagnostic options. *DIAGNOSTIC IMAGING-SAN FRANCISCO*-, 24(12):83–94, 2002.
- [42] E. Merz, G. Weber, F. Bahlmann, and D. Miric-Tesanic. Application of transvaginal and abdominal three-dimensional ultrasound for the detection or exclusion of malformations of the fetal face. *Ultrasound in Obstetrics & Gynecology*, 9(4):237–243, 1997.
- [43] P. Milgram and F. Kishino. A Taxonomy of Mixed Reality Visual Displays. *IEEE TRANSACTIONS on Information and Systems*, 77(12):1321–1329, 1994.
- [44] L. Moody, A. Waterworth, A.D. McCarthy, P.J. Harley, and R.H. Smallwood. The feasibility of a mixed reality surgical training environment. *Virtual Reality*, 12(2):77–86, 2008.
- [45] L. Naimark, E. Foxlin, and I.S. Inc. Circular data matrix fiducial system and robust image processing for a wearable vision-inertial self-tracker. *Mixed and Augmented Reality, 2002. ISMAR 2002. Proceedings. International Symposium on*, pages 27–36, 2002.
- [46] A.J. Newens. Bodyworks for windows v5. 0. *Intensive & Critical Care Nursing*, 12(5):312–312, 1996.
- [47] G.L. Nieder, J.N. Scott, and M.D. Anderson. Using QuickTime virtual reality objects in computer-assisted instruction of gross anatomy: Yorick - the VR Skull. *Clinical Anatomy*, 13(4):287–293, 2000.
- [48] S.O. Onuh and Y.Y. Yusuf. Rapid prototyping technology: applications and benefits for rapid product development. *Journal of Intelligent Manufacturing*, 10(3):301–311, 1999.
- [49] R. Pabst. Gross anatomy: an outdated subject or an essential part of a modern medical curriculum? *Anat Rec*, 237:431–433, 1993.



- [50] B. Panchaphongsaphak, R. Burgkart, and R. Riener. Three-Dimensional Touch Interface for Medical Education. *Information Technology in Biomedicine, IEEE Transactions on*, 11(3):251–263, 2007.
- [51] G. Papagiannakis, G. Singh, and N. Magnenat-Thalmann. A survey of mobile and wireless technologies for augmented reality systems. *Computer Animation and Virtual Worlds*, 19(1), 2008.
- [52] E.S. Pretorius and E.K. Fishman. Volume-rendered Three-dimensional Spiral CT: Musculoskeletal Applications 1, 1999.
- [53] Primal pictures 3d human anatomy medical software  
<http://www.primalpictures.com/>.
- [54] David Robertson, William Johnston, and Wing Nip. Virtual Frog Dissection: Interactive 3D Graphics Via the Web. In *Proceedings, The Second International WWW Conference '94: Mosaic and the Web*, 1994.
- [55] J. Rolland, L. Davis, F. Hamza-Lup, J. Daly, Y. Ha, G. Martin, J. Norfleet, R. Thumann, and C. Imielinska. Development of a training tool for endotracheal intubation: distributed augmented reality. *Stud Health Technol Inform*, 94:288–94, 2003.
- [56] J. Rolland, C. Meyer, L. Davis, F. Hamza-Lup, J. Norfleet, C. Imielinska, and K. F. Kerner. Merging augmented reality and anatomically correct 3d models in the development of a training tool for endotracheal intubation. In *Biomedical Imaging, 2002. Proceedings. 2002 IEEE International Symposium on*, pages 895–898, 2002.
- [57] J. Rolland, C. Meyer, L. Davis, F. Hamza-Lup, J. Norfleet, C. Imielinska, and KF Kerner. Merging augmented reality and anatomically correct 3D models in the development of a training tool for endotracheal intubation. *Biomedical*

- Imaging, 2002. Proceedings. 2002 IEEE International Symposium on*, pages 895–898, 2002.
- [58] J. P. Rolland, D. L. Wright, and A. R. Kancherla. Towards a Novel Augmented-Reality Tool to Visualize Dynamic 3-D Anatomy. *Medicine Meets Virtual Reality: Global Healthcare Grid*, 1997.
- [59] Chris Rorden. DICOM introduction and free software <http://www.sph.sc.edu/comd/rorden/dicom.html>.
- [60] H.G. Schmidt. Problem-based learning: rationale and description. *Medical Education*, 1983(17):11–16, 1983.
- [61] M. Schnaider, B. Schwald, H. Seibert, and T. Weller. Medarpa—a medical augmented reality system for minimal-invasive interventions. *Stud Health Technol Inform*, 94:312–4, 2003.
- [62] W. Schroeder, K. Martin, and B. Lorensen. *The Visualization Toolkit An Object-Oriented Approach To 3D Graphics*. Kitware Inc., 2004.
- [63] M. Seitel, L. Maier-Hein, U. Rietdorf, S. Nikoloff, A. seitel, A. Franz, H. Kengott, M. Karck, R. De Simone, I. Wolf, and H.-P. Meinzer. Mixed reality for cardiac surgery planning.
- [64] B.E. Shelton and N.R. Hedley. Using augmented reality for teaching earth-sun relationships to undergraduate geography students. In *First IEEE International Augmented Reality Toolkit Workshop, Darmstadt, Germany*. Citeseer, 2002.
- [65] Jimmy Lee Shreeve. New laws lead to a shortage of human cadavers for use in medical research - independent news article <http://www.independent.co.uk/life-style/health-and-wellbeing/health-news/new-laws-lead-to-a-shortage-of-human-cadavers-for-use-in-medical-research-450864.html>.

- [66] T. Sielhorst, C. Bichlmeier, S.M. Heining, and N. Navab. Depth Perception-A Major Issue in Medical AR: Evaluation Study by Twenty Surgeons. *Lecture Notes in Computer Science*, 4190:364–372, 2006.
- [67] Susan Standring, editor. *Gray's Anatomy: the Anatomical Basis of Clinical Practice*. Churchill Livingstone, 40<sup>th</sup> edition, 2008.
- [68] J.E. Swan II, A. Jones, E. Kolstad, M.A. Livingston, and H.S. Smallman. Ego-centric depth judgments in optical, see-through augmented reality. *IEEE Trans Vis Comput Graph*, 13(3):429–42, 2007.
- [69] Patrick W Tank. *Grant's Dissector*. Lippincott Williams & Wilkins, 14<sup>th</sup> edition, 2008.
- [70] Bruce Thomas, Ben Close, John Donoghue, John Squires, Phillip de Bondi, Michael Morris, and Wayne Piekarski. Arquake: An outdoor/indoor augmented reality first person application. In *ISWC '00: Proceedings of the 4th IEEE International Symposium on Wearable Computers*, page 139, Washington, DC, USA, 2000. IEEE Computer Society.
- [71] RY Tsai. A efficient and accurate camera calibration technique for 3D machine vision. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR 86)*, pages 364–374, 1986.
- [72] F.P. Vidal, N.W. John, A.E. Healey, and D.A. Gould. Simulation of ultrasound guided needle puncture using patient specific data with 3D textures and volume haptics. *Computer Animation and Virtual Worlds*, 19(2):111–127, 2008.
- [73] virtusMED <http://www.virtusmed.info>.
- [74] The national library of medicine's visible human project [http://www.nlm.nih.gov/research/visible/visible\\_human.html](http://www.nlm.nih.gov/research/visible/visible_human.html).



- [75] Visible korean human <http://vkh.ajou.ac.kr/vkh.ajou.ac.kr/index.htm>.
- [76] VOXEL-MAN - surgery simulators and virtual body models <http://www.voxel-man.de>.
- [77] P.A. Webb. A review of rapid prototyping (RP) techniques in the medical and biomedical sector. *Journal of Medical Engineering & Technology*, 24(4):149–153, 2000.
- [78] Diana F Wood. ABC of learning and teaching in medicine: Problem based learning. *BMJ*, 326(7384):328–330, 2003.
- [79] Paul A. Yushkevich, Joseph Piven, Heather Cody Hazlett, Rachel Gimpel Smith, Sean Ho, James C. Gee, and Guido Gerig. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*, 31(3):1116–1128, 2006.
- [80] W.D. Zimmer, T.H. Berquist, R.A. McLeod, F.H. Sim, D.J. Pritchard, T.C. Shives, L.E. Wold, and G.R. May. Bone tumors: magnetic resonance imaging versus computed tomography. *Radiology*, 155(3):709–718, 1985.
- [81] Zygote Media Group, Inc. 3D Models and 3D Animations of the Male and Female Anatomy <http://www.3dscience.com/>.