Sheffield Hallam University

A Sheffield Hallam University thesis

The Use of Consumer Depth Cameras for Calculating Body Segment Parameters

Sean Alan Clarkson

A thesis submitted in partial fulfilment of the requirements of
Sheffield Hallam University
for the degree of Doctor of Philosophy

June 2015

ABSTRACT

Body segment parameters (BSPs) are pivotal to a number of key analyses within sports and healthcare. Accuracy is paramount, as investigations have shown small errors in BSPs to have significant impact upon subsequent analyses, particularly when analysing the dynamics of high acceleration movements. There are many techniques with which to estimate BSPs, however, the majority are complex, time consuming, and make large assumptions about the underlying structure of the human body, leading to considerable errors. Interest is increasingly turning towards obtaining person-specific BSPs from 3D scans, however, the majority of current scanning systems are expensive, complex, require skilled operators, and require lengthy post processing of the captured data. The purpose of this study was to develop a low cost 3D scanning system capable of estimating accurate and reliable person-specific segmental volume, forming a fundamental first step towards calculation of the full range of BSPs.

A low cost 3D scanning system was developed, comprising four Microsoft Kinect RGB-D sensors, and capable of estimating person-specific segmental volume in a scanning operation taking less than one second. Individual sensors were calibrated prior to first use, overcoming inherent distortion of the 3D data. Scans from each of the sensors were aligned with one another via an initial extrinsic calibration process, producing $360°$ colour rendered 3D scans. A scanning protocol was developed, designed to limit movement due to postural sway and breathing throughout the scanning operation. Scans were post processed to remove discontinuities at edges, and parameters of interest calculated using a combination of manual digitisation and automated algorithms.

The scanning system was validated using a series of geometric objects representative of human body segments, showing high reliability and systematic over estimation of scan-derived measurements. Scan-derived volumes of living human participants were also compared to those calculated using a typical geometric BSP model. Results showed close agreement, however, absolute differences could not be quantified owing to the lack of gold standard data. The study suggests the scanning system would be well received by practitioners, offering many advantages over current techniques. However, future work is required to further characterise the scanning system's absolute accuracy.

ACKNOWLEDGEMENTS

PERSONAL BIBLIOGRAPHY

Clarkson, S. et al., 2015. Assessment of a Microsoft Kinect-based 3D scanning system for taking body segment girth measurements: a comparison to ISAK and ISO standards. Journal of Sports Sciences.

Clarkson, S. et al., 2014. Assessing the Suitability of the Microsoft Kinect for Calculating Person Specific Body Segment Parameters. In 4th IEEE Workshop on Consumer Depth Cameras for Computer Vision, European Conference on Computer Vision. Zurich, Switzerland: IEEE.

Clarkson, S. et al., 2013. Distortion Correction of Depth Data from Consumer Depth Cameras. In 3D Body Scanning Technologies. Long Beach, California: Hometrica Consulting, pp. 426–437.

Clarkson, S. et al., 2012. Calculating Body Segment Inertia Parameters from a Single Rapid Scan Using the Microsoft Kinect. In 3D Body Scanning Technologies. Lugano, Switzerland: Hometrica Consulting, pp. 153–163.

CONTENTS

iv

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

Abbreviations

FOV	Field of view
ICC	Intra-class correlation coefficient
TEM	Technical error of measurement
ISAK	International Society for the Advancement of Kinanthropometry
RGB-D	A sensor capable of obtaining colour data (RGB) and 3D depth (D)
NUI	Natural user interaction
BSP	Body segment parameter
FPS	Frames per second
ASIS	Anterior superior iliac spine
OEM	Original equipment manufacturer
ICP	Iterative closest point
PCA	Principal component analysis
CSV	Comma separated values
RMS	Root mean square
BMI	Body mass index
BVI	Body volume index

Symbols

uv	2D image space coordinates, comprising horizontal and vertical components
uvZ	Coordinates comprising 2D horizontal and vertical components, and a 3D depth
XYZ	3D real world components, comprising horizontal, vertical, and depth components

1   INTRODUCTION

The following chapters document the culmination of a three year programme of study into the development of a low cost 3D body scanning system, with which person specific segmental volume can be estimated.

1.1   Motivation for the research

Within biomechanical and healthcare communities, reliable estimates of person specific body segment parameters (BSPs) are desirable for a number of analyses (Lerch et al. 2006). For example, biomechanics communities rely on the availability of accurate measures of segment volume, which are used in conjunction with inverse dynamics models to calculate joint force and power (Lerch et al. 2006; Piovesan et al. 2011). Accuracy is paramount (Durkin et al. 2002), as small changes in parameter values have been shown to greatly influence subsequent calculations, particularly when high accelerations are involved (Damavandi et al. 2009; Pearsall & Costigan 1999; Rao et al. 2006; Piovesan et al. 2011).

Other applications lie within elite sport, where anthropometric assessment for monitoring and assessing fitness and performance of athletes is ever increasing in popularity (Williams & Reilly 2000; Pienaar et al. 1998; Mohamed et al. 2009; Bullock et al. 2009; Hoare & Warr 2000). For example, anthropometric assessment can be used within training environments to track the change in an athlete's physique over time, and in response to specific training interventions (Kerr et al. 1995). However, current manual measurement techniques are time consuming, and can be prohibitive to normal activities (Schranz et al. 2010). 3D scanning systems offer the ability to obtain such measurements in a much quicker manner, whilst also providing much more information about the athletes, such as BSPs. Scans collected over longitudinal periods can also be overlaid, allowing analysis of both the magnitude and location of any change in body morphology.

Healthcare communities are also increasingly demanding accurate measures of body segment volume (Robinson et al. 2012) to use in conjunction with the newly developed body volume index (BVI) (Rahim & Barnes 2009), designed to improve upon the

currently used body mass index (BMI). Other applications include the treatment of lymphedema patients who suffer from localised fluid retention throughout the body, and are often subject to lengthy anthropometric measurement processes (Armer & Stewart 2005). The experience of such patients would be immensely improved with a quick and accurate non-contact surface scanning system, which is able to obtain measures and distribution of segmental volume. Similarly, breast surgeons are currently seeking low cost, accurate, and non-invasive techniques of measuring breast volume (Cardoso et al. 2012): enabling a more objective method of selecting the correct implant to use for an operation (Cardoso et al. 2012). There are many other applications of segmental volume and anthropometric measures within the healthcare sector, including weight loss and management, child growth monitoring and tracking, tracking the growth of a baby during pregnancy, and more widely, the use of 3D scanning for diagnosing and tracking the onset or development of particular medical conditions (including conditions such as curvature of the spine and body asymmetry).

Previous studies have used medical imaging and scanning systems (DEXA, MRI and CT) to obtain accurate subject specific BSP estimates (Bauer et al. 2007; Wicke et al. 2008; Pearsall et al. 1996; Cheng et al. 2000; Martin et al. 1989). However, the required investment (Martin et al. 1989), lengthy scan time of MRI (Martin et al. 1989), and health risks of DEXA and CT (Pearsall et al. 1996) have led to their criticism as viable methods.

Data tables (Dempster 1955) and regression equations based upon cadaver data (Zatsiorsky 1983; Leva 1996) have proven a popular method of estimating BSPs, owing to their quick, easy, and cost effective techniques. However, the use of models and historical data results in BSPs of an inherently generic nature (Gittoes & Kerwin 2006) that are not specific to the individual being studied. Such methods are also criticised as the data underpinning the models typically comes from small sample groups (Pearsall & Reid 1994) that have a lack of gender and racial diversity (Cheng et al. 2000) - typically originating from elderly male Caucasians - (Zatsiorsky 1983; Leva 1996), and are therefore unrepresentative of the wider population.

Geometric modelling techniques (Hanavan 1964; Yeadon 1990b) involve more measurements of the body, requiring up to 30 minutes of manual measurement.

However, they have been shown to offer significant accuracy improvements (Challis 1999). Wicke and Dumas (2010) and Challis (1999) suggested that limb segments can be reliably modelled using geometric shapes, but the trunk segment is less reliably represented. For example, participants with a large stomach may not be well represented by the stadium solid shapes used in Yeadon's model (Yeadon 1990b), leading to an underestimation of volume. The complexity of approximating the trunk segment is further increased due to its likelihood of changing shape during the breathing cycle (Challis 1999), leading to poor accuracy and reliability. These limitations have left researchers seeking alternative methods of estimating BSPs that are quick, offer greater accuracy, take into account the very specific nature of body segment shape, and are low cost: in order they can be used within typical research laboratory, clinical, and training environments.

Handheld laser scanners have previously been used to obtain subject specific BSPs within laboratory (Sheets et al. 2010; Lerch et al. 2006) and training environments (Schranz et al. 2010). Although offering high point accuracy (Nikon Corp 2011) without associated health risks (Henderson & Schulmeister 2003), their conventionally cited accuracy may be reduced when scanning living humans due to the possibility of involuntary movement over the lengthy scanning duration (~ 30 minutes). Full body scanners based on laser scanning (Vitronic 2011) offer shorter scan times (< 20 seconds), but are prohibitively expensive for the majority of sports and healthcare research laboratories (Weiss et al. 2011). Recently, (Wicke & Dumas 2010) suggested that scanning systems based upon structured light would be a suitable low cost alternative technique with which to obtain 3D scans, enabling estimation of person specific BSPs.

The recent interest in natural user interaction (NUI) has led to the development of low cost (in the region of £200 (Amazon 2012)) depth cameras such as the Microsoft Kinect® (Microsoft Corporation, Redmond, USA) and Asus Xtion Pro (ASUSTeK, Taipei, Taiwan), able to capture human motion in 3D (Boehm 2012). Commonly using a combination of 3D structured light scanning (Shpunt & Zalevsky 2009) and computer vision techniques (Shotton et al. 2011), the underlying 3D vision system can be utilised to obtain 3D point cloud scan data at a rate of 30Hz (Khoshelham 2010): providing a low cost method of 3D scanning. Their launch has led to significant interest in a range

of communities including: robotics (Henry et al. 2012), body scanning (Boehm 2012), healthcare (Labelle 2011), graphics (Izadi et al. 2011) and apparel (Stampfli et al. 2012).

Despite this, there have been few studies investigating the accuracy, reliability, and validity of raw measurement data provided by such devices. Recent studies typically focus on simple measurements (such as Euclidian distances and plane fitting residuals) from single sensors (Boehm 2011; Menna et al. 2011; Khoshelham 2010), whereas a scanning system for anthropometric assessment would typically comprise multiple sensors (Boehm 2012), and involve more complex measurements (such as girths and surface distances), possibly leading to a compounding of error. Nevertheless, the results suggest potential that accurate and reliable person specific BSPs could be calculated from the 3D scans obtained from such sensors.

Research has suggested there is a need for a new method of obtaining accurate and reliable person specific BSPs. BSPs calculated with the system should meet sector requirements for reliability, whilst accuracy should exceed that offered by existing BSP estimation techniques. The scanning method must be timely, taking less than 10 minutes per person: representing a significant time saving when compared to current techniques and hence permitting use within sports training environments without being prohibitive to normal activities. The method must also be low cost, costing in the order of £2000, rather than in the region of £20,000 as is typical of most commercial scanning systems. A cost of this order would increase the likelihood of healthcare and biomechanics communities accepting the system as a valid alternative to the freely available current techniques.

It appears that a 3D scanning system based upon depth cameras, capable of obtaining person specific BSPs from 3D scans would meet this need. Assuming the scanning system meets the requirements highlighted above, it is anticipated such a system would be greatly appreciated and accepted by both biomechanics and healthcare communities. Firstly, a scanning system holds the potential to improve reliability and accuracy of the BSPs offered by current techniques, owing to their person specific nature. A scanning system may also improve the experience of participants, patients, or athletes which must undergo anthropometric assessment and profiling exercises, taking far less time, and

being less invasive than conventional manual measurement techniques. Furthermore, the significant time savings offered by a scanning system in comparison to current techniques suggests measurement and assessment exercises could be completed in environments where it is currently infeasible due to time constraints. For example, anthropometric assessment could feasibly become further integrated within elite sport training and competition environments, providing coaches and athletes with much richer information.

For the purposes of this programme of study, it was decided to solely concentrate on estimating segmental volume, forming a first step towards calculating the full range of BSPs from person specific 3D scans. Once a technique to reliably and accurately calculate segment volume has been demonstrated, then further work may focus upon obtaining the range of additional BSPs.

1.2 Aims and objectives

Aim
- To develop a 3D body scanning system that is capable of obtaining estimates of person specific segmental volume with accuracy greater than current techniques, reliability in excess of clinical requirements, takes less than 10 minutes to obtain a scan of an individual, and costs less than £2000 in hardware.

Objectives
- To investigate and critique the range of 3D scanning systems currently available as commercial products or academic research output.
- To develop a 3D body scanning system which is simple to use, meaning that it can be used by non-specialists
- To develop a range of user application software and data analysis algorithms to allow person specific segmental volume to be calculated from the 3D scans.
- To perform initial system validation studies using geometric objects representative of human body segments.
- To test the system on a range of living human participants, comparing BSPs obtained from the system to those obtained using currently accepted techniques.

The overall aim of the research is therefore to answer the question: "Can low cost consumer 3D surface scanning techniques be used to obtain estimates of person-specific segmental volume with accuracy greater than current techniques and a reliability meeting clinical requirements."

1.3 Thesis structure

The chapters in this thesis form four main parts, namely: an introduction and background to the area of BSPs and 3D scanning, the system development process and testing methodology, results of validation studies, and final discussions.

The first stage of the project was to develop a simple to use and low cost 3D scanning system, in conjunction with a range of algorithms and techniques in order that person specific segmental volume could be calculated from the 3D scans. These stages are discussed in chapter 3, 4, and 5. Before testing the scanning system on living participants, a series of validation tests were carried out to evaluate the scanning system and data analysis techniques. Validation studies were carried out using geometric objects representative of typical body segments, discussed in chapter 6. Analysis of the results provided an indication of accuracy and reliability obtained under ideal conditions and limited external influential factors, providing a fundamental assessment of performance with which the results of subsequent studies could be compared. The system was then tested on a large cohort of living human participants, comparing volume estimated with the scanning system to that calculated using a currently accepted geometric modelling technique. Results were also compared to those previously obtained as part of the validation study, providing background to likely real world accuracy of the system, and differences in reliability which may be attributable to human influential factors. Implications of the system upon current practice, and the likelihood of acceptance of the system in relevant sectors are discussed in chapter 7 and 8.

2 LITERATURE REVIEW

2.1 Introduction

The aim of this program of study is to develop a low cost 3D scanning system, capable of calculating segmental volume and representing a first step towards calculating the full range of BSPs from person-specific 3D scans. In order to fulfil this aim, knowledge of existing BSP estimation techniques and their shortcomings is required. Knowledge of suitable 3D scanning systems and techniques for handling and processing 3D scan data is also required. This chapter provides an overview of relevant literature in such areas.

2.2 Body segment parameters

In the study of human motion, the relationship between forces, moments, and motion is of common interest. Such analysis – referred to as the kinetics of human motion – can either be approached as a forward or inverse dynamics problem. Inverse dynamics is used to calculate the joint force, torque, and power required to achieve an observed sequence of human motion. Similarly, forward dynamics uses known joint forces, torques, and powers to infer the expected motion. In each case, knowledge of segmental volume, mass, moment of inertia, and centre of mass is required: properties collectively referred to as body segment parameters (BSPs).

The ability to obtain accurate BSPs is paramount (Durkin et al. 2002) as many studies have shown errors to have a large effect upon subsequent kinetic analysis (Challis 1996; Rao et al. 2006; Piovesan et al. 2011), particularly in movements involving high accelerations (Kwon 1996; Chiu & Salem 2005; Arampatzis et al. 1997). For these reasons, users of BSP data are continually seeking methods of obtaining BSPs of increased accuracy and reliability.

Importantly, the use of BSPs spans a number of applications within both sports and healthcare environments. For example, BSPs may be used within sports environments to determine the maximal forces, torque, and power which can be exerted by an athlete in order to determine whether training goals have been met. Similarly, BSPs may be

used for anthropometric profiling as part of talent identification programmes (Williams & Reilly 2000; Pienaar et al. 1998; Mohamed et al. 2009), or tracking the effectiveness of training routines (Kerr et al. 1995). In healthcare applications, BSPs may be used to profile the gait of an individual who has undergone an injury or surgery, in order to ensure they are not susceptible to future injury (Pearsall & Costigan 1999). Use of segmental volumetric information is integral to the newly developed body volume index (BVI), designed to replace the currently used body mass index (BMI): improving accuracy and relevance (Robinson et al. 2012). Other healthcare applications of volumetric data include the diagnosis and characterisation of lymphedema (Hughes & Lau 2008), and measurement of breast volume for pre and post-operative assessment (Choppin et al. 2013).

2.3   Techniques for estimating body segment parameters

Early attempts to obtain estimates of body segment parameters (BSPs) date back to the 1800s, and the work by Harless (1860) and Braune and Fischer (1889). Numerous techniques have been developed in the years following, designed to overcome limitations in previous techniques, and improve accuracy and reliability. BSP estimation has been approached using a wide variety of methods, which can be grouped into the following categories.

2.3.1   Predictive models from cadaver studies

Early studies focussing on body segment parameters date back to the work by Harless (1860) and Braune and Fischer (1889). Harless dissected two male cadavers (29 year old, decapitated ex-prisoners) into eighteen individual segments, defined by key joint pivotal axes. Mass was calculated using a balance scale, and centre of mass (CoM) determined by placing the segments on a simple balance board. The volume of each segment was calculated using a postulated total body specific gravity. Harless later verified his early work using seven additional cadavers of mixed age and gender. The same methodology was followed, with the exception that volume was calculated using a water immersion technique and the principals of Archimedes: known as hydrostatic weighing. Harless concluded that age and gender had an effect on the position of CoM, and density varied from the postulated value between segments. The findings of

Harless' study were presented as a component model, allowing users to estimate segmental mass, volume, and CoM knowing basic measurements such as total body mass and height.

Similarly, Braune and Fischer developed a fourteen segment model for calculating mass and CoM, based upon three male cadavers (average build German soldiers who had committed suicide). Braune and Fischer criticised the work by Harless, suggesting the loss of blood and fluids from the segments - accounted for via an average offset factor – will have degraded the accuracy and reliability of the results. In response, the cadavers used by Braune and Fischer were kept frozen throughout the investigation to prevent loss of blood and fluid. The cadavers were segmented about key joints using approximate centres of rotation. Segmental mass was calculated using a balance scale, and CoM calculated by vertically suspending the segment about three axes and the intersection of three externally fixed planes: assumed to be the centre of mass. The thorough nature of the study meant it was used as standard laboratory practice for over half a century (Bjornstrup 1995), however, significant questions have been raised regarding the reliability of the joint centres used for segmentation and the impact that freezing the segments had upon CoM (Dempster & Gaughran 1889).

Further advancement was made by Dempster (1955), as part of the US Air Force's investigations into human segmental properties and the impact upon cockpit design. Dempster conducted the largest cadaveric study at the time, dissecting eight male deceased war veterans (smaller and weighing less than an average male, but representative of their age). Before segmentation about joint centres of rotation, Dempster flexed the segments to mid-range, citing this would provide a more even mass distribution. Like Braune and Fischer, Dempster also froze the limb segments after segmentation. Mass was calculated using a balance scale, and volume calculated using hydrostatic weighing. CoM was calculated using a knife edge balance technique, and moments of inertia (MOI) calculated using a pendulum technique. Like previous studies, Dempster also created a component model, allowing mass to be estimated from total body mass, whilst CoM and MOI could be estimated from segment length.

Barter (1957) later combined the data from (Braune & Fischer 1889) and (Dempster 1955) to produce regression equations with the use of stepwise regression analysis.

Barter believed these equations would provide more reliable segmental mass estimates than the component models presented in previous research.

Clauser et al (1969) and Chandler et al (1975) suggested the approximate joint centres and axes of rotation used as the segmentation points in previous studies were hard to locate, and potentially introduced large errors in previous studies. Segmenting six and thirteen cadavers respectively, they instead chose to dissect the segments about clearly palpable bony landmarks, citing this would lead to greater accuracy and reliability.

Whilst cadaveric studies have proved a popular and effective method of developing data tables and models with which to estimate BSPs of living participants, all the previous studies have significant shortfalls. For example, previous studies all use very small sample sizes – due to the difficulty in obtaining cadavers, and the cost and intricacy of the techniques – meaning results are likely unrepresentative of the wider population. Similarly, the cadavers are typically of elderly Caucasian males, likely making extrapolation of the results to different populations, ages, and genders very inaccurate (Behnke 1959; Hinrichs 1985; Pataky et al. 2003; Cheng et al. 2000). Unfortunately all the studies adopted their own segmentation techniques, preventing the data being combined to produce more representative models.

Cadavers offer additional limitations, due to the loss of blood and fluid during segmentation (Reid & Jensen 1990), and the differences between living and deceased tissue (Pearsall & Reid 1994). Furthermore, the way the person died will likely have a large impact upon mass and CoM measurements (Reid & Jensen 1990). For example, Harless (1860) used cadavers which had been decapitated and are likely to have suffered large losses of blood and fluid at the point of death: consequently weighing much less than the cadaver of someone who died from natural causes.

2.3.2 Person specific direct measurement techniques

Limitations in BSP predictive models derived from cadaveric studies and the resulting effect upon accuracy and reliability are commonly accepted by researchers, which has led to the development of techniques to directly measure person specific BSPs using in vivo kinematic and kinetic techniques. The use of person specific measurements to

estimate BSPs – rather than relying on models or assumptions – is likely to greatly improve reliability, as model assumptions, generalisations, and bias are eliminated.

Water immersion techniques (Hughes 2005; Hughes & Lau 2008; Piovesan et al. 2011; Bjornstrup 1995) relying on Archimedes principals have been a popular (Drillis et al. 1966; Dempster & Gaughran 1889; Bjornstrup 1995; Pearsall & Reid 1994) method of estimating person specific segmental volume. For example, Plagenhoef (1983) calculated segmental inertia parameters of one hundred and thirty six living participants (one hundred females and thirty five males) using a combination of water immersion techniques – to obtain volume measures – and Dempster's equations (1955) to estimate mass, CoM, and radius of gyration from the volumetric data. Results were published in the form of a gender specific segmental inertia parameter model, enabling others to use Plagenhoef's values within their studies. As the study involved a large number of participants of both male and female genders, it likely provides greater representation of the population than many previous studies. However – like many other studies - Dempster's uniform segment density values were used as part of the process to obtain segmental mass from the volumetric data. This approach is often criticised as the segment density values were derived from cadavers, and therefore unlikely to be representative of living tissue (Behnke 1959; Mungiole & Martin 1990). Furthermore, there are a number of practical issues associated with suspending segments in water receptacles within a laboratory environment for example.

Other person specific in vivo techniques typically make use of the motion capture and force plate systems (Chen et al. 2011; Pataky et al. 2003; Kingma et al. 1995; Fuschillo et al. 2012; Damavandi et al. 2009) typically found within biomechanics laboratories to estimate person specific BSPs. For example, Chen et al (2011) compared BSP estimates calculated using the models by Dempster (1955), Zatsiorsky (1983), and Cheng (2000) to those obtained using a force plate and motion capture system. Twelve participants were recruited, aged 24 ± 2 years. Participants were asked to displace body segments as far from the body as possible whilst standing on a force plate, with the body movement recorded using a motion capture system. These data allowed segmental mass, CoM, and moments of inertia to be calculated. Results showed comparable segmental masses and moments of inertia derived using the four techniques, however large differences were observed in CoM (up to 50.4% of segment length). Ostensibly,

BSPs derived using Chen et al's method produced better estimates of whole body centre of pressure and ground reaction force than the three investigated BSP models. As discussed in previous studies, results highlighted performance of the predictive models to be highly dependent upon the physique of the participants (Durkin & Dowling 2003).

Hatze (1975) developed a simple technique for measuring BSPs, based upon an oscillation technique. The method requires a body segment to be set into oscillation with an instrumented spring, ensuring the muscles are relaxed. This creates a spring damper system, enabling mathematical equations based upon small oscillation theory (Gregory 2006) to be used to calculate the required inertial parameters. Importantly however, the method cannot be used on the trunk segment as it cannot be fully separated from its adjoining segments.

Studies focussing upon BSPs derived using direct measurement techniques suggest they are able to produce more accurate and reliable BSPs than those derived using model based approaches. However, such techniques are inherently time consuming, require specialist equipment, specialist training, and are often impractical. Therefore, such techniques appear unsuitable for use within sports training and healthcare environments, largely owing to the time demands that would likely be prohibitive to normal activities (Schranz et al. 2010).

2.3.3 Medical imaging and scanning

Limitations in the previously discussed approaches, combined with a wider availability of scanning systems, has led to an increased interest in the use of medical imaging and scanning techniques (such as MRI, DEXA, and CT) to obtain BSP estimates, offering a number of advantages. Firstly, 3D scans offer individual specific BSP estimates that do not make significant assumptions about the shape of the body. Secondly, such scanning techniques provide information about the structure of the segments underneath the skin. This means uniform density assumptions do not have to be used when calculating mass, instead allowing use of varying density values derived directly from the person being scanned. However, use of uniform density assumptions has been shown to have little impact upon subsequent calculations (Wicke & Dumas 2010). Thirdly, they eliminate the need for any form of anthropometric measurement, thereby eliminating the inherent

error in anthropometric measurements (O'Haire & Gibbons 2000). However, it is important to note the health risks associated with medical scanning techniques such as DEXA and CT (Pearsall et al. 1996) have led to their criticism as viable methods.

Medical magnetic resonance imaging (MRI) scanners are arguably the most technologically advanced medical scanning system, based upon the principal of a large magnetic field. The body naturally contains a large number of hydrogen atoms, each spinning in their own magnetic field. When the MRI scanner's magnetic field is turned on, the atoms line up, pointing in either a north or south direction. A radio frequency pulse is then applied to the body, which causes all the hydrogen atoms to become aligned in the same direction. When the radio frequency pulse is turned off, the atoms return to point in their original direction after the magnetic source was enabled. This rotation causes energy to be emitted from the atoms, previously absorbed from the radio frequency pulse. The emitted energy is detected by coils within the scanner, enabling scans of the body to be produced. MRI scanners produce scans in the form of multiple cross-sectional 'slices' (figure 2.1) which can be examined in detail to look at a specific area of the body. A scan can take anything between fifteen and ninety minutes, depending upon the area of the body being investigated (Gould & Edmonds 2010).

Figure 2.1 – A typical MRI scan cross-sectional view of the brain (Dilmen, 2006)

Martin et al (1989) developed a method of obtaining person-specific BSP estimates from MRI scans. Although the method was tested upon eight baboon cadavers, the method is directly applicable to living humans. The MRI-derived cross sectional slices were used in conjunction with scan-derived density values to produce BSP estimates. Centre of mass and moments of inertia showed good agreement (mean differences of - 2.4% and 4.4% respectively) between those derived from criterion techniques (balance (Drillis et al. 1966) and oscillation techniques (Dempster 1955) respectively) and the MRI scans. Volume estimates calculated from the MRI scans were found to be systematically greater (mean of 6.3%) than those calculated using criterion techniques (hydrostatic weighing).

Computed tomography scanners (CT) work on a similar principal to that of conventional x-ray scanners, also using an x-ray source and detector. Unlike x-ray scanners, the x-ray source and detector is moved around the body in a $360°$ rotation, as well as moving along the length of the body. Multiple scans are captured in the different positions, which are later aligned with one another to produce a model of the body's internal and external structure. CT scanners also work by capturing the scans in cross-sectional 'slices' (figure 2.1), enabling change in the bodies shape and internal structure to be analysed throughout its length. CT scans are comparable to MRI in terms of time required to collect the scan, taking between twenty minutes and an hour (Harris 2002).

Pearsall et al (1996) performed a similar study to that of Martin et al (1989), instead using CT imaging, and testing the method upon four living humans. CT slices were collected at 10 mm intervals, allowing estimation of BSPs whilst also providing segment-specific density values derived from the pixel intensities of the returned scans. Whilst the technique proved effective, Pearsall et al choose to validate the obtained BSPs against those derived from previously published techniques (including cadaver studies, MRI scanning, gamma and x-ray scanning, water immersion, and photographic techniques) rather than the criterion measures used in other comparative studies. Results showed discrepancies when compared to previous techniques (mass differences of up to 10% of whole body mass for example), however Pearsall et al expected this was due to differences in the way the body was segmented, rather than fundamental problems with the methods. Pearsall et al concluded by proposing a standardised

technique of segmenting the human body for the assessment of BSPs, defining a consistent set of segmentation landmarks that should be used in future comparative studies.

Dual energy x-ray absorptiometry (DEXA) scanners are based upon x-ray sources, and are typically used in medical environments for determining the density and health of bones. The person being scanned typically lies on an x-ray detector, with an x-ray source located above them. The x-ray source emits x-rays, which are passed through the body and attenuated by differing amounts owing to the density of the material they are passed through. The intensity of the detected x-rays is recorded by the detector and later converted into a colour coded intensity image, providing an indication of the body's external and internal structure. Some systems use multiple x-ray sourcés of different intensities in order to calculate density, cited to achieve greater accuracy (Derrer 2013). DEXA scans are typically quicker than MRI and CT, taking between ten and twenty minutes dependent upon the area being scanned (Derrer 2013).

Durkin et al (2002) conducted a similar study to that of Pearsall et al (1996), instead developing a technique based upon DEXA scanning. The DEXA scans provided both surface geometric information and underlying density values, enabling calculation of person specific BSPs. The method was tested upon eleven living male participants, with the scan derived segmental masses summed and compared to whole body mass derived from criterion weighing techniques.

In addition to the eleven living participants, a plastic cylinder and human cadaver leg were scanned with the system, enabling comparison between centre of mass and moments of inertia derived from the scans and criterion techniques (using a knife edge balancing technique and an oscillation technique respectively). Results showed a mean percentage difference of -1.05% ± 1.32% in whole body mass between the scan-derived values and those derived from criterion techniques. Differences for the cylinder and cadaver leg were under 3.2% for all BSPs apart from the moment of inertia, (14.3% and 8.2% for the cylinder and leg respectively). However, this was attributed to uncertainty in the pendulum balancing criterion technique. When compared to criterion geometric calculations, the average error in the cylinder moments of inertia fell to 2.63%. Similar to previous studies, results again suggest that medical scanning techniques are able to

15

obtain accurate person-specific BSP estimates, however the interpretation and comparison of results is limited due to differences in the way the studies have been conducted.

Durkin and Dowling (2003) later conducted a study to compare BSPs derived from four human body BSP models (Dempster 1955; Hanavan 1964; Zatsiorsky 1983; Zatsiorsky & Seluyanov 1985) to those derived from DEXA scans. Durkin and Dowling recruited participants from four different human populations (male and females, aged 19-30, and aged 55+), expecting to see significant differences in the BSPs reported by each technique, owing to the data used to produce the underlying models. Results showed significant differences between the BSPs reported for the four populations (mass differences up to 13% of whole body mass, CoM differences up to 42% of segment length, and radius of gyration differences up to 33% of segment length), further highlighting the problems previously discussed, and identifying the need for person specific BSPs rather than generic models. Another potential source of discrepancy between techniques was the inclusion of person-specific density values in the DEXA-derived BSPs, whilst the predictive BSP models typically used segmental uniform density assumptions. However, other studies suggest this would only account for small differences (Wicke & Dumas 2010).

Cheng et al (2000) recognised the problems associated with the use of BSP models for calculating BSPs of human populations that are different to the elderly Caucasians upon which the models are typically based. For this reason, Cheng obtained BSP estimates of eight male Chinese participants (aged 26 ± 4) derived from MRI scans. In addition, Cheng obtained equivalent BSPs using the models of Dempster (1955), Clauser et al (1969), Martin et al (1989), and Pearsall et al (1994). Based upon the findings of previous studies, Cheng expected to observe significant differences between BSPs derived from the five methods. Segmental masses showed the most variation, likely owing to the considerably different segment definitions used in the different studies. Cheng et al's segmentation protocol was most comparable to that used by Dempster (1955), ostensibly, results of the two studies showed the closest agreement but still showed considerable differences. For example, larger percentages of whole body mass were suggested to exist in the upper arm (average increase of 4%) and thigh (average increase of 13.6%) when compared to previous techniques. Cheng et al suggested the

examined BSP models were not suitable for application to the Chinese population, instead proposing the developed MRI based approach should be used.

Wicke and Dumas (2008) used a combination of DEXA scanning and photographic techniques to obtain density and volume estimates of 25 female and 24 male participants. The use of DEXA scanning in conjunction with photogrammetric techniques enabled the raw measurements from photographs to be supplemented with density values from the DEXA scans, obtaining segmental mass values. The photogrammetric technique provided volume estimates with less than 5% error, and DEXA scanning provided mass per unit area errors of less than 1%. However, issues with data alignment were found when combining the two data sets in order to obtain segmental mass.

Despite the accuracy and reliability reported in previous studies, medical imaging and scanning techniques are not viable options for widespread BSP studies within research laboratories, healthcare, and sports training environments: owing to their cost, associated health risks, and time required to obtain the scans. A lower cost alternative of equivalent accuracy without the associated health risks is therefore required.

2.3.4 Geometric models

Geometric techniques for estimating BSPs represent the bodies segments using a series of geometric shapes, varying in complexity between the range of available techniques. The shapes are dimensioned based upon a series of anatomical measurements, and segmental properties calculated from the shapes. Unlike previously discussed techniques based upon regression equations and data tables, geometric techniques are more sensitive, and therefore often able to detect small changes in person specific volume and shape (Kingma et al. 1996). As a result, they are often cited to be the most accurate BSP estimation technique (Nigg 1999).

Hanavan (1964) developed a geometric model of the human body, representing the bodies segments using fifteen simplified geometric solids (figure 2.2), dependent upon the segment being modelled. The dimensions and properties of the shapes are calculated using anthropometric measurements of the person being modelled, tailoring

the shapes to the individual. As the shapes are geometrically simple, only twenty five anthropometric measurements of the body are required to appropriately scale and dimension the shapes, meaning the time required with participants is limited.
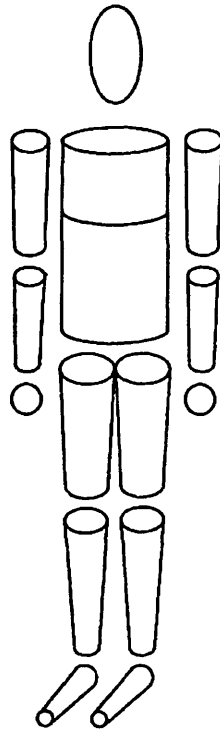
Figure 2.2 - The 15 segment BSP model of Hanavan (1964)

Hanavan assumed a uniform density in all of the bodies segments, greatly reducing the simplicity of the ensuing calculations. Results showed an average of 9.7% error in total body mass when compared to that derived from a balance scale. An error of this magnitude likely arises from two sources. Whilst the geometric shapes representing the bodies segments are scaled and dimensioned to fit the individual using anthropometric measurements taken directly from the body, the shapes are inherently very simple, meaning they are likely to misrepresent or incorrectly report key areas of body mass. Error is also undoubtedly introduced due to the uniform density assumption, which will vary considerably between lean and fat people. However, this is expected to be a smaller contributor to the overall estimation error than the geometric misrepresentation (Wicke & Dumas 2010).

Jensen (1978) developed a geometric technique of representing the human body, choosing to represent it as a series of 20 mm elliptical cross-sections (represented by the

dashed lines in figure 2.3): a technique commonly used in conjunction with MRI and CT scans.
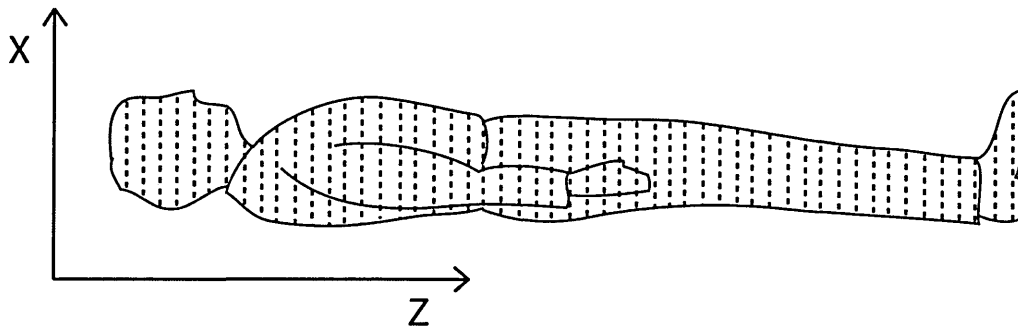
Figure 2.3 - The elliptical body BSP model of Jensen (1978)

Ostensibly this requires a large number of body measurements be taken in order to correctly dimension the elliptical representations. For this purpose, Jensen developed a photographic digitisation technique. Instead of the participant having to undergo a lengthy anthropometric measurement process, the required anthropometric measurements are instead derived from photographs. Photographs are taken from the sagittal and coronial planes (sagittal plane shown in figure 2.3), providing the width and breadth dimensions required to dimension the representative ellipses. Like many other anthropometric techniques (Stewart et al. 2011), Jensen only modelled one arm and leg, assuming symmetry on the other side of the body.

Whilst the time required with the participant is very short, the time required to digitise the two photographs takes around two hours per participant. Like Hanavan, Jensen assumes a uniform density throughout the bodies segments, using the density values published by Dempster (1955). Results show around 2% error in model-derived total body mass when compared to that derived from balance scales, representing a big improvement upon the model of Hanavan (1964). The improvement in accuracy likely arises from the large number of anthropometric measurements, meaning small cross-sectional anatomical features of the body are detected and modelled by the elliptical zones, rather than being averaged throughout the length of a segment by the shapes used in Hanavan's model. In contrast, errors in Jensen's model likely arise from some parts of the body being inaccurately represented by ellipses, leading to an under or over estimation of mass. As before, error is also likely partially introduced by the uniform density assumption.

19

Hatze (1980) developed a geometric model for estimating BSPs, based upon two hundred and forty two anthropometric measurements divided amongst seventeen segments of the body, figure 2.4.
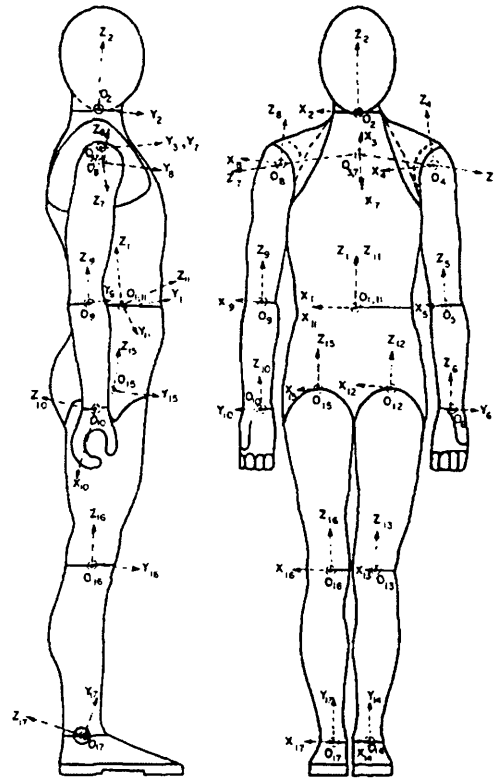
Figure 2.4 - The 17 segment BSP model of Hatze (1980)

Hatze's model offers many advantages over existing techniques. Firstly, the seventeen segments are divided into small mass elements of different geometric structures, allowing the shape of a segment to be modelled in detail. For example, intricate areas of the body such as the shoulders are treated as separate entities rather than including them in the arms or torso. Unlike previous techniques (Hanavan 1964; Jensen 1978), there are no direct assumptions regarding the underlying shape of the body, instead relying solely upon the anthropometric measurements. Hatze's model also differentiates between males and females, defining differing mass distributions for example. Importantly, Hatze's model does not make a uniform density assumption, meaning density is dependent upon the segment being modelled, whilst also including a subcutaneous fat indicator to account for differences in densities of pregnant and obese individuals for example.

Results show the accuracy of whole body mass to be more than 3% on average, with a maximum error of around 5% when compared to balance scales. Hatze suggests a large portion of this accuracy is attributable to the anthropometric measurements being derived from the person directly, rather than digitising photos. Whilst this means a long time with each participant, it results in the accuracy reported above. Although the anthropometric measurement process is very long, the ensuing data processing time is very low, taking a computer program around half a second to produce the BSPs relating to a series of input anthropometric measurements. It is interesting to note that whilst Hatze's model theoretically offers numerous accuracy improvements - owing to the lack of assumptions in the underlying data - it actually produced a greater difference in comparison to criterion techniques than that of Jensen (1978), whilst suffering from the drawback of requiring a significant amount of time with each participant.

Yeadon (1990b) developed a geometric model of the human body with which to calculate BSPs. Yeadon represented the human body using a series of forty geometric shapes, known as stadium solids, figure 2.5.
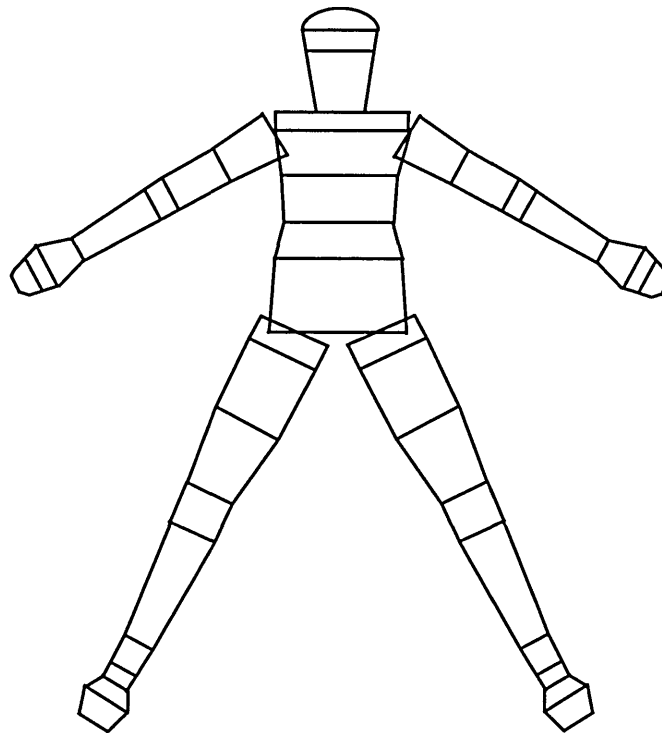
Figure 2.5 - The stadium solid BSP model of Yeadon (1990b)

Each stadium solid is defined between two anatomical landmarks, with the distance between the two landmarks, and a circumference and breadth measurement taken at either end of the segment used to appropriately scale the stadium solids, figure 2.6. This results in ninety five anthropometric measurements per participant. Importantly, Yeadon's model allows the left and right arms and legs to be modelled individually, rather than relying upon the symmetry assumptions of previous techniques. Yeadon uses published density values (Dempster 1955) for each segment, which are appropriately varied amongst the stadium solids which form a particular segment.
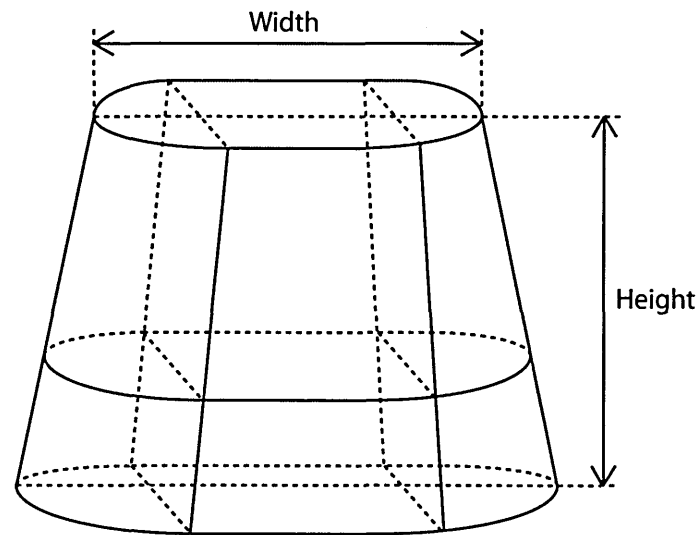
Figure 2.6 - Yeadon's stadium solid used to represent segments of the body (Yeadon 1990b)

Results suggest around 2% error in total body mass when compared to that derived from balance scales, in line with the model of Jensen (1978). Like previous studies, the difference between techniques likely arises from the uniform density assumptions, and the stadium solid segment representations. Whilst the stadium solids represent a significant improvement upon the simple shapes used in previous techniques (Jensen 1978; Hanavan 1964), they are still likely to introduce a degree of error.

Although Yeadon's model requires ninety five anthropometric measurements, it is still quicker than the techniques of Jensen (1978) and Hatze (1980), whilst producing comparable accuracy. For this reason, Yeadon's geometric BSP model appears to be the best technique currently available for use within sports and healthcare research laboratories, training environments, and clinical environments.

22

Importantly however, the accuracy of geometric models has been reported to vary considerably between segments, likely due to the geometric shapes used to represent the body being unrepresentative of reality (Wicke & Dumas 2010). This is likely to be most apparent in the torso segment, having the most potential to considerably change shape as a result of excess abdominal fat (Wicke & Dumas 2010). In the case of Yeadon's geometric model, this may lead to under or over reporting of volume, dependent upon the shape of the individual. Care must therefore be taken when analysing previously reported accuracy and reliability results to ensure the study includes a variety of somatotypes. Other problems associated with the torso segment also originate from the likelihood it contains a higher amount of fat in comparison to other segments (Wicke & Dumas 2010). Therefore, anatomical landmark identification is likely to be harder and more prone to error than in other segments where anatomical landmarks are usually located on bony prominences and not surrounded by fat (Huijbregts 2002). The presence of fat also increases the likelihood of soft tissue depression during anatomical landmark identification, which is likely to lead to further errors.

Gittoes et al (2009) recently improved upon Yeadon's BSP method, developing a technique to reduce the time required with each participant, instead deriving the ninety five anthropometric measurements from three digitised photographs (coronial, left, and right sagittal plane). Results were comparable to those derived using direct measurement techniques, showing a 2.87% mean difference in whole body mass between the measured value (weighing scales) and the image technique, versus a 2.1% mean difference between the measured value and Yeadon's direct measurement techniques (tape measure and callipers). Results therefore suggest the practicality of Yeadon's BSP model can be further improved with this technique, suffering from only a slight reduction in accuracy. The practicality of the manual anthropometric measurement process should therefore be validated against the time savings offered by the photographic technique and the trade off in accuracy before deciding upon the most applicable technique for a specific application.

BSP models appear to offer increased accuracy and reliability when compared to previous techniques based upon cadaver studies. The improved accuracy and reliability typically arises as a result of the increased number of anatomical measurements which

are required to scale the shapes representing the body. Ostensibly the anatomical measurement process can be lengthy, making such techniques unsuitable for a number of applications requiring immediate or quick access to data. Accuracy and reliability has also been shown to vary considerably between models, depending upon the simplicity of the shapes used to represent the body. Furthermore, geometric models are often too insensitive to model person-specific anatomical features: leading to considerable measurement errors when used to calculate the properties of some body types.

2.3.5 3D Surface scanning systems

Laser scanners have previously been used as a method of obtaining 3D scans of the human body from which BSPs can be calculated, likely owing to their widespread availability, high point accuracy (Nikon Corp 2011), and a lack of risks to health (Henderson & Schulmeister 2003).
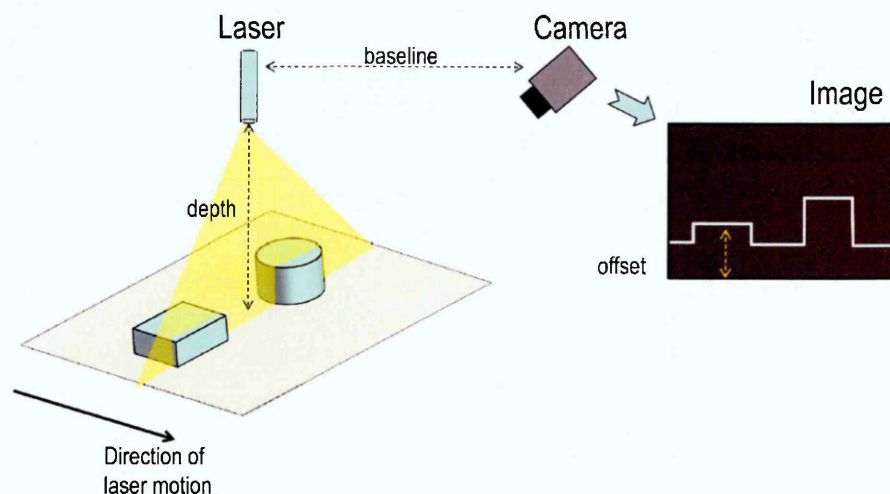
Figure 2.7 – The principals of a laser scanning system (Voosen 2011)

Laser scanners designed for small areas are typically based upon the theory of triangulation, emitting a stripe of laser light from a laser projector, which is projected onto the scene to be scanned. Objects in the scene deform the shape of the laser stripe and as a result change the way the stripe appears to the camera, figure 2.7. Objects that are taller offset the stripe from the origin base line by an amount equal to their height. The offset from the baseline can be classed as a disparity from the norm, and used to calculate real world depth: i.e. the height of the object. Measurements are only obtained

along the narrow line illuminated by the laser projector, therefore, the laser stripe is moved across the full object to obtain a full geometric profile of the object being scanned. Due to the nature of laser light, such systems are able to produce 3D models of very high accuracy (Voosen 2011; Lerch et al. 2006).

Although typically offering high point accuracy (Nikon Corp 2011), the conventionally cited accuracy of handheld laser scanners is likely reduced when scanning living humans. In order to obtain a 3D scan of a full human body it is likely the person will have to stand still for periods of up to thirty minutes, providing significant potential for involuntary movement due to involuntary postural sway and breathing, having an impact upon the geometry returned from the laser scan.

Outram et al (2011) recently used a handheld laser scanner (Model Maker D100, Nikon Metrology) to obtain 3D scans of three male participants. BSPs were calculated directly from the 3D scans using Pro-Engineer CAD software (PTC, Massachusetts, USA), as well as extracting relevant anthropometric measurements. The anthropometric measurements were used in conjunction with Yeadon's geometric model to calculate equivalent BSPs. Results showed Yeadon's model to overestimate segmental mass (assuming uniform density) by around 4% on average when compared to the scan derived value. Importantly, the anthropometric measurements used in conjunction with Yeadon's model were derived from the 3D scan and were therefore not susceptible to the usual manual measurement variation (O'Haire & Gibbons 2000). Were they obtained using normal measurement techniques (tape measure and callipers), then greater disparity between the two techniques would be expected.

In a similar study, Sheets et al (2010) obtained 3D scans of four male participants using a full body laser scanner (Cyberware WBX, California, USA), providing a 4 mm scanning resolution. The 3D scans were morphed to fit a generic fifteen segment model of the body, enabling segmentation of the scans and calculation of relevant BSPs. Scan-derived BSPs were compared to those calculated using Dempster's (1955) and Clauser et al's BSP models (1969). Results showed the laser scan to overestimate whole body mass by around 2.3% when compared to mass derived from balance scales. This may arise in part from errors introduced when morphing the 3D scans to fit the segmental model of the body. Alternatively, it may be due to the 3D scan detecting small

variations in body morphology that the models by Dempster and Clauser are not sensitive enough to detect.

Norton et al (2002) used a full body laser scanner (Hamamatsu body lines scanner, Hamamatsu Photonics, Hamamatsu City, Japan) to obtain 3D scans of ten participant's legs. Scan derived volume measurements were compared to those derived using a water displacement technique. Results showed good agreement between the two techniques, with an average difference between techniques of 0.61% and a lack of systematic bias.

The studies by Outram et al (2011) and Sheets et al (2010) show the scan-derived volume measurements to overestimate in comparison to values derived from commonly used BSP models. It is likely the scan-derived values are nearer the 'true' value, as they are likely to detect small variations in body morphology that may not be detected by the BSP models. This is supported by results of the study by Norton et al (2002), showing scan-derived volume measurements to be very close to those obtained via a water immersion technique. The water immersion technique is likely more sensitive to small variations in body morphology than the BSP models, and therefore closer to the laser scanner.

Full body laser scanners appear more popular than handheld laser scanners, and are also often used in anthropometric studies of athletes (Schranz et al. 2010). Although offering a lower scanning resolution and accuracy (typically by a factor of ten (Nikon Corp 2011; Vitronic 2011)), full body laser scanners offer much shorter scan times (around twenty seconds versus around thirty minutes to scan a full body with a handheld laser scanner (Vitronic 2011)), hence providing far less potential for errors to occur in the 3D scans as a result of involuntary movement or breathing during the scanning duration. For these reasons, it would not be recommended to use handheld laser scanners for any form of body morphology study.

Although more widely available than MRI and DEXA scanners, full body laser scanners still represent a considerable investment (around £100,000 (Vitronic 2011)) and are likely prohibitively costly for the majority of sports and healthcare research laboratories (Weiss et al. 2011). It is important to reiterate that despite their high cost, such scanners are only 'surface scanners', providing only the outer geometry of observed surfaces.

Unlike CT and DEXA scanners they are unable to obtain the densities of the underlying tissue and bone, meaning uniform density assumptions must be used in a similar manner to those used in conjunction with the BSP models discussed above.

Many other commercial 3D surface scanning systems are available, noteworthy is that by 3dMD (3dMD 2014) utilising a projector-camera 3D vision system and primarily designed for medical and healthcare sectors. Although offering high accuracy, resolution, and reliability (Paul et al. 2009), such systems also represent a considerable investment (around £100,000 (3dMD 2014)) and are likewise prohibitively costly for the majority of sports and healthcare research laboratories (Weiss et al. 2011).

Considerably cheaper commercial 3D scanning systems are available, including those by TC^2 (TC2 2011) and Sizestream (Sizestream 2014) , costing around £15,000 (TC2 2011). Accuracy and reliability is typically lower than more expensive systems, but is still suitable for many body scanning applications. However, most are typically designed towards applications within the apparel sector, and as such the data capture software only allows limited access to the raw scan data and measurements (Sizestream 2014), making them unsuitable for other specialist applications such as that considered here.

2.3.6 Summary

The literature suggests that BSP estimation techniques based upon cadaver data are unsuitable for application to the general population as the underlying data typically lacks any form of gender, age, or racial diversity: often being derived from the cadavers of elderly Caucasian males (deceased war veterans). For example, (Cheng et al. 2000) showed such techniques to be unsuitable for estimating the BSPs of participants representative of the Chinese population. Similarly, the BSPs of a young female athlete are likely to be very different to that of an elderly Caucasian male. Although such techniques require few measurements of the human body – and are therefore very quick – they are rarely used within research studies as they are unable to calculate BSPs that are representative of the underlying participants.

Geometric BSP estimation techniques have proved a popular method of estimating BSPs within research studies, offering the best trade off in accuracy, time, and cost. They offer the advantage of taking into account more person-specific features when calculating BSPs, but the accuracy of current models varies considerably, dependent upon the complexity of the geometric shapes used to represent the body. Accuracy and reliability has been shown to vary between segments, with the greatest potential for errors being in the torso segment due to the potential for abdominal fat. However, studies have shown geometric estimation techniques to produce more accurate BSPs than those derived from cadaver studies. Geometric estimation techniques suffer from the drawback of requiring a large number of anthropometric measurements in order to correctly dimension the representative geometric shapes, and as a result the measurement process can be lengthy (taking in excess of forty minutes per person in some cases). Whilst this is acceptable within dedicated research studies it is impractical for sports training or competition environments.

Attention is turning towards methods of obtaining person specific BSPs derived from kinetic measurements and surface representations, rather than relying upon modelling techniques or equations. As a result, a number of in vivo BSP estimation techniques have been developed, making use of the equipment already found within a typical gym or biomechanics research laboratory: such as force plates. However, such techniques are often criticised for making large assumptions about the shape or movement properties of the bodies segments, leading to variable accuracy. They also suffer from the drawback of being highly complex and time consuming, often requiring specialist training. For this reason, such techniques are proving increasingly unpopular, and typically constrained to a small number of specialised research studies.

Medical scanning systems (such as DEXA, MRI, and CT) have been used in a number of research studies to obtain person specific BSPs. Scanning systems of this form offer the unique advantage of being able to obtain person specific density values, rather than relying upon the generic values used in the majority of studies. However, this has been shown to have a limited effect upon further analyses (Wicke & Dumas 2010). Despite their unique advantages, they offer a number of drawbacks and limitations. Firstly, the availability of such systems is limited, meaning widespread use would be unrealistic. The data capture and post processing steps are also highly complex and time

consuming, meaning specialist training is required, and making the whole process unrealistic to perform upon large cohorts of participants. Such systems are highly expensive, meaning they could not be installed within typical sports training or competition environments. However, most importantly, MRI and CT scanners pose significant health risks – particularly if used repeatedly – making widespread use infeasible.

Three dimensional surface scanning systems, such as those based upon laser and projector-camera systems have been used in a number of studies to calculate person specific BSPs. They offer the advantage over geometric and cadaver based techniques of being able to obtain full 3D geometric representations of the individual, enabling person specific BSPs to be calculated based upon real world 3D shapes. However, such techniques typically suffer from the drawback of taking a long time to obtain the 3D scans, requiring the person being scanned to stand still for long periods of time and increasing the possibility for errors to arise due to postural sway and breathing. However, previous studies have shown such techniques to produce accurate and reliable BSPs. Importantly, such scanners still represent a considerable investment, require specialist training, and involve lengthy and specialised post processing of the returned 3D data. For this reason, such techniques have been used in a limited number of studies to date.

Considering previously published BSP estimation techniques, reported results, and discussions, it appears the range of currently available BSP estimation techniques are unable to provide BSPs that are both accurate and reliable, and hence likely to have large impacts upon subsequent dynamics analysis. For these reasons, sports and healthcare communities are currently seeking BSP estimation techniques offering greater reliability and accuracy. Researchers and practitioners appear to be seeking techniques of obtaining BSPs that are specific to an individual and not based upon generic assumptions or heavily simplified models. Additionally, the technique should be low cost (as current geometric and cadaver based techniques are free to use) in order that it can be used in a large number of sports and healthcare environments. The technique should also be as simple and quick as possible, in order that it can be used in a large number of environments (such as sports training) without being prohibitive to normal activities, and does not require lengthy training processes for practitioners.

2.4 The influence of body segment parameter estimation errors upon dynamics analyses

A number of studies have focussed on the influence of errors in BSP estimation upon subsequent biomechanical analyses, such as dynamics analysis. Conclusions differ largely between studies, suggesting the impact of errors in BSPs upon dynamics analysis and the importance of high accuracy and reliability depends greatly upon the type of movement being investigated.

For example, Challis (1996) suggested that segmental moment of inertia errors of up to 8% only had small effects on calculated joint moments during activities such as walking, vertical jumping, and rapid elbow extensions. RMS differences in joint moments for walking and vertical jumping activities were reported to vary by < 2% between perturbed and non-perturbed values (see (Challis 1996) for the magnitude of the perturbations). Similarly, 5% perturbations in forearm moments of inertia were shown to produce RMS differences of 4.1% in joint moments during the rapid elbow extension activity. Importantly however, Challis (1996) recognised the influence of perturbations in BSPs may not have been prominent in the studied activities, owing to the relatively slow accelerations involved.

Kwon (1996) investigated the observed differences in dynamics analysis due to the BSPs calculated from ten different BSP estimation techniques, split into four categories: cadaver based, medical scanning, and geometric modelling. Importantly, the activity for which the dynamics were calculated (double somersault with full twist H-bar dismounts) involved higher acceleration movements than those studied by Challis (1996). For example, calculated CoM differed between methods by up to 3.5% of body height, leading mean airborne angular momentum to vary by up to 10.4%.

A comparable study by Chiu and Salem (2005) calculated the BSPs of a male weight-lifter using a DEXA scanner and the regression equations of Dempster (1955). Calculated BSPs were used to determine knee and hip joint moments during a snatch-pull activity. Results showed knee joint moments to vary by 5%, and hip joint moments to vary by 10%, due to differences in the initial BSPs. Importantly, Chiu and Salem found the largest differences between techniques to occur during the second pull phase

of the activity, where the accelerations of the shank and thigh are at their greatest: affirming the thoughts of Challis (1996) and findings of Kwon (1996).

Arampatzis (1997) studied differences in calculated joint moments during dynamic jumping activities, owing to differences in BSPs calculated using the techniques of Zatsiorsky and Seluyanov (1985) and Hanavan (1964). Results showed differences in net joint moments at the ankle, knee, and hip to vary between $\pm$ 3%, $\pm$ 5%, and $\pm$ 8% respectively. Despite this, Arampatzis et al concluded that differences in BSPs didn't have a significant impact upon calculated net joint moments. However, given the magnitude of the observed differences, the high acceleration movement involved, and the agreement with previous studies, it is hard to agree with such conclusions.

Considering other measures, Pearsall and Costigan (1999) showed that a 20% increase in thigh mass had the potential to increase the distal-proximal hip force by 14.7% during a quiet stance activity. Similarly, a 40% increase in thigh mass was shown to increase the distal-proximal hip force by 29.3% when compared to the norm. Pearsall and Costigan also showed similar results within the low acceleration swing phase of normal walking gait. Whilst a 40% increase in thigh mass may sound a considerable error, given the relatively low total segment mass of the thigh (around 7kg (Pearsall & Costigan 1999)), a 40% increase in mass would only represent an absolute mass increase of 2.8kg. Importantly, Pearsall and Costigan showed that a 40% difference in thigh mass was typical of that expected across commonly used BSP estimation models, further supported by Kingma et al (1996).

Research suggests currently available BSP estimation techniques have a varying effect upon subsequent dynamics analyses, dependent upon both the estimation technique and the type of activity for which the dynamics are being investigated. For example, studies have shown errors in BSPs to have a limited impact upon the calculated dynamics of slow movements (walking for example). However, a much greater impact upon dynamics has been found when higher acceleration movements are involved (such as those typically found within running and tennis). For these applications, current BSP estimation techniques appear unsuitable, suggesting a clear demand for methods of calculating BSPs to a greater level of accuracy and reliability. Whilst current

techniques may be suitable for slower movements, practitioners would undoubtedly benefit from simpler and quicker methods of obtaining BSPs.

2.5   Low cost 3D surface scanning systems

Section 2.3.5 discussed a number of 3D scanning techniques which have previously been used to obtain person-specific 3D scans from which BSPs were calculated. However, the majority of systems are prohibitively costly for widespread use. Section 2.5 therefore discusses a number of lower cost 3D scanning techniques suitable for widespread use, some of which having being previously used to obtain person-specific BSPs.

Recently, Wicke and Dumas (2010) suggested that structured light 3D scanning systems offer a low cost technique for obtaining person-specific 3D scans from which BSPs can be calculated. However, the majority of current systems are still relatively costly (around £20,000 (Rossi et al. 2013; Artec 2014)), require specialist training, and involve lengthy skilled post processing techniques.

The recent interest in natural user interaction (NUI) has led to the development of low cost (in the region of £200 (Amazon 2012)) depth cameras such as the Microsoft Kinect® (Microsoft Corporation, Redmond, USA) and Asus Xtion Pro (ASUSTeK, Taipei, Taiwan), able to capture human motion in 3D (Boehm 2012), figure 2.8. Commonly using a combination of structured light (Shpunt & Zalevsky 2009) or time of flight (TOF) techniques in conjunction with computer vision algorithms (Shotton et al. 2011), they are also capable of capturing 3D scan data at a rate of 30Hz (Khoshelham 2010). This provides a low cost method of obtaining person-specific 3D scans, from which BSPs can be calculated.

Figure 2.8 – Typical NUI sensors

Their launch has led to significant interest in a range of communities including: robotics (Henry et al. 2012), body scanning (Boehm 2012), healthcare (Labelle 2011), and apparel (Stampfli et al. 2012) to name only a few.

2.5.1 Structured light scanners

Structured light 3D scanning systems work by projecting a known light pattern (typically comprising a series of stripes) onto a 3D scene and subsequently capturing it using a camera. The light pattern is distorted by the objects in the scene, meaning the pattern observed by the camera is different to the one originally projected. By determining the shift of each pixel in the known projection pattern in comparison to the captured pattern, the depth of any point in the cameras field of view can be determined. A diagrammatic overview of such a system is shown in figure 2.9.
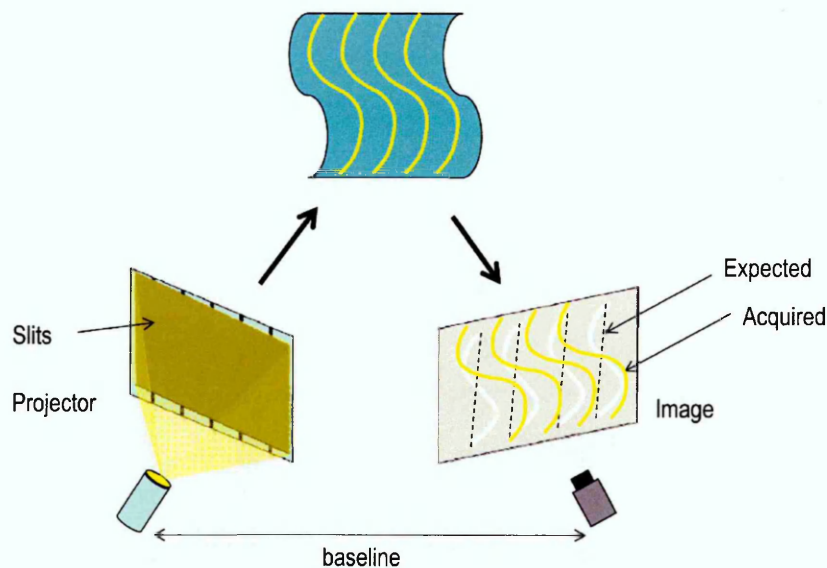
Figure 2.9 – A typical structured light scanning system (Voosen 2011)

Due to the nature of the projected stripes they are very small in size, meaning like laser scanners, they are only able to calculate depth over a limited area in a single scan. In order to scan an entire object, structured light scanners typically project a range of different light patterns in different orientations in order to successfully determine the depth at each point in the scene. For this reason, scanning with structured light systems is relatively slow and therefore typically suited to scanning stationary objects.

A number of 3D motion sensors designed primarily for gesture recognition and natural user interaction have recently been released into the market. A large number rely upon the underlying 3D vision technology developed by Primesense (Primesense 2011), including the Primesense Carmine (Primesense Ltd 2012), Asus Xtion Pro (ASUSTeK Computer Inc 2012), Kinect for Windows (Microsoft 2014a), and the Fotonic P70 (Fotonic 2012), whilst others – such as the Intel RealSense (Intel 2014) - use proprietary 3D vision technology

The Primesense 3D vision technology is essentially a structured light 3D vision system, comprising a projector and camera. A significant difference between the Primesense system and conventional structured light is the ability to capture 3D data at a rate of thirty frames per second, owing to its singular projection pattern and advanced computer vision techniques. The principals behind this derivative of structured light are discussed in more detail in chapter 3.

The Asus Xtion Pro and Kinect for Windows appear to be the most popular of these devices, likely owing to their widespread availability and widely available developer resources. Many studies have investigated such sensors, considering fundamental performance (Menna et al. 2011; Boehm 2012; Khoshelham 2010) and application specific suitability (Weiss et al. 2011).

Boehm (2012) investigated both the accuracy and reliability of the Asus Xtion. Boehm reported good reliability, providing depth measurements within 1 mm of one another when observing a sphere located around 0.5 m from the device. Accuracy was tested in accordance with the VDI/VDE 2634 standard, comprising ten inter-sphere centre measurements of a test object. The test object was scanned with ten different sensors, and results compared to a laser scan. Results showed the ten sensors to vary considerably, having distance measurement errors between 4 mm and 19 mm. Given the high reliability and accuracy offered by some of the sensors, Boehm suggested that such sensors are capable of producing accurate and reliable measurements, but the accuracy is highly dependent upon the manufacturer's calibration process. Therefore, an additional before-use metric calibration process may be required to improve accuracy. Given their low cost, multiple sensors could instead be purchased, tested, and the best performing sensors used for a particular application.

Molnar et al (2012) conducted a similar study using the Microsoft Kinect, but tested the device over a greater range. Molnar et al collected scans of a plane at 0.1 m increments between a distance of 0.5 m and 5 m from the device. Reliability was calculated over the 4.5 m range by calculating the variation of ten thousand 3D scan points consistently lying upon the plane. Results showed a comparable reliability of ± 1 mm at the 0.5 m distance tested by Boehm. However, it is important to note the tests by Molnar et al calculate the reliability based upon ten thousand points rather than the one point used by Boehm. Molnar et al showed measurement uncertainty to increase considerably with distance, becoming as large as ± 22 mm at 5 m from the device. Accuracy was tested by scanning two 0.3 m diameter spheres at 0.1 m increments, over a range of 0.7 m to 4 m from the device. Spheres were fitted to the scans, the scan-derived radius calculated, and the inter-sphere distance calculated. Results showed good performance up to a distance from the sensor of around 2 m, having an average radius error of around ± 4

mm. Molnar et al only tested one sensor, so it is impossible to compare the results to those produced by Boehm. However, given Boehm's results, it is expected the device tested by Molnar et al must have had a 'good' manufacturer's calibration to achieve the results discussed above.

Binney and Boehm (2011) conducted similar tests to those by Molnar et al (2012), instead using the Asus Xtion. Results were highly comparable despite slight differences in the experimentation protocol, affirming thoughts that devices with the underlying Primesense 3D vision system will have comparable performance.

Gonzalez et al (2013) investigated performance of both the Microsoft Kinect and Asus Xtion, performing sphere centre spacing tests. Results showed comparable performance across both devices, with a reported reliability of 2 mm – 6 mm for objects located 1 m from the device: comparable to the values reported by Molnar et al (2012) and Boehm (2012).

Menna et al (2011) also investigated the accuracy of the Kinect using sphere centre spacing distances. Accuracy was comparable to that reported in other studies (Molnar et al. 2012; Boehm 2012; Binney & Boehm 2011; Gonzalez-Jorge et al. 2013), reporting accuracy of around ± 2 mm, despite the test object being located slightly further away from the device at 1.5 m.

Studies into the accuracy and reliability of depth cameras suggest they would be suitable for obtaining accurate and reliable person specific 3D scans. Research suggests that an additional device-specific calibration process may be required prior to use of the devices, owing to their original intended application and hence the comparatively inaccurate and unreliable manufacturer's calibration process. Of the devices relying upon the Primesense 3D vision system, the Microsoft Kinect appears the best overall option, owing to its widely reported operating characteristics, widespread availability, low cost, and support of a commercial software development kit (SDK).

Various studies investigate the use of the Microsoft Kinect and Asus Xtion for human body scanning (Weiss et al. 2011; Boehm 2012; Wan et al. 2013). However, the suitability of a 3D sensor for a specific application should be based upon tests as close

to reality as possible, and not rely solely upon generic geometric tests or application specific tests (Robinson et al. 2012). With that in mind, the relatively simple geometric tests conducted above cannot solely be used to determine the suitability of such devices for obtaining person specific 3D scans. Prior to the work published as part of this program of study, only one study (Wheat et al. 2011) had investigated the Kinect as a method of calculating person specific BSPs. Wheat et al obtained a scan of a trunk segment manikin using a single stationary Kinect. The manikin was placed in eight stationary positions, with a scan collected in each position. The location of the manikin in relation to the Kinect was determined at each of the eight positions using an electromagnetic tracking system. The scans were aligned via a post processing technique, calculating a series of transformation matrices based upon the tracking data to align the scans with one another. BSPs were calculated from the scans using commercial software (Geomagic, 3D Systems, California, USA; Pro-Engineer, PTC, Massachusetts, USA) after segmentation into areas of interest defined by the anatomical landmarks of Yeadon (1990b). BSPs derived from the Kinect scan were compared to those derived in the same manner from a laser scan (Model Maker D100, Nikon Metrology, New York, USA) of the manikin, considered a gold standard. Results showed errors of -1.9 ± 1.6%, 0.5 ± 0.4%, -3.2 ± 2.7%, 2.8 ± 2.3%, and -3.0 ± 2.8% in volume, CoM, I_{xx}, I_{yy}, and I_{zz} respectively. As the study suggests good performance of the Microsoft Kinect in an application specific study, in addition to the promising results presented in geometric tests, it appears the Microsoft Kinect would indeed be suitable for obtaining person specific 3D scans from which BSPs can be calculated.

There are also many other 'conventional' structured light scanning systems that may be used for scanning human bodies, relying upon the capture of multiple light patterns in a variety of orientations in order to determine 3D shape. Stancic et al (2013) recently developed a structured light scanner, comprising of a separate projector and camera, and capable of scanning the human body for calculating BSPs. The forearm of eight participants was scanned with the system. Scan-derived volume estimates were compared to those derived using a water immersion technique. Results showed good agreement between the two techniques, with a mean difference of around 0.14 ml ± 3.67 ml. A significant disadvantage of the method proposed by Stancic et al and indeed any structured light scanning system relying upon multiple projection patterns is the time required to obtain a 3D scan, taking around 4.2 seconds with the system developed by

Stancic et al. Ostensibly, this introduces significant potential for involuntary movement of the participant during the scanning duration. Importantly, despite the scanning system being constructed from off the shelf consumer products (projector and camera), the system cost was in the region of £1000.

Various other consumer structured light 3D scanners are currently available, including the handheld 'Sense 3D' scanner (Cubify 2014b) for direct attachment to a computer via USB, the 'iSense' (Cubify 2014a) and Structure 3D (Occipital 2014) for attachment to an iPad or tablet computer. Whilst both devices are suitable for scanning human bodies in 3D, at the time of writing there is a lack of published accuracy or reliability specifications. Being handheld devices, they also suffer from the same drawbacks as handheld laser scanners.

There are also a number of commercially available structured light scanners, such as the Artec L 3D scanner (Artec Group, Luxemburg), having been used in previous studies to obtain BSPs of human participants (Rossi et al. 2013). Although having a reported resolution of 1 mm and an accuracy of 0.2 mm (Artec 2014), such devices typically cost in the region of £20,000 (Artec 2014), making them overly costly for many sports and healthcare applications.

2.5.2 Time of flight (TOF) scanners

Typically costing slightly more than comparable structured light scanners, TOF cameras have proved a popular method of 3D scanning within many industrial applications (Voosen 2011). As a result, there are many commercially available TOF scanners that are suitable for 3D scanning of the human body.

TOF cameras conventionally work by illuminating a scene with a pulsed light source. The time taken for the light to travel from the source, be reflected by an object, and travel back to the camera is proportional to the distance between the object and camera. This allows the 3D location of each point in the scene to be calculated, forming a complete 3D model (Voosen 2011). Other variations include those based on a continuous light source modulated with an FM or AM carrier wave. The change in

frequency or amplitude after reflection is respectively analysed, and used to form a 3D model of the scene (Voosen 2011).

The accuracy and reliability of the data from TOF cameras largely depends upon the cost of the device, being highly sensitive to manufacturing errors which may result in noise, distortion, and propagation errors (Cui et al. 2012).

Figure 2.10 shows the principals of a pulsed light source TOF scanning system. Light is projected by the source and the time taken for it to travel to the surface, be reflected, and travel back to the detector calculated. As the speed of light and time taken is known, the distance between the light source and all objects in the scene can be calculated.
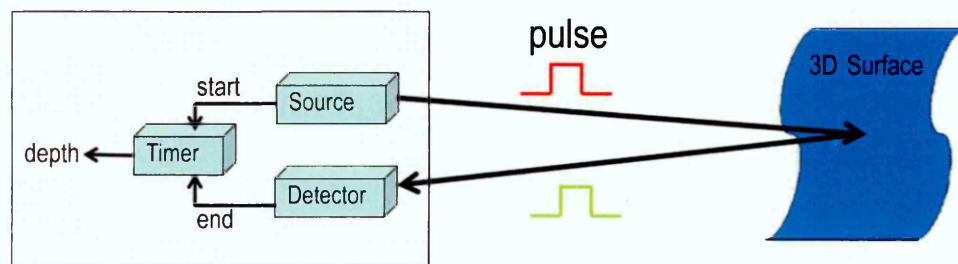
Figure 2.10 – Operating principals of a pulsed light source TOF scanning system (Voosen 2011)

Each pixel in the 'camera' has its own detector timer, meaning the depth of each pixel in the field of view is determined directly. Providing a light source capable of illuminating the entire scene is used, then it is possible to determine the depth of all points in the scene with one image capture: offering many advantages over conventional structured light scanning systems.

Modulated TOF camera systems work in a very different manner but are similar in principal, as shown in figure 2.11.
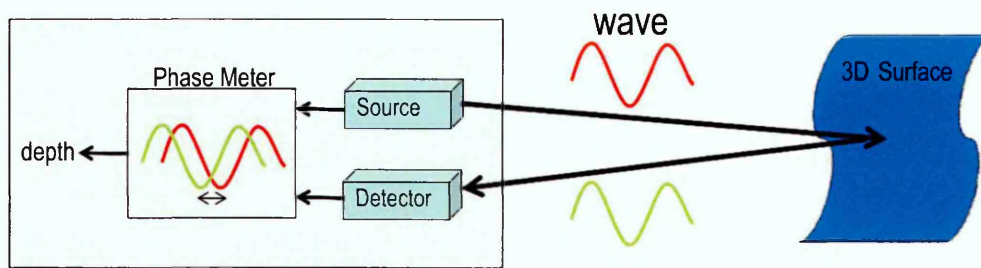
Figure 2.11 – Operating principals of a modulated light source TOF scanning system (Voosen 2011)

The system consists of a continuous light source that is frequency or amplitude modulated (FM or AM) in correspondence with an RF carrier wave, resulting in a light source of sinusoidal form with a known frequency or amplitude. When the modulated light comes into contact with an object in the field of view then the wave is phase or frequency shifted as a result of the shape of the objects surface. The wave is then reflected back from the object, with a detector used to determine the phase or frequency shift. The phase or frequency shift in the reflected light is proportional to the distance between the source and the object, enabling the depth of each point in the scene to be determined. In the same manner as the pulsed light derivative considered above, the depth of the entire scene can be determined with a single image capture.

TOF cameras have the significant advantage over laser scanners in that they are very quick and easy to setup. Unlike traditional structured light scanners, TOF cameras are able to deliver dynamic scanning and hence a reasonably high frame rate. Dependent upon the light source and camera used, they are also able to deliver a very wide field of view and hence scanning range. Unfortunately, the accuracy of TOF cameras can vary considerably. Due to noise, distortion, and propagation errors associated with the transmitted light, they are often unable to return the same level of accuracy as would be observed with a comparable structured light scanning system for example.

The Swissranger SR4000 (Mesa Imaging 2011) has proved popular within the industrial sector, providing a 69° x 55° field of view (FOV) and returning 3D data at a rate of 30 fps. However, accuracy is relatively low, with a measurement uncertainty of ± 10 mm and a typical reliability of 5 mm. With such a large measurement uncertainty, such a

device would likely be comparable to many of the current model based BSP estimation techniques, and hence unable to detect small features of the human body.

Another TOF sensor designed for the industrial sector is the PMD Camcube 3.0 (PMD Technologies 2010), providing a 40° x 40° FOV and returning 3D data at a rate of 40 fps. Reliability is reported to be < 3 mm (PMD Technologies 2010). However, its high cost (~ £700) and considerable weight suggest it is unsuitable for this application.

The Depthsense 311 (Soft Kinetic 2011) is a TOF sensor aimed towards the consumer sector, costing much less than other TOF cameras (around £150 (Soft Kinetic 2011)), and designed for tracking human motion within the application of natural user interaction. The FOV is 57.3° x 42° and provides 3D data at a rate of 60 fps, however the manufacturer provides no accuracy or reliability values.

The Kinect v2 (Microsoft Corporation, Redmond, USA) (Microsoft 2014d) has recently been released, relying upon a TOF system for 3D vision rather than the pseudo structured light technology used in the previous Kinect for Windows. The Kinect v2 offers a larger FOV than the previous device of 70° x 60°, whilst maintaining a 3D data rate of 30 fps. At the time of writing (December 2014) the Kinect has only recently been released, meaning there is little published information regarding the accuracy and reliability of the device. Like the Depthsense 311 discussed above, testing would be required in order to determine suitability for this application.

2.5.3 Camera based scanning techniques

In addition to the projector-camera sensors discussed above, it is also possible to obtain 3D scans using camera-camera systems.

Peyer et al (2014) developed a body scanning system comprising eighteen networked cameras, and used commercial software to stitch together the multiple viewpoints, forming a 3D point cloud scan. A convex hull was produced from the 3D point cloud scans, and segmented in order to calculate segmental inertial parameters. Whilst Peyer et al present scan-derived BSPs and compare them to those calculated using Zatsiorsky's model, the participants used in the study were fully clothed for the scans,

41

introducing significant overestimation of volume and mass. Similarly, the overall cost of the system including the eighteen cameras, commercial software, and a computer with sufficient processing power is likely to be high (in excess of £6,000), making the system prohibitively expensive for many sports and healthcare environments.

Similar functionality may be achieved using pre-calibrated stereo camera pairs, such as the Point Grey Bumblebee (Point Grey Research 2014), producing 3D data directly from the device, However, the high hardware cost (around £1,400), and care which must be taken when handling the devices makes such a system appear unsuitable.

2.5.4 Summary

Since Wicke and Dumas's (2010) original suggestion to use structured light cameras as a method of obtaining person-specific BSPs, a number of comparable low cost 3D scanning systems have been developed, based upon structured light, TOF, and camera based techniques. In terms of accuracy, reliability, and cost, a number of currently available devices appear suitable for the purpose of obtaining accurate and reliable person-specific BSPs from 3D scans.

The range of currently available industrial TOF cameras appear unsuitable for this application largely owing to their size, weight, and cost, meaning a practical and low cost system could not be developed. Current lower cost consumer level TOF sensors appear inaccurate, owing to their inherently lower cost manufacturing processes and tolerances, which lead to large errors in the returned 3D data. However, the recent introduction of the Kinect for Windows v2 may offer the ability to develop a low cost 3D scanning system based upon TOF technology. At the time of writing there is little information available regarding the accuracy and reliability of the device.

Camera based systems appear to be a viable alternative technique. However, the system developed by Peyer et al (2014) showed that a large amount of expensive hardware and software is required in order to develop such a system, making it too costly for the majority of sports and healthcare environments. Additionally, the lengthy data processing technique (around thirty minutes per scan) makes the system unsuitable for a number of environments requiring immediate access to BSP data. A number of

procedural techniques used by Peyer et al mean their results are unrepresentative of reality, meaning the accuracy and reliability of such a system cannot be truly determined from their study. Firstly, the convex hull approach was shown to introduce considerable errors into the 3D models. Secondly, the participants wore loose fitting whilst being scanned, likely making the results unrepresentative of reality.

Three dimensional scanners based upon projector-camera structured light systems appear the most suitable. There are a number of structured light scanners currently available which would be suitable, owing to their low cost, reported accuracy and reliability, and availability of software development kits. Of the devices currently available, those based upon the Primesense reference design appear most suitable, with all the devices having comparable accuracy and reliability, owing to their shared underlying 3D vision technology. The Microsoft Kinect for Windows appears the most suitable device, with the largest number of studies focussing upon the accuracy and reliability of the device in a variety of applications. In addition, the Kinect is supported by a full commercial SDK, meaning bespoke software can be developed for capturing and processing 3D data.

Ostensibly, previous studies (Wheat et al. 2011) suggest it is not possible to obtain a full 360° 3D model of a human body segment from a single scan with the Microsoft Kinect. For this reason, further research is required into the practicalities of developing a 3D body scanning system based upon the Microsoft Kinect, including research into techniques for obtaining multiple scans from different viewpoints in order that a full 360° 3D model can be produced.

2.6 Scan alignment techniques

Investigations suggest self-contained projector-camera structured light scanners such as the Microsoft Kinect are best suited to the application of 3D human body scanning. However, such devices have a limited field of view, meaning they are unable to obtain a full 360° scan of an object in a single scanning operation.

Therefore, a technique of aligning multiple scans collected from different viewpoints is required. Potential techniques are discussed below, combined with critical appraisals, and selection of the most appropriate technique.

2.6.1 Correspondence matching

A common technique used in many studies (Cui & Ngan 2010; Cui & Stricker 2011; Wan et al. 2013; Izadi et al. 2011) involves the object remaining stationary and the scanner being moved around the object being scanned. Owing to the 30 fps 3D data rate and the comparatively slow speed in which the scanner is moved around the device, the 3D scene changes very little between successive frames. With this in mind, correspondence matching techniques can be used to align successive scans with one another, creating a series of transformation matrices bringing the scans into alignment.

A common correspondence matching technique uses the iterative closest point algorithm (ICP), which seeks to align 3D data frames with one another. There are many software implementations available with which to achieve this functionality, with KinectFusion (Izadi et al. 2011) being popular amongst developers of software utilising the Microsoft Kinect. Put simply, the KinectFusion algorithm first performs a basic feature matching technique to produce an initial rough alignment between two successive frames: although the amount of movement should be very little in reality. Next, an iterative process begins, in which the algorithm seeks to align the frames with one another, consistently optimising transformation matrices which minimise the distance between all the points in the two frames. Ostensibly, the correct alignment is the transformation which delivers the smallest total inter-point distance error. The ICP process can either be performed in real time upon dedicated GPU hardware, or via a post processing technique.

Numerous studies have used this technique to obtain 3D scans of human bodies (Cui & Ngan 2010; Cui & Stricker 2011; Wan et al. 2013), including those focussing specifically upon obtaining anthropometric measurements directly from the 3D scans (Wan et al. 2013). Wan et al (2013) developed a 3D body scanning system based upon a single Kinect and a KinectFusion based scan alignment technique. Scans of a manikin were collected and post processed, taking around three minutes per scan. Results

showed an average difference of 2.32% in body circumference measurements when compared to known measurements of a standardised manikin, and a difference of 3.16% in the volume of a cube when compared to manual measurements and standard geometric equations.

Importantly, all the scans collected by Wan et al were based upon static objects, meaning there would be no movement over the scanning duration. In reality, a living human is likely to make considerable involuntary movements due to breathing or postural sway throughout the scanning process, likely leading to errors in the returned 3D models. In reality, this makes such techniques impractical. In addition, environments such as sports training and competitions typically do not have time to dedicate to scanning activities of this nature, as the time requirement would likely be prohibitive to usual activities (Schranz et al. 2010). For this reason, the scanning process needs to be as quick as possible.

Additionally, the ICP algorithm is very computationally intensive, particularly on dense point clouds such as that of a human body. For this reason the algorithm typically runs upon dedicated graphics processing hardware, which can be very costly. If a post processing scan alignment technique is adopted, the time to align the scans with one another can also be lengthy, reducing the practicality of the system, and meaning person specific BSPs could not be viewed directly after the scanning process has taken place.

Considering all of these factors, it appears such techniques are unsuitable for the application discussed herein.

2.6.2 Rigid body alignment

Other studies have used alignment techniques based upon rigid bodies, assuming the sensor collecting the scans remains stationary. For example, Wheat et al (2011) obtained a scan of a trunk segment manikin using a single stationary Kinect. The manikin was moved into eight stationary positions, with a scan collected in each position. The location of the manikin in relation to the Kinect was determined at each of the eight positions using an electromagnetic tracking system. The scans were therefore linked to one another through a series of rotations and translations. The scans

were aligned to create a watertight model using commercial software (Geomagic, 3D Systems, California, USA). BSPs were then calculated from the model using Pro-Engineer (PTC, Massachusetts, USA). Whilst this technique produced good agreement with values derived from an equivalent laser scan, it is impractical for use with living human participants as it is unlikely a participant would maintain the same stance and posture when standing in multiple positions, having significant impact upon their body shape. An alternative technique would be to rotate the person being scanned upon a turntable (Newcombe et al. 2011) however the rotating motion in addition to the lengthy scan duration (around one minute) is likely to lead to movement, also introducing errors into the 3D scans.

Another technique commonly published in the literature relies upon multiple sensors, creating a FOV sufficient to achieve a full 360° view of the object being scanned. For the full human body, this typically results in a total of eight sensors (Kinect/Asus Xtion) (Boehm 2012; Kilner et al. 2012). The sensors are affixed to a rigid frame (Boehm 2012; Kilner et al. 2012) and assumed to remain in the same position between the calibration and data capture process. An initial calibration process is used to determine the extrinsic relationship between the sensors, producing transformation matrices which can be used to align the multiple viewpoints with one another, creating a complete 3D model.

Kilner et al (2012) developed a 3D scanning system comprising eight stationary Kinects. Collected scans were aligned using transformation matrices derived from an initial calibration process. A calibration object comprising six interconnected cubes was placed into the centre of the capture volume, and a scan collected with each sensor. An ICP algorithm was used to determine the location of each Kinect with respect to one another, producing the required inter-sensor transformation matrices. These transformation matrices were later used to align the 3D scans obtained from each sensor. Although Kilner et al introduce a multi-frame smoothing technique to account for noise in the 3D data, this still results in a quick scanning time of less than one second, offering significant advantages over previously discussed techniques and almost eliminating the problem of involuntary movement of the body during the scanning process.

Boehm (2012) adopted a similar technique, developing a 3D body scanning system comprising eight stationary Asus Xtion sensors. Similarly, the sensors were affixed to a rigid framework and assumed to remain stationary between calibration and capture. The necessary inter-sensor transformation matrices were obtained via an initial calibration process. A calibration object comprising six spheres was placed into the centre of the capture volume and a scan collected with each sensor. Although the specific technique is not disclosed in the paper, it is expected the six spheres provided 3D points common to all eight sensors, which could later be used to calculate transformation matrices to align the scans from each sensor with one another.

The two highlighted multi-sensor techniques appear the most suited to this particular application, largely owing to the quick scanning time they offer (around one second). Although the use of multiple sensors increases the overall cost of the system it significantly reduces the scanning time, likely to result in more accurate 3D scans. Within the ensuing development process, effort will be made to develop a calibration technique which does not rely upon the ICP algorithm, and instead uses a simpler, less computationally intensive registration algorithm. This is likely to eliminate the need for dedicated GPU hardware, keeping the overall system cost as low as possible.

2.7 Measurement and Validity standards

Despite the widespread use of 3D scanning systems, there are few recognised standards for assessing their accuracy and reliability. Relevant standards are identified below, along with a summary of standards with which the developed scanning system will be compared to.

2.7.1 VDI/VDE 2634 (The Association of German Engineers (VDI) 2008)

The VDI/VDE 2634 standard was one of the first standards setup to assess the quality of measurements from 3D scanning systems. The standard provides a generic assessment of measurement quality based upon a series of simple measurement tests, rather than an application specific testing procedure.

The standard is split into three parts, with part three 'Multiple view systems based on area scanning' being relevant to the system developed here. Part three describes a series of measurement tests to perform upon a multi-view 3D scanning system. This includes measurements of planarity and sphere centre spacing, which must be taken from at least three different positions within the calibrated capture volume (The Association of German Engineers (VDI) 2008).

2.7.2 ISO 20685-1 (International Standards Office 2010)

The ISO 20685-1 standard (International Standards Office 2010) defines data collection protocols and acceptable reliability standards for body measurements from 3D body scanners. However, there are few commercial 3D body scanners which claim adherence to the standard.

The standard states that validation studies must use the same hardware and software that would be used in reality, agreeing with the earlier suggestion by Robinson et al. (2012). Adherence to the standard is determined by a system's ability to reach a range of 'maximum mean differences', table 2.1. Difference between measurements is classed as the difference between those from the scanning system and an experienced anthropometrist using traditional manual measurement and landmarking techniques.

Table 2.1 - Maximum mean differences in body measurements (adapted from ISO 20685-1 (International Standards Office 2010))

Measurement Type	Permissible ± mean difference (mm)
Large circumference (e.g. chest circumference)	9
Small circumference (e.g. neck circumference)	4

The 'mean difference' is calculated as follows:

- Each participant is scanned and manually measured at least once
- The difference between measurements from the scanner and the anthropometrist are calculated for each participant
- The mean of these differences are grouped for each measurement type from the pool of participants

- The mean difference, standard deviation, sample size, and 95% confidence interval are reported for each measurement type
- If the 95% confidence interval for a measurement type falls within the range defined in table 2.1 then the scanning system is said to meet ISO 20685-1 standards.

2.7.3 ISAK Standards (Stewart et al. 2011)

In response to widespread interest in anthropometric assessment within the sporting community – and hence a need for standardisation - the International Society for the Advancement of Kinanthropometry (ISAK) was formed. ISAK define a unified protocol for measuring the human body, reliability and agreement standards, and a series of accreditation and training courses (Stewart et al. 2011; Stewart & Sutton 2012).

Reliability and agreement of body measurements taken by an ISAK assessor is based upon the technical error of measurement. This is calculated based upon two repeated measurements performed on a number of participants, (Perini et al. 2005). Alternatively, TEM can also be calculated based upon more than two repeated measurements performed on a number of participants. In this case, TEM is calculated using a two way ANOVA (Norton & Olds 1996)

ISAK define two measurement pro-formas: a restricted pro-forma comprising thirteen measurements (including five circumferences) and a full pro-forma comprising forty two measurements (including thirteen circumferences) (Stewart & Sutton 2012). ISAK accredited practitioners are required to meet technical error of measurement (TEM) scores dependent upon their level of accreditation. Level one practitioners are trained in taking the restricted measurement pro-forma, and are required to demonstrate an inter-tester TEM of < 2.5% and intra-tester TEM of < 2% at examination, with the intra-tester TEM falling to < 1.5% post examination. Similarly, level two practitioners are trained in taking the full measurement pro forma, and are required to demonstrate an inter-tester TEM < 1% and intra-tester TEM of < 1.5% at examination, with the intra-tester TEM falling to < 1% post examination. Intra-tester TEM is asses by comparing all measurements from a single practitioner with one another, whilst inter-tester TEM is

assessed by comparing measurements from level one and two practitioners to those obtained by an ISAK level three criterion practitioner, considered to obtain 'true' values (Stewart & Sutton 2012). Further information on the formulae for calculating intra-tester TEM and inter-tester TEM can be found in Perini et al (2005)

Maximum TEM values for inter and intra-tester errors are summarised below in table 2.2. The criteria is dependent upon the level of assessor, and hence the measurements they are trained in taking.

Table 2.2 - ISAK TEM Criteria (adapted from Perini et al. (2005))

Level	Assessment	Maximum TEM (not including skinfolds)
1	Intra assessor (post exam)	1.5%
1	Inter assessor (post exam)	2.5%
2-4	Intra assessor (post exam)	1.0%
2-4	Inter assessor (post exam)	1.0%

2.7.4 Summary and conclusions of validity standards

Robinson et al (2012) suggested that a scanning system should be tested as it would be used in reality. Given this suggestion, it appears the generic nature of the tests defined in the VDI/VDE 2634 standard means it is unsuitable for assessing the performance of the system presented here.

The ISO 20685-1 standard defines a comprehensive protocol and range of acceptable measurement errors, specifically focussed upon measurements of the body obtained from 3D body scanners. The standard focusses on assessing the relative accuracy of a 3D scanning system, comparing measurements from a scanner to those obtained using traditional manual measurement techniques (tape measures and callipers). Comparing measurements from a scanner to those obtained with manual measurement techniques (particularly as there is no requirement to repeat measurements) presents significant possibility of misrepresenting measurement differences between the two techniques, owing to the well cited problems with manual measurement techniques (O'Haire & Gibbons 2000). Furthermore, the standard provides no explicit way of assessing the

repeatability of a scanning system. There is the possibility of repeating measurements and including them in the set of 'mean differences', however this simply combines differences between measurement techniques with random variability, meaning it is impossible to determine the root metrics for reliability and relative accuracy.

Although not directly focussed on body measurements from 3D scanning systems, the ISAK standard presents a commonly accepted technique of assessing the quality of anthropometric measurements within the field of sports. As this is an established sector standard, ostensibly, a 3D scanning system providing equivalent body measurements should seek to meet the same standards. The ISAK standard defines a range of acceptable TEM values for a variety of measurements taken from the body, assessing the measurement reliability. Accuracy is also assessed by the ISAK standard, but in the form of relative accuracy: measured as the agreement between a level one or two ISAK examiner and a higher qualified level three examiner. This is logical, as it is unlikely that gold standard anthropometric values would be available to permit assessment of absolute accuracy. However, the availability of gold standard measurements within this study means that both absolute reliability and accuracy can be reported.

2.8 Chapter summary

Research has shown current BSP estimation techniques have a number of limitations and problems, often leading to poor accuracy and reliability of calculated BSPs and lengthy anthropometric assessment sessions. Current techniques suffer from a range of issues including: data underlying BSP models being biased towards specific population types, small cohorts of participants being used to form BSP models, time consuming and complicated techniques, techniques requiring expensive equipment, and a lack of sensitivity to person-specific anatomical features.

Whilst some BSP estimation techniques are able to eradicate some of these problems – such as medical scanning systems – the considerable cost of such equipment prevents their use within typical research laboratories and training environments. As a result, the majority of BSP assessment currently relies upon the range of model based estimation techniques discussed above.

A number of studies have investigated the impact of inaccurate BSPs upon subsequent dynamics models: a typical application of BSPs within the sports community. Studies have led to a variety of conclusions, however, researchers appear to agree on two common conclusions. Firstly, that current BSP estimation techniques typically used in research laboratories are likely sufficient for dynamics analysis of slow movements, owing to the minimal reported impact of BSP perturbation upon the results of dynamics analyses. Secondly, that current laboratory based BSP estimation techniques are inadequate for dynamics analyses of movements involving high accelerations – such as those found within the area of sports – as BSP perturbation has been found to have a significant impact upon the results of subsequent dynamics analyses.

Despite these findings, there appears consensus amongst practitioners over a need for an improved technique of obtaining BSPs, which is accurate, able to detect person-specific anatomical features, and is reliable, quick and simple, and low cost. Whilst current techniques may be sufficient for slow movements, it is likely such assessment exercises would benefit from a simpler and quicker technique of obtaining BSPs. On the other hand, the sports community would undoubtedly benefit from a more accurate and reliable technique for calculating BSPs when used to infer the dynamics of fast movements.

Recent interest appears to have been focussed upon 3D scanning systems, able to produce 3D models of the human body from which BSPs can be calculated, and eliminating the potential for errors becoming introduced owing to the inherent errors in anthropometric measurements (O'Haire & Gibbons 2000). However, the majority of current 3D scanning systems are prohibitively costly. In 2010, Wicke and Dumas suggested that structured light scanning systems would be a viable lower cost alternative. Since then, there have been a number of low cost structured light 3D scanners launched into the market. Of these, the Microsoft Kinect for Windows appears the most suitable for this application. The Kinect is low cost, supported by a full software development kit, has widely published accuracy and reliability values from a variety of 3D scanning studies, and has previously been used to successfully obtain BSPs from 3D scans (Wheat et al. 2011). With this in mind, it has been decided to focus upon the Microsoft Kinect for the ensuing development process.

It is apparent that a single Kinect sensor is unable to achieve a full 360° of the human body, in order to produce a full 3D scan of the body. Previous studies have used a single Kinect and rotated the object being scanned, however that appears unsuitable for scanning human participants owing to the potential for errors to become introduced due to involuntary movement throughout the lengthy scanning process (Daanen et al. 1997; Schranz et al. 2010). Similarly, techniques using a single Kinect sensor which is moved around the person being scanned appear unsuitable for the same reasons. The best option appears to involve multiple Kinect sensors, able to achieve a full 360° view of the body from a single scanning operation. In order to align the individual viewpoints with one another and produce a single 3D scan, a rigid body alignment technique should be used: assuming the Kinects remain in a stationary position between an initial calibration procedure and the collection of 3D scans. This results in a quick scanning process (around one second), reducing the potential for errors due to involuntary movements.

The findings of this literature study enable a number of decisions regarding the requirements and design of an improved technique for estimating person-specific segmental volume, which can later be extended to include the full range of BSPs. The ensuring development and implementation process is discussed in detail in chapter 3. Chapter 7-9 later discuss the developed BSP estimation technique, and compares its performance to a currently accepted and commonly used BSP estimation technique.

3 SYSTEM DEVELOPMENT

3.1 Introduction

As identified in the literature review, there is a need for a low cost method of collecting person specific 3D scans from which segmental volume can be calculated. As also identified, the system should collect the scans in a timely manner: preventing problems associated with participant movement. This chapter discusses the development and data capture process of such a system.

3.2 Aims and objectives

Aim

- To develop a 3D body scanning system costing less than £2000, and able to obtain person-specific 3D scans in less than 10 minutes, later capable of estimating segmental volume.

Objectives

- Develop a process for aligning 3D scans collected from multiple viewpoints to create a complete 360° scan.
- Develop a software application capable of collecting the required 3D scans, produce and display a complete 360° 3D scan, and perform the necessary data post processing and calculation steps.
- Allow palpated anatomical landmarks to be viewed on the 360° 3D scan and digitised using the software application.

3.3 Background to the Microsoft Kinect

As discussed in chapter 2, the Microsoft Kinect was chosen for this study owing to its low cost, favourable results presented in previous studies, and support of a freely available official SDK. Section 3.3.1 and section 2.5.1 discuss the overall design of the Kinect, operating principals, and format of the data output streams.

The Xbox Kinect was originally released in November 2010, designed as a games controller add on for the Xbox 360 games console (Adafruit 2010; Amazon 2012). The Xbox Kinect tracks the movement of people in its FOV, calculates the position of key anatomical points in 3D, and sends this information to an attached Xbox. The Xbox then uses the 3D points to map the movement of observed people onto on screen avatars for example (Microsoft 2014b).

Although designed for use with the Xbox 360, the Xbox Kinect features a standard USB 2 plug and external power supply, meaning it can also be attached to a computer. Very shortly after its launch, the Xbox Kinect was 'hacked', and unofficial driver's released for Linux and Windows (Adafruit 2010).

Following the widespread release of drivers and software development tools, the 3D motion tracking and underlying 3D vision technology received ever growing interest from a range of communities including: robotics (Henry et al. 2012), apparel (Stampfli et al. 2012), healthcare (Labelle 2011), and hobbyists alike. Microsoft responded to the ever growing interest, later releasing a freely available official SDK and drivers, in addition to a Kinect for Windows sensor (Microsoft 2012a). The two devices are cosmetically the same (apart from different branding), figure 3.1, however the Kinect for Windows features a range of additional functionality and can be used for commercial projects when combined with the official Microsoft Kinect SDK (Microsoft 2012b).
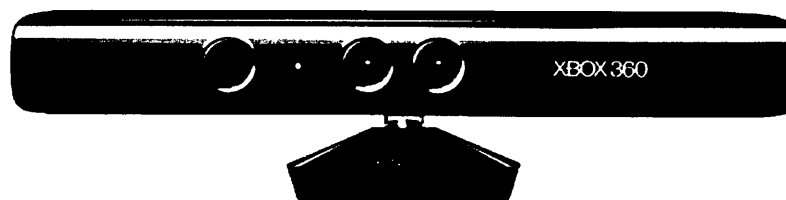
Figure 3.1 – The Xbox Kinect (Microsoft 2014c)

Both devices feature the following components (Microsoft 2014c), figure 3.2:

- A USB 2 interface and external mains power supply.

- An infra-red projector, used to emit an infra-red speckle pattern that forms the basis of the Kinect's 3D vision system.

- An RGB camera, providing a 1280 x 960 colour image stream at 30 fps.

- An IR camera providing a 640 x 480 image stream at 30fps, forming another essential part of the Kinect's 3D vision system.

- An array of four microphones, which can be used for locating people by their voice.

- A tilt motor, allowing the camera to move and follow people as they move.

The field of view offered by the devices is relatively limited, being 58° and 45° (Microsoft 2014c) in the horizontal and vertical direction respectively.
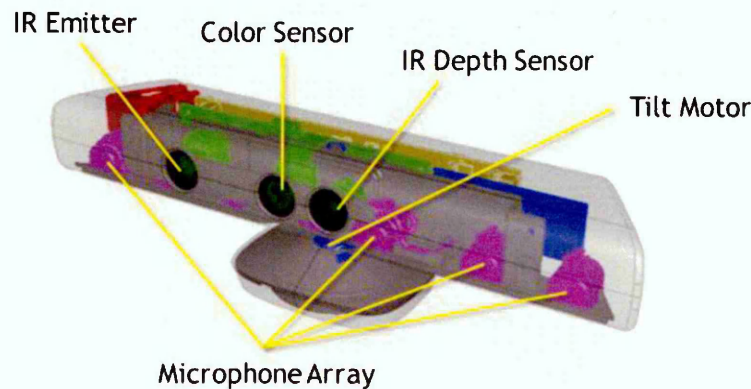
Figure 3.2 – Breakdown of the Kinect's integral components (Microsoft 2014a)

One of the main differences between the two devices is their operating range, due to their originally intended purpose. The Xbox Kinect was originally designed for tracking movement of humans within a typical home environment, and consequently has an operating range of 0.8 m and 4.0 m (Microsoft 2012c). As well as the 'normal' operating range the Kinect for Windows includes an additional operating mode 'near mode', which can be enabled via software control. This reduces the devices maximum operating distance to 3.0 m, but reduces the minimum distance to 0.5 m (Microsoft 2012c). Microsoft included this additional feature to allow the Kinect to be used for smaller scale applications, such as tracking the movement of hands and fingers.

From here on, only the Kinect for Windows will be referred to and used, owing to the additional features, flexibility, and support.

3.3.2 The Kinect's data output stream

The official Kinect for Windows SDK exposes a range of raw data streams from the Kinect, all output from the device at 30fps. For example, the Kinect allows access to the raw 640 x 480 image feeds from the RGB and IR cameras, figure 3.3.

Figure 3.3 – RGB and IR images from the Kinect

One of the most useful data streams from the Kinect is its depth image stream, figure 3.4, providing a 640 x 480 image of the observed scene with each pixel colour coded based upon distance from the device. The depth image stream is produced by a 'virtual camera', formed from analysing the projected speckle pattern with the IR camera and calculating the real world distance between the Kinect and a particular point in space (Herrera et al. 2012).

Figure 3.4 – Depth image from the Kinect

For motion tracking applications, the Kinect outputs real world 3D locations of key anatomical points tracked on the body of a person moving within its FOV (Microsoft 2014b). As per the original Xbox Kinect, this data can be used for mapping the movement of a person onto an on-screen avatar. Figure 3.5 shows the points of the body tracked by the Kinect.
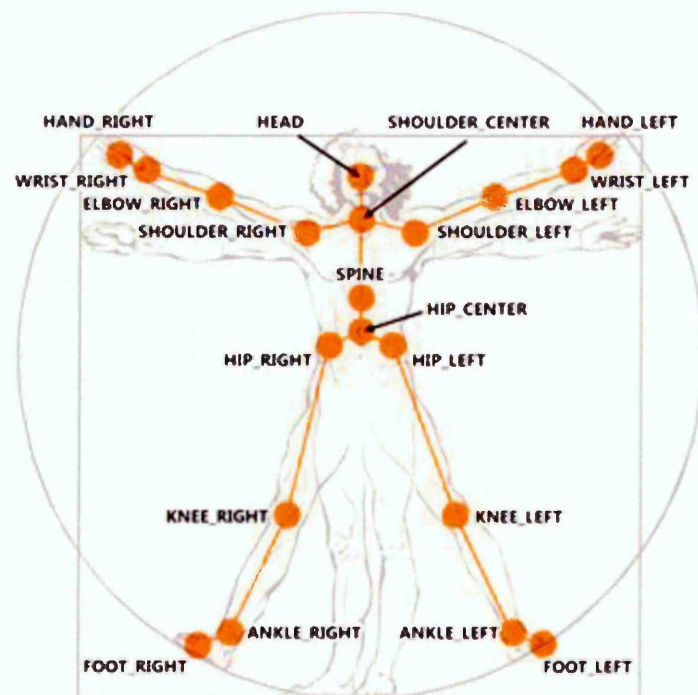
Figure 3.5 – Skeletal data points tracked by the Kinect (Microsoft 2014b)

Although used far less, the Kinect also has an audio stream from its four integral microphones (Klug 2010). For example, this can be used in conjunction with some of the software tools provided with the Kinect SDK to locate people within a room based upon their voice (Microsoft 2012b).

Perhaps one of the reasons the Kinect has received more interest than other consumer depth cameras is owing to range of coordinate mapping functions provided as part of the Kinect for Windows SDK. These enable coordinate system transformations to be performed within the device, transforming one image space into another. For example, a function exists to transform the skeleton stream into RGB image space. This produces 2D equivalents of the identified 3D points, allowing a 'stick man' equivalent of the person observed by the Kinect to be overlaid onto the RGB image stream.

Of relevance to this project is a function to convert the depth image stream to skeleton space. This converts each pixel in the uvZ depth image into 3D space, generating real world equivalent co-ordinates. It is this function which enables the Kinect to be used as a 3D scanner, producing a 3D point cloud scan of the observed scene comprising of 307200 XYZ points (assuming an original depth image of 640 x 480 pixels) at a rate of 30 fps. Additionally, a function exists to convert the RGB image stream into skeleton space, meaning each point in the point cloud scan can be rendered with its real world colour. These two transformation functions are integral to the data capture process of the scanning system, discussed in more detail in section 3.6.1.

3.3.3 The Kinect's 3D vision system

Originally designed for tracking the motion of humans in three-dimensions, the Kinect contains a highly sophisticated 3D vision system which can be utilised for the purpose of 3D scanning. The Kinect's 3D vision system is based upon the Primesense reference design, protected by numerous patents ((Cohen et al. 2011; Freedman, Shpunt, Machline, et al. 2010; Shpunt & Zalevsky 2009) to name but three), and hence little affirmed information is available about its operation principals and technology contained within. However, the information contained within the numerous patent applications provides sufficient information to build a fundamental understanding.

Put simply, the Kinect's IR projector outputs an infra-red 'speckle pattern' of thousands of dots, which is in turn projected onto the observed scene. The Kinect's IR camera images the scene illuminated the speckle pattern and compares the dots in the projected pattern to a reference image stored in the Kinect at manufacture. The difference between the two images allows the three-dimensional depth at each pixel to be calculated, in turn producing the depth image shown in figure 3.4, colour coded with distance between the objects and the Kinect. This principal is discussed below in more detail.

The basis of the Kinect's 3D vision system is the projected infra-red speckle pattern, shown in figure 3.6 as imaged with a consumer camera.
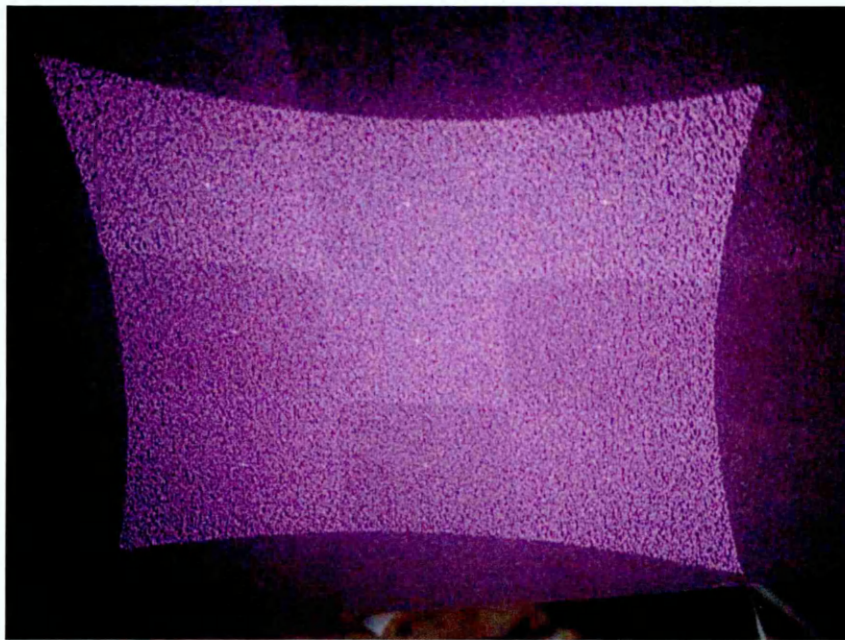
Figure 3.6 – The Kinect's projected IR speckle pattern (Stern n.d.)

A number of characteristics are immediately apparent from the imaged speckle pattern shown in figure 3.6. The pattern is clearly comprised of a 3x3 grid of projected dots, with each of the grids having a different light intensity and containing a centre bright dot. Research by Reichinger (2011) has shown the complete speckle pattern is comprised of a random projection pattern, which is repeated and tiled in a 3x3 grid.

The projected pattern is optimised to different depths (Freedman, Shpunt & Arieli 2010; Reichinger 2011; Voosen 2011; Shpunt & Zalevsky 2011), meaning some points are

used at greater depths and not at shorter depths: likely explaining the varying intensity of the speckle pattern.

The projected speckle pattern is produced by a combination of a constant source IR projector, and two diffractive optical elements (DOEs) or diffraction gratings (Shpunt 2009; Shpunt & Pesach 2010). The first DOE creates the random speckle pattern and central bright dot (Shpunt & Zalevsky 2011), whilst the second grating multiplies the generated pattern and creates the tiled 3x3 grid pattern observed in figure 3.6 (Shpunt 2009; Shpunt & Pesach 2010). Figure 3.7 shows an abstract of the relevant patent, showing a schematic of the projection system and repeated projection grid.
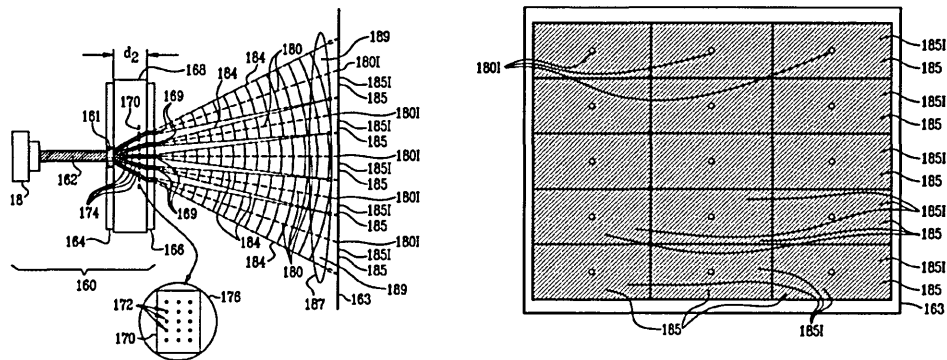
Figure 3.7 – Overview of the Kinect's IR projection system (Shpunt 2009)

Once the speckle pattern has been projected onto a scene, the Kinect images the scene with its IR camera. The received image is then analysed to produce the depth image shown in figure 3.4.

The Kinect uses a combination of two techniques to analyse the received images and determine depth, namely: depth from focus (Maccormick 2011), and depth from stereo pattern (Maccormick 2011; Shpunt et al. 2010).

Given a fixed focal length lens, depth from focus relies on the principal that objects located further away from a lens will be blurrier than objects located closer to the lens. The Kinect builds on this principal with use of an astigmatic lens (Maccormick 2011). An astigmatic lens has different focal lengths in the x and y directions, meaning that circular objects become ellipsoid shaped, with the orientation of the ellipsoid being dependent upon depth (Maccormick 2011). Subject to an initial calibration process, the

Kinect uses this principal to determine depth from the projected speckle pattern, which is initially produced by the DOE as a series of circular points (Shpunt 2009).
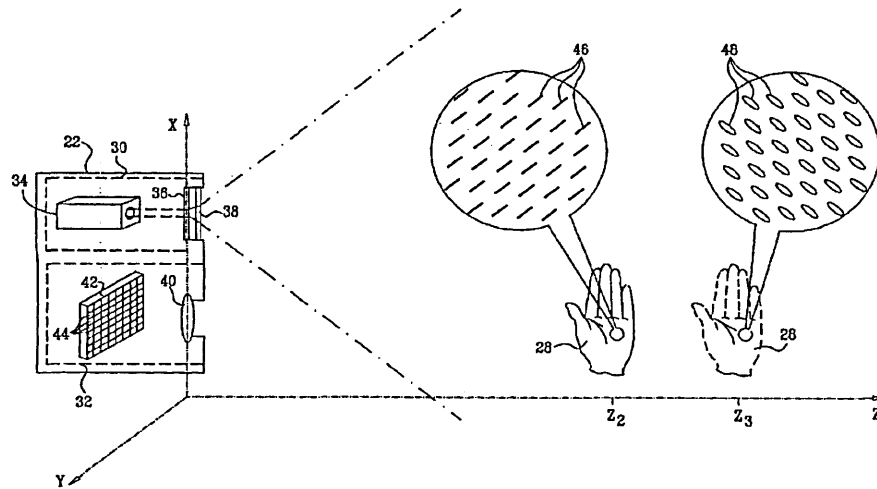
Figure 3.8 – The principal of the Kinect's astigmatic lens (Freedman, Shpunt & Arieli 2010)

The Kinect also uses the principal of depth from stereo to determine depth from the projected speckle pattern (Maccormick 2011; Shpunt et al. 2010). Figure 3.9 illustrates the principal of stereo triangulation and depth from stereo, given a projected speckle pattern and imaging camera.
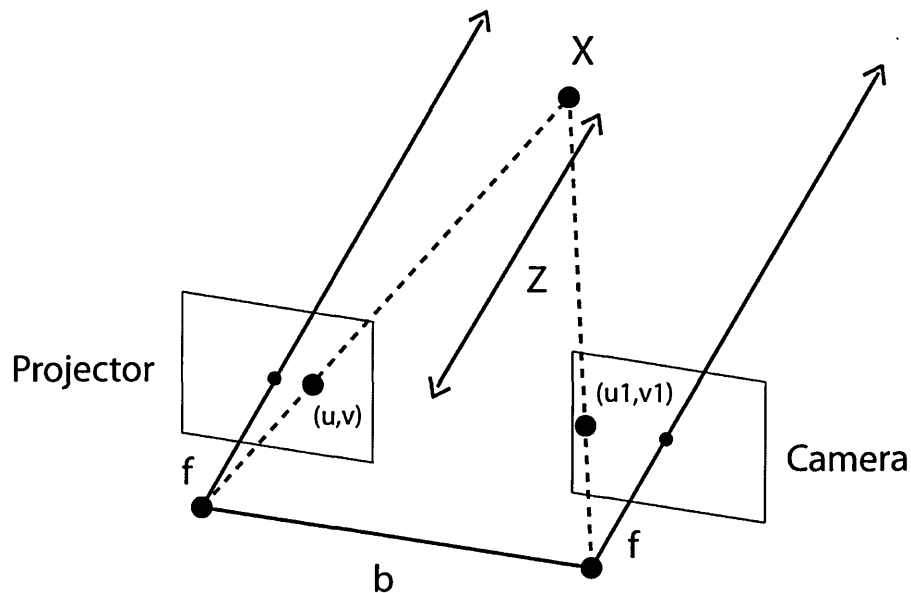
Figure 3.9 – Principal of depth from stereo given a projector and camera

In figure 3.9 a reference speckle pattern – stored in the Kinect at manufacture (Shpunt et al. 2010) - is projected by a projector, producing a point in space, X. The same point in space is then imaged by the camera, separated from the projector by a fixed base distance, b. Owing to the different position of the camera in relation to the projector, the point in space, X, will appear in a different position in the image of the IR camera, u^1, as opposed to the reference projection image, u, (Mihelich n.d.). Calculation of the disparity, d, i.e. the pixel shift between the two images is then a simple process, as shown below in equation 3.1 (Mihelich n.d.).

$$d = u - u^1$$

Equation 3.1

Knowing the fixed focal length, f, of the IR camera and projector, the calculation of the real world distance, Z, between the Kinect and object in space is then a simple process, as shown below in equation 3.2 (Mihelich n.d.).

$$Z = \frac{fb}{d}$$

Equation 3.2

When objects are placed in the scene upon which the pattern is projected the principal is exactly the same. Importantly, objects that are closer to the Kinect will shift the speckles of the pattern further to one side than objects that are further away (Maccormick 2011), as illustrated below in figure 3.10.
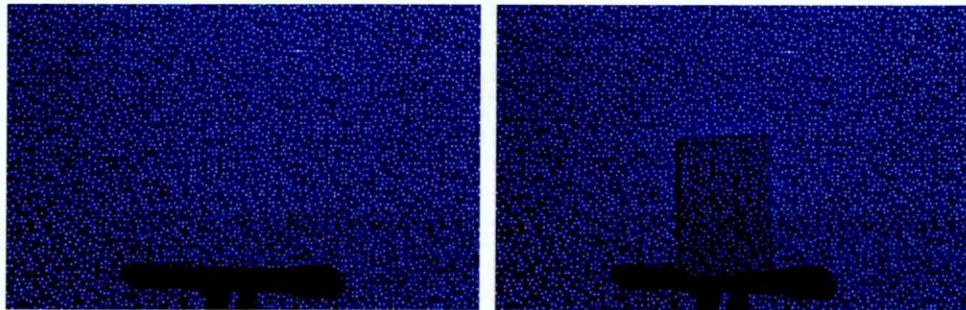
Figure 3.10 – Principal of disparity change due to objects in the scene (Future Picture n.d.)

Whilst this may appear a relatively simple technique, an important factor is the identification of the same point in both the reference speckle pattern and image from the IR camera.

In order to achieve this, the Kinect appears to use a combination of the 9 bright central spots and a 9x9 correlation window (Mihelich n.d.). For a given image from the IR camera the Kinect first localises itself using the 9 bright central spots (Mihelich n.d.) to determine which of the 9 repeated patterns it is working within. For a given pixel in the image, the pixels surrounding that pixel in a 9x9 window are extracted, known as a correlation window. The correlation window is then used to search the reference IR pattern to determine the equivalent pixels in the original projected pattern (Mihelich n.d.). Once identified, the Kinect then performs a further sub pixel (1/8 pixel) optimisation to determine the disparity between the IR image and reference pattern (Mihelich n.d.). This process is repeated for each of the white spots in the projected pattern (figure 3.6), allowing the disparity of each white spot to be calculated with use of equation 3.1. The real world depth of each of the white spots is then calculated using equation 3.2.

With this in mind, it is important to emphasise that the actual depth will only be calculated at each of the white spots shown in figure 3.6. The Kinect must then perform some form of nearest neighbour's interpolation to determine the depth at the black pixels shown in figure 3.6.

Although not immediately apparent from figure 3.4, it is important to note that the Kinect has an 8 pixel wide null band along the leftmost vertical side of the depth image. This is a factor of the 9x9 correlation window used to determine correlation between the IR image and reference speckle pattern (Mihelich n.d.). This is unlikely to cause problems in most applications, and care should be taken to ensure objects do not reside at the very edges of the depth image.

There are a range of drivers and SDKs available for the Kinect, namely the official Microsoft Kinect SDK, OpenNI framework, and Freenect to name but three. The Microsoft Kinect SDK and Open NI drivers expose the Kinect's depth information in real world measurements (metres), ready for use in most applications. However, the Freenect drivers expose the depth information in raw disparity units, meaning a conversion to real world measurements must be applied (Mihelich n.d.) before the measurements are of practical use.

Importantly, all conclusions and discussions herein relate to the official Microsoft Kinect SDK and associated drivers.

3.4 Design of the scanning system

As discussed in the aims and introduction to this chapter, there is a need to develop a scanning system capable of obtaining person specific 3D scans as quickly as possible. As highlighted in chapter 2, this makes the many single sensor techniques unsuitable as it takes a considerable amount of time to move a single Kinect sensor around the body, also introducing significant potential for involuntary motion artefact in the 3D scans. Instead, the 3D scans must be captured from multiple Kinects which observe the person being scanned from multiple viewpoints. Single snapshot scans can be collected from each Kinect in a near simultaneous manner, resulting in a considerably shorter scan time. After collection, the 3D scans can be aligned with one another to produce a complete 360° scan.

3.4.1 Physical layout

The Kinect's triangulation base distance is only 75 mm (distance between IR camera and projector), limiting the distance at which depth can be perceived accurately (Boehm 2012). A general rule of thumb proposed by Waldhausl and Ogleby (1994) suggests an acceptable base to distance ratio of 1:16, meaning the distance between the Kinect and object being scanned should not exceed a distance of 1.2 m.

Experimentation showed that four Kinect sensors located at the corners of a 0.8 m x 0.8 m square was sufficient to obtain a 360° scan of a human body. Were a full body required to be scanned at once, this would mean the Kinects would need to be located at a distance from the person greater than the limit suggested above. With this in mind, eight Kinect sensors would instead be required, aligned via an additional calibration process.

The scanning system comprises four Microsoft Kinect sensors affixed to a supporting framework, able to capture a single anatomical segment in one scan, figure 3.11. The Kinect sensors were affixed to the framework in a vertical orientation, maximising the

field of view in the vertical direction (58° instead of 45°). This means the Kinects can be located closer to the centre of the capture volume, improving accuracy by reducing the triangulation distance (Waldhausl & Ogleby 1994). Initial experimentation showed this setup was sufficient to produce a 360° scan of a human torso, whilst allowing sufficient flexibility to cope with different sized people, figure 3.11 and figure 3.12.
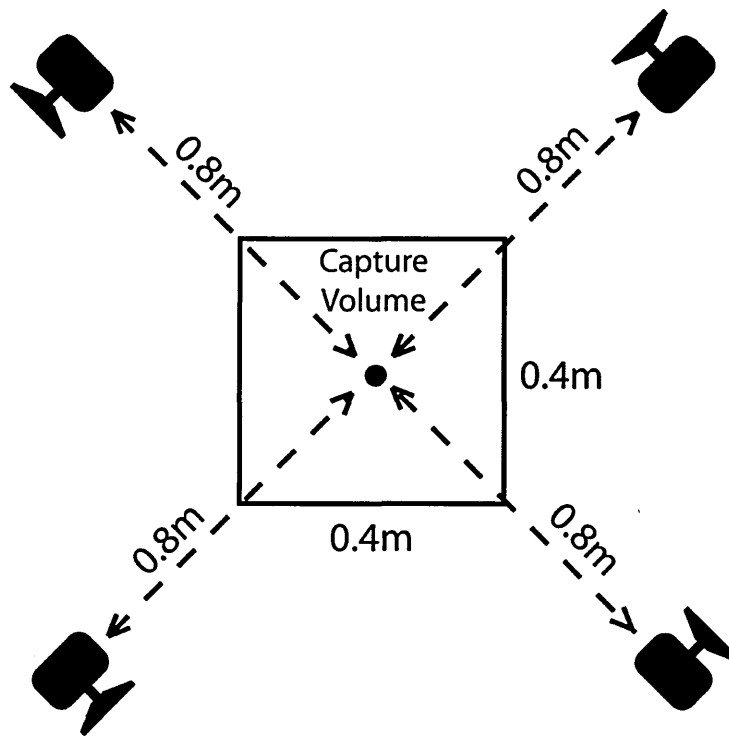
Figure 3.11 – Layout of the scanning system
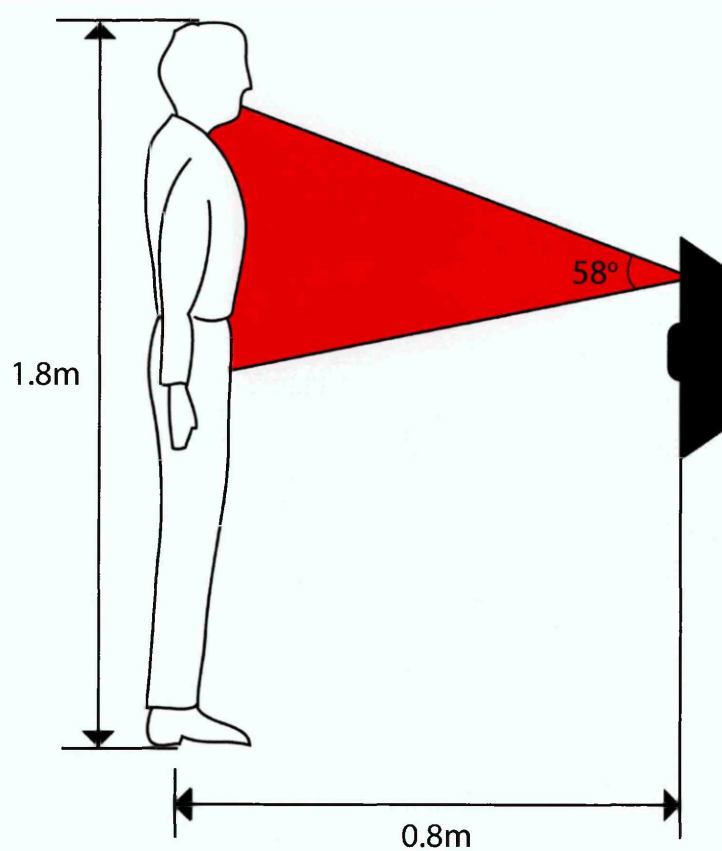
Figure 3.12 - Scanning field of view provided by each Kinect sensor

3.4.2 Sensor configuration

The 0.8 m distance between the centre of the capture volume and each Kinect sensor meant the Kinects could be placed into their 'near' operating mode, maximising the number of projected points used for depth measurement, and hence accuracy (Draelos 2012).

When multiple Kinect sensors are used to observe the same object their projected speckle patters interfere with each other, causing a significant number of 'holes' to appear in the returned 3D scans (Anderson et al. 2012; Boehm 2012). Within the proposed scanning system each Kinect sensor has two areas of overlap from neighbouring Kinects, likely to introduce a significant number of holes in the final 3D scans. Therefore, each Kinect's projector was turned off prior to scanning.

When a scan was collected, each Kinect's projector was turned on individually, a scan collected, and the projector turned off again. This process was repeated for each of the four Kinects until a scan had been collected from each.

This process also eliminates the problem that depth measurements from the Kinect become unreliable over extended periods of use, due to heating of the IR projector (Binney & Boehm 2011; Fiedler & Heinrich 2013). As the IR projectors are only enabled for very short periods of time this problem is negated.

However, research has shown the Kinect's depth measurements need time to stabilise after initialisation of the IR projector (Anderson et al. 2012). Experimentation has shown this to be the case, taking around 200 milliseconds rather than the twenty seconds proposed by Anderson et al. (2012). Furthermore, Anderson et al. (2012) provide no ground truth measurements in their testing, suggesting the observed transient effect may instead be related to heating of the Kinect's IR projector rather than an initialisation time. Within the developed scanning system, a delay of two hundred milliseconds was used between enabling each IR projector and collecting 3D data: allowing sufficient time for the depth data to stabilise, whilst limiting any effects due to heating.

3.5 Comparison of extrinsic calibration techniques

Each of the four sensors in the scanning system observes a different part of the body, which needs to be aligned with a single coordinate system to produce a complete 360° scan of the human body.

As the sensors are affixed in the same place they remain in the same position throughout the system initialisation and scanning process. This means a rigid body calibration algorithm can be used to find the position of the sensors with respect to one another, enabling the four individual scans to be aligned with a global coordinate system.

3.5.1 2D Checkerboard calibration

Zhang's checkerboard calibration technique (Zhang 2000) was investigated as a method for extrinsically calibrating the Kinects. Zhang's technique is a commonly used method for calibrating the intrinsic and extrinsic properties of a camera, relying on multiple images of a black and white checkerboard.

Zhang's technique has previously been used by Burrus et al. (2011a) to refine the intrinsic parameters of the Kinect's IR and RGB cameras, as well as extrinsically calibrating two cameras with respect to one another. This has included calibrating two IR cameras to align scans from different viewpoints, as well as calibrating an IR and RGB camera to overlay colour data onto 3D data to produce colour rendered 3D scans. Burrus et al. (2011a) showed all approaches to work well.

The work discussed in section 3.5.1 was conducted using the open source OpenNI drivers for the Kinect (OpenNI 2011), as opposed to Microsoft's official SDK as they were not available at the time of investigation (Microsoft 2012b).

To investigate this technique, three Kinects were located at 90° to one another and a capture volume defined such that the checkerboard could be observed by two cameras at all times, as shown in figure 3.13. As this was only an initial investigation only three Kinects were used, simplifying the calibration process.
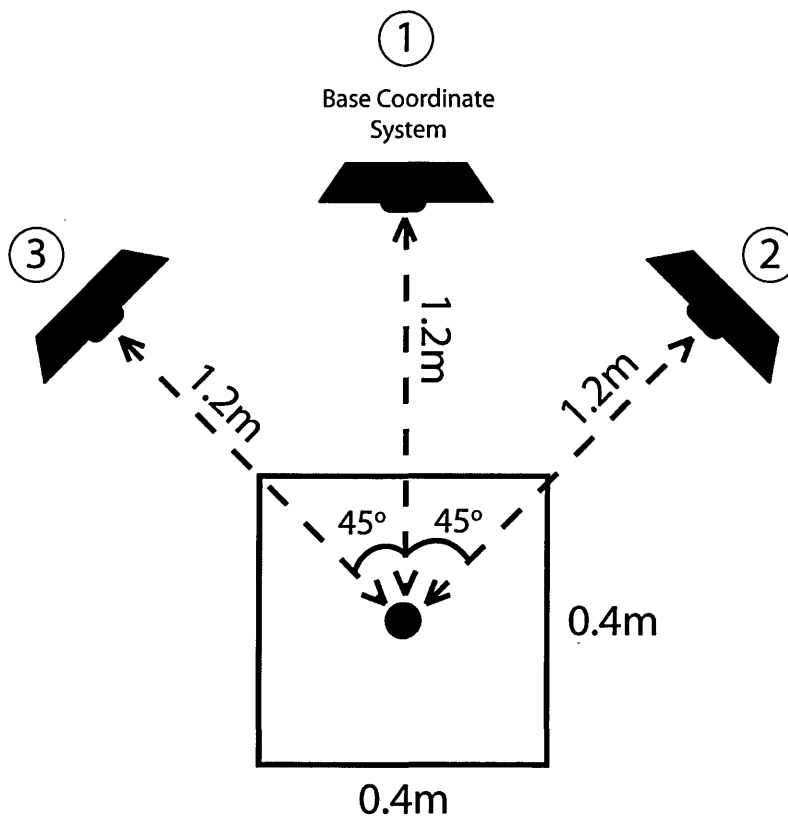
Figure 3.13 – Schematic layout of the Kinects for calibration using a checkerboard

Zhang's calibration technique only allows cameras to be extrinsically calibrated in pairs, therefore a 'base' Kinect was initially defined, which would become the global coordinate system after calibration (Kinect 1 in figure 3.13). The outer two cameras were then calibrated in two pairs, comprising one of the outermost cameras and the base Kinect. This meant the calibration process produced transformation matrices to align the scans from Kinect two and three into a global coordinate system aligned with Kinect one, producing a single aligned scan from the three alternative viewpoints.

For each Kinect pair, an image of the checkerboard was captured with both IR cameras. In normal operation the IR camera observes the projected speckle pattern, enabling 3D vision. However, the projected speckle pattern means the image of the checkerboard is unclear, and likely to cause problems in later calculations. Therefore the IR projector was covered with a piece of card and the scene illuminated using a 40w incandescent bulb. After collection of the IR images, the IR projectors were uncovered, the light turned off, and a 3D scan collected of the scene with each Kinect.

This process was repeated for each of the two Kinect pairs, collecting checkerboard images in thirty five different positions and orientations: ensuring the entire capture volume was covered with checkerboard images. Figure 3.14 shows the checkerboard position and relative camera positions for one of the Kinect IR camera pairs.
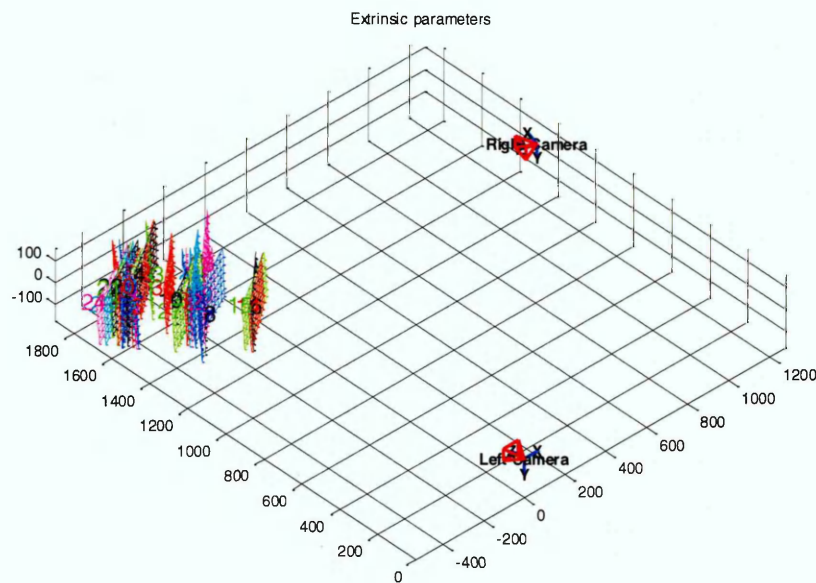
Figure 3.14 – Checkerboard positions used as part of Zhang's calibration technique

After collection, the IR images were input to Bouguet's Matlab camera calibration toolbox (2010), based upon the algorithms by Zhang (2000). This produced transformation matrices to align the Kinects with one another. Due to the lack of information regarding the Kinect's intrinsic properties – and the impact that any intrinsic calibration may have upon the Kinect's 3D vision system – correction and recalculation of intrinsics was not performed, ensuring the original input images were unchanged.

Separately, each IR image of the observed checkerboard was analysed using the OpenCV 'FindCheckerboardCorners' function to identify the 2D checkerboard corner locations. The 3D equivalent of these 2D points were then identified in the collected 3D scans, producing a series of 3D checkerboard corner points. These points were used to assess the performance of the transformation matrices, producing real world error metrics.

71

The calibration process was repeated three times, in order to assess reliability of the technique. Table 3.1 shows a summary of the results from the three calibrations. The pixel re-projection error shows the average error in alignment of a pixel between a pair of IR images. Similarly, the RMS error in millimetres shows the root mean square re-projection error of the checkerboard corners in real world 3D coordinates.

Table 3.1 – Re-projection errors after calibration using Zhang's technique

Trial Number	Kinect 2- Kinect 1 RMS (mm)	Kinect 2- Kinect 1 Error (px)	Kinect 3- Kinect 2 RMS (mm)	Kinect 3- Kinect 2 Error (px)
1	60.89	0.30	146.39	0.46
2	111.52	0.30	137.25	0.30
3	67.59	0.30	106.63	0.30

Results show the pixel re-projection errors to be very small, suggesting a good calibration on all three occasions. However, the RMS errors in real world coordinates (millimetres) are shown to be very high, suggesting poor alignment between Kinect pairs. In order to investigate further a 3D scan of a torso manikin was collected with each Kinect. The transformation matrices formed from the calibration process were used to align the scans from each Kinect with one another, further assessing the real world alignment.
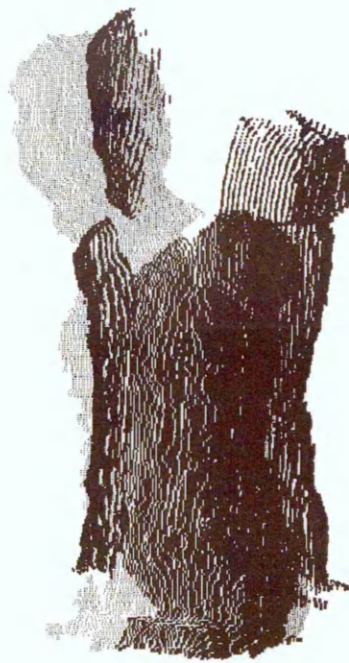
Figure 3.15 – Scans of a torso manikin after alignment using results from Zhang's calibration technique

Figure 3.15 clearly shows very poor alignment of the three scans, agreeing with the RMS values reported in table 3.1, and demonstrating the unsuitability of the calibration technique.

The considerable difference between the re-projection errors in pixel and metric units suggests the Kinect must apply some form of integral camera model or transformation matrix to data from the IR camera, meaning there is not a single direct transformation between points in the IR camera and 3D space: despite the IR camera being the basis for the Kinect's 3D vision system. This suggests a 2D calibration approach cannot be used as the basis for extrinsic calibration when using the Kinect's integral coordinate system transformation functions for converting between IR image space and 3D space. As Burrus (2011b) worked with the Kinect's raw disparity units and formed his own coordinate system transformation functions for converting to 3D space, it appears the data were not vulnerable to the problems identified here.

A considerable drawback of this calibration technique is the amount of time required to collect the necessary calibration images, owing to the need to calibrate cameras in pairs. The amount of time required would be further increased were four cameras calibrated using this approach. In this case, one camera would need to use two transformation

matrices to align it with a base Kinect (forming the global coordinate system), likely meaning one Kinect will have greater misalignment than the other three.

3.5.2 3D Checkerboard rigid body calibration

Due to the identified problems with 2D calibration techniques it was instead decided to investigate potential 3D calibration techniques. The same coordinate space is therefore used for both calibration and data capture, meaning the problems of spurious alignment as observed previously should not be apparent.

The first approach investigated was based upon the black and white checkerboard used by Zhang's calibration technique (Zhang 2000). Four Kinects were used for this investigation, representative of the final system, and laid out as shown in figure 3.16 with a central capture area of 0.4m x 0.4m.
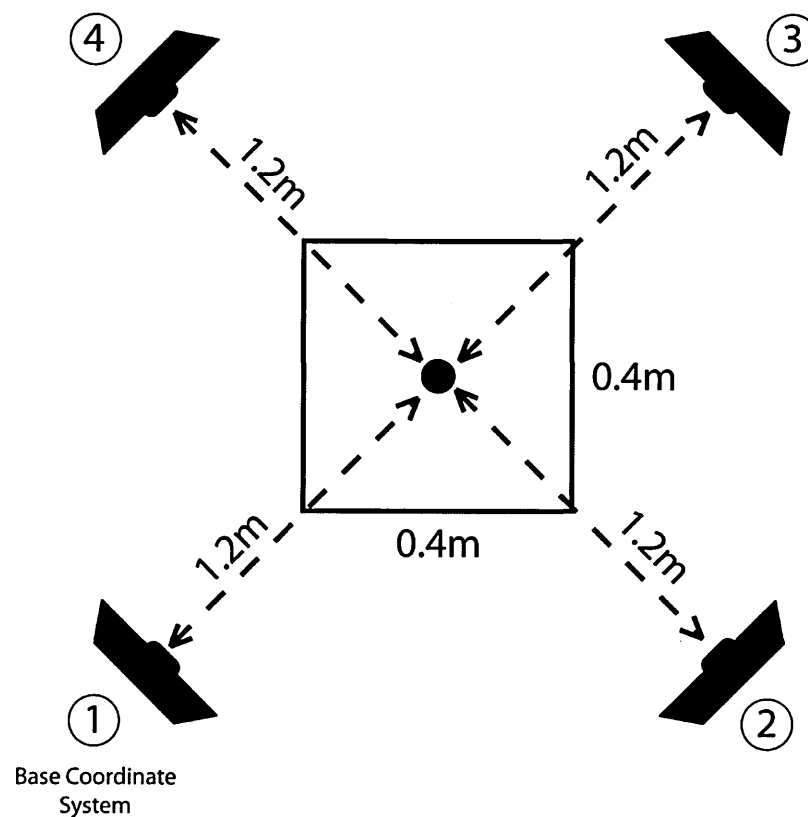
Figure 3.16 – Layout of the Kinects for calibration using a 3D checkerboard technique

Like the 2D method, a significant drawback of this approach is that only two Kinects are able to observe the calibration checkerboard at once. This means the Kinects must be calibrated with respect to one another in pairs, assuming a reference to a single Kinect which becomes the global coordinate system. The drawback of this technique is that one Kinect must undergo two transformations in order to align its data with the base Kinect, meaning one Kinect is likely to be further out of alignment than the others.

A simple software application was written to control and capture the required data from the Kinects, based upon the Microsoft Kinect SDK. For a given calibration image collection, the checkerboard was positioned so that its corner points were visible by both cameras in the pair. Firstly, an image of the checkerboard was captured with the IR camera and its checkerboard corner locations identified in 2D using the OpenCV 'FindCheckerboardCorners' function (WillowGarage 2012), producing sixty 2D points per Kinect, figure 3.17. With the checkerboard in the same place, a 3D scan of the checkerboard was collected with the Kinect. This enabled the 2D points previously identified to be converted into 3D, producing sixty 3D points per checkerboard. As before, the Kinect's IR projector was covered during IR image capture and the scene illuminated with an external incandescent light source. When the 3D scan was captured, the IR projector was uncovered and the external light source switched off.
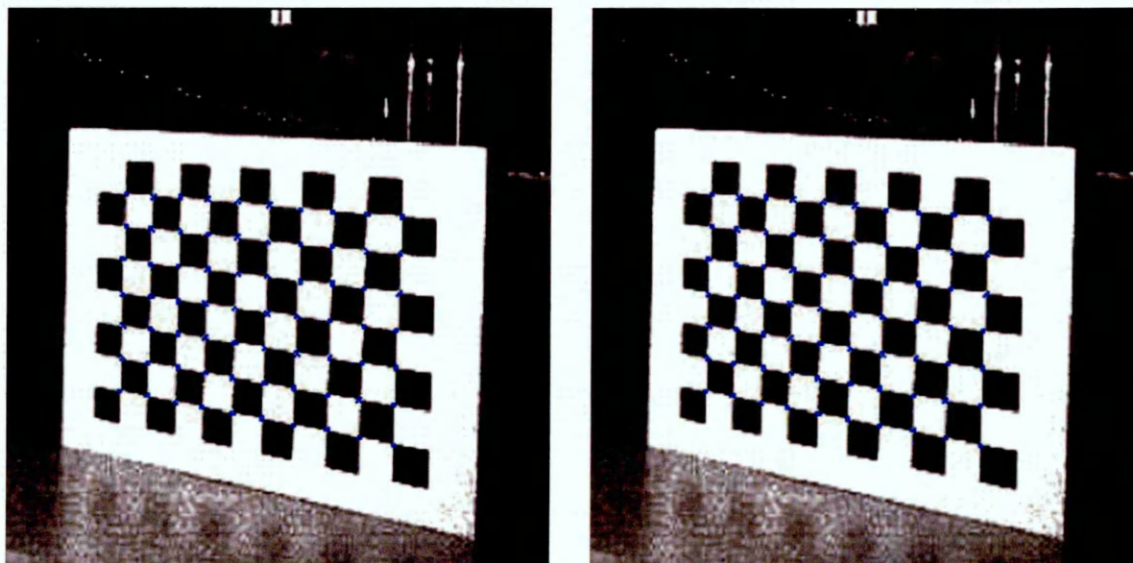
Figure 3.17 – Checkerboard corner locations identified in the 2D image before conversion to 3D

The checkerboard was placed in twenty positions covering the capture volume of the Kinect pair, and the above process repeated for each position. This produced a total of 1200 3D point correspondences per Kinect pair. Transformation matrices were then calculated to align the Kinects with one another.

As previously discussed, the scanning system is considered a rigid body as the position of the Kinects does not move between calibration and data collection. Therefore a rigid body transformation algorithm can be used to produce the transformation matrices to align one Kinect with another. This is shown in equation 3.3, where P_n' is the transformed version of P_n, linked by a 3 x 3 rotation matrix (R) and the 3 x 1 translation vector (v): forming the 4 x 4 transformation matrix.

$$P_n' = RP_n + v$$
<div align="right">**Equation 3.3**</div>

Given the two sets of N (1,200 in the case of this study) corresponding 3D points (p) in each Kinect's local coordinate system (P_1 and P_2 respectively), the rotation matrix (R) and the translation vector (v) were obtained using a common approach based upon singular value decomposition (SVD) (Challis 1995).

The mean location (m_1 and m_2 respectively) of the N points in each of the Kinect's local coordinate systems were first calculated and subtracted from the point locations (P_1 and P_2), separating the rotation and translation, shown in equation 3.4.

$$m_n = \frac{1}{N} \sum_{n=1}^{N} P_n$$
<div align="right">**Equation 3.4**</div>

The rotation was then estimated by first generating matrix A, shown below in equation 3.5, where $\bar{P}_1$ and $\bar{P}_2$ are 3 x N matrices containing the corresponding 3D points with the respective mean positions subtracted.

$$A = \bar{P}_1 (\bar{P}_2)^T$$
<div align="right">**Equation 3.5**</div>

The singular value decomposition of A was then calculated, equation 3.6, where U, D, and V are the products of the SVD operation.

$$UDV^T = A$$

<div align="right">**Equation 3.6**</div>

Subsequently, this enabled calculation of R, the 3 x 3 rotation matrix, as shown below in equation 3.7.

$$R = VU^T$$

<div align="right">**Equation 3.7**</div>

With the value of R known, the translation vector, v, was calculated as shown below, in equation 3.8, where m_1 and m_2 are the mean locations of the 3D corresponding points in the local coordinate systems of Kinect 1 and Kinect 2 respectively.

$$v = m_2 - Rm_1$$

<div align="right">**Equation 3.8**</div>

The calibration process was repeated three times in order to test the reliability of the approach. Results of the three repeated calibrations are shown in table 3.2. The reported RMS values represent the re-projection error of all one thousand two hundred points per Kinect pair after application of the respective transformation matrix. In order to align scans from Kinect three with Kinect one (the adopted global coordinate system) it is evident that the scans have to undergo a transformation from Kinect three to Kinect two, and then a transformation from Kinect two to Kinect on one, leading to a higher RMS in this device.

Table 3.2 – Calibration re-projection errors using the 3D checkerboard approach

Trial Number	Kinect 1 – Kinect 4 RMS (mm)	Kinect 1 – Kinect 3 RMS (mm)	Kinect 1 – Kinect 2 RMS (mm)
1	3.50	7.36	2.56
2	2.10	5.40	2.88
3	2.38	5.05	2.25

The results shown in table 3.2 demonstrate much lower RMS values than those previously reported in table 3.1, suggesting a much better alignment of the 3D scans would be observed. As before, a 3D scan of a torso manikin was captured with the four Kinects and the requisite transformation matrices used to align the four scans, forming a

complete 360° model. This is shown in figure 3.18, demonstrating considerably better scan alignment than previous investigations.
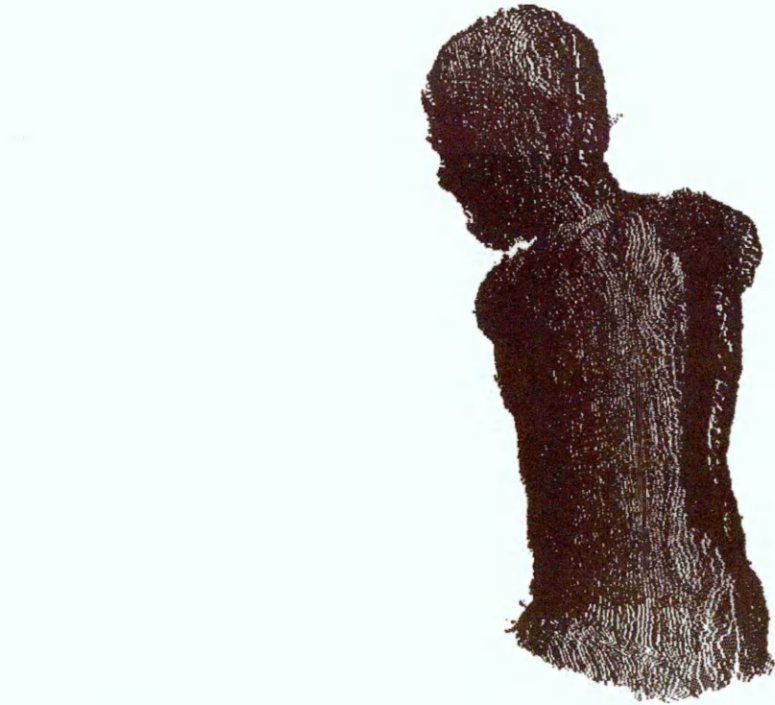
Figure 3.18 – 3D Scans after alignment using the transformation matrices from the 3D checkerboard technique

Table 3.2 shows comparable RMS values across the three repeats, suggesting high reliability of the calibration technique and comparable alignment on all occasions.

An area of potential improvement focusses on the conversion of coordinates from 2D to 3D space. Presently, the direct 3D equivalent of a point in the 2D image is identified for each of the checkerboard corners. As the Kinect's depth data is prone to be noisy, there is the potential that the depth of a singly identified pixel is spurious in relation to its neighbours, possibly introducing errors into the calibration process. Therefore, averaging the depth around the points surrounding the identified checkerboard corner may lead to greater performance of the calibration technique.

The software application was modified to return the depth of each identified checkerboard corner averaged over a varying window size. The calibration procedure was repeated, and the results for each window size input to the rigid body transformation algorithm. As before,

table 3.3 reports the RMS re-projection errors in the 1200 identified checkerboard corner locations per Kinect pair. For simplicity, only a single Kinect pair (Kinect 1-3) was used as the basis for investigation.

Table 3.3 - Calibration re-projection errors using the 3D checkerboard approach with a depth averaging window

Window Size (px)	Repeat 1 RMS (mm)	Repeat 2 RMS (mm)	Repeat 3 RMS (mm)
±20	10.90	22.98	5.67
±15	3.69	3.70	3.65
±10	4.00	4.81	3.36
±5	3.16	3.38	3.55

Table 3.3 shows a considerable difference between the four window sizes, with the ± 5 pixel window consistently returning the lowest RMS errors. This is logical, as on occasions the checkerboard may be located at high angles of incidence with relation to the Kinect. On occasions such as this, the reported depth over a window as large as 40 pixels will vary considerably, leading to the high RMS values reported above in table 3.3. With a ± 5 pixel window producing a better result than that without a window, it was expected there was an optimum window size somewhere in between, whereby the window size was large enough to smooth typical noise in the Kinect's depth data, but small enough so the depth data wasn't affected by the angle of the checkerboard.

With this in mind – and the expectation that the optimum window size may vary between calibrations – the calibration code was further improved, adding a RANSAC optimisation algorithm to determine the most accurate calibration. At the point of calibration, multiple calibration files were produced, containing the depth of each identified corner location averaged over a window size of ± 20 pixels to ± 0 pixels (i.e. just the point) in one pixel increments. The requisite transformation matrices were calculated for each window size, and used to calculate the post transformation RMS re-projection error of the calibration points. Simply, the optimum window size and hence the transformation matrices returned by the RANSAC operation and used for the ensuing calibration process was the window size delivering the lowest RMS point re-projection error.

In summary, this technique produces notably better alignment than the previously investigated techniques: suggesting a 3D extrinsic calibration technique is most applicable to the Kinect. The depth averaging window appears to further improve performance of the calibration technique, leading to smaller re-projection errors. However, as previously discussed in section 3.5.1, the checkerboard calibration technique suffers from the drawback that one Kinect must undergo two transformations in order to align it with the global coordinate system, leading to higher RMS re-projection errors.

3.5.3 3D Sphere rigid body calibration

Investigations showed the 3D checkerboard approach was considerably better than the previously investigated 2D techniques, affirming thoughts that 3D techniques were the best to follow. However, the 3D checkerboard technique suffers from the same problems as the 2D checkerboard approach, namely that the Kinects must be calibrated in separate pairs, resulting in one Kinect requiring two transformations to become aligned with the global coordinate system. As a result an alternative technique was investigated, simultaneously calibrating all four Kinects and hence eliminating the need for one Kinect to undergo a double transformation.

Four Kinects were again used for this investigation, laid out as shown in figure 3.16 and figure 3.19.

Figure 3.19 – Setup of the Kinect for calibration using the 3D spheres technique

A calibration rig was manufactured, figure 3.20, comprising three spheres (120 mm diameter) mounted on a base plate and suspended at different heights and positions. The rig is constructed in such a way that the three spheres are always visible by the four Kinects, regardless of the rig's position in the capture volume.
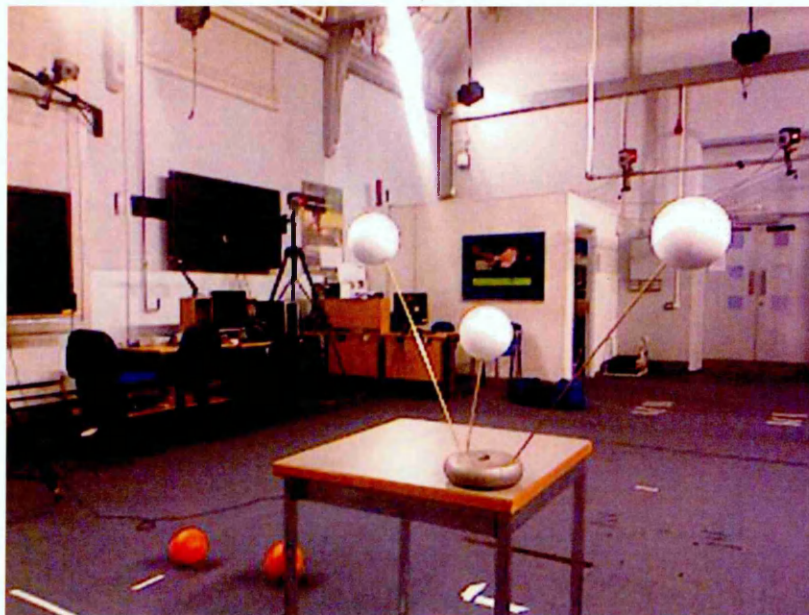
Figure 3.20 – Spheres calibration rig

A simple software application was written to control and capture the required data from the Kinects, based upon the Microsoft Kinect SDK. Data processing, image processing,

and calculation of calibration parameters was performed using scripts written using the MATLAB scripting language.

Firstly, the software was used to define depth limits for each Kinect, removing background objects from view and leaving each FOV focussed upon the central capture volume. Next, the calibration rig was positioned within the capture volume and a 3D scan of the rig was obtained from all four Kinects. As identified in 3.4.2, scans were collected from each Kinect one after the other: due to the need to turn on/off the Kinect's IR projectors to prevent interference. For the purpose of calibration, a scan comprised a depth image (saved as a JPEG image file, figure 3.21), and a 3D point cloud (saved as a list of XYZ points in a CSV file). Owing to the earlier defined depth limits, the depth image only contained the spheres and supporting tripod as white areas, and the background as black.
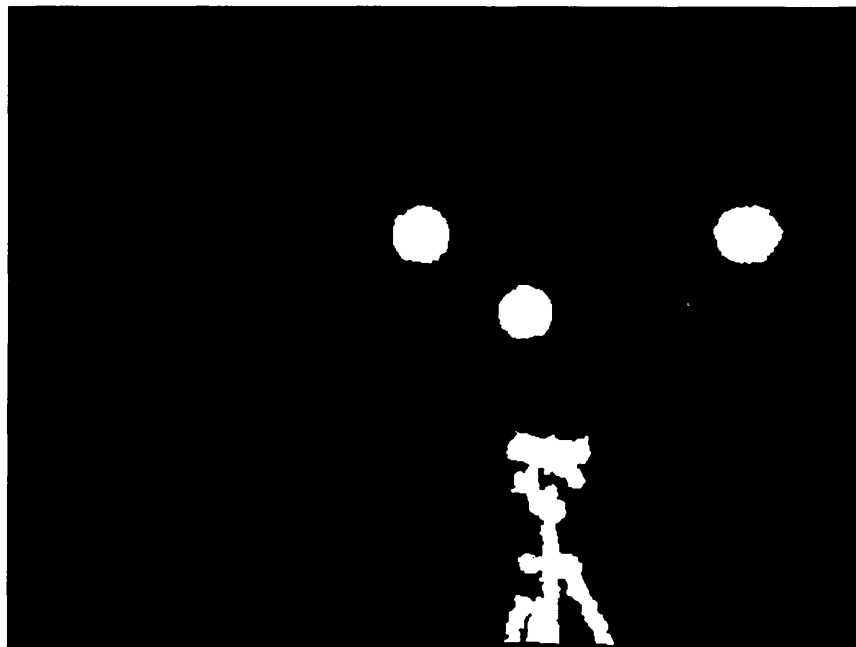
Figure 3.21 – Depth image of the spheres calibration rig and supporting tripod

This process was repeated, placing the calibration rig in twelve random positions throughout the capture volume and a scan captured in each position. The captured calibration data was then analysed using the MATLAB scripts.

Focusing on one Kinect at once, the depth images were first loaded in by the scripts. Prewitt edge detection (sensitivity of 0.4) was used to convert the depth images to black

and white binary edge images, showing the outline of the three spheres. Hough transforms (Ballard 1981), initialised with the number of pixels a sphere was expected to occupy, were then used to identify the 2D centre of each sphere in the binary edge image, figure 3.22.
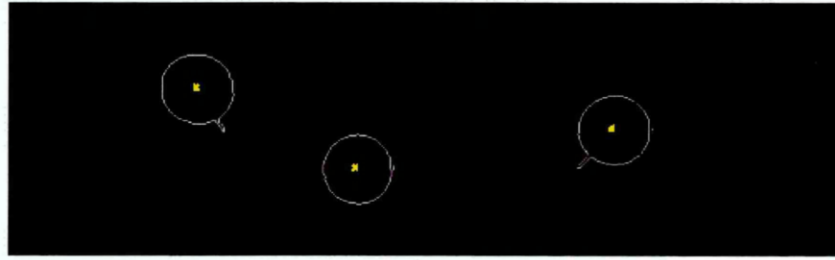
Figure 3.22 – Binary edge image of the calibration rig with the sphere centres identified in 2D

The corresponding point cloud was then loaded in by the scripts. The 3D equivalent of the sphere centres previously located in 2D were then identified from the point clouds: locating the sphere centres on the front faces of the spheres, figure 3.23.
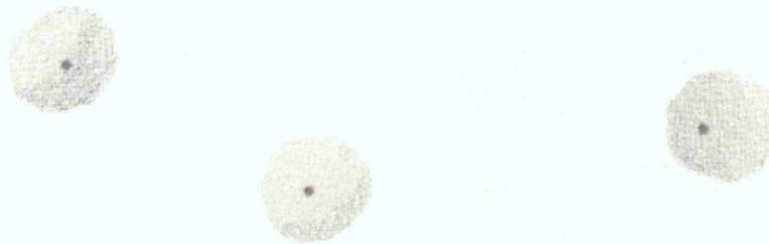
Figure 3.23 – The initial 2D sphere centre locations converted into 3D points on the front face of each sphere

Points laying within a search window of the known sphere radius plus 20%, centred on each sphere centre were then identified. The sphere centres previously identified on the front face of the spheres were projected in a positive Z direction by the known sphere radius, providing an initial estimate of the actual sphere centre.

A nonlinear minimisation of Euclidian point to point standard deviation was then used to identify the actual 3D sphere centres based upon the restricted point collections. In other words, the actual sphere centre was defined as the point having the lowest standard deviation of the distance between the point and all the points in the restricted

collection. For reference, identified sphere centres were then plotted on the original point cloud, figure 3.24.

Figure 3.24 – The actual sphere centres identified in 3D

This process was repeated for all of the calibration images corresponding to the four Kinects, creating four arrays of 3D sphere centre locations. Regardless of the different viewpoints of the spheres, the sphere centres remain the same: meaning the four arrays contain points which are common to all four Kinects.

In the same manner as previously discussed, pairs of point arrays were formed, maintaining reference to one Kinect which became the global coordinate system. The point arrays were then used in conjunction with the algorithm by Challis (1995), shown in equation 3.3 to equation 3.8, calculating the transformation matrices between pairs of Kinects. Importantly, the fourth Kinect (figure 3.16) which previously had to undergo two separate transformations only requires one transformation, owing to the sphere centres being common to both the reference Kinect and all other Kinects.

The calibration process was repeated three times, positioning the calibration rig in twelve positions which cover the capture volume. In the same manner as before, table 3.4 shows the results of the three calibrations, represented as the RMS of the re-projection error between the thirty six calibration points (twelve positions, three sphere centres).

Table 3.4 – Spheres calibration RMS re-projection errors

Trial Number	Kinect 1 – Kinect 4 RMS (mm)	Kinect 1 – Kinect 2 RMS (mm)	Kinect 1 – Kinect 3 RMS (mm)
1	4.65	3.10	2.98
2	4.55	3.07	2.76
3	4.40	3.27	3.11

Table 3.4 shows an improvement in calibration re-projection RMS errors in comparison to the results previously presented in table 3.2 and table 3.3. RMS errors are comparable across the three different calibrations, suggesting a reliable technique. Importantly, all three Kinects requiring transformation have comparable RMS errors, eliminating the previously discussed problem of one Kinect undergoing a double transformation. As before, a 3D scan of a torso manikin was captured with the four Kinects and the requisite transformation matrices used to align the four scans, forming a complete 360° model, figure 3.25.

Figure 3.25 – 3D Scans after alignment using the spheres calibration technique

Visually, there appears little difference between the two scans of the manikin as shown in figure 3.18 and figure 3.25. However, close inspection of the aligned scans in figure 3.25 shows a notable improvement in the alignment.

However, one drawback of this technique is the amount of time taken to move the calibration rig throughout the capture volume and collect the required number of calibration scans/images. This is further exacerbated due to the calibration rig only having three points, in comparison to the sixty of the calibration checkerboard. In previous investigations the number of calibration points has varied considerably, based upon positions which cover the capture volume and the number of calibration points available in the calibration object. However, this should be optimised, providing enough points to reliably constrain the calibration process and cover the capture volume, without collecting too many points and positions so the calibration process becomes overly arduous.

In his paper proposing the rigid body transformation algorithm used as part of the calibration process, Challis (1995) suggests a minimum of eight point correspondences are required to reliably constrain the calculation of transformation matrices. Inclusion of additional points only has minimal impact upon the quality of the calibration, figure 3.26.
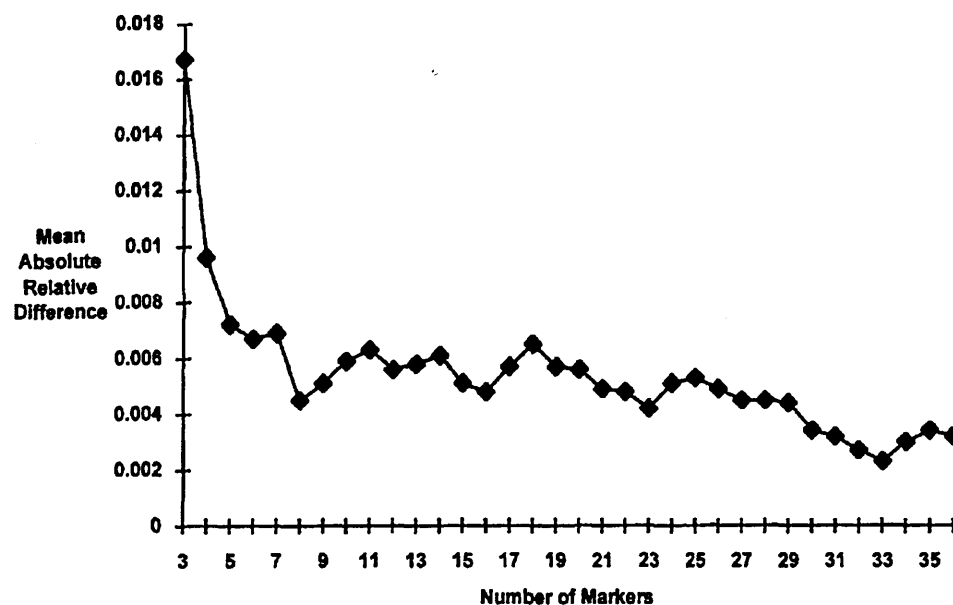
Figure 3.26 – Variation in the mean absolute relative difference between points after transformation versus the number of points used to calculate the transformation (Challis 1995)

It was proposed that nine positions was sufficient to cover the capture volume with calibration points, ensuring the calibration would not become optimised to only a small area of the entire capture volume. Using the calibration rig with three spheres, this resulted in a total of twenty seven points per Kinect.

Referring back to figure 3.26, it is apparent when given more than eight calibration points there is potential for the calibration to be degraded on some occasions, reaching local minima and maxima. With this in mind, there is a need to perform some form of optimisation on the calibration points, ensuring the optimum number of calibration points are used to calculate the transformation matrices.

A RANSAC calibration point discrimination algorithm was developed. Firstly, a random set of eight points is selected and the transformation calculated. All points in the original data set are then transformed using the calibration results and the RMS re-projection error calculated. The calculated transformation matrix and associated re-projection error are then stored in an array. This process is repeated up to a maximum number of iterations. Once complete, the transformation matrix returning the lowest re-projection error is returned. Figure 3.27 illustrates this iterative process.
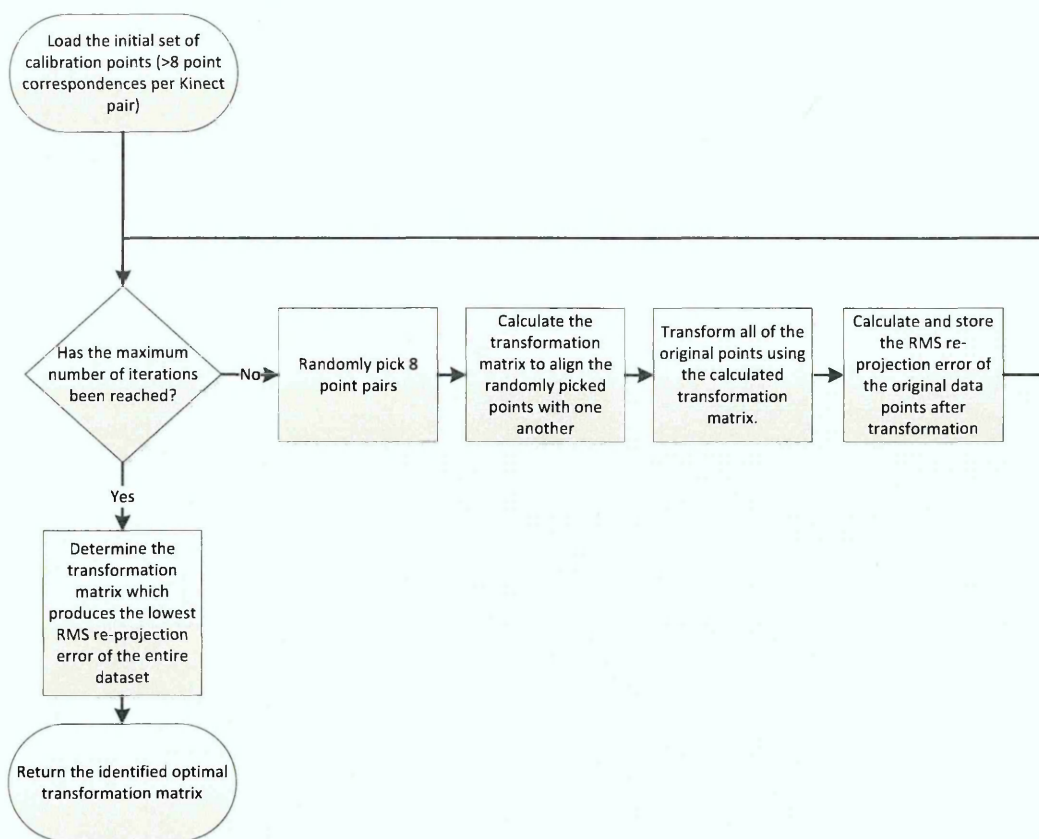
Figure 3.27 – RANSAC calibration optimisation process

In order to test this technique, the calibration data collected as part of the results presented in table 3.4 was re-analysed. The calibration point arrays were processed through the RANSAC algorithm and the optimised RMS re-projection errors calculated. Table 3.5 shows the RMS values calculated from the RANSAC optimised calibration, which should be read in comparison with the initial RMS values shown in table 3.4.

Table 3.5 – Spheres calibration RMS re-projection errors before and after the RANSAC optimisation process

Trial Number	Kinect 1 – Kinect 4 RMS optimised (mm)	Kinect 1 – Kinect 2 RMS optimised (mm)	Kinect 1 – Kinect 3 RMS optimised (mm)
1	3.98	2.45	2.32
2	3.70	2.49	2.25
3	3.72	2.42	2.43

Table 3.5 shows the RANSAC optimisation process produces an improvement in RMS re-projection error on all occasions and all Kinects, whilst also improving reliability of the calibration process, indeed showing it is an important process.

Initial investigations were conducted using custom MATLAB processing scripts. Running on an Intel i5 computer, it took around twenty minutes to calculate the optimised transformation matrices for the three Kinects requiring transformation: assuming thirty six calibration points. With this in mind, the original MATLAB scripts were integrated within bespoke calibration software, developed using C# and the Microsoft Kinect SDK (discussed in more detail in section 3.5.4). This considerably reduced the calibration time, instead taking around thirty seconds to calculate the optimised transformation matrices.

3.5.4 Summary

Four extrinsic calibration approaches were investigated in order to establish a suitable technique for aligning the 3D scans obtained from the four different viewpoints within the scanning system. The 2D calibration technique of Zhang (2000) appeared unsuitable for this application, seemingly owing to the Kinect applying its own integral transformation algorithms when converting from 2D points in the IR camera to 3D points, meaning a direct extrinsic transformation does not exist. With this in mind, two 3D based extrinsic calibration techniques were developed.

The 3D checkerboard calibration technique produced acceptable results, further optimised with application of the depth averaging window to eliminate problems associated with noise in the Kinect's depth data. However, it suffered from the problem that one Kinect had to undergo two transformations to align it with the base coordinate system, leading to greater misalignment in one of the Kinects.

The 3D spheres calibration technique further improved upon the results from the 3D checkerboard technique, producing higher accuracy and reliability. Although a number of calibration points are required throughout the capture volume, the technique is simpler and quicker to achieve than the checkerboard approach. Furthermore it

eradicates the need for one Kinect to undergo two transformations, meaning all Kinects have comparable RMS re-projection errors.

With this in mind, the 3D spheres calibration technique was decided upon as the basis for the scanning system. In order to speed up the calibration process an improved calibration rig was developed, figure 3.28.

Figure 3.28 – Improved spheres calibration rig

The base of the rig measures 0.4 m x 0.4 m, equivalent to the capture volume of the scanning system as identified in section 3.4.1. The pegs in the base define nine positions in which the spheres are placed, ensuring the entire capture volume is covered with calibration points.

A bespoke software application was written using the C# programming language and Microsoft Kinect SDK (Microsoft Corporation, Redmond, USA), implementing the previously discussed Matlab scripts and data collection software into a single application, figure 3.29 and figure 3.30.
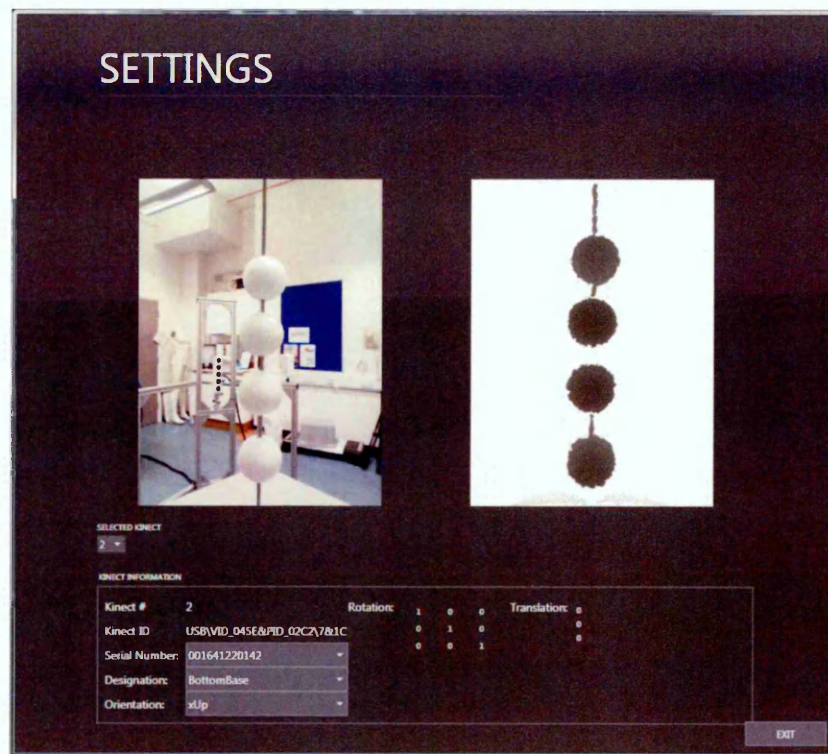
Figure 3.29 – Settings window in the calibration software
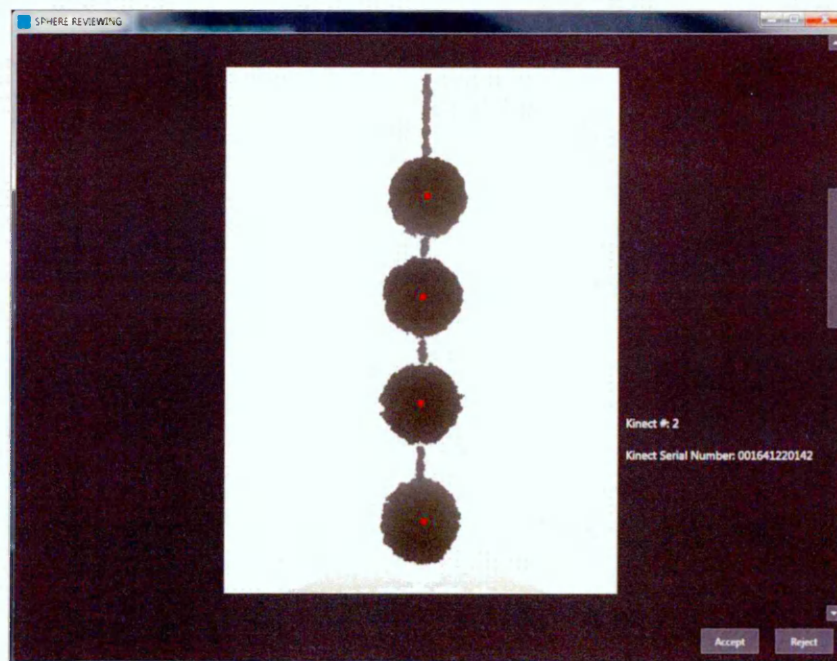
Figure 3.30 – Image review window in the calibration software

After collection of the calibration data, calculation of the transformation matrices for each Kinect (including RANSAC optimisation) takes around thirty seconds. The

software reports the RMS re-projection error of each Kinect pair, based upon all of the original calibration points, figure 3.31.
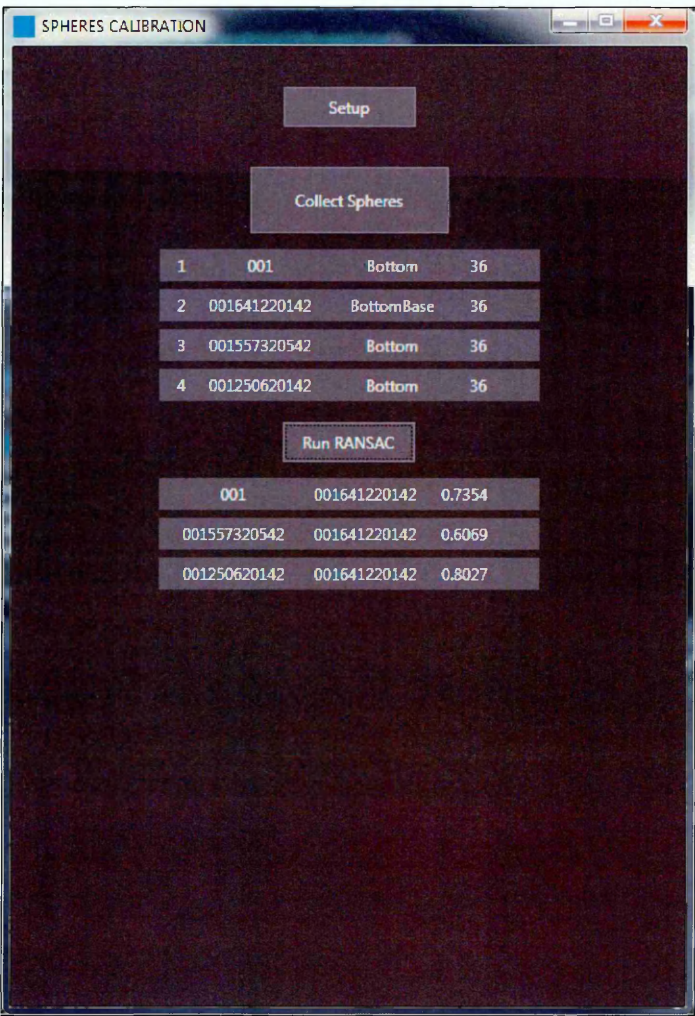
Figure 3.31 – Calibration review window in the calibration software

3.6   Data capture process

3.6.1   Data capture and handling

A bespoke software application was written to handle the capture, processing, and display of the 3D scan data from the system, figure 3.32.   The software was written using the C# programming language and Microsoft Kinect SDK (Microsoft Corporation, Redmond, USA).

Figure 3.32 – The developed scanning software application

Assuming the scanning system has already been calibrated (discussed in section 3.5.4), the data collection and processing software automatically loads in the calibration file, loading in the extrinsic calibration and distortion correction information used as part of the data collection process.

The software application enables the user to carry out the following tasks:

- View the colour and depth data streams from each Kinect sensor.
- Check the transformation matrices generated as part of the extrinsic calibration process.
- Collect and save 3D scans, applying the relevant data post processing techniques: distortion correction (discussed in chapter 4) and fringe removal (discussed in chapter 5).
- Load in and manipulate previously collected 3D scan files, discussed in 3.6.2.
- Manually digitise points palpated on the human body, discussed in chapter 5
- Analyse the collected 3D scan files and produce values of interest, discussed in chapter 5.
- Export the 3D scan files in a variety of industry standard file formats

Figure 3.33 shows the process of data capture, post processing, and storage of the 3D scan data. Individual data processing steps are discussed in more detail in chapter 5.

```
┌─────────────────┐
│ Scan button     │
│ clicked in the  │
│ software        │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ A beep is       │
│ sounded to      │
│ inform the      │
│ participants to │
│ remain still    │
└─────────────────┘
         │
         ▼
      ◇ 3D Scans       ┌──────────────┐   ┌──────────────┐   ┌──────────────┐   ┌──────────────┐
      collected  ──No──│ Turn on the  │──▶│ Delay for    │──▶│ Simultaneous-│──▶│ Turn off the │
      from all 4       │ IR projector │   │ 200ms to     │   │ ly collect   │   │ IR projector │
      Kinects?◇        │ of the next  │   │ allow the    │   │ uvZ and RGB  │   │              │
         │             │ Kinect in    │   │ data to      │   │ data         │   │              │
        Yes            │ the list     │   │ stabilise    │   │              │   │              │
         │             └──────────────┘   └──────────────┘   └──────────────┘   └──────────────┘
         ▼
┌─────────────────┐
│ A beep is       │
│ sounded to      │
│ inform the      │
│ participants the│
│ scanning        │
│ operation is    │
│ complete        │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ Apply the fringe│
│ removal filter  │
│ to the uvZ data │
│ from each Kinect│
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ Apply the depth │
│ distortion      │
│ correction model│
│ to the uvZ data │
│ from each Kinect│
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ Convert all uvZ │
│ data to XYZ data│
│ using the       │
│ Kinect's        │
│ integral        │
│ transformation  │
│ functions       │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ Align the XYZ   │
│ data with the   │
│ global          │
│ coordinate      │
│ system using    │
│ data from the   │
│ extrinsic       │
│ calibration     │
│ process         │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ Find the XYZ    │
│ point           │
│ correspondences │
│ for the RGB data│
│ from each Kinect│
│ using the       │
│ Kinect's        │
│ integral        │
│ transformation  │
│ functions       │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ Transform the   │
│ point           │
│ correspondances │
│ to the global   │
│ coordinate      │
│ system using the│
│ transformation  │
│ matrices from   │
│ the calibration,│
│ producing a     │
│ colour rendered │
│ 3D model        │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ Save the XYZ,   │
│ RGB, and        │
│ extrinsic       │
│ transformation  │
│ matrices in a   │
│ binary .scan    │
│ file            │
└─────────────────┘
```

Figure 3.33 – Scanning system data capture process

95

Beeps emitted by the software tell the participant the duration of time they must stand still for, aiding the control of involuntary postural sway. The time in between the beeps is equal to the time taken to obtain a scan from each Kinect, meaning the entire scanning operation takes around 0.8 seconds to complete, discussed in more detail in 3.4.2. With such a low scanning time, it is feasible that the aim of a 10 minute measurement exercise per person could be achieved.

After the data has been collected, the distortion correction algorithm and fringe removal filter is applied to the data in uvZ coordinate space, discussed in more detail in chapter 4 and 5 respectively. The data is then converted into real world 3D XYZ coordinates using the Kinect's integral coordinate mapping function, discussed in more detail in section 3.3.2. After conversion to XYZ points, the transformation matrices calculated as part of the calibration process are applied to the data from each Kinect, aligning the scans with a global coordinate system and producing a 360° scan, discussed in more detail in section 3.5.3 and 3.5.4.

The RGB data collected from each Kinect is then converted to 3D space using the Kinect's integral transformation functions, producing 3D point correspondences for the colour data in the RGB images. The extrinsic transformation matrices are then applied to the 3D point correspondences, producing RGB correspondences for each 3D point in the transformed point cloud scan. Finally, the respective RGB data is stored alongside the 3D points, creating a colour rendered point cloud.

After all these steps are complete, the XYZ, RGB, and transformation matrices corresponding to each Kinect are saved in a bespoke binary file format. This considerably reduces the file size, and is necessary due to the considerable amount of data which must be saved. A typical scan file is therefore around 30 MB in size.

3.6.2 Data display and interpretation

After collection, the binary files can be loaded into the software application for further viewing and analysis. Upon loading a scan file the XYZ and RGB data from each Kinect is converted into a colour rendered triangular mesh. The four individual meshes are then transformed into the global coordinate system using the individual

transformation matrices, creating a 360° model (figure 3.32). Although the meshed model is not watertight and 'stitched together', the good alignment of the four scans creates the same appearance.

Meshing the scans creates greater visual appeal for users of the software, whilst also making digitisation of palpated points much easier. However, it is important to note that it is only the points in the point cloud which are used for subsequent calculations, as discussed in chapter 5. Meshing the Kinect's 3D data is a trivial process, and doesn't require traditional computationally intensive meshing algorithms. Each Kinect returns its data as a rastered 640 x 480 grid of points, which is simple to connect with one another using a series of triangles, figure 3.34, creating a meshed surface.
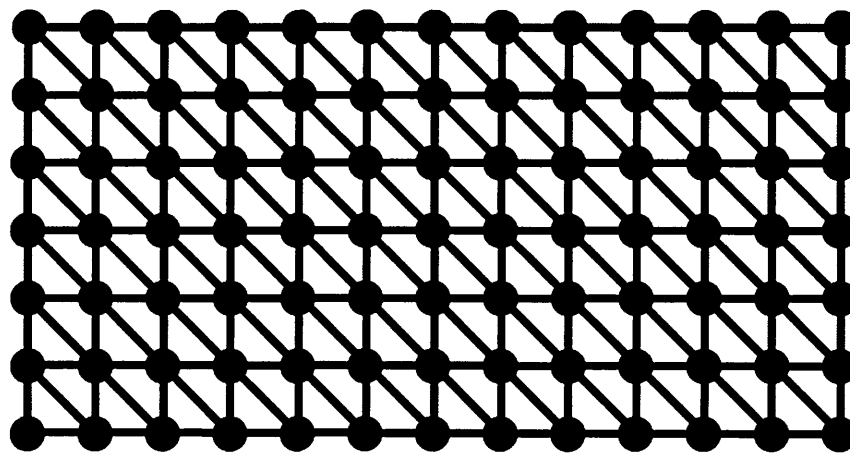
Figure 3.34 – Meshing process for 3D points from the Kinect

3.7 Chapter summary

A low cost 3D body scanning system has been developed, capable of obtaining person specific 3D scans of the body's torso segment. The hardware cost of the entire system is around £1500 (including Kinects, mounting hardware, and computer), meeting the original aim of the project. Subsequent data processing steps enable use of the 3D scans as the basis for calculating person specific BSPs.

A bespoke software application capable of capturing the 3D scans, performing the necessary post processing steps, viewing and manipulating the scans has been developed. The scanning time of around 0.8 seconds suggests that the original aim of a

97

measurement exercise not exceeding 10 minutes per person could easily be met. The software application also enables calculation of numerical parameters of interest (BSPs), based upon anatomical landmarks which can be digitised on the 3D scans. Subsequent chapters discuss the data post processing and calculation steps in more detail.

A bespoke calibration technique and associated software application has also been developed. This provides a simple and reliable technique for aligning the 3D scans from the four Kinects with one another, producing a single 3D scan. One drawback of the calibration approach is the need to collect calibration data from nine positions within the capture volume, which can be a time consuming process despite the bespoke calibration rig, figure 3.28. However, the process is simpler and quicker than the other investigated techniques. Possible improvements to this technique include the collection of only a few initial calibration points which are used to obtain a rough alignment of the scans, which is later optimised via ICP type algorithms. However, this approach is likely to considerably increase the required post processing time (Rusinkiewicz & Levoy 2001).

Figure 3.35 shows the final scanning system, along with the bespoke data collection and analysis software application.
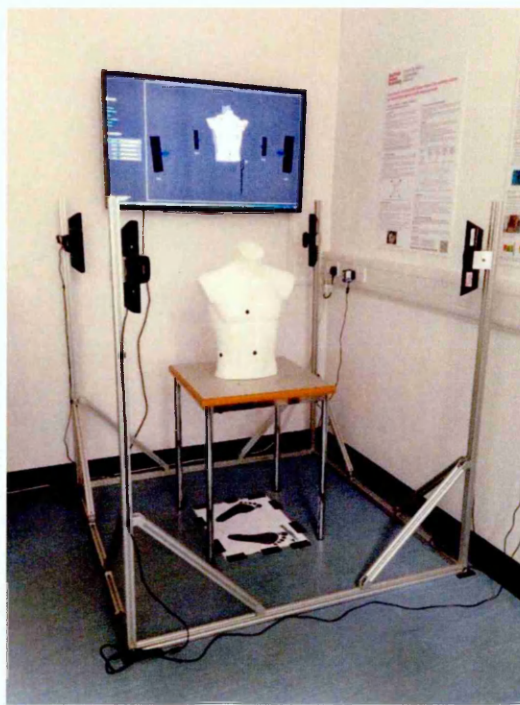
Figure 3.35 – The completed scanning system and associated software

4 METHODOLOGY FOR INDIVIDUAL DEVICE CALIBRATION

4.1 Introduction

The inherent low cost nature of consumer RGB-D sensors likely means they provide poorer quality data when compared to higher cost devices. Inaccuracies can typically be attributed to a number of factors: pseudo random noise, point resolution degrading with distance (Menna et al. 2011), depth binning errors, and distortion of the depth data (Herrera et al. 2012).

A number of accuracy studies focus on the Microsoft Kinect RGB-D sensor, and demonstrate significant distortion of its depth data (Herrera et al. 2012; Smisek et al. 2011; Chow et al. 2012; Khoshelham 2010; Molnar et al. 2012). Smisek et al. (2011) showed the errors to be approximately radially symmetric (figure 4.1) throughout the depth range, resulting in a plane appearing convex about the centre of the depth image. Menna et al. (2011) investigated this distortion, and reported non planarity errors of ± 20 mm when scanning a plane filling the FOV.
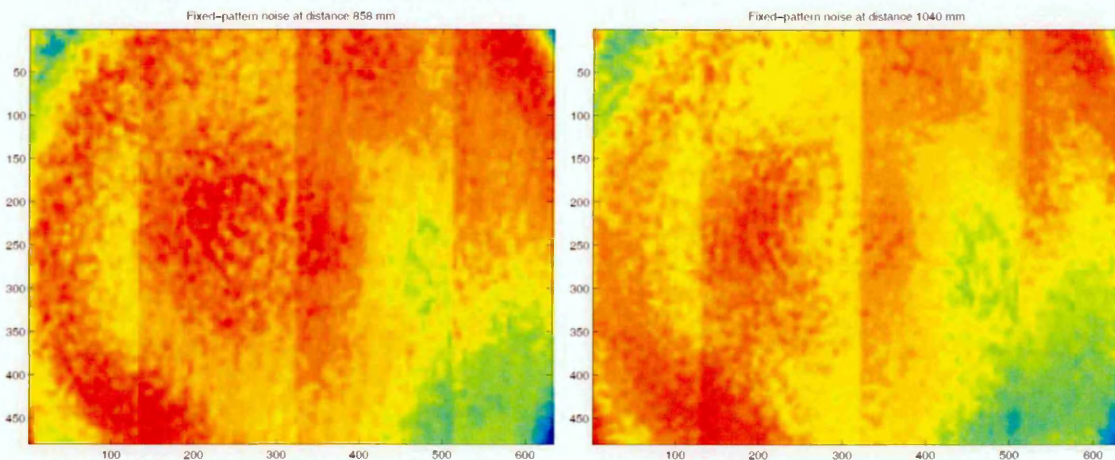
Figure 4.1 - Distortion of the Kinect's depth data at varying depths (Smisek et al. 2011)

The literature suggests the distortion pattern to be approximately equivalent between devices (Herrera et al. 2012), but the magnitude of the resulting error to vary considerably (Boehm 2011). This suggests the problem to be a function of the Kinect's hardware, and the variability due to a factory calibration technique or manufacturing tolerances (Herrera et al. 2012).

As the Primesense technology underpinning the Kinect's 3D vision system is proprietary and closed, there is little information available regarding the factory device calibration procedure. However, it is well reported that devices are calibrated during manufacture and the calibration parameters stored in on board memory (Computer Vision Group- Technische Universitat Mumchen 2013; Herrera et al. 2012; Smisek et al. 2011). This calibration is sufficient for the original intended purpose of the device (as a natural user interface), but appears inadequate for applications such as scanning which require greater accuracy.

4.2 Aims and objectives

Aim

- To characterise and correct the distortion present in the Kinect's depth data.

Objectives

- To conduct a literature search to identify existing methods of correcting distortion of depth data from RGB-D sensors.
- To identify the advantages and disadvantages of previous techniques, and use as the basis for developing a suitable correction technique.
- To perform practical investigations to identify and characterise distortion apparent in the Kinect's depth data.
- To develop a practical method of correcting distortion of depth data from RGB-D sensors, preferably compatible with a range of sensors.

4.3 Investigation of existing techniques

Existing techniques for correcting depth distortion of TOF and structured light RGB-D sensors can typically be split into the following categories:

- Global correction models, considering both depth (Z) and image space parameters (uv) in one process
- Two step correction models, considering depth and image parameters individually. There are also some models which only consider either depth (Z) or image space parameters (uv).

The range of existing techniques are summarised below, before discussing the relative advantages of each approach.

4.3.1 Two step correction models

The simplest correction technique proposed by Smisek et al. (2011) focusses on the Microsoft Kinect. They assume all points with a given v coordinate in the Kinect's depth image (a horizontal row of pixels) have a consistent depth error, regardless of depth within the global space. Smisek calculated the depth errors by collecting scans of a plane over a range of 0.69 m to 1.22 m and plotting the residual error in plane fitting for each v coordinate in the depth image. The depth error for a given v coordinate was calculated as the mean of all the depth errors apparent in that particular v coordinate over the 0.53 m calibration range. Depth values are then corrected in metric units. Smisek reported the standard deviation in plane fitting to reduce from 2.18 mm to 1.54 mm with application of the correction technique. This is inherently a very simple technique, making no consideration of how the depth error may vary over a range of u coordinates corresponding to a single v coordinate. Similarly it makes no consideration how the depth error varies throughout the global depth space, with previous research showing errors in the Kinect's depth data to increase with depth (Khoshelham 2010).

Herrera et al. (2012) improved upon the work by Smisek et al. (2011), taking into consideration how the error varies with depth and image space coordinates, citing this will lead to greater accuracy. Hererra et al. collected plane scans at a range of 0.56 m to 1.24 m between plane and sensor and normalised the depth images to allow analysis of the errors independent of depth. A 2D spatial distortion model (in terms of u and v) was formed from this data, and assumed to be consistent throughout the Kinect's range of operation. The median normalised error for each depth was then plotted, producing an exponential depth error decay model. When correcting depth measurements the calculated offset is therefore based upon a two-step process: depth correction in terms of 2D image space using the spatial distortion model, and a multiplication factor in terms of depth within global space using the error decay model. The correction factor is applied in raw Kinect disparity units rather than metric units, which Herrera et al. suggested leads to greater accuracy. Importantly, independent two step models have

102

been criticised in the literature for providing poor per pixel depth error correction (A Belhedi et al. 2012).

Lindner et al. (2006) propose a two-step method, focussed on correcting depth errors in TOF cameras. They highlight the need to take into consideration both position in the image space, and distance from the sensor (depth). They model the global depth adjustment using a B spline, citing that modelling this error using a model permitting high complexity allows the use of a much simpler (linear) model for per pixel error correction. Given the high number of pixels in a typical depth image, this results in greater storage and processing efficiency. However - as previously stated - such two-step procedures are criticised in the literature (A Belhedi et al. 2012).

4.3.2 Global correction models

Belhedi et al. (2012) review existing depth correction techniques. They suggest that authors often make significant assumptions about the characteristics of the depth errors when producing models, leading to a loss of information, incorrect error models, or restricting the technique to a particular device. Belhedi et al. propose a global correction model to correct depth distortion in TOF cameras (also applicable to structured light cameras) which makes limited assumptions about the characteristics of the depth distortion errors, suggesting this will lead to greater accuracy than existing techniques which comprise two discrete steps and make large assumptions about the characteristics of the depth errors. Unlike previous authors, Belhedi et al. correct depth error in 2.5D (uvZ) space, before later converting to full 3D space. A significant drawback of Belhedi et al.'s method is the need to obtain ground truth measurements, which are notoriously hard to obtain (Amira Belhedi et al. 2012). However, it offers the advantage of correcting fundamental measurement inaccuracies in addition to depth distortion.

Belhedi et al. (2012) later recognise the requirement for a large number of ground truth measurements is practically very challenging and time consuming. They propose an improved method which uses a combination of plane scans and only a few ground truth measurements. A 3D thin plate spline was used to build the distortion model, allowing sufficient complexity in the model whilst making no underlying assumptions about the

characteristics of the errors. As before they employ a global depth distortion model, but apply the corrections in 3D XYZ space, and not the 2.5D uvZ space previously used. Assuming a simple pinhole camera model, they show how these two spaces are equivalent.

The method proposed by Herrera et al. (2012) makes no assumption about the relationship between depth error and actual depth, only that it obeys a calculated exponential depth decay model. However, this means that many calibration plane scans are required to sufficiently constrain the error model. Raposo et al. (2013) suggest the number of plane scans required to reliably constrain the depth error model would be a very time consuming and impractical process. They instead rely upon far fewer plane scans, using a global correction model and open loop processing technique to constrain the error model and estimate the appropriate parameters, citing this will lead to an equivalent level of accuracy with a significant reduction in computation time.

Likewise, Yamazoe et al. (2012) suggest the approach by Herrera et al. is impractical, instead proposing a global correction model to estimate the depth decay model. Rather than relying upon plane scans alone they also estimate the intrinsic parameters of the sensors IR camera and projector, citing this leads to a reduced requirement for calibration plane scans.

However, Belhedi et al. (2012) suggests that whilst such techniques reduce the required collection and computation time, the introduction of global models is likely to over simplify the complex distortions apparent in the 3D data from RGB-D sensors (as shown in figure 4.1), likely reducing the level of accuracy apparent in Herrera et al.'s previous work (2012).

4.3.3 Discussion

The work by Smisek et al. (2011) appears too simple to model the complex depth distortions apparent in RGB-D sensors, as it makes no differentiation between individual u coordinates for a given v coordinate: possibly adding an inherent smoothing or error factor to objects of complex geometry. Similarly it makes no

consideration of how error varies with depth, previously shown to be an important factor (Boehm 2011).

Herrera et al. (2012) make considerable improvements over this early work, considering the relationship between pixel position, depth in global space, and depth error. However, the error correction is calculated in the form of a two-step model which have been criticised for providing poor per pixel error correction (A Belhedi et al. 2012).

The technique proposed by Yamazoe et al. aims to improve the practicality of Herrera et al.'s method, reducing the number of plane scans required by including intrinsic parameters when forming the global correction model. However, Belhedi et al. (2012), suggests the limited number of required plane scans and inherent assumptions made by the model is likely to degrade the performance of Herrera's et al.'s original method.

The most effective technique appears to be that proposed by Belhedi et al. (2012), as the relationship between pixel position, depth in global space, and depth error is considered in a single model. Unlike other global correction models, the method proposed by Belhedi et al. does not make significant assumptions about the characteristics of the depth distortion, allowing sufficient mathematical freedom to consider all possible attributing factors. However, a significant drawback of this approach is the need to obtain ground truth measurements to reliably constrain the model. Importantly, the correction of ground truth within the application of 3D body scanning as presented here is not required, as any measurement errors are taken into consideration as part of the global calibration process. Therefore, such a method could be adapted to apply the necessary correction in a simpler and less time consuming manner.

Correcting depth error in raw Kinect disparity units appears to be the most common approach, and is cited to provide the greatest accuracy (Herrera et al. 2012). However, Belheidi et al. demonstrated that depth distortions can be accurately corrected in 2.5D (uvZ) (A Belhedi et al. 2012) and 3D (XYZ) (A Belhedi et al. 2012) space. They show the two spaces are equivalent to one another, as a known transformation must exist between the two.

4.3.4 Conclusions

The literature suggests that models which make significant assumptions about the underlying characteristics of the depth distortion offer the advantage of requiring significantly fewer plane scans or ground truth measurements to constrain the correction model, also offering significant time savings. However, they appear unable to reliably model the complex depth distortions apparent in RGB-D sensors, owing to the irregular shape of the distortion which changes with distance between the sensor and an object. This may lead to an over generalisation of the errors, and hence a loss of information. Models which make only limited underlying assumptions appear better suited, utilising techniques such as spline fitting which offer sufficient freedom to accurately model the relationship between pixels, depth error, and actual depth.

Two step models are criticised in the literature, suggesting they result in inaccurate per pixel error coefficients. Global error models that jointly consider pixel position and reported depth to calculate the depth error appear to perform the best.

It appears best to correct the depth errors in raw disparity units. However, such information is not available from all RGB-D sensors. Similarly, when working with the Microsoft Kinect and official Microsoft SDK this information is not exposed to the developer. Belheidi et al. (2012) demonstrate an alternative solution, showing that depth errors can instead be corrected in 3D (XYZ) or 2.5D (uvZ) space, therefore providing a practical alternative.

The literature search has not yielded a single existing method which is ideally suited to correcting the complex depth errors apparent in RGB-D sensors such as the Microsoft Kinect. Therefore a method of correcting such distortions will be developed, using knowledge gleaned from the literature.

A global depth correction model will be developed, considering pixel position, depth in global space, and depth error as one. The model will not make significant assumptions about the characteristics of the depth distortion, allowing sufficient freedom to ensure they are accurately modelled. Depth correction will be applied to the Z component in

2.5D (uvZ) space. The 2.5D depth corrected coordinates will then be converted to 3D space using the Kinect's integral coordinate mapping functions.

4.4 Identifying and characterising depth distortion

Before proceeding with development of the depth error correction technique, it was decided to further identify and characterise the distortion apparent in the Microsoft Kinect's depth data.

4.4.1 Methodology

A single Kinect sensor was mounted vertically on a supporting frame, approximately parallel to a planar surface (a taught projection screen), figure 4.2. The Kinect was incrementally moved away from the planar surface, ensuring the entire field of view remained filled with the plane.

Figure 4.2 – Experimental setup for investigating depth distortion

Scans were collected at known distances between the Kinect and plane, allowing the depth distortion to be characterised over the Kinect's typical operating range.

Principal component analysis (PCA) was used to fit a plane to the collected 3D point clouds. Non planarity errors were calculated as the difference between the recorded position and the PCA plane's position at each pixel.

4.4.2 Results

Figure 4.3 and figure 4.4 show the non-planarity errors apparent in the Kinect's 3D data at two distances between the Kinect and the plane. The depth errors shown in figure 4.3 and figure 4.4 are expressed in 2.5D space (uvZ), equivalent to the Kinect's depth image. The magnitude and direction of the planar deviations are indicated by the colour scale, with negative numbers representing a planar deviation towards the Kinect (smaller reported depth) and positive numbers representing a planar deviation away from the Kinect (larger reported depth).
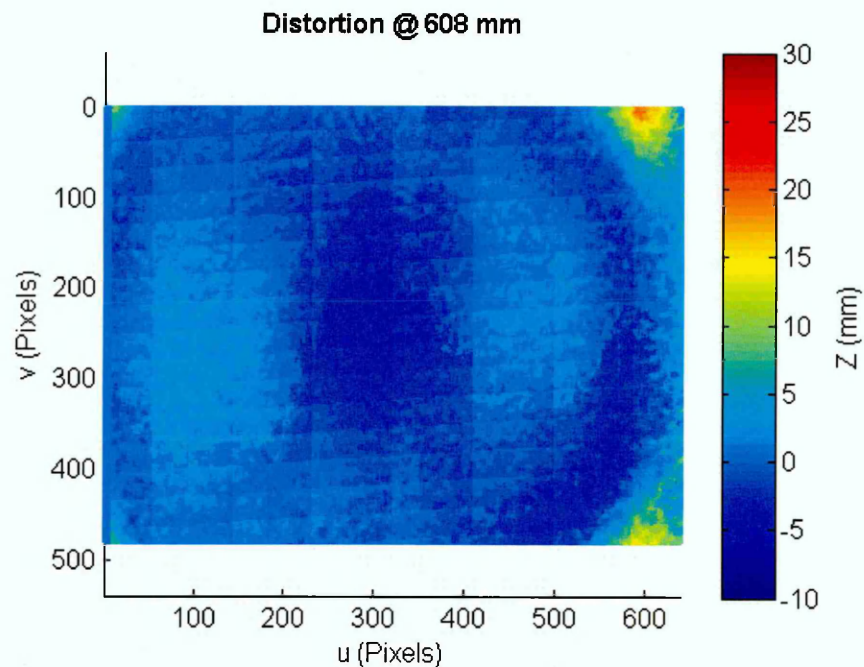
Figure 4.3 – Depth distortion at 0.61m
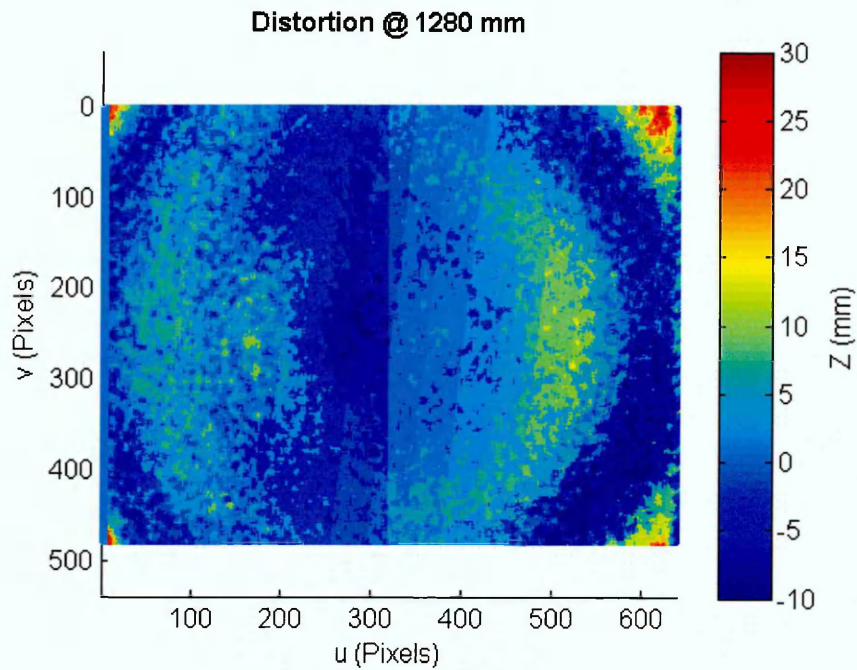
Figure 4.4 – Depth distortion at 1.28m

4.4.3 Summary and conclusions

Figure 4.3 and figure 4.4 show the shape of the depth distortion pattern to be partially radially symmetric, with the greatest errors at the centre of the FOV. As previously discussed, this results in scanned planes appearing convex shaped.

However, the shape of the error pattern changes considerably with an increasing depth (figure 4.4), and varies in magnitude considerably from one side of the depth image to the other. Interestingly, figure 4.4 appears to show the depth errors become coarser at increased depths, however this is expected to be related to the size of the Kinect's depth quantisation 'bins' increasing with distance.

Considering these findings, it is immediately apparent it would be hard to model the error distortion pattern using a model which employs significant underlying assumptions about the shape and characteristics of the depth error. This affirms the previous decision to model the complex distortions using a model which offers a high level of mathematical freedom and makes limited assumptions about the underlying characteristics of the depth error.

109

4.5 Methodology development

This section discusses the development of a technique for correcting the depth distortions apparent in the Microsoft Kinect. However, the methodology described below is applicable to other TOF and structured light cameras with only few modifications.

Section 4.5.1 discusses the development of a technique for capturing and characterising the depth distortions apparent in the Kinect, whilst section 4.5.2 uses this data to develop a distortion correction model. Finally, the distortion correction model is tested in section 4.5.3.

4.5.1 Data collection protocol

A data capture technique similar to that described earlier was used to characterise the depth distortion apparent in the Microsoft Kinect. A single Kinect sensor was mounted on a supporting framework in a vertical orientation, and used to collect scans of a plane filling the FOV, figure 4.2. The vertical orientation had no implication on the distortion correction, but best maximised the available space. The Kinect was placed into its 'near' operating mode, representative of the actual setup in the 3D body scanning system.

Scans of the plane were collected at 10 mm intervals over a distance of 0.5 m to 1.2 m between the Kinect and the plane. This range of calibration was chosen as previous experiments had shown this to be representative of the required capture range within the 3D body scanning system. Accurate real world measurement of the spacing between scans was not important, but a spacing of around 10 mm ensured sufficient constraint when subsequently developing correction models.

Software based upon the Microsoft Kinect SDK (Microsoft Corporation, Redmond, USA) was used to control the Kinect during calibration and capture the necessary data. The software only enabled the Kinect's IR projector whenever a scan was collected, in order to be representative of the body scanning system, and as previous research has shown the Kinect's depth data to become unstable after long periods of IR projector

illumination (Fiedler & Heinrich 2013). For each scan, two sets of data were simultaneously collected: a 3D XYZ point cloud and a corresponding 2.5D uvZ depth image.

The data collection and processing pipeline followed the stages shown in figure 4.5, and are described in more detail beneath. For the purpose of explanation, the flow chart and ensuing explanation consider the processing steps for a single calibration scan. In practice, this process was repeated for each calibration scan throughout the calibration range.
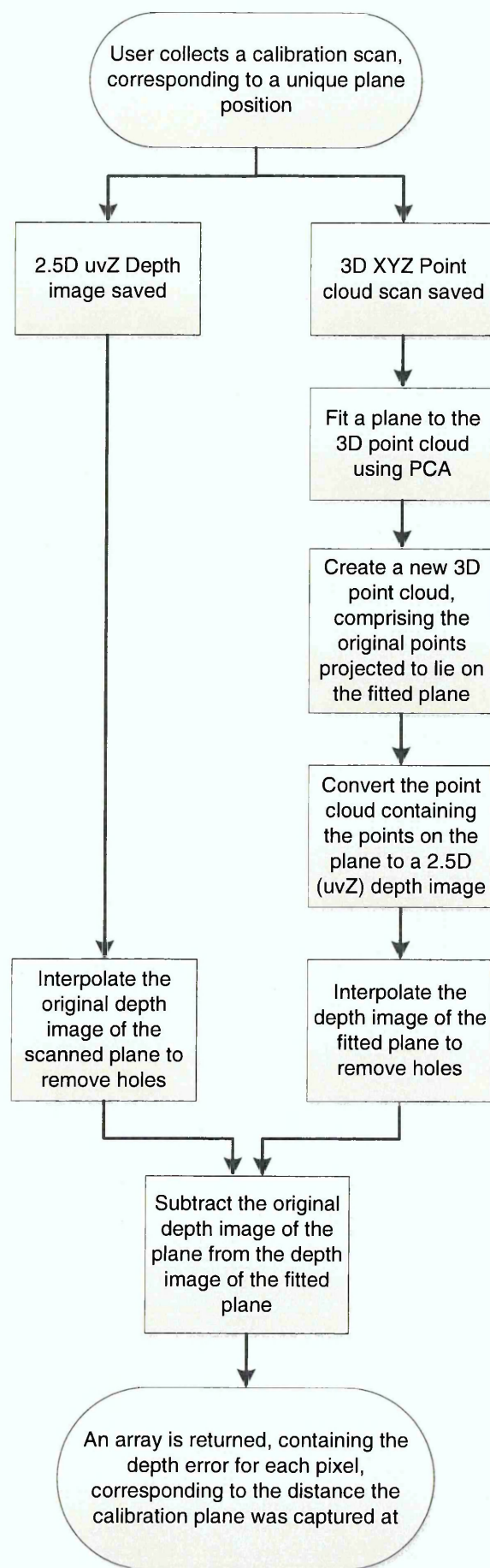
Figure 4.5 - Distortion correction data capture and processing pipeline

Data processing scripts and software based upon the MATLAB language (Mathworks Inc, Natick, USA) and the Microsoft Kinect SDK (Microsoft Corporation, Redmond, USA) were used to process the captured calibration data. After data collection, principal component analysis (PCA) was used to fit a plane to the point cloud scan of the observed plane, figure 4.6, providing the basis for the subsequent depth error calculation.
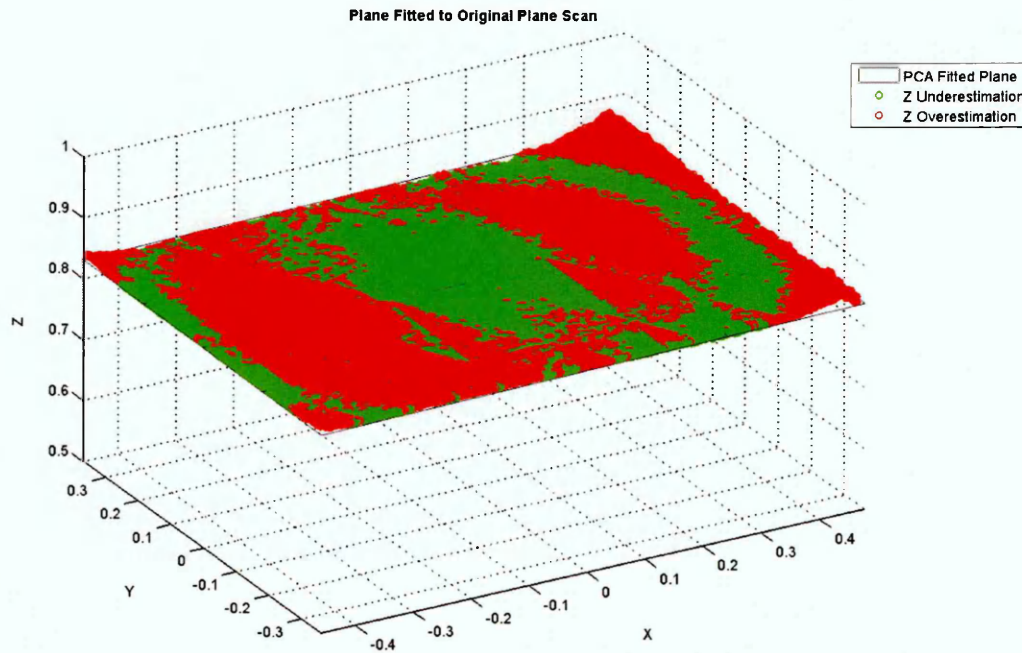
Figure 4.6 – Plane fitted to the 3D scan

As shown in figure 4.6, points from the original point cloud scan typically lie above and below the fitted plane, corresponding to depth being over and under reported by the Kinect respectively.

Points from the original point cloud were projected back to the plane, forming a cloud of points which lie on the fitted plane. This point cloud was imported back into the Kinect collection software, and converted into a 2.5D uvZ depth image. Conversion to a depth image using the Microsoft Kinect SDK (Microsoft Corporation, Redmond, USA) and associated coordinate transformation functions ensured the manufacturer's calibration parameters corresponding to a particular device being calibrated were used to perform the coordinate system transformation.

113

The newly created depth image and the original depth were interpolated to remove any 'holes' in the depth information, ensuring each pixel could subsequently have an associated depth error. The depth images were interpolated using a simple nearest neighbour's interpolation algorithm. For each pixel without valid depth data, the interpolated depth was considered to be the mean of the nearest pixels depth in the positive and negative u and v directions.

After interpolation, the original depth image of the scanned plane (leftmost figure 4.7) was subtracted from the depth image of the fitted plane (rightmost figure 4.7), creating the distortion map (lowermost, figure 4.7). It is important to note the diagonal shading in the rightmost of figure 4.7 is due to way MATLAB has rendered the figure, caused by the fitted plane not being completely parallel to the front face of the Kinect.
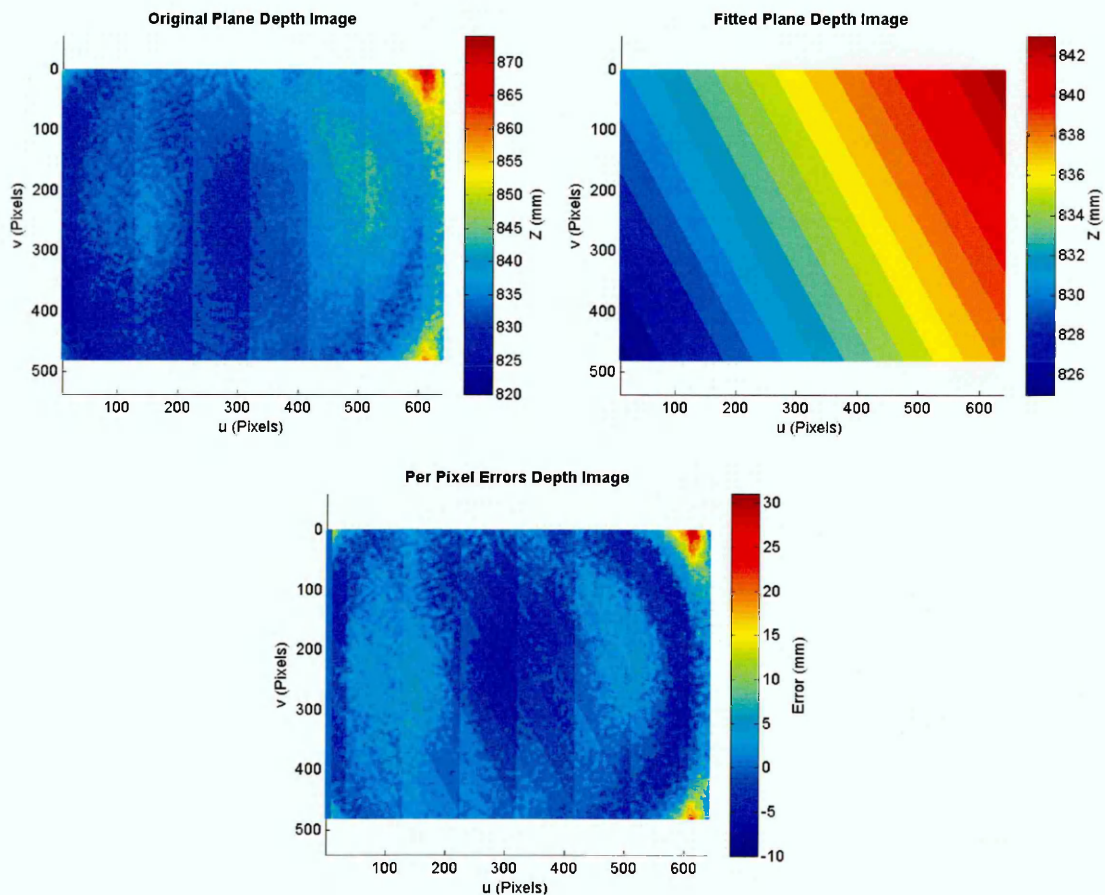
Figure 4.7 – Distortion correction image subtraction process

The error deviation plot (figure 4.7) shows the depth error for each of the Kinect's pixels corresponding to that particular calibration scan. Negative values indicate areas

114

where the Kinect underestimated depth, and positive values indicate areas where the Kinect overestimated depth.

This process was repeated for each of the calibration scans collected over the calibration region. This resulted in n matrices of the form shown in equation 4.1, where u and v are the pixel coordinates of the depth image, Kd is the depth reported by the Kinect for a given uv pixel, *Offs* is the calculated depth error for a given uv pixel at the depth in Kd, and n is the number of captured calibration scans. Each matrix has the size of 307200 x 4, corresponding to the four variables and the 640 x 480 pixels in the depth image for which depth correction information must be stored.

$$\begin{bmatrix} u \\ v \\ Kd \\ Offs \end{bmatrix}$$

Equation 4.1

4.5.2 Error model development

After characterisation of the depth error over the operating range, a depth correction model was developed. The matrices described in equation 4.1 were split into a matrix for each pixel in the Kinect's depth image of the form shown in equation 4.2. This resulted in p $(n$ x $2)$ arrays, where p is the number of pixels in the Kinect's depth image (307200), and n is the number of collected calibration scans.

$$\begin{bmatrix} Kd \\ Offs \end{bmatrix}$$

Equation 4.2

The matrices for each pixel in the depth image - as shown in equation 4.2 - contain depths at which calibration data was collected (Kd), and the depth offset corresponding to that particular depth (*Offs*).

It is important to note the values of Kd corresponding to a particular pixel relate to the actual depth reported by the Kinect for a given calibration scan. They are not obtained from ground truth measurements or average plane depths; the purpose of the defined calibration range and spacing is to ensure calibration scans are collected at regular intervals over the operating region. Therefore, it is highly likely that two successive

115

reported depths for a given pixel will not be exactly spaced by 10 mm due to non-planarity between the Kinect and plane: negating differences in depth error and random noise. Similarly, it is likely that two neighbouring pixels will not have the same reported depth in the calibration data.

A typical error array for a single pixel - as shown in equation 4.2 - is plotted below in figure 4.8, showing the relationship between Kinect reported depth and depth error.

Pixel u = 300, v = 200
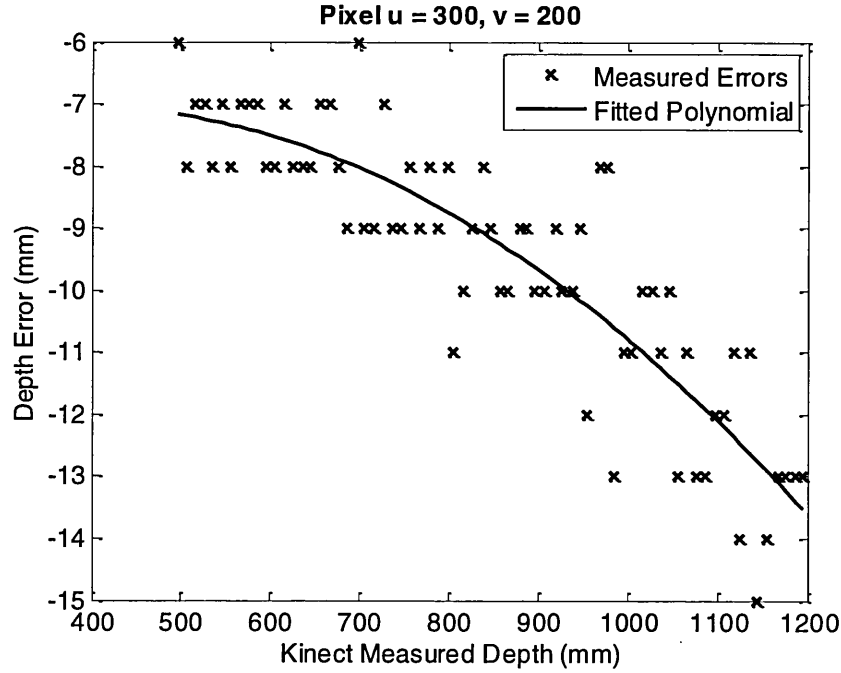
Figure 4.8 – Relationship between Kinect measured depth and depth error

To enable correction of the Kinect's depth error, per-pixel models of Kinect depth and depth error were developed. To enable their development, a number of different mathematical models were considered. Many existing techniques use splines to model the complex distortions apparent in the Kinect's data, however, this technique differs in that each individual pixel has an associated error model: rather than a single model being shared across many pixels. This increases the amount of noise likely present in the training data, possibly leading to errors in the developed model. For this reason it was chosen not to consider splines to model the complex distortions, and instead to solely concentrate on polynomial models. The use of polynomials introduces a high level of constraint (greater than a cubic spline's smoothing parameter for example), restricting the developed model to take a relatively constrained shape and hence

reducing the influence of noise in the training data. Visual inspection of figure 4.8 and other equivalent plots suggested the relationship consistently approximated a second order polynomial. However, this was investigated further.

Table 4.1 summarises the polynomial models which were considered to model the relationship between Kinect depth and depth error. These models were evaluated on the distortion data relating to two Kinects. A mean error and standard deviation (shown in brackets) is presented for each model and each Kinect, representing the error in the models fit, and hence the magnitude of the model-error that would be introduced into the distortion correction process.

Table 4.1 – RMS error values for the models of Kinect depth and depth error

Mathematical Model	Kinect 1- Mean Error and Standard Deviation (mm)	Kinect 2- Mean Error and Standard Deviation (mm)
1st Order Polynomial	1.38 (0.33)	1.43 (0.34)
2nd Order Polynomial	1.36 (0.32)	1.35 (0.31)
3rd Order Polynomial	1.36 (0.32)	1.34 (0.31)
4th Order Polynomial	1.35 (0.31)	1.33 (0.30)

Considering the results presented in table 4.1, it appears that the 2nd, 3rd, and 4th order polynomials performed better than the 1st order polynomial. Given the increased flexibility with greater polynomial order, and hence their increased ability to accurately model the underlying data, this is an expected characteristic. The 2nd, 3rd, and 4th polynomials all performed similarly, with the 4th order polynomial performing slightly better than the 2nd order polynomial: again likely attributable to the increased flexibility offered by the 4th order polynomial, and hence the likelihood of any noise being well approximated in addition to the core data. Given the very small differences in the 2nd, 3rd, and 4th order polynomials, it was decided to use a 2nd order polynomial to model the relationship between the Kinect depth and depth error. The higher order polynomials (3rd and 4th order) did not offer a noticeable increase in goodness of fit, but did result in a significantly longer computation time, and an increased possibility of the developed models being affected by noise in the underlying data.

A second order polynomial was therefore fitted to each of the arrays shown in equation 4.2, providing a relationship between Kinect reported depth and associated depth error for each pixel. This creates a correction model which treats pixel, depth error and reported depth in one model, whilst allowing sufficient freedom. A typical fitted 2nd order polynomial for a single pixel is shown in figure 4.8. Formation of the polynomial error models allowed a calibration file to be produced for each Kinect. Importantly, each calibration file is unique to a particular Kinect and is identified by the manufacturers serial number found on the underside of the device.

The calibration files take the form of a CSV file with a size of 307200 x 5. Each file contains the information shown in equation 4.3, where *abc* are the coefficients of the 2nd order polynomial and *uv* are the pixel coordinates.

$$\begin{bmatrix} a \\ b \\ c \\ u \\ v \end{bmatrix}$$
Equation 4.3

The calibration file is used within the data collection software (discussed previously) at the point of data capture. For a given 3D scan the Kinect initially returns a 2.5D uvZ depth image, with each unique pair of uv coordinates corresponding to a unique line in the calibration file. Using the scanned depth reported by the Kinect for a given pixel, the required depth offset can be calculated using equation 4.4, where *abc* are the polynomial coefficients, *Kd* is the depth reported by the Kinect, and *Offs* is the required depth offset.

$$Offs = a(Kd)^2 + b(Kd) + c$$
Equation 4.4

After calculation of the offset, it can be easily applied to the Kinect reported depth using equation 4.5, where *Kd* is the depth reported by the Kinect, *Offs* is the required depth offset, and *KdCor* is the corrected Kinect depth.

$$KdCor = Kd - Offs$$
Equation 4.5

118

This process is repeated for each of the 307200 pixels in the Kinect's depth image, creating a corrected 2.5D uvZ depth image. The Kinect's integral coordinate transformation algorithms are then used to convert the corrected depth image into a corrected 3D point cloud scan.

Belhedi et al. (2012) demonstrated that depth distortions could be corrected in 2.5D (uvZ) or 3D (XYZ) space, providing a known transformation exists between the two coordinate systems. As it may seem easier and more intuitive to correct depth distortions in 3D space, it is important to provide the rationale for correcting depth distortions in 2.5D space, as discussed above. The 2.5D uvZ image produced by the Kinect prior to calculating 2D XYZ coordinates originates from a 'virtual camera', produced using uv data from the IR camera and interpretation of the projected speckle pattern to calculate Z. Previous investigations have shown the IR camera to contain very little radial distortion, meaning the uv coordinates it returns are correct, it is simply the calculated Z component which is incorrect, containing the discussed depth error. Therefore, it is only the Z component which should be corrected as part of the distortion correction process.

When converting between 2.5D and 3D spaces using the Kinect's integral transformation functions, the Kinect uses the Z measurement in conjunction with the uv coordinates to calculate the corresponding XYZ coordinates. Therefore, simply correcting the Z coordinate in 3D XYZ space would have the effect of shifting the underlying uv coordinates, which is clearly incorrect. It may appear possible to correct Z in 3D space, and apply a corresponding correction to XY, as proposed by Belhedi et al. (2012). However, the Kinect's transformation between 2.5D and 3D space is not revealed to the user, making this process impractical.

Therefore when working with the Kinect, it is important to correct for depth distortion in 2.5D space by only applying the necessary correction to Z, leaving uv unchanged. The Kinect's integral transformation functions and factory calibration parameters can then be used to convert the corrected 2.5D coordinates to 3D coordinates.

After development of the calibration technique, a typical sensor was calibrated and tested to verify performance. An additional set of six plane scans were collected at random depths across the calibrated region not included in the original training data set. At the point of data capture, both distorted and undistorted scans were collected.

PCA was used to fit a plane to each of the scans. The projection of each point to the plane, i.e. the planar deviation, was analysed to determine the depth error of each pixel. The sum squared error (SSE) of these residuals provided a good measure of planar deviation, allowing the effectiveness of the distortion correction technique to be quantified. Table 4.2 shows the SSEs of the planes fitted to the six distorted and undistorted scans.

Table 4.2 – SSEs of planes fitted to the distorted and undistorted plane scans

Plane distance (m)	SSE Distorted (m)	SSE Undistorted (m)
0.637	1.01	0.12
0.678	1.10	0.13
0.739	1.27	0.16
0.778	1.33	0.16
0.818	1.35	0.21
0.848	1.90	0.24

The smaller SSEs post distortion correction (table 4.2) indicate small residuals in the plane fitting, and hence a better approximation of the original plane. Importantly, it is evident that the SSEs of both undistorted and distorted scans increase slightly with depth. This is possibly due to the Kinect's depth quantisation 'bins' increasing in size with distance, rather than the depth distortion increasing with distance. The increase in the size of the depth quantisation 'bins' means there is a greater distance between the quantised depths to which a point can be allocated, hence increasing the possibility and magnitude of noise in the scan data (Boehm 2011).

Simply evaluating the performance of the distortion correction by examining the SSEs in plane fitting is not sufficient, as the distortion correction model may perform better in

some areas of the scan than others. Of particular interest are the areas of high distortion visible in figure 4.3 and figure 4.4. Therefore, it is important to visually interpret and compare the distorted and undistorted scan data. Figure 4.9 and figure 4.10 show the distorted and undistorted scans for the 608 mm plane distance respectively. It is worth noting that figure 4.9 and figure 4.10 are displayed on different scales to allow better interpretation of the data.
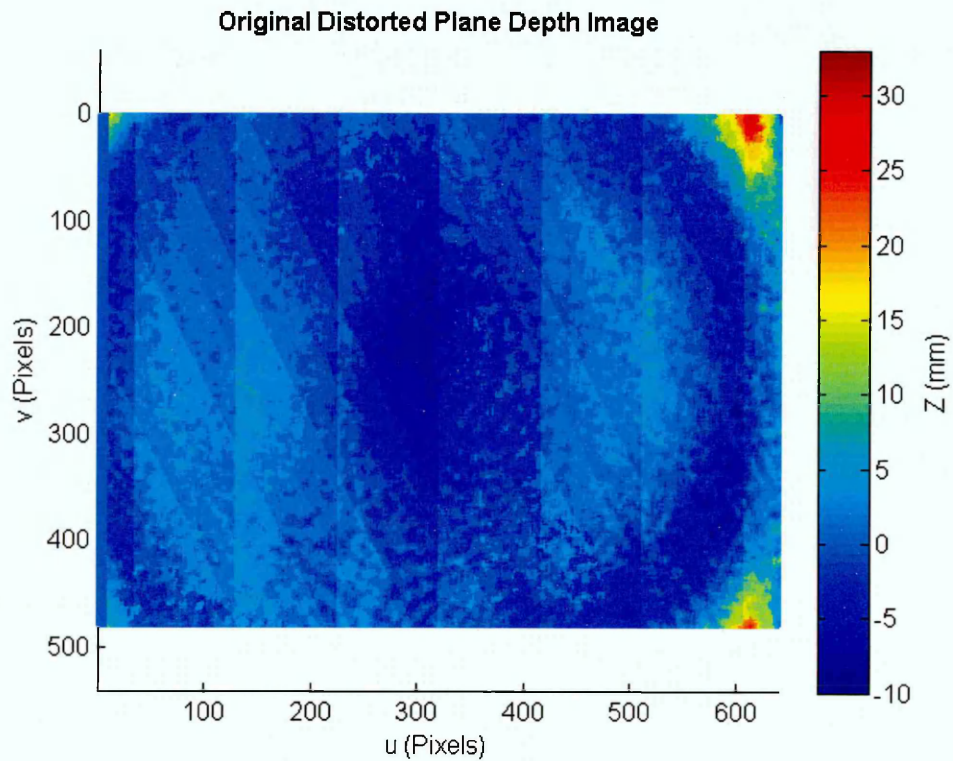
Figure 4.9 – Original distorted plane scan

Undistorted Plane Depth Image
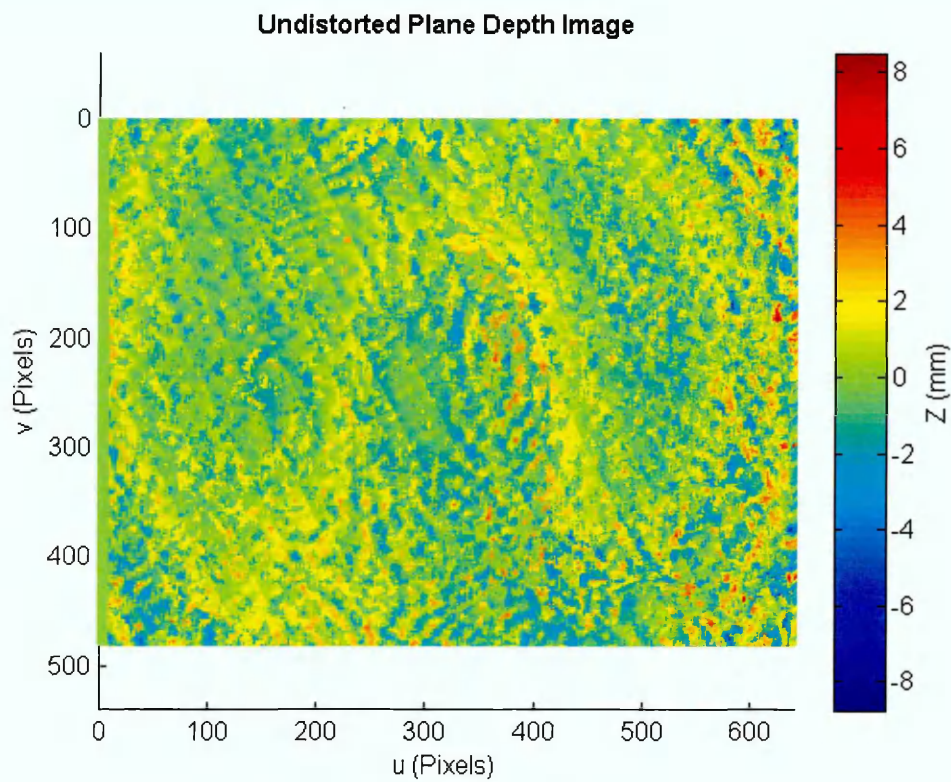
Figure 4.10 – Undistorted equivalent of figure 4.9

The significant reduction in depth distortion is immediately apparent in figure 4.10 when compared to figure 4.9. Importantly, the three areas of significant distortion visible in the distorted scan (figure 4.9) are no longer apparent in the undistorted scan (figure 4.10), suggesting the distortion correction works well across the entire FOV. The majority of residual depth errors are zero, representing planarity of reported depth. Importantly, other errors are randomly distributed throughout the plane, having no clear systematic shape. There are a few pixels having errors of -5 mm, however these are likely related to depth binning errors. Importantly, spurious points such as these could be easily removed with simple smoothing or noise reduction techniques.

4.6 Chapter summary

A simple and effective method of correcting the depth distortion found in many consumer RGB-D sensors is presented here. Unlike the majority of existing work, this method relies only on raw 2.5D and 3D data, permitting compatibility with a wide range of RGB-D sensors in addition to the Microsoft Kinect.

A correction method which makes limited assumptions about the underlying characteristics of the depth errors and considers uvZ in one model was developed, cited in the literature to provide the best overall performance. Results showed the method to significantly reduce depth distortion across a range of depths, leaving random errors of only a few millimetres, and eliminating the previously apparent areas of high distortion. Spurious errors apparent in the corrected scans and the nature of SSEs increasing with depth are possibly a factor of the depth quantisation 'bins' used in low cost RGB-D sensors. Unfortunately this is a limitation of such devices, and little work can be done to avoid this problem.

The pixel-specific correction model assumes the depth distortion apparent in each pixel approximates a second order polynomial. This adds an inherent level of constraint to the model, but still requires a high number of plane scans to sufficiently constrain the polynomial fit. However, this means that calibrating multiple sensors over a large distance would likely be a time consuming process. A possible improvement would be to calibrate multiple sensors at once, but this would require a larger calibration plane. Other improvements lay around automating the calibration process, leaving the plane stationary and moving the sensor(s) away from the plane with use of a linear stepper motor. It is important to note that whilst the Kinect's relationship between per pixel depth and depth error approximates a second order polynomial, other RGB-D sensors may portray different relationships. Therefore, if this method is adapted to suit other sensors, such relationships should be initially visually inspected before deciding upon the correct function to apply.

A limitation of the method presented here is the lack of constraint between neighbouring pixels, meaning the distortion data corresponding to each pixel is treated independently. This can lead to spurious points in the corrected data, or two neighbouring pixels having distinctly different distortion correction models due to noise in the original scan data. Whilst current tests have found no problem with this approach, it is an area of potential further work.

Importantly, a limitation of this study is that the calibration method relies on scans of a planar object, and the performance of the method was evaluated using plane scans: although at depths not part of the original training set. This may lead to different

performance when scanning objects of more complex geometry, or surfaces with greater curvature. Therefore, further testing will take place to verify comparable performance when scanning such objects.

5 METHODOLOGY FOR POST PROCESSING AND PARAMETER ESTIMATION

5.1 Introduction

Once the 3D scans have been obtained (discussed in more detail in chapter 3), further data processing steps are necessary to obtain the parameters of interest. These include the identification and digitisation of body segments of interest, segmentation of the 3D point clouds, and calculation of the required parameters. The data processing steps must be easy and timely, producing results to users of the system without the need for lengthy post processing of the captured data.

Initial investigations with the Microsoft Kinect suggested the 3D data provided by the system are prone to a variety of random errors arising from a number of potential sources. Such errors were investigated further to identify their influence upon measurements from the system. Correction techniques were then developed to limit the influence of the identified error sources.

Development of the post processing and parameter estimation techniques are discussed below, with an evaluation of previous work in this area.

5.2 Aims and objectives

Aims

- To develop methods of improving the quality of raw data from the Microsoft Kinect.
- To develop algorithms and methods for obtaining a range of numeric properties (including circumference and volume) from the 3D scans in an easy and time effective manner.

Objectives

- Scan a variety of objects with the Microsoft Kinect, each having a range of geometric properties. Visually inspect the 3D scans to identify typical errors in the 3D data from the Microsoft Kinect.

125

- Where appropriate, develop suitable data post processing techniques.

- Develop a quick and simple method of digitising and segmenting a 3D point cloud scan.

- Investigate existing methods of obtaining circumference and volume measurements from 3D point cloud scans. Ideally these methods should be iterative, allowing continuous assessment of reliability and accuracy as complexity of the metric is increased.

- Subsequently develop a data processing algorithm capable of obtaining such measurements in a timely manner.

5.3 Point cloud post processing

5.3.1 Introduction

In order to obtain a better understanding of the 3D scan data provided by the Microsoft Kinect, a variety of objects were scanned with the system:

- An aluminium cylinder covered with a non-reflective coating, measuring 373 mm in diameter and 113 mm in length

- A plastic torso manikin, measuring approximately 0.9 m in height

- A cardboard box, measuring 0.4 m x 0.2 m x 0.2 m

The system was setup and calibrated in exactly the same way as discussed in section 3.4, representative of how the system would be used in reality. Scans were collected of the objects in various positions and orientations throughout the calibrated 0.4 m x 0.4 m footprint.

Assessment of the 3D scans was limited to qualitative visual inspection only, as chapter 6 provides an in depth quantitative investigation of the scanning system's accuracy and reliability. A typical scan is shown in figure 5.1.
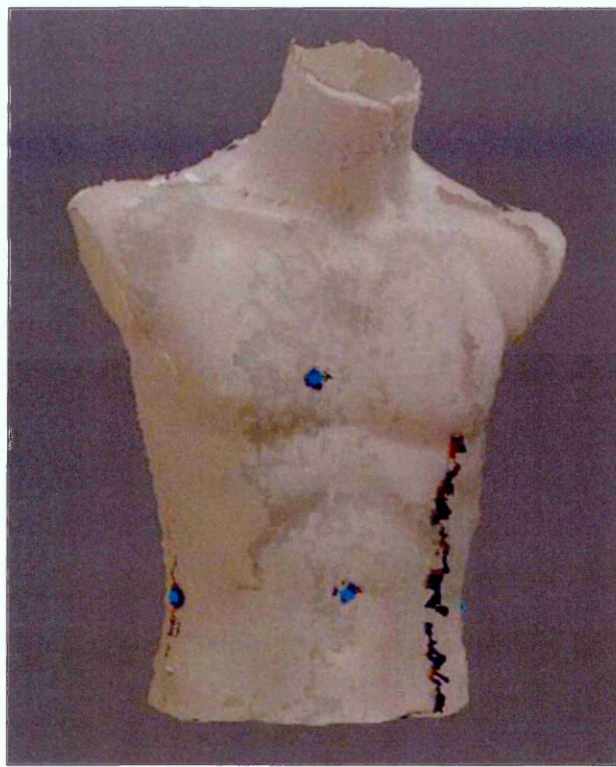
Figure 5.1 – Typical raw 3D scan provided by the scanning system

A number of undesirable artefacts were apparent in all of the 3D scans, which can be grouped under the following three headings.

5.3.2 3D Point fringes

Apparent in all of the 3D scans were continued fringes of points in areas of high curvature, sharp angles, or high angles of incidence between the object and one of the four Kinect sensors. These artefacts were highly visible in the 3D scans, appearing as 3D points representing a continuation of a surface when it does not exist in the underlying object. Although not immediately apparent from figure 5.1, the dark coloured band down the front of the torso manikin is such an area. It is most visible when examining the scans of the cylinder. Figure 5.2 shows a typical scan of the cylinder whilst figure 5.3 shows a simulated cross section, visualising such artefacts.

Figure 5.2 – Typical 3D scan of the aluminium cylinder
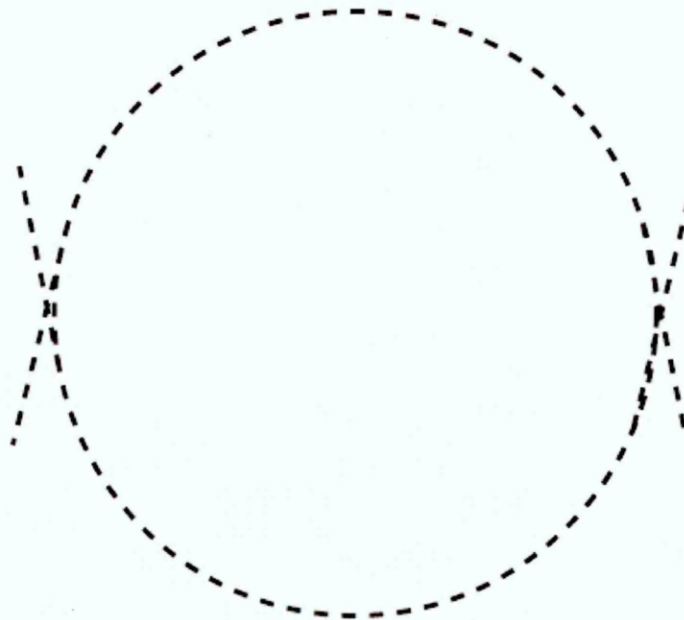
Figure 5.3 – Representation of a cross section of the cylinder scan

Visual inspection of figure 5.2 shows a light coloured band on the right hand side of the scan. This is the same as the dark coloured band on the front of figure 5.1, adding points representing a continuation of the surface which does not exist in reality. Fringes

128

can be easily identified in the 3D scans as they are mostly coloured differently to that of the underlying object. The discolouring is most probably an artefact of the Kinect's integral coordinate transformations: the Kinect identifies a given point as being close to a surface when in fact it is located a long way away, typically being a surrounding object, wall, or surface.

Figure 5.3 better demonstrates the problem, visualising typical point fringes in a scan of a cylinder. Initial inspection may suggest such areas are an artefact of misalignment between the Kinect sensors. However, figure 5.2 shows the rest of the scans to be very well aligned, suggesting this is not the case. Similarly, the fringes are approximately consistent in their shape and size throughout the calibrated footprint. Were they attributed to misalignment of the Kinect sensors they would be expected to change considerably in shape, size, and position throughout the calibrated footprint.

Research showed this was a commonly accepted problem with the Kinect's depth data (Khoshelham & Elberink 2012; Chow et al. 2012), and indeed with other 3D vision systems such as TOF (Soudarissanane et al. 2011) and high end laser scanners (Hancock et al. 1997; Huber et al. 2010). Chow et al. (2012) reported the Kinect to suffer problems resolving depth at high angles of incidence, resulting in increased non-planarity errors when scanning planes that are not parallel to the Kinect sensor. Chow et al. (2012) suggests this is caused by the dots within the Kinect's projected speckle pattern becoming overly elongated at high angles of incidence and hence providing inaccurate depth measurements. However, they provide no suggestion for methods of correcting such errors. Khoshelham and Elberink (2012) reported similar problems at high angles of incidence, however, they too provide no suggestion of an appropriate correction technique. Soudarissanane et al. (2011) reported similar problems in TOF sensors and demonstrate a suitable correction technique. However, this technique is specific to TOF sensors, relying on the fundamental properties of light wave propagation.

The findings of this brief study agree with Chow et al. (2012) and Khoshelham and Elberink (2012), showing the Kinect's projected speckle pattern to become distorted in areas of high object-projector incidence angle. This leads to depth measurement errors, appearing in the form of 3D point fringes. It is evident such areas need to be removed

from the returned 3D scans before proceeding to take measurements from the scans, as the additional points are likely to have impact upon measurements such as circumference and volume.

Whilst complex surfaces with fluctuating surface detail are likely to present high angles of incidence throughout their surface, within the application of body scanning this is unlikely to be the case. As identified above, the fringes are likely to occur at the edges of possible vision for each of the four Kinect sensors. There is potential for point fringes to occur around the spinal area for example (having a sudden and high level of surface curvature), however this is expected to be very limited, and could be easily dealt with via other data processing steps. The cross sectional slice of a typical torso manikin scan shown in figure 5.4 demonstrates this case, showing significant fringes in areas where the four scans overlap, and very limited fringing around the spinal area. Therefore, a fringe removal filter concentrating on removing points with high angular incidence about the edges of the scan was developed.
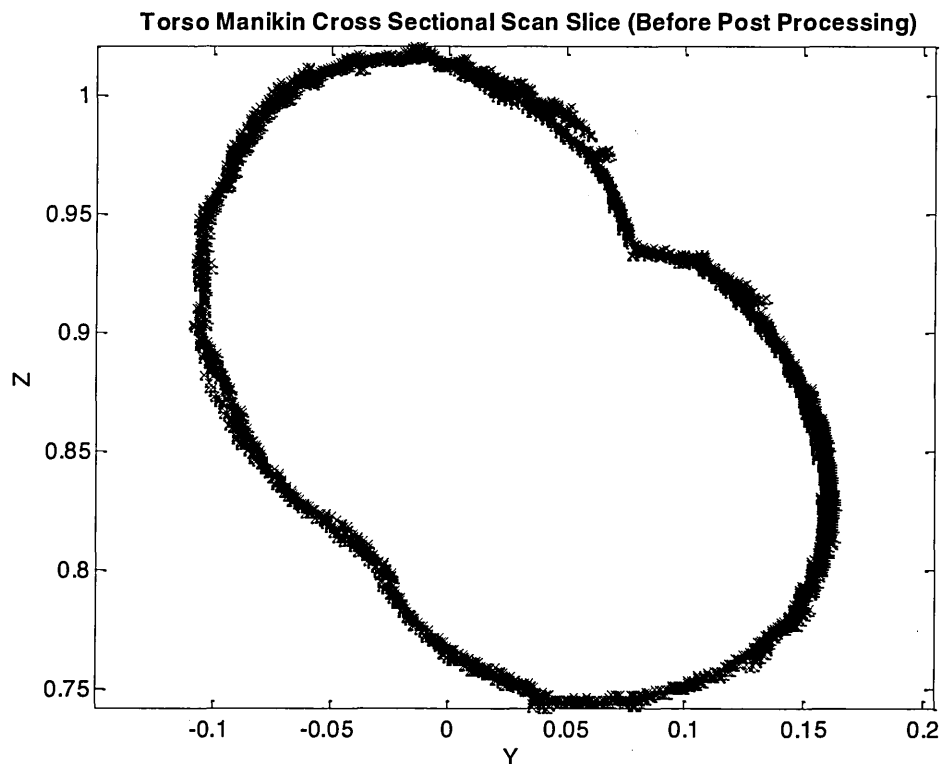
Figure 5.4 – Cross sectional point 'slice' of the torso manikin scan

It was decided to develop the correction technique based upon the Kinect's 2.5D coordinate system, with the fringes being removed in 2.5D space before converting the remaining points to 3D via the Kinect's integral coordinate mapping function.

Each Kinect in the scanning system initially returns a 640 x 480 depth image, colour coded based upon distance, with each pixel therefore having an associated real world depth measurement (2.5D). The returned data are rastered, making examining the data much easier. Figure 5.5 shows a simplified version of how the depth image appears when scanning an object such as a cylinder. Black coloured points represent points on the cylinder, dark grey points represent potential areas of fringes, whereas light grey points represent those located outside the area of interest, representing the surrounding objects and surfaces. The Kinect was assumed to be oriented vertically, as per the system design discussed in chapter 3.
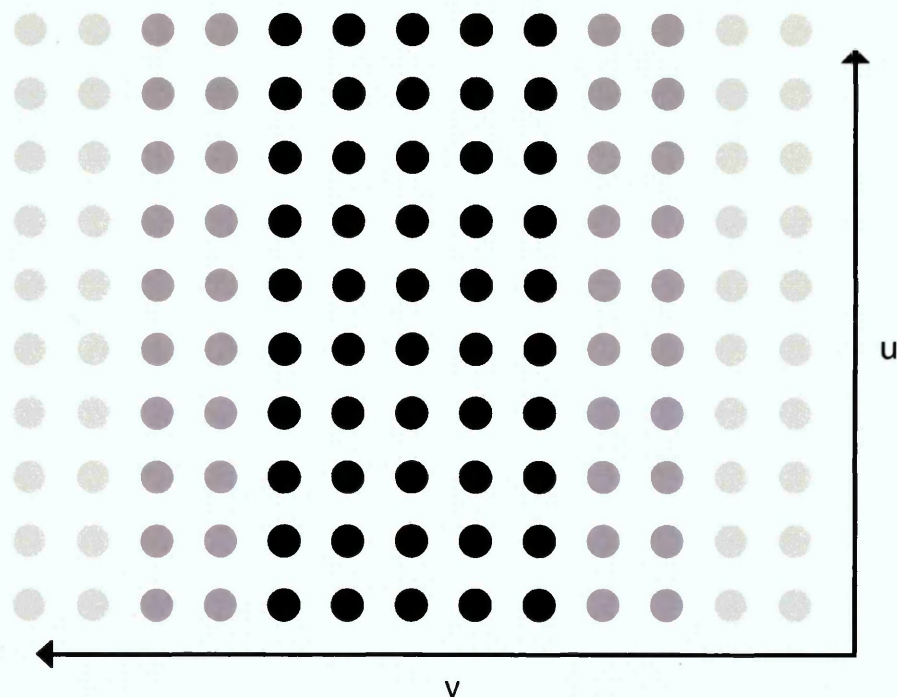
Figure 5.5 – A simplified representation of the Kinect's rastered depth image

To remove the points forming the fringes, the points were searched using a fourteen pixel window. For a given pixel of interest, the window searched seven pixels in the positive v direction and seven pixels in the negative v direction whilst maintaining a consistent u coordinate. The size of the search window was determined as part of an iterative development process, visually inspecting scans using different window sizes

131

and forming a compromise between adequately removing the fringes and not removing too much data.

All 480 v coordinates in the depth image were searched, cycling through all 640 possible corresponding u values. When near the edges of the depth image and it is not possible to achieve a fourteen pixel search window, the window was shrunk to only search seven pixels in a single direction depending whether it is at the top or bottom of the image.

Points lying in the window were examined to determine if they were outside of the depth range of interest, if they were, they were added to a counter. If more than 30% of the points in the window were outside of the depth range then the point the search window is centred upon is likely at an edge, rather than being a spurious hole in the depth data. If this is the case the point is likely a fringe and should be removed: the depth of that point was therefore set to -1, representing no depth. This process was repeated for each point in the u direction, importantly comparing points the original depth image rather than the iteratively refined image. This process is demonstrated in figure 5.6, showing the acceptance decision of a variety of points around the edge of a scanned object. Black points represent those accepted by the filter, grey points represent those rejected by the filter, whilst gaps (white) represent holes in the underlying 3D scan.
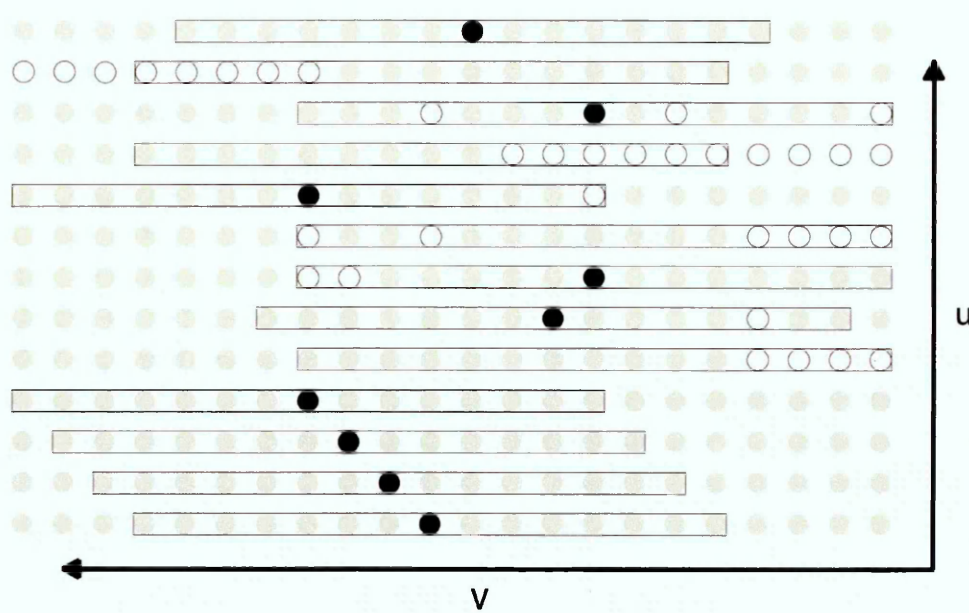
Figure 5.6 – Examples of the fringe removal filter search window

The 30% acceptance threshold was also determined via an iterative development process, also forming a balance between sufficiently removing the fringe and not significantly degrading the amount of data. This effectively removes ten pixels at the end of each returned column of scan points (assuming a vertical Kinect orientation) and accounts for broken or intermittent fringes, as only points representing a continuous crisp edge will be left in the refined depth image. Continued development and testing showed this was sufficient to remove typical fringes in the Kinect's returned 3D data.

The depth image corresponding to each Kinect was processed in this way before conversion to 3D coordinates using the Kinect's integral coordinate transformation functions. As discussed in chapter 3, the global transformation matrices were then applied to the individual 3D scans in order to produce a single 3D scan.

Figure 5.7 and figure 5.8 show scans of a cylinder and trunk manikin respectively, with the original 3D scans on the left and the same scans after processing through the fringe removal filter on the right. Similarly, figure 5.9 shows a topological point cloud slice of the cylinder scan before and after application of the fringe removal filter.

133
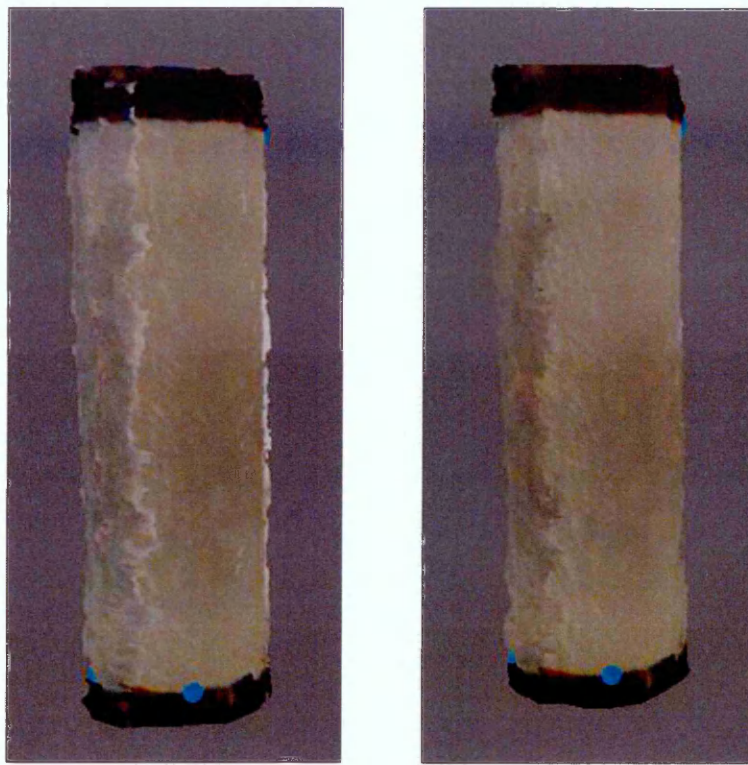
Figure 5.7 – Cylinder scan before and after application of the fringe removal filter
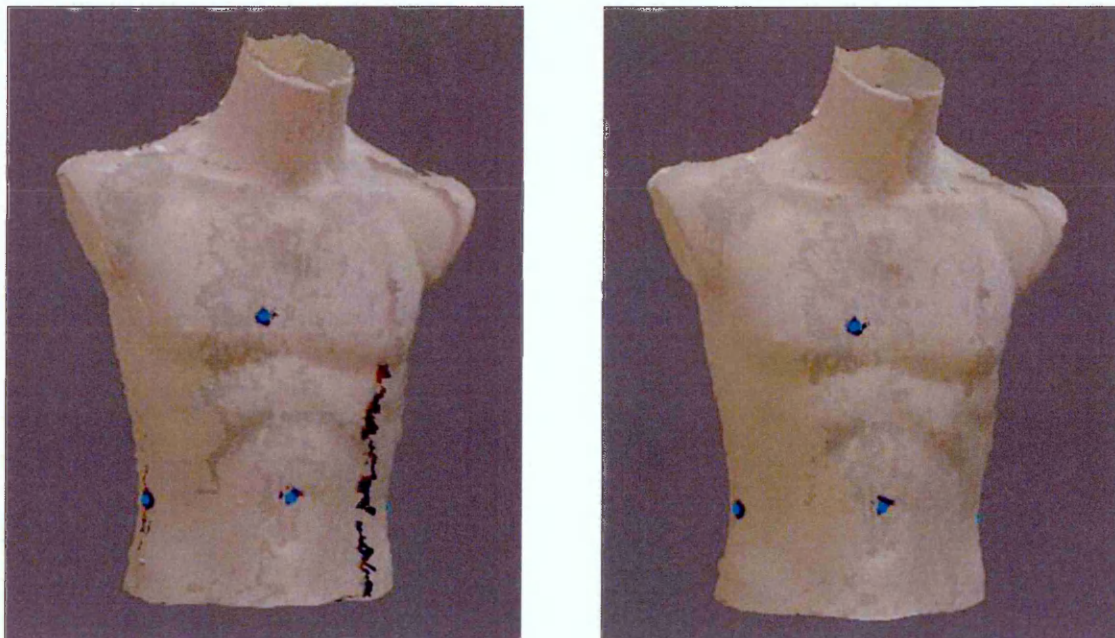
Figure 5.8 – Torso scan before and after application of the fringe removal filter
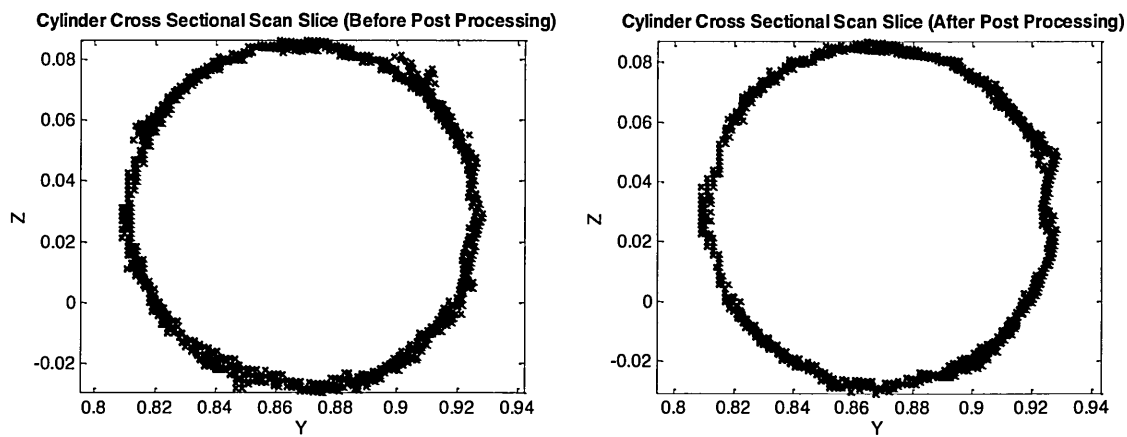
Figure 5.9 – Point cloud slice of the cylinder before and after application of the fringe removal filter

Visual inspection of figure 5.7, figure 5.8, and figure 5.9 demonstrate the effectiveness of the fringe removal filter, removing the light coloured bands in the cylinder scan and the dark coloured bands in the manikin scan. The topological point cloud slice shown in figure 5.9 demonstrates this further, showing a lack of fringing at the edges of the scans: producing cleanly overlapping areas.

The effectiveness of the fringe removal filter was assessed as part of the study discussed within chapter 6, determining the accuracy and reliability of the scanning system when scanning objects representative of typical body segments.

5.3.3 Pseudo random noise

Inspection of the Kinect's 3D scans show the 3D data to portray a variety of random noise, appearing as holes in the 3D data or general noise in the surface of objects. This is well reported (Khoshelham & Elberink 2012; Menna et al. 2011; Boehm 2012), and an accepted attribute of the Kinect's low cost hardware. A variety of techniques have been proposed for reducing such artefacts, which can be applied to the Kinect's 3D data or depth images (Camplani et al. 2012; Chen et al. 2012; Matyunin et al. 2011).

However, for the purpose of this application it was decided not to apply any form of filtering, noise removal, or smoothing to the raw 3D data provided by the Kinect. Instead it was decided to handle (and correct where necessary) such artefacts at the point of calculating parameters of interest from the 3D scans. This allows techniques

specific to the parameter of interest to be employed, possibly resulting in a better quality final measurement. This is discussed in more detail in section 5.5.1.

5.3.4 Colour bleeding

The 3D colour rendered scans showed the Kinect struggled to resolve surface colour in areas where there is a high contrast between two colours. This was particularly noticeable about the interface between the black and white areas of the cylindrical object scanned above. Figure 5.10 shows a close up view of such an area.
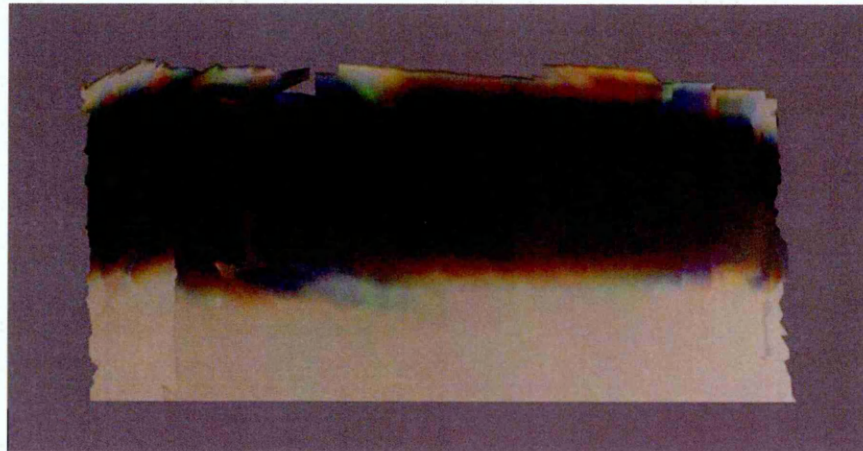
Figure 5.10 – Close up of the interface between the black and white areas of the cylinder

Figure 5.10 shows an area of colour uncertainty about the border between the black and white coloured areas, colouring points in a variety of random colours. Research suggests this problem to be a limitation of the low cost colour camera used in the Kinect (Cambridge in Colour 2013).

However, within the context of human body scanning this is unlikely to be a problem. As discussed in chapter 3, anatomical points of interest will be marked using coloured circular markers (10 mm Ø). The findings of this chapter suggest the colour at the edges of the markers will be distorted, however the centre of the marker will remain intact. This means the markers can still be digitised with ease, allowing reliable digitisation of anatomical points of interest.

5.4 Point cloud segmentation

A fundamental prior step to calculating parameters of interest from the 3D scans includes splitting up the scans to only include specific body segments, based upon the location of key anatomical landmarks. As discussed in chapter 3, anatomical landmarks are marked on the human body using 10 mm diameter coloured markers which are easily identifiable in the colour rendered 3D scans, figure 5.1.

5.4.1 Manual digitisation process

A manual digitisation process was decided upon, rather than automatic marker identification or a 3D feature identification algorithm. Given the intended application for the system it was expected this would lead to greater reliability and accuracy.

Collected 3D scans are manually digitised by clicking directly on the 3D scan displayed in the software discussed in section 3.6. This process is repeated for each point that requires digitising, returning a set of 3D points defining limits of a body segment of interest which can later be used with the data processing algorithms discussed in section 5.5. Figure 5.11 shows a typical digitised human torso segment, with blue circular markers automatically placed on the surface of the 3D model to verify the points have been digitised correctly.

137
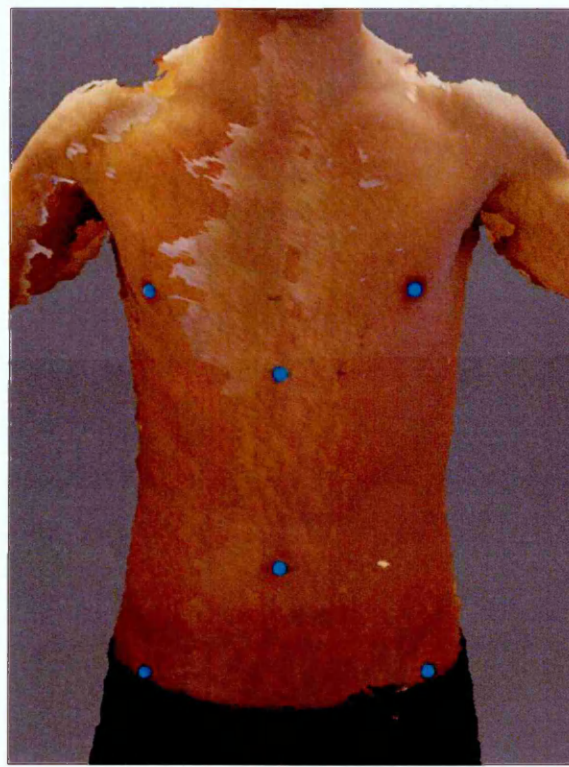
Figure 5.11 – Typical digitised human torso segment

5.4.2 Point cloud clustering

A further step required before calculating the numeric parameters of interest is to separate the 3D scan to only include the segment of interest. For example, when interested in the torso segment this step involves the removal of points on the left and right arms, as well as points laying on the rest of the body.

The principle of identifying and separating clusters of points is not a new principal, and is indeed non-trivial. The k-means clustering algorithm is a very popular technique (ALGLIB 2014) of clustering points based upon their distance to one another, making use of a Euclidian distance metric. Initialisation parameters such as the number of clusters known to exist in the scan are used to separate the data into the appropriate number of groups (ALGLIB 2014). Initial investigations were conducted using the ALGLIB C# implementation of the k-means clustering algorithm (ALGLIB 2014). However, the computation time was found to be very high, and the algorithm overly complex for the relatively trivial clustering problem presented here. Therefore a bespoke clustering algorithm was developed.

The 3D scan was first split up into 2 mm thick cross sectional slices, and the clustering process conducted on each slice. This process and its rationale are discussed in more detail in section 5.5.1.

Each cross sectional slice represents a series of 3D points with the Y component removed to allow plotting on a 2D axis. Importantly, the Y component can be added back in at the end of the computation process to return a 3D point cloud in the same form as the original.

As discussed in chapter 3, the scanning system presented here is designed to focus on a single segment at a time. Therefore it was assumed the segment of interest will always lie approximately at the centre of the scan. A weighted distance function was used to find the centre of each scan slice and subsequently transform all the points so the centre of the slice lied at (0,0), which should also be located at the centre of the segment of interest. The Cartesian coordinates were converted to polar coordinates, with the rotation centre located at the centre of the segment of interest, figure 5.12.

Cross Sectional Slice of the Torso Segment (Polar Coordinates)
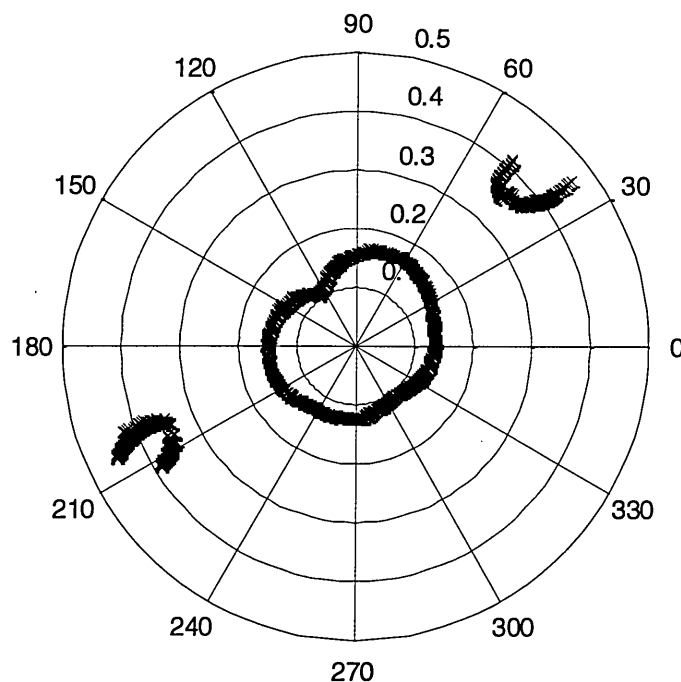
Figure 5.12 – Cross sectional scan slice in polar coordinates

139

The points were progressed through circularly, starting at the equivalent XZ location (converted to polar coordinates) of the digitised landmark. The Euclidian distance between the start point and the next point (progressing in a circular clockwise direction) was calculated, and if the distance was less than a pre-determined threshold the point was accepted as a valid point. This process was repeated around the scan, subsequently using the accepted valid points as new start points. If a point was detected outside the threshold, the start point was kept the same and the search through the list of points continued until a valid point was found. Experimentation showed a threshold of 20 mm provided the necessary level of acceptance.

Cross Sectional Slice of the Clustered Torso Segment
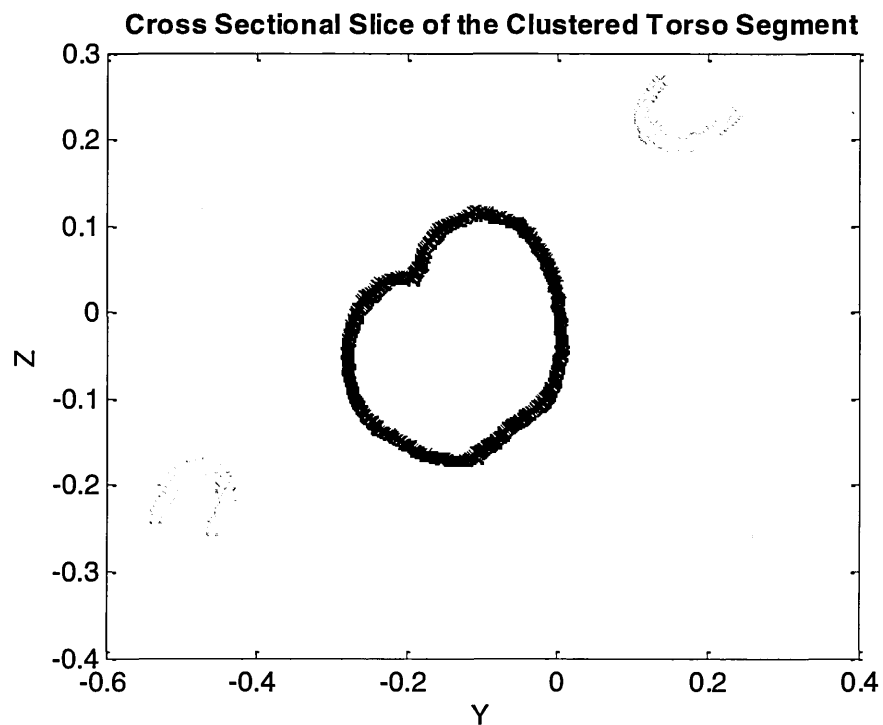
Figure 5.13 – Cross sectional scan slice after the clustering procedure

Once the scan had been completely looped through the accepted points were converted back to 2D Cartesian coordinates, figure 5.13, and then to 3D Cartesian coordinates. For the purpose of interpretation, figure 5.13 shows the accepted points in black as well as the points excluded by the clustering operation in grey. The cluster of accepted points can subsequently be used by the parameter calculation algorithms discussed in section 5.5.

5.5 Point cloud properties calculation

Once the 3D scans have been collected and post processed there are a number of steps necessary to determine the numeric parameters of interest. As discussed in chapter 3, the ultimate goal of the scanning system is to produce estimates of segmental volume, meaning there is a need to develop an algorithm for calculating the volume enclosed by a specific area of a 3D scan. For other analyses it was expected a circumference calculation algorithm would also be useful.

In order to segment a 3D scan to only include the torso segment for example, two segmentation planes are required: one at the inferior and one at the superior segment boundaries. Two techniques were initially considered, the first relied on the user digitising three points at the top and bottom of the segment of interest, creating segmentation planes directly from the three points. The second relied on only two digitised points: one at the inferior and one at the superior segment boundary. The participant was assumed to be stood parallel to the global coordinate system (chapter 3), allowing the segmentation planes to be constructed using the single digitised point, combined with knowledge about the system's global coordinate system.

Investigations showed the method using three digitised points to be very unreliable, as it was difficult to palpate, mark, and digitise three planar points located across the front and back of the body. This led to the angle of the segmentation plane changing with respect to the segment, leading to poor reliability of returned volume measurements. The method using a single digitised point was found to be much more reliable, with the participant's position being controlled procedurally to ensure they were parallel to the global vertical. With this in mind, the single point segmentation method was used from then on.

Figure 5.14 shows a typical torso scan, segmented to contain a particular area using two planes, each formed from a single digitised point.
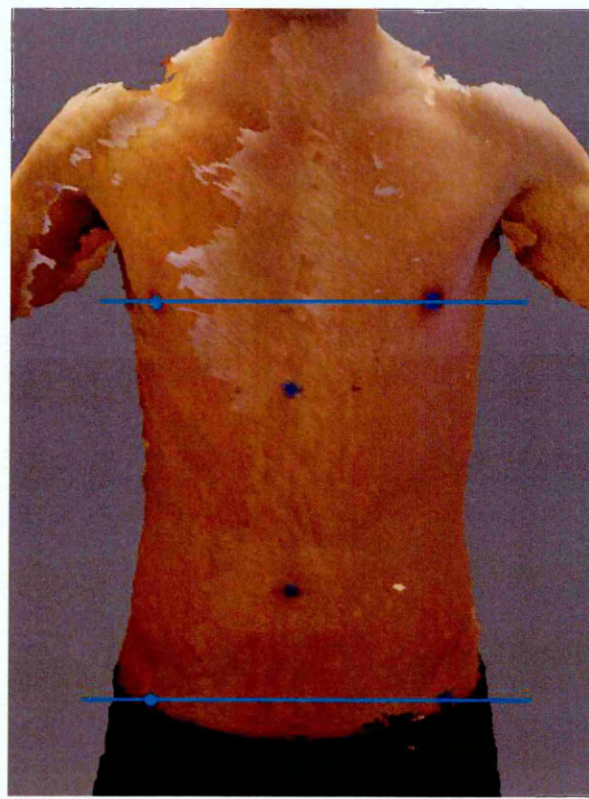
Figure 5.14 – Typical segmented torso scan

5.5.1 Circumference calculation

The segmented area of interest was first split up into a series of point 'slices', 2 mm in height. This allowed the circumference to be calculated throughout the length of a scanned object – such as a cylinder – enabling analysis of measurement error throughout the scanners FOV. The 2 mm slice height was determined by experimentation, small enough to ensure typical body features are captured, and large enough to ensure a slice contains sufficient points to enable calculation of circumference: limited by the Kinect's scanning resolution. Figure 5.15 illustrates the slicing process for a 3D scan of a cylinder, with the digitised upper and lower points shown as solid black circles, the upper and lower segmentation planes shown as darker planes, and the points corresponding to a particular slice located between the lighter coloured planes. For the purpose of illustration, far fewer point slices are shown than are used in reality.
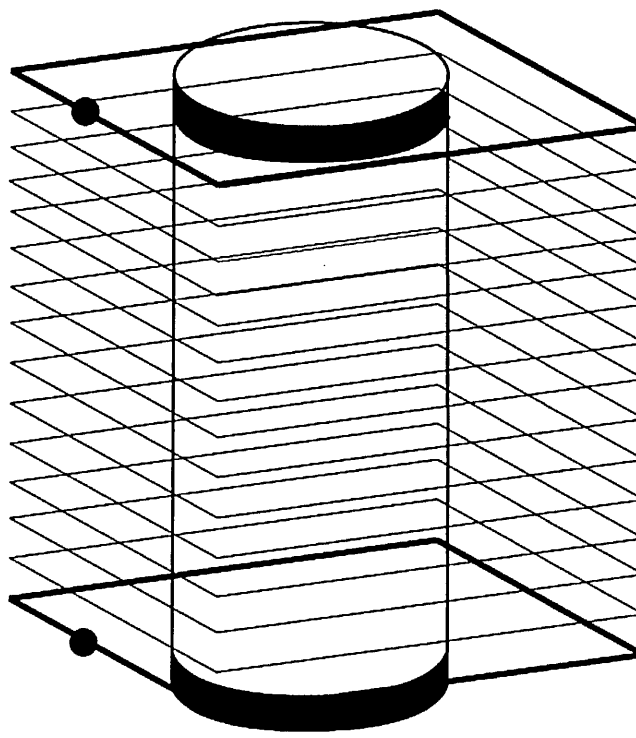
Figure 5.15 – Cylinder 3D scan split up into a series of point 'slices'

The circumference of each slice was then calculated. Points forming the slice were first converted to a 2D coordinate system, disregarding the Y component of the global coordinate system (chapter 3) and assuming all points in the slice lie on a plane. The 2D points were then converted to a polar coordinate system and transformed so there centre lay at the centre of the cylinder. A penalised regression spline (ALGLIB 2014) was then fitted to the transformed 2D points, creating a series of points representing the underlying 3D scan whilst smoothing random variation in the original data, figure 5.16. A penalised regression spline was chosen as it allows independent control over flexibility, nonlinearity and noise suppression, using two initialisation parameters N and ρ. N controls the number of nodes or basis functions in the spline, i.e. the degrees of freedom. N must be large enough to provide sufficient flexibility, and small enough for good performance. However, N can be infinitely large, as a penalty function prevents the spline from being too flexible. ρ is the regularization coefficient for the spline, controlling the linearity and hence the smoothing applied by the spline. Initial investigations showed N = 50 provided sufficient flexibility whilst maintaining an acceptable level of performance. Similarly, ρ = 0 was found to provide sufficient smoothing whilst not overly smoothing anatomical features typically found on a human body segment.

Typical Scan Slice with Fitted Spline
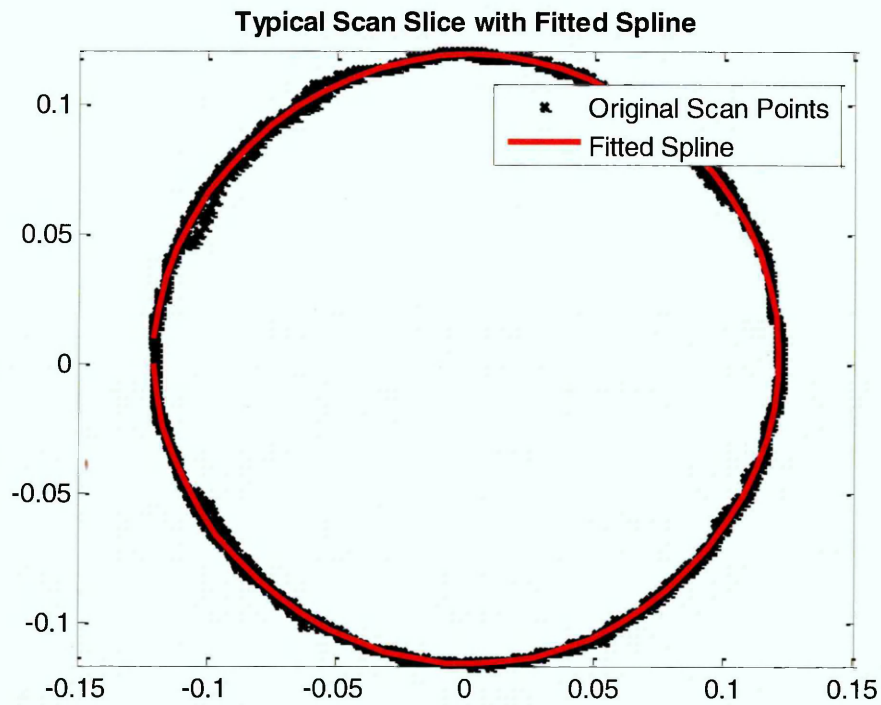
Figure 5.16 – Typical point 'slice' with fitted smoothing spline

Points were created on the spline at 1° intervals around the polar coordinate system and subsequently converted back to a 2D Cartesian coordinate system. The inter-point distance of all the points created on the spline was then calculated. As the points are arranged in a circular order this corresponds to the total distance around the slice, in turn corresponding to the circumference of the slice.

This process was repeated for each slice in the region of interest, creating a series of circumference measurements. If the circumference of a particular slice could not be calculated – for example the slice contained too many holes to enable reliable fitting of the spline – then simple linear interpolation between the last available circumference and the missing circumference was used to calculate the missing value. Importantly. investigations have shown this situation to hardly ever arise.

Another problem which may occur is if the slice height between the top segmentation level and the upper limit of the area of interest is less than 2 mm, and hence insufficient to enable reliable fitting of the spline. In this case it was simply assumed the circumference of the uppermost slice was the same as the penultimate slice. Given the

very small distance between the two, the difference in circumference was assumed to be negligible.

Reliability and accuracy of the circumference calculation algorithm is considered in chapter 6 as part of the initial system validation study.

5.5.2 Volume calculation

A number of techniques have been developed for calculating the volume enclosed by 3D scans (Mirtich 1996; Dellen et al. 2013). For example, Mirtich (1996) proposed a technique based upon volume integrals, requiring an input 3D mesh. However, creating a watertight mesh from an original 3D point cloud is a non-trivial process, being very computationally intensive and time demanding. If incorrectly constrained, the meshing process has the potential to introduce significant errors into underlying 3D models (such as under/over smoothing features) and hence produce unrepresentative volume estimates.

Furthermore, one of the original objectives of the chapter was to create a series of iterative data processing algorithms. When determining accuracy of the algorithms this allows them to be 'stacked', evaluating performance as complexity is increased. The introduction of a distinctly different technique for calculating volume goes against this principal, potentially introducing unknown sources of error.

Crisco and McGovern (1998) proposed a method based upon Green's theorem (Wrede 2010) for calculating inertial properties from cross sectional slices of CT scans. In a similar method to that outlined above in section 5.5.1, the CT scan was initially split into a series of slices. The volume of each slice was calculated according to Green's theorem (Wrede 2010), before summation of all the individual volumes in order to calculate the total enclosed volume. Crisco and McGovern (1998) showed the method to work accurately and quickly, both important objectives of this study.

In line with the original objectives of the chapter, a volume calculation algorithm was developed which is iterative, building upon the slicing method used as part of the circumference algorithm. The method is inspired by that of Crisco and McGovern

(1998), but further simplifies some of the techniques to reduce the computation time as much as possible.

Like the circumference calculation algorithm discussed in section 5.5.1, the volume calculation algorithm also works on a slice by slice basis. With the point cloud slice represented in 2D Cartesian space, the surface area contained by the points was first calculated. Triangles were formed within the region bounded by the spline, with their vertices located at the centre of the slice and two successive points on the fitted spline. This process was repeated throughout the slice, creating a series of triangles covering the surface area. This process is depicted in figure 5.17, however, for the purpose of illustration the size of the triangles is increased significantly. The original points from the 3D scan are shown in white, the red line is the fitted spline, and the white lines represent the fitted triangles.
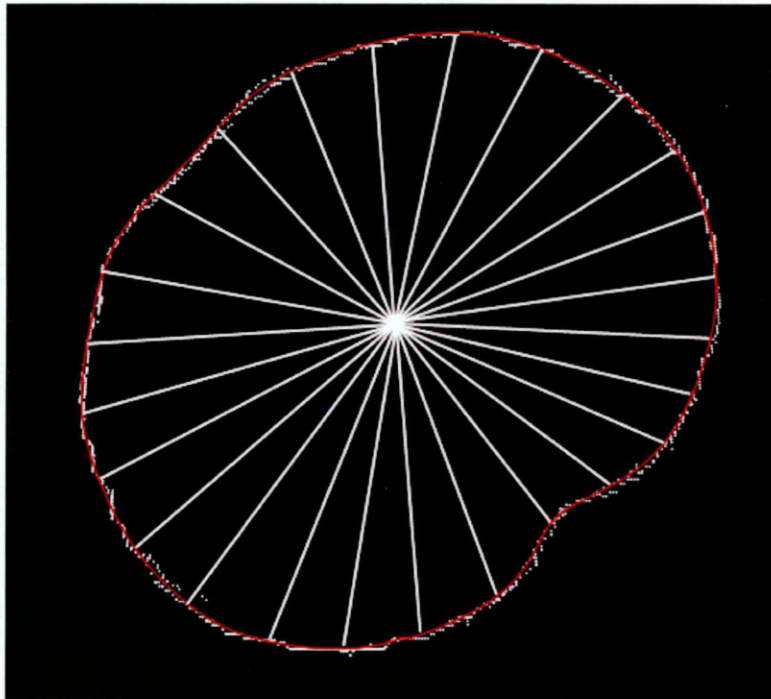
Figure 5.17 – Technique for calculating the volume of a point cloud slice

The area of the triangles was calculated and summed, estimating the surface area of the slice, equation 5.1.

$$Slice\ Surface\ Area = \sum Area\ of\ Individual\ Traingles \qquad \textbf{Equation 5.1}$$

The estimated surface area of each slice was then multiplied by the fixed slice height of 2 mm, returning the estimated volume of the slice, equation 5.2.

$$Slice\ Volume = Slice\ Surface\ Area * Slice\ Height$$ **Equation 5.2**

This process was repeated for each slice in the area of interest, and the volumes summed to estimate the total volume enclosed by the region of interest, equation 5.3.

$$Total\ ROI\ Volume = \sum Slice\ Volumes$$ **Equation 5.3**

Similar to the circumference calculation algorithm, problems may occur if the points forming a slice are not sufficient to form the triangles required to calculate surface area, and hence volume. In a similar manner to inserting missing circumference values, simple linear interpolation between the last available volume estimate and the missing volume estimate was used to determine the missing volume value. Investigations have again shown this situation to hardly ever arise.

Again in a similar manner, a problem may occur if there are insufficient points in the uppermost slice to allow formation of the triangles required to calculate surface area, and hence volume. In this case it was assumed the uppermost slice has the same surface area as the penultimate slice: which should be approximately the same due to their very close proximity. The slice height of the uppermost slice was then multiplied by the estimated surface area to estimate the volume in the uppermost slice.

5.6 Chapter summary

This chapter discusses common issues likely to be encountered when using the Kinect as a 3D scanner. Where these are likely to have an impact upon subsequent calculations, error correction techniques and algorithms have been developed. Finally, techniques and algorithms for calculating numeric parameters of interest from 3D scans have been developed.

The Kinect was found to struggle resolving an object's 3D surface profile in situations where the incidence angle between the object and Kinect is high, introducing areas of point 'fringes'. This is a common problem of 3D scanning systems, discussed heavily in the literature. However, none of the current correction techniques are suitable for use with the Kinect. A simple yet effective algorithm was therefore developed for removing such areas from the raw 3D scans.

Potential problems due to the Kinect's low cost RGB camera were identified, introducing a region of colour uncertainty in areas with a sharp change in colour: typically occurring when digitisation markers are placed on an item being scanned. However, within the final application of human body scanning this is unlikely to be problem. It is expected that body segments will be palpated using small coloured circular markers, visible in the colour rendered 3D scans. Whilst the edge of the markers may be subject to a degree of colour uncertainty the centre will remain reliable, meaning it can be reliably digitised with ease.

Part of the scan segmentation process involves the removal of points relating to other segments. For example, scans of the torso are likely to include points from the arms which need to be removed before calculating numeric parameters of the torso. A quick and effective technique for removing such areas was developed. One potential drawback is that the algorithm makes the assumption the segment of interest is the point cluster located closest to the centre of the scan, removing any clusters or random points laying away from this region. Within the intended application of this scanning system (focussing upon BSPs of a single segment at once) this is not a problem, however, if a single scan of the upper body was collected with the intention of calculating BSPs of the arms and torso simultaneously the current algorithm would not be suitable. This is therefore an area of potential further work.

An algorithm capable of calculating the volume of a 3D scan was also developed. Many algorithms for calculating the volume of 3D scans already exist, however the majority are very complex, time consuming, and introduce significant potential for errors in the numeric outputs. Furthermore, one of the original objectives was to develop a series of data analysis algorithms which are iterative in terms of complexity, meaning one algorithm can be used to infer the expected performance of another

148

algorithm (which may not be transparent to the effects of some influential factors). The size of the triangles used for calculating surface area of the point slices is an important consideration. The current size appears suitable for calculating the volume of typical human body segments, however, if objects are scanned that have a range of small features around their outer surface then this parameter may need adjusting in future. Importantly, the nature of the surface area calculation technique means that objects having an outer surface which loops back upon itself cannot be used with the volume calculation algorithm: in order to form the triangles to calculate surface areas (and hence volume) there must be a clear path between each point and the centre of the scan. Within the application of human body scanning this is unlikely a problem, but is an important factor which should be considered when scanning distinctly different objects.

6   VALIDATION OF THE SCANNING SYSTEM USING UNIFORM OBJECTS

6.1   Introduction

The recent introduction of low cost NUI sensors has led to significant interest in a range of communities including: robotics (Henry et al. 2012), body scanning (Boehm 2012), healthcare (Labelle 2011), and apparel (Stampfli et al. 2012). However, there have been few studies investigating the accuracy and reliability of measurements derived from such devices.

Recent studies typically focus on simple measurements (such as Euclidian distances and plane fitting residuals) from single sensors (Boehm 2011; Menna et al. 2011) whereas a scanning system for measurement of body morphology typically comprises multiple sensors (Boehm 2012; Clarkson et al. 2012; Clarkson et al. 2013), and involve more complex measurements: possibly leading to a compounding of error.

Robinson et al. (2012) states that "assessment of application specific suitability should be as close to reality as physically possible and not rely on simple geometric tests". Furthermore, Robinson et al. (2012) goes onto say that reference measurements should have a measurement uncertainty of at least five times better than the system with which they are being compared. In many cases, this means that one 3D scanning system cannot credibly be used to validate another 3D scanning system.

An initial system validation study was conducted to determine the system's performance under typical operating conditions, whilst limiting external influential factors (such as object movement during scanning) as much as possible. The purpose of the study was threefold: to determine the accuracy and reliability of scan-derived measurements, to determine the system's suitability for calculating person specific BSPs, whilst inherently evaluating performance of the data processing algorithms previously developed.

6.2 Aims and objectives

Aim

- To determine typical accuracy and reliability of measurements obtained from the scanning system.

Objectives

- To limit external influential factors as much as possible in order to determine the system's fundamental operating characteristics, whilst being representative of reality.
- To compare system performance with established industry standards to provide an indication of performance and likely acceptance of the system.
- Perform multiple system calibrations and scans from different positions within the capture volume, considering typical changes in operating characteristics.

6.3 Experimental design

The system was validated using measurements taken from machined aluminium cylindrical objects, chosen to be representative of typical human body segments.

The cylinders enabled assessment of the system's reliability and accuracy without the possibility of factors – such as involuntary movement of the body – to degrade the accuracy and repeatability of the system or mask fundamental measurement errors. Secondly, the dimensions of the cylinders can be accurately obtained using conventional measurement techniques, such as calibrated engineer's digital callipers. This provides accurate gold standard measurements, with which measurements from the system can be compared.

Results of the study provide baseline 'ideal world' operating characteristics for the scanning system with which subsequent measurements from 'non-ideal' objects (such as human body segments) can be compared, essential for further investigations. For example, if reliability is found to decrease when scanning human body segments, then the decrease can be isolated to influential factors from the participant, such as involuntary movement of the body during scanning. Without this initial assessment it

151

would be impossible to determine the source of such a decrease, or subsequently provide any form of quantification.

Subsequent investigations involve the calculation of volume; however, volume alone is not a good metric with which to assess system performance, having potential to mask considerable measurement inaccuracies. Therefore, it must be supplemented with a range of additional measurements. For example, a 3D scan of a cylinder may have a series of bulges and holes which are dimensionally equivalent. In this case the enclosing volume reported by the system would be correct as the bulges and holes cancel each other out. However, the underlying 3D scan would be unrepresentative of reality, having significant implications upon further BSP calculations such as centre of mass and moment of inertia.

Therefore, for the purposes of initial analysis of system performance, it was decided to supplement the volume measurements with circumference measurements taken throughout the length of the cylinders. This negates the problems highlighted above, as measurement variation throughout the length of the cylinder scans will be detected by the circumference measurements.

6.4   Experimental setup

Following the recommendations by Robinson et al. (2012), the system was setup and calibrated essentially the same way it would be were human body segments being scanned, as discussed in chapter 3. For the purpose of the investigation, the Kinects were mounted on tripods and positioned to be dimensionally equivalent to the framework as discussed in chapter 3: creating a calibrated scanning volume of approximately 0.4 x 0.4 x 1.2 m. To make the testing process easier, the scanning area was raised from the floor to coincide with a table upon which the cylinders could be placed, figure 6.1.
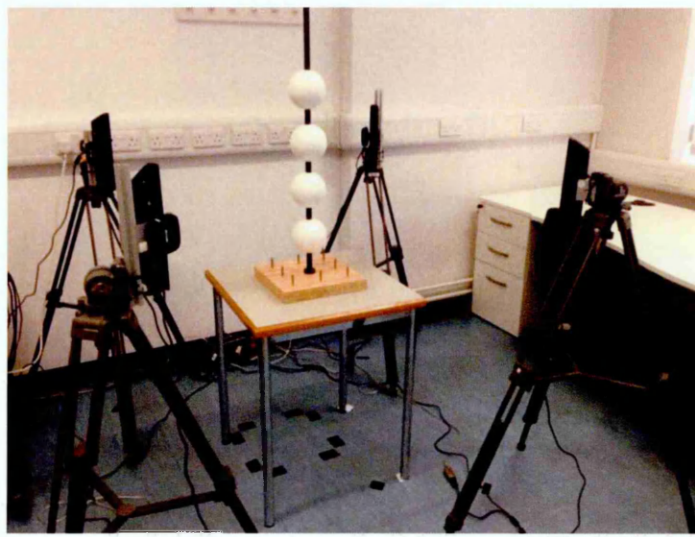
Figure 6.1 – Experimental setup for scanning cylinders representative of typical body segments

6.5 Experimental procedure

6.5.1 Test objects

Cylinders representing common body segments (table 6.1 and figure 6.2) were manufactured from solid aluminium section using a V290 centre lathe (Harrison Colchester, Heckmondwike, UK). The use of cylinders was inspired by the National Physical Laboratory (Teddington, Middlesex, UK) 'Phantom Man', an anthropomorphic collection of rigid, metal, prismatic shapes 1.8 metres tall and designed to represent the human body (Robinson et al. 2012). As each body segment is a simple geometric shape, it can be manufactured and measured very accurately, enabling the performance of scanning systems to be assessed at an appropriate scale.

Table 6.1 - Cylinders representing typical body segments

Cylinder	Representative body segment	ROI Length (mm)	Diameter (mm)	Circumference (mm)	Volume (m³)	ISO 20685-1 Circumference Category
1	Lower arm	272	88	276	1.656×10^{-3}	Small
2	Upper arm and lower leg	373	113	355	3.757×10^{-3}	Small
3	Upper leg	373	162	509	7.695×10^{-3}	Large
4	Torso	350	227	713	1.419×10^{-2}	Large

A uniform coating of satin black spray paint was applied around the ends of each cylinder. After curing, the black paint was covered with a single layer of masking tape. The cylinders were placed on a measurement table and a DHG-300 vertical height gauge (Baty, Sussex, UK) - accurate to ± 0.01 mm – used to cut the tape at a consistent height of 25 mm from the table top. After removing the offcut tape, a uniform coating of white powder was applied to the cylinders, providing a non-reflective surface for scanning. The remaining tape was removed, leaving a black band at either end of the cylinders, figure 6.2. The white area in-between the black bands defined a region of interest (ROI), which could later be identified and segmented from the colour rendered scans in a similar manner to that discussed in chapter 5.

Figure 6.2 – Cylinders representing typical body segments

The ROI of each cylinder was measured by a single skilled engineer using pair of digital engineer's callipers (Kennedy, Leicester, UK), accurate to ± 0.01 mm. These measurements enabled calculation of gold standard circumference values and object classification based on ISO 20685-1 (International Standards Office 2010), table 6.1.

6.5.2 Data capture and post processing

The cylinders were scanned in five positions within the capture volume, figure 6.3, representing the centre and extremes of the calibrated footprint. Each cylinder was raised from the scanning surface and approximately centred within the vertical FOV. This allowed the effectiveness of the device specific calibration procedure (as discussed in chapter 4) to be assessed, as previous work has shown this region to exhibit the greatest errors (Clarkson et al. 2013).
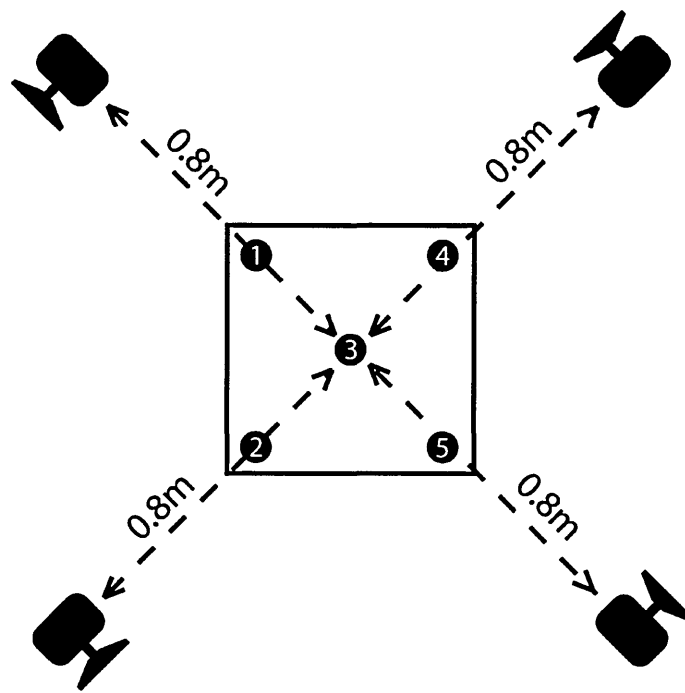
Figure 6.3 – The five positions for scanning the test cylinders

The system was setup and calibrated four times. For each calibration, the cylinders were scanned once in each of the five positions.

Scanning followed the procedure discussed in chapter 3, with the Kinect's IR projectors tuned on individually and simultaneous IR and RGB data captured in a scanning operation totalling around one second. Similarly the scans were post processed as described in chapter 5, including the important process of 'fringe' removal.

A single operator manually digitised three points on the upper and lower black/white interfaces using the data collection software described in chapter 3. These points were used to create segmentation planes, identifying the upper and lower limits of the ROI. Data processing scripts based on the MATLAB programming language (Mathworks, Cambridge, UK) were used to create parallel segmentation planes at 10 mm intervals throughout the length of the ROI. Points in between the segmentation planes were identified, splitting up the scan into point cloud 'slices'. More detail on this process is included in chapter 5.

The ensuing process of calculating circumference of the point cloud 'slices' followed the process described in chapter 5. To account for differences in cylinder length, ten

point slices equally spread throughout the length of the ROI were considered the circumference measurements of each cylinder.

In a similar manner, the total volume enclosed by the digitised ROI was calculated from all of the individual point 'slices' using the method described in chapter 5.

For the purpose of this study, measurements from the scanning system will be assessed with both the ISO 20685-1 standard and the ISAK standard: providing recognised measures of accuracy and reliability (discussed in more detail in section 2.7). As gold standard measurements are available for the cylindrical test objects, both standards will report absolute accuracy, rather than relative accuracy. Similarly, the gold standard measurements will be used for the assessment of inter-tester error according to ISAK standards, therefore reporting absolute inter-tester TEM. In addition to such comparisons, absolute measurements of accuracy and reliability will also be presented: forming the basis of subsequent discussions regarding the suitability of the scanning system for calculating person specific BSPs.

6.6   Investigation of scan-derived measurements and system validity

System performance was assessed in two stages, firstly, scan-derived circumference measurements were analysed and interpreted for the reasons as discussed in section 6.3. With this background, scan-derived volume measurements were then analysed and interpreted.

6.6.1 Accuracy and reliability of scan-derived circumference measurements

Mean absolute accuracy was calculated by comparing circumference measurements from the 3D scanning system to the gold standard measurements, also enabling comparison to ISO 20685-1. The absolute values presented in table 6.2 are based upon two hundred measurements per cylinder: ten representative cross-sectional slices, four repeated calibrations, and five scanning positions. Relative technical error of measurement (TEM) was calculated using the method by Norton and Olds (Norton & Olds 1996)

Average relative intra-tester TEM was calculated for the 10 girth measurements per cylinder across the five positions and four calibrations, with the average value reported. Similarly, the relative inter-tester TEM was calculated by comparing all the measurements for a given cylinder to the gold standard manual measurement.

Values in brackets within table 6.2 represent the corresponding standard deviation values.

Table 6.2 - Accuracy and reliability of scan-derived circumference measurements for the test cylinders

Measure	Cylinder 1	Cylinder 2	Cylinder 3	Cylinder 4
Gold standard circumference (mm)	276	355	509	713
Mean scan-derived circumference (mm)	284 (4)	363 (4)	516 (3)	722 (2)
Mean circumference error (mm)	7 (4)	7 (4)	7 (3)	9 (2)
95% Confidence interval (mm)	± 8	± 8	± 6	± 5
Relative intra-tester TEM (%)	1.05 (0.44)	0.82 (0.38)	0.42 (0.18)	0.28 (0.12)
Relative inter-tester TEM (%)	2.08	1.71	1.04	0.92

Figure 6.4 shows the mean error for each slice (based upon the five positions and four calibrations), combined with the associated 95% confidence interval for each slice position.
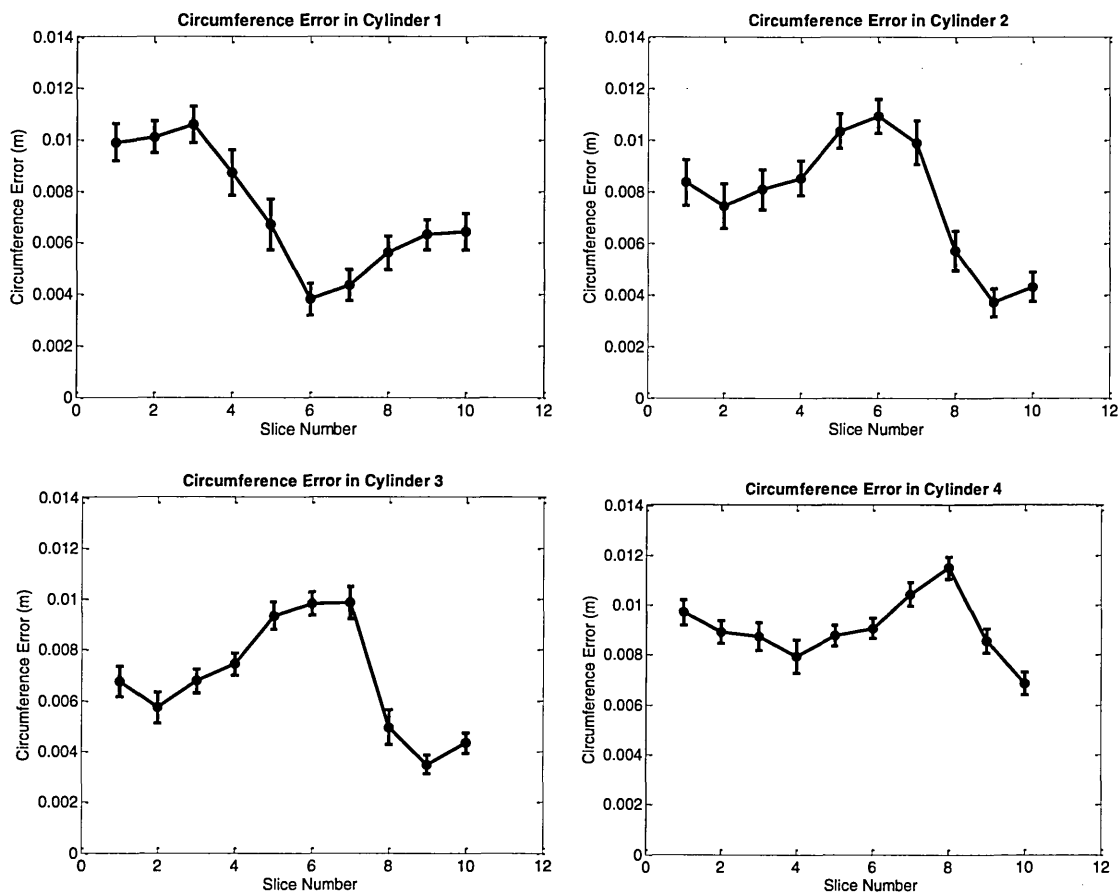
Figure 6.4 – Accuracy and reliability of measurements from the test cylinders

Absolute accuracy was comparable across the four cylinders, table 6.2, with the largest cylinder (cylinder 4) showing a slight increase in mean error. Nevertheless, results show the scanning system to consistently overestimate circumference, suggesting presence of a systematic measurement bias.

Reliability throughout the length of all four cylinders was also comparable, table 6.2. Reliability was slightly higher in the two larger cylinders (Δ 1-2 mm). However a Kendall's tau test ($\tau = 0.112$, $p < 0.05$) showed a lack of interaction between cylinder circumference and error, indicating a lack of proportional bias.

Circumference measurements of cylinders three and four met the requirements of the ISO 20685-1 standard (large circumference, 95% confidence interval of $\pm$ 8 mm (International Standards Office 2010)), whereas cylinders one and two did not (small circumference, 95% confidence interval of $\pm$ 4 mm (International Standards Office 2010)).

Average relative intra-tester TEM varied considerably (a difference of 0.77%) across the four cylinders (table 6.2), suggesting circumference measurements are more reliable throughout the length of the larger cylinders. Results show the scanning system to meet both inter and intra tester post-examination TEM requirements for an ISAK level one practitioner in all four cylinders. Inter and intra tester post-examination TEM requirements of a level two practitioner were met in the largest cylinder only (cylinder 4).

6.6.2 Discussion of scan-derived circumference measurements

The aim of the first part of this study was to determine typical accuracy and reliability of scan-derived circumference measurements, factoring in variability possibly attributed to repeated calibrations and measurements taken from different positions within the capture volume. Circumference measurements of cylindrical test objects were chosen as the basis for initial assessment of system accuracy and reliability, as they enable a greater depth of understanding when subsequently analysing scan-derived volume measurements.

Scan-derived circumference measurements were shown to comply with ISO standards for larger body segments - such as the torso and upper legs – but to be non-compliant when scanning smaller segments, such as the lower arms. Scan-derived circumference measurements were shown to meet the requirements of a level one ISAK practitioner in all four cylinders, and the requirements of a level two practitioner in the largest cylinder only. However, it is important to note the conclusions drawn from this study relate to a relatively small sample size of circumference measurements only: forming only a small part of the ISAK measurement pro-forma. Further studies with larger cohorts of living human participants are required to verify the system is able to adhere to ISO and ISAK standards across the range of measurements included in the ISO and ISAK measurement pro-formas, covering a variety of the bodies segments.

TEM and standard deviation was shown to decrease with an increasing cylinder diameter. One possible explanation for this would be that the Kinect cannot accurately resolve the surface of small objects, or objects with surfaces having high incidence angles, resulting in increased noise in the 3D scans. Another explanation is that as

cylinder diameter decreases, the physical distance which the cylinder must be moved when moving between positions within the capture volume is increased, leading to greater potential for inter scan variability. Such variability can be easily reduced within the application of body scanning, as participants can be requested to stand on feet markers to ensure they stand in the same position within the capture volume across repeated scans.

Results showed the scanning system to overestimate circumference in all four cylinders, suggesting presence of a systematic measurement bias. The circumference overestimation was approximately consistent in all four cylinders (negating random error), indicating a lack of proportional bias. The overestimation was investigated further to determine whether it was a factor of the circumference calculation technique, or related to the underlying 3D scans. Figure 6.5 shows a typical cross sectional point 'slice' from cylinder two plotted against the known diameter of the cylinder.
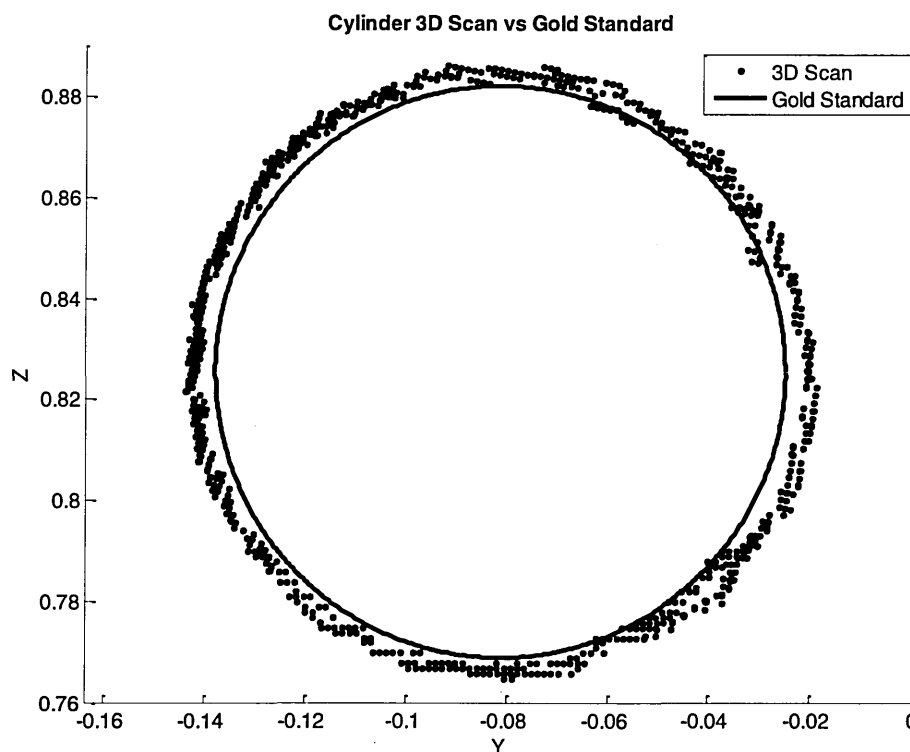
Figure 6.5 – Cross sectional point 'slice' of cylinder 2 plotted against the known diameter

Inspection of figure 6.5 showed the 3D scan clearly over reported the size of the underlying cylindrical object, as reflected in the over reported circumference values.

This characteristic is apparent in other examined cross-sectional slices, verifying the systematic overestimation of circumference observed in this study is as a result of fundamental measurement errors, and not due to the circumference calculation algorithm. Unfortunately, without knowing more about the Microsoft Kinect's 3D vision system it is impossible to definitively conclude where the measurement error arises from.

Importantly, these results suggest that volume estimates of equivalent objects would also demonstrate systematic overestimation. The lack of significant random measurement errors suggests the returned volume would be representative of the underlying geometry (in this case, systematically larger), and would not mask significant measurement errors of the form previously discussed. With this in mind, it is expected that other BSP calculations such as moment of inertia and centre of mass would also offer high accuracy.

6.6.3 Accuracy and reliability of scan-derived volume measurements

The accuracy and reliability of scan-derived volume measurements were next investigated, with the underlying knowledge that circumference is overestimated by the scanning system. With this in mind, it was expected that volume estimates would also be proportionately overestimated in accordance with the circumference overestimation.

Mean scan-derived volume was calculated as the average volume reported from the twenty scans per cylinder, along with the associated standard deviation. Mean absolute accuracy was calculated by comparing the scan-derived volume estimates to the gold standard volume measurements shown in table 6.1. The average volume error and 95% confidence interval associated with each cylinder is expressed both in metric units (m^3) and as a percentage of the gold standard volume.

Results are shown in table 6.3. Where appropriate, standard deviations associated to a mean value are shown in brackets adjacent.

Table 6.3 - Accuracy and reliability of scan-derived volume measurements for the test cylinders

Measure	Cylinder 1	Cylinder 2	Cylinder 3	Cylinder 4
Mean scan-derived volume (m^3)	1.767×10^{-3} (3.968×10^{-5})	3.969×10^{-3} (6.731×10^{-5})	8.014×10^{-3} (9.985×10^{-5})	1.474×10^{-2} (9.770×10^{-5})
Mean volume error (m^3)	1.112×10^{-4} (3.968×10^{-5})	2.120×10^{-4} (6.731×10^{-5})	3.193×10^{-4} (9.985×10^{-5})	5.461×10^{-4} (9.770×10^{-5})
Mean volume error (%)	6.71 (2.40)	5.64 (1.79)	4.15 (1.30)	3.85 (0.69)
95% Confidence interval $(m^3 / \%)$	$\pm 7.776 \times 10^{-5}$ ± 4.70	$\pm 1.319 \times 10^{-4}$ ± 3.51	$\pm 1.957 \times 10^{-4}$ ± 2.54	$\pm 1.915 \times 10^{-4}$ ± 2.54

Accuracy of scan-derived volume measurements was found to vary considerably across the four cylinders, with the largest cylinder appearing to have the greatest accuracy. As expected, the system was shown to overestimate the volume of the cylinders in all cases, agreeing with the previously reported overestimation of circumference.

Reliability was comparable in the four cylinders, with the larger cylinders again appearing to more reliable.

6.6.4 Discussion of scan-derived volume measurements

Analysis of the scan-derived circumference measurements (discussed in section 6.6.2) identified the presence of a systematic measurement bias, with the system over estimating circumference in all cases. Owing to the three dimensional nature of volume measurements, a given overestimation in cylinder circumference will have a greater apparent impact upon the volume of a larger cylinder than a smaller cylinder. Therefore, it is hard to analyse the results shown in **Error! Reference source not found.** as they are being influenced by an additional factor, which is masking the true characteristics of the scan-derived volume measurements.

For this reason, reference volume measurements for the cylinders were re-calculated, based upon the mean systematic overestimation in circumference identified in section 6.6.2. Factoring the circumference overestimation into calculation of the reference volume negates its influence when comparing to scan-derived volumes. In this case,

discrepancies between scan-derived volume and the reference volume must be attributable to random measurement errors and errors in the volume calculation algorithm.

Results reported in table 6.3 were therefore re-calculated by comparing the scan-derived volume measurements to the re-calculated reference volumes. The re-calculated results are shown in table 6.4.

Table 6.4 – Cylinder volume recalculated using systematic overestimation in circumference

Measure	Cylinder 1	Cylinder 2	Cylinder 3	Cylinder 4
Mean scan-derived volume (m^3)	1.767×10^{-3} (3.968×10^{-5})	3.969×10^{-3} (6.731×10^{-5})	8.014×10^{-3} (9.985×10^{-5})	1.474×10^{-2} (9.770×10^{-5})
Recalculated Cylinder Volume (m^3)	1.740×10^{-3}	3.907×10^{-3}	7.908×10^{-3}	1.455×10^{-3}
Mean volume error (m^3)	2.631×10^{-3}	6.252×10^{-3}	1.061×10^{-3}	1.858×10^{-3}
Mean volume error (%)	1.51	1.60	1.34	1.28
95% Confidence interval (%)	± 4.47	± 3.38	± 2.47	± 1.32

Raw scan-derived mean volume errors were shown to vary slightly in the four cylinders, however, they appear to lack any sign of proportional bias of the nature previously observed in table 6.3. As already discussed, a given circumference error will have a greater impact upon the longer cylinders than the shorter cylinders. Therefore, it is expected the raw volume errors will differ slightly in the four cylinders.

In this situation it is therefore better to interpret performance of the system by examining the relative mean volume errors shown in table 6.4 (represented by the reported percentage errors), as the errors are then proportional to the size of the cylinder. Relative accuracy is shown to be comparable in all four cylinders, suggesting all the cylinders experience approximately the same errors.

Reliability of the scan-derived volume measurements followed the same trends as those apparent in the scan-derived circumference measurements, in that reliability appears to increase with an increasing cylinder diameter. However, as discussed in section 6.6.2, this is expected to be related to the nature of the scanning positions and the Kinect's inability to reliably resolve the surface of small objects.

These findings suggest that a systematic overestimation in scan-derived volume would be observed when scanning human body segments. As the raw data from the system will still be susceptible to the seemingly systematic measurement error identified previously in section 6.6.2, it is expected the volumetric accuracies will be comparable to those presented in table 6.3. However, the results of this study will prove essential in analysing the results of ensuing studies with living human participants, providing more information about the system's operating characteristics and hence enabling a greater depth of analysis.

6.7 Discussion

Cylinder shaped objects were chosen for this study, enabling gold standard measurements to be obtained using first principal measurement techniques, whilst reducing the effects of external influential factors: enabling an accurate assessment of the system's fundamental operating characteristics. The cylinders were manufactured to be dimensionally similar to common body segments (International Standards Office 2010; Robinson et al. 2012), but their simple shape means they lack the complex curves and contours typical of the human body. Furthermore the surface of the cylinders was optically 'ideal', providing a non-reflective and non-attenuating surface for scanning. Further investigations with human participants are therefore required to determine equivalent performance when scanning real body segments of greater complexity.

The use of static objects eliminated the potential for motion artefact in the 3D scans. However, involuntary movement of the human body during scanning is likely, despite the system's short scanning duration (less than one second). Such movement would increase measurement errors, having an impact upon both reliability and accuracy. Further investigations with human participants are therefore required, using simple techniques such as light touch (Kouzaki & Masani 2008; Lackner et al. 2001), visual

focus upon a fixed target (Vuillerme & Nafati 2007; Paulus et al. 1984), and hand holds to minimise involuntary postural sway.

It is important to note the purpose of this study was not to identify interactions between object position, calibration, and measurement error: it was to assess the system's performance as a whole, considering all possible attributing factors. The results and validity analysis presented above therefore take into consideration variability as a result of random noise, inter calibration measurement differences, and measurement differences throughout the capture volume. If the interaction between measurement error and such parameters were to be investigated individually, then this should be factored into the initial design of the study, taking more repeated measurements for each condition.

6.8   Chapter summary

Results from a study investigating the typical accuracy and reliability of the scanning system are presented here. Cylindrical test objects representing typical human body segments were scanned with the system and scan-derived measurements compared to gold standard measurements. This enabled determination of the system's fundamental operating characteristics, eliminating external influential factors such as involuntary participant movement.

Results showed the system to systematically overestimate circumference in all four cylinders, which could be corrected with a simple offset factor. Requirements of a level 1 ISAK examiner (TEM $\leq$ 1.5%) were met by the system, and on average the requirements of a level two examiner (TEM $\leq$ 1%). ISO standards for circumference measurements were met in the two larger cylinders only (large circumference, (95% CI of $\pm$ 8 mm). Subject to further studies producing comparable results, these initial findings suggest the system would be well received by practitioners currently using equivalent manual measurement techniques.

As expected, results also showed the scanning system to consistently overestimate volume. The relative volume overestimation appeared approximately systematic in all four cylinders, suggesting the scanning system would produce comparable accuracy on

all body segments, regardless of their size. Assuming the values presented in table 6.4 are representative of the system's performance if the seemingly apparent systematic overestimation is corrected, then the accuracy of the scan-derived volume measurements represents a significant improvement over current BSP estimation techniques. This suggests adherence to one of the initial aims of the project, however further work with more body segments is required. Reliability of scan-derived volume measurements followed the same trends as the scan-derived circumference measurements, exceeding sector requirements, and therefore meeting one of the initial aims of the project.

The use of cylindrical test objects permitted analysis and knowledge not possible with other objects, especially human participants. However, the results of this study provide 'ideal world' operating characteristics. The system is likely to perform differently when scanning human participants, introducing inherent problems such as motion artefact from involuntary movement during the scanning process. This is an area of future work.

Therefore, this study suggests the scanning system presented here would be suitable for obtaining accurate and reliable volume estimates of human participants. Subsequent investigations will consider this further, determining comparable performance when scanning living participants. Knowledge gained as result of this study will permit a deeper understanding of the results, enabling greater informed conclusions.

7 VALIDATION OF THE SCANNING SYSTEM USING LIVING PARTICIPANTS AND COMPARISON TO CURRENT TECHNIQUES

7.1 Introduction

The initial system validation study assessed performance of the system using objects representative of human body segments. The study suggested the system had potential to produce person specific 3D scans from which accurate and reliable BSPs could be obtained. However, the study highlighted a number of potential sources of error which are likely to be apparent when scanning living humans, and would not have been apparent in the initial system validation study. For this reason the system was also tested on living human participants.

For the system to be accepted by clinicians and researchers as a valid alternative to current techniques it must be compared to current practice. This is further supported by Robinson et al (2012), stating that a system should be tested as it would be used rather than relying upon generic accuracy and reliability tests.

It was decided to evaluate performance of the scanning system by obtaining scans and manual measurements of human body trunk segments, owing to the reported problems of assessing trunk volume with current BSP estimation techniques (discussed in more detail in chapter 2). For example, the trunk segment contains a relatively high amount of fat when compared to limb segments (Wicke & Dumas 2010), which could be compressed and change shape between scans, affecting both accuracy and reliability. Additionally, the trunk segment has the greatest potential to differ considerably in shape and size between participants, meaning it is commonly accepted as the hardest segment to estimate and model using current techniques (Huijbregts 2002). Other problems are associated with the likelihood of the segment changing shape and size during and in between scans, owing to the effects of breathing.

Only scan-derived volume estimates were investigated, as they could be compared to the values obtained as part of the investigation described in chapter 6.

7.2 Aims and objectives

Aim

- To compare scan-derived volume estimates of living human participants to those obtained from a geometric BSP model

Objectives

- To determine and compare the reliability of volume estimates.
- To determine the differences in reported volume reported from both techniques, explaining possible reasons.
- To investigate the influence of involuntary participant movement or breathing upon reliability, developing breathing and movement control techniques where applicable.

7.3 Experimental design

It was decided to compare volume estimates obtained from the scanning system to those obtained from a geometric BSP model (Yeadon 1990b), owing to the comparatively high accuracy and reliability when compared to other techniques (discussed in more detail in chapter 2).

7.3.1 Experimental setup

The scanning system was setup in the same form as discussed in chapter 3 and 6, giving an equivalent scanning FOV to that shown in figure 3.12. The experimental setup, figure 7.1, was determined from previous investigations, and was sufficient in size to contain a participant's trunk segment within the calibrated volume without the outer extremities of the body becoming too close to the Kinects to prevent reliable resolution of depth.

169
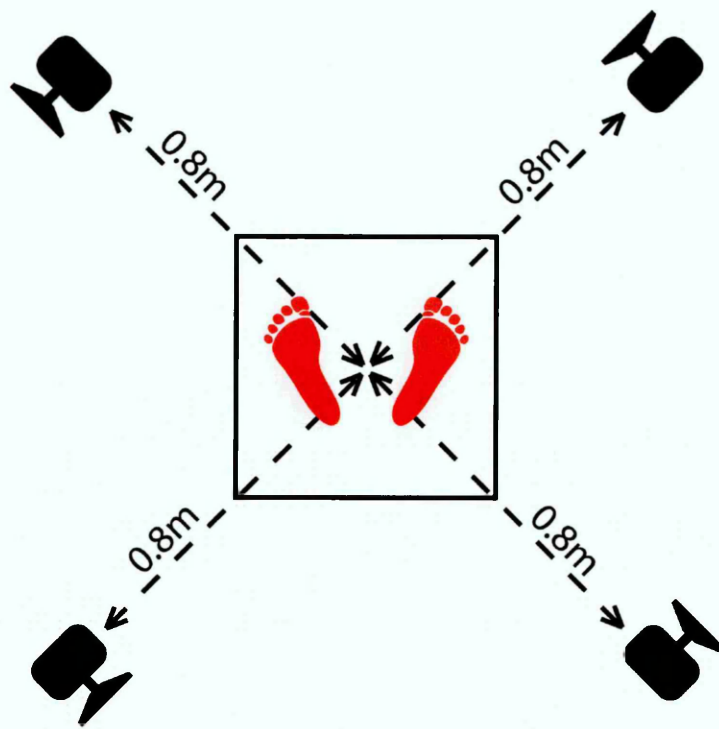
Figure 7.1 – Setup of the scanning system for scanning human participants

The system's extrinsic calibration software (discussed in chapter 3) was used to calibrate the scanning volume before use. Owing to the length of the data collection process it was not possible to maintain one system calibration, as the system may have been inadvertently knocked or moved whilst the scanning system was not in use. For this reason the scanning system was calibrated at the start of each day, with care taken throughout the day to ensure it was not knocked out of alignment. The calibration re-projection RMS errors were checked after every calibration to ensure the values were comparable with those previously reported. Additionally, a scan of a plastic torso manikin was collected with the system immediately before each participant was scanned, serving as a quick 'quality check' to ensure there was not anything immediately amiss with the system setup.

Once calibrated, the data collection software discussed in chapter 3 was used to control the Kinects, capture the required data, and perform the necessary analysis.

7.3.2 Participants

Only male participants were recruited for this study, owing to ethical restrictions and the differences in the torso of male and female participants, otherwise preventing direct comparison of results.

After obtaining institutional ethics approval, a total of forty one living male participants (Mass 77.3 ± 9.1 Kg, Height 1.81 ± 0.06 m, BMI 23 ± 2) were recruited for the study. All the participants were recreationally active and free from any condition which would prevent them from standing in a stationary upright position, hence preventing them from taking part in the study.

7.3.3 Definition of measurements and landmarks

A slightly modified version of the anatomical landmarks proposed in Yeadon's geometric model (Yeadon 1990b) were used to define the area of interest, enabling more reliable palpation. This allowed Yeadon's stadium solids to be formed within the area to be scanned, enabling direct compatibility between techniques. Figure 7.2 shows the anatomical landmarks used to define the proximal and distal limits of the area of interest. In addition, the dashed lines and associated anatomical landmarks in figure 7.2 illustrate the three stadium solids which Yeadon defines to represent the torso segment.

Nipple- 2 Points

Xiphoid Process- 1 Point

Umbilicus- 1 Point

Anterior Superior Iliac
Spine (ASIS)- 2 Points
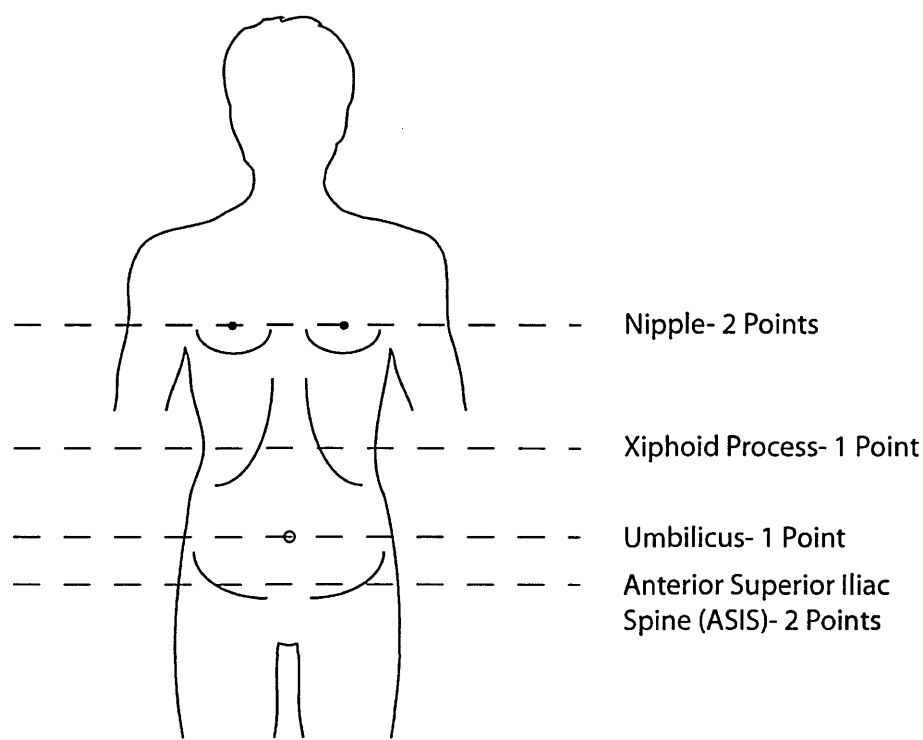
Figure 7.2 – Anatomical landmarks and segmentation process

The anatomical landmarks shown in figure 7.2 were manually identified and palpated on each participant, discussed in more detail below in 7.4.1 .

7.4 Data collection protocol

The data collection process formed two parts, namely the manual measurement process and the scanning process. As the scanning process relied on procedures carried out as part of the manual measurement process, i.e. the palpation and marking of anatomical landmarks, the manual measurements were always taken first. Given the potential for the markers to move between the manual measurements and the scans, the location of the markers was verified by an ISAK trained practitioner prior to the start of each scan.

7.4.1 Manual measurement protocol

Upon arrival of each participant, their height and mass was recorded to enable later classification of the participants based upon BMI. Participants were asked to remove clothing on their upper body, wearing only a pair of close fitting non-compressive lycra shorts throughout the duration of the data collection.

172

The anatomical landmarks defined in figure 7.2 were palpated by a trained ISAK level one practitioner and marked using 10 mm diameter blue self-adhesive markers. Circumference and breadth measurements were taken at each segmentation level using anatomical tape and digital callipers (Kennedy, Leicester, UK) respectively. The height of each segment was also measured using a pair of digital callipers (Kennedy, Leicester, UK). Each measurement was repeated three times and an average value taken. Measurements were recorded in a spreadsheet by a second person with no access permissible by the ISAK practitioner, preventing 'memory effects' (Hencken & White 2006) in the case of repeated measurements. These measurements were used in conjunction with Yeadon's formulae (Yeadon 1990b) to construct the three stadium solids forming the human torso, and hence calculate the overall volume of each participants torso segment using currently accepted techniques.

Throughout the manual measurement process, the same body control techniques as those used for the scanning process (balance and position) were adopted by the participant. This ensured the body was in the same state for both measurement exercises, meaning equivalent measurements were taken in each case. This is discussed in more detail in section 7.4.2.

7.4.2 Scanning protocol

After manual measurement and palpation, the participants were asked to enter the scanning area. Each participant was scanned a total of four times, including one initial scan to allow the data to be checked and the participant to become familiarised with the data collection protocol. This initial scan was discarded and not used as part of the final data set. Each scan took around one second to complete, owing to the delay in turning on/off the Kinect's IR projectors in sequence (discussed in chapter 3). A break of one minute was included between each scan, with the participants asked to leave and re-enter the scanning area after each scan.

Participants were asked to adopt a modified version of the scanning pose defined by ISO 20685-1 (International Standards Office 2010) for the duration of the scans, figure 7.3. The scanning pose was adapted so the participant's arms were externally rotated by 35° instead of 20° with reference to their trunk. This was necessary to ensure the

underarm area of each participant was included in the scans, as these data are required to define the distal limit of the region of interest. Care was taken when defining the data collection protocol to ensure the smallest possible angle of rotation was used, as the position of the arms has potential to considerably change the shape of the upper torso. Furthermore, participants were asked to adopt this position when equivalent measurements were collected as part of the manual measurement process to ensure the collected measurements were comparable. It was not practical to use reference tripods to ensure an exact angular rotation of $35°$ was adopted during the manual measurement process (as used for the scanning process), instead the correct position was verified by visual inspection.
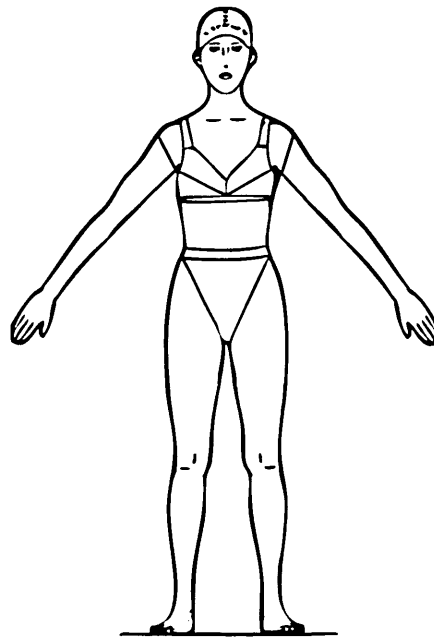
Figure 7.3 – The anatomical pose used for the scanning process (adapted from ISO 20685-1)

Care was taken to ensure the participants remained as stationary as possible throughout the duration of the scanning process. Footprints were placed in the centre of the capture volume, figure 7.1, to ensure participants stood in the correct place, which also aided with improving inter-scan variability. The position of the feet markers was determined with reference to Kirby et al (1987), who studied the impact of foot position upon balance. They concluded the most stable position had the feet separated by a mediolateral distance of 150 mm (mid heel-heel distance) and rotated externally by $25°$, figure 7.4. The feet markers were therefore placed in this position.

174
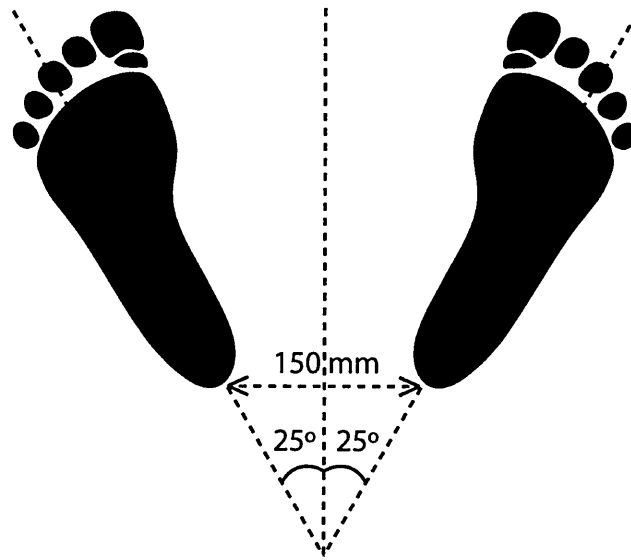
Figure 7.4 – Position of the feet for the scanning process

Whilst this may aid with balance, it is unlikely to significantly reduce involuntary postural sway. In order to achieve this, techniques to balance the upper body were utilised. There are numerous research publications which suggest involuntary postural sway can be reduced by lightly touching a fixed surface with the index finger (Kouzaki & Masani 2008; Lackner et al. 2001), known as light touch stabilisation. The purpose is not to provide any form of mechanical support, but simply to provide a tactile reference surface (Kouzaki & Masani 2008). Other publications suggest balance can be improved by maintaining focus on a fixed visual target (Paulus et al. 1984). However, there appears a lack of studies which directly compare the two techniques in order to assess which is the most effective technique.

It was decided to adopt the light touch stabilisation technique, as regardless of balance, the participants needed something to ensure their arms were held in the defined scanning pose, figure 7.3. The same supports were therefore used for both purposes. Upon entering the scanning area for the first time the participants were asked to hold their arms at approximately 35° external rotation, with the position verified using a goniometer. Participants were then asked to clench their fingers and extend their index fingers. Tripods were then moved up to their index fingers until they were just in contact, whilst ensuring their arms were still in the correct position: therefore providing the necessary level of light touch stabilisation. Figure 7.5 shows the data collection

175

setup, complete with floor markers for feet positioning and supporting tripods for body stabilisation.
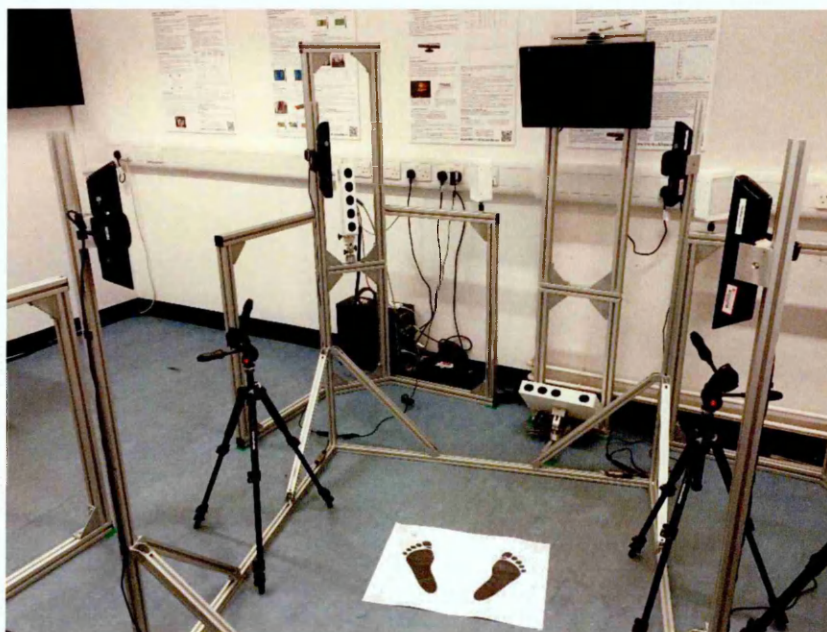
Figure 7.5 – The setup of the scanning system including the feet markers and tripods for stabilisation using light touch

As previously discussed, the typical human breathing cycle has potential to significantly change the shape and hence size of the human torso: particularly the lower of the torso into which the diaphragm moves during the breathing cycle (Daanen et al. 1997). For this reason it was important to control breathing during the scanning process, ensuring repeated scans were always taken during the same part of the breathing cycle. Following the techniques used in previous similar studies (Schranz et al. 2010), participants were asked to hold their breath at the end of the expiration cycle (end-tidal expiration) throughout the short scanning process (around one second), This ensured the body was in a repeatable and relaxed state, as the diaphragm remains in its normal position throughout the duration of the scans. As manual measurement of the torso took significantly longer than the scanning process - around twenty minutes - participants were instead asked to simply maintain normal breathing throughout the manual measurement process. To ensure manual and scan-derived measurements were as comparable as possible, the ISAK practitioner ensured all manual measurements were also taken at the end of the expiration cycle.

7.5 Data post processing and analysis

After collection, each 3D scan was manually digitised by a single operator using the techniques discussed in chapter 5. Unlike Yeadon's technique which models the torso as three separate shapes, the 3D scan includes a complete geometry of the torso, and was therefore treated as a single object. Two markers were digitised on each scan, one of the ASIS markers, and one of the nipple markers: defining the proximal and distal limits of the area of interest. For consistency, the participant's left most markers were always digitised. Segmentation planes were constructed from the two digitised markers using the techniques discussed in chapter 5, returning the enclosed volume of the region between the two digitised markers.

Equivalent volume measurements using Yeadon's technique were obtained using the manual measurements in conjunction with Yeadon's formulae (Yeadon 1990b). This enabled construction of the three stadium solids forming the human torso, allowing the overall volume of each participants torso segment to be calculated. Volume estimates obtained from the 3D scans and Yeadon's technique were then compared to one-another to determine relative accuracy of the scanning system.

Agreement between methods was assessed using limits of agreement (LOA) (Bland & Altman 1986), enabling identification of systematic and random noise differences in the two techniques. Ordinary least products regression (Ludbrook 1997) was used to identify fixed and proportional measurement bias between the two techniques, combined with an r^2 value to indicate random differences between the two techniques.

Reliability was assessed using a combination of techniques. Typical error was calculated for each of the two trial pairs in order to identify presence of order bias (Smith & Hopkins 2011). The average overall typical error was calculated for all 3 repeats and all participants in order to determine test-test variation. Relative accuracy was assessed by calculating intra-class correlation coefficients (ICC) (Shrout & Fleiss 1979) using a two way random effects model with single measures accuracy (ICC (2,1)) (Shrout & Fleiss 1979; de Vet et al. 2006). Average relative technical error of measurements (TEM) was also calculated (Perini et al. 2005; Stewart & Sutton 2012). Both ICC and TEM are established methods of assessing clinical reliability (Munro

2000), meaning results from the study can be compared to previously published values (Shrout & Fleiss 1979; de Vet et al. 2006; Perini et al. 2005; Stewart & Sutton 2012) in order to determine validity and likely acceptance of the system presented herein.

7.6 Results

Data relating to two participants were removed from the study, owing to problems which occurred in the data collection process and were not identified at the point of collection. Therefore, the results shown below relate to a total of thirty nine participants.

7.6.1 Agreement with current techniques

Figure 7.6 shows the agreement between scan-derived and Yeadon-derived volume, assessed using limits of agreement (Bland & Altman 1986). Data values represent the mean volume calculated from the three repeated scans and three repeated anthropometric measurements. Error bars represent the range of the reported differences either side of the mean.
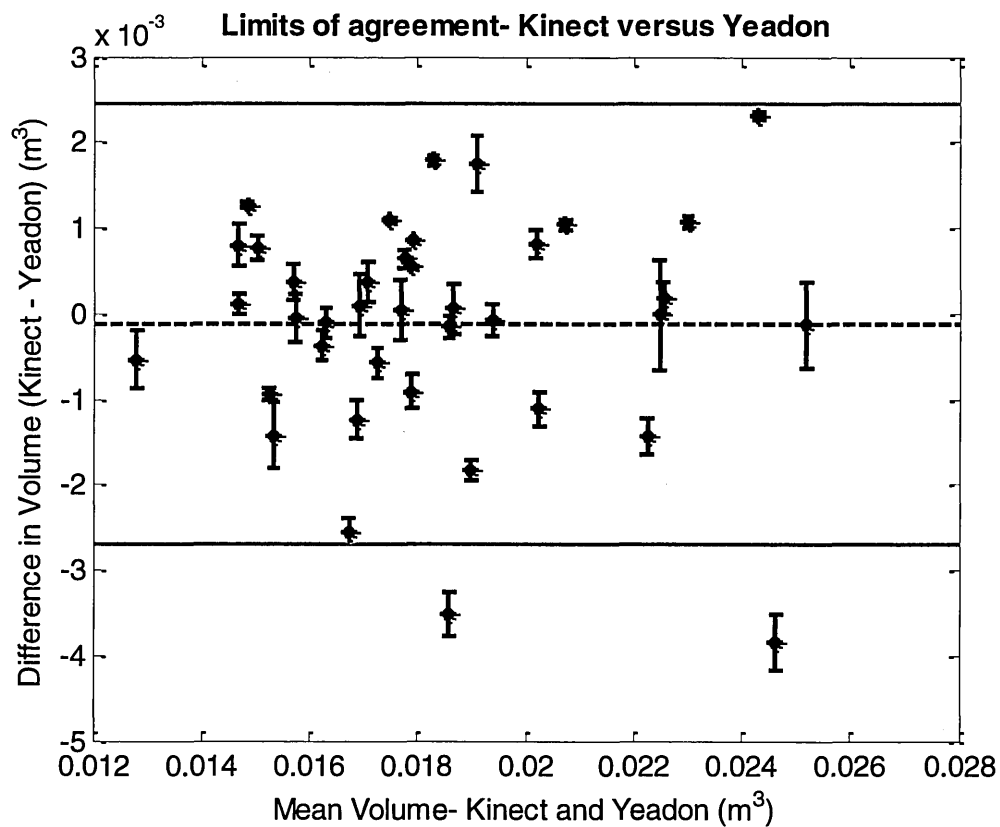
Figure 7.6 – Limits of agreement between scan-derived volume and Yeadon-derived volume

Mean and standard deviation of differences (Kinect-Yeadon) $-1.25 \times 10^{-4} \pm 13 \times 10^{-4}$ m^3

Mean and standard deviation of differences (Kinect-Yeadon) -0.44 ± 6.6 %

Figure 7.6 shows the mean difference between techniques to be close to zero, suggesting good agreement between techniques. However, the 95% confidence interval suggests presence of variable differences between techniques, which results in a mean difference of approximately zero across the thirty nine participants. In some participants the agreement between methods is very close – being close to zero - whereas in other participants there is a considerable difference, as large 16% in some participants. This difference may be owing to errors in the scan derived measurements, or due to errors in the Yeadon-derived measurements as previously discussed. This therefore requires further investigation.

Interestingly, this suggests the previously identified systematic overestimation of scan-derived volumes is not apparent in the results presented above.

Proportional bias between techniques was assessed using ordinary least products regression (OLP) (Ludbrook 1997). Figure 7.7 shows the OLP plot, complete with corresponding a and b values to identify fixed and proportional bias respectively. An r^2 value also provides a measure of the random differences between the two techniques. The plotted data values represent the average of the volume calculated from the three repeated scans and three repeated anthropometric measurements for each participant. Error bars represent the deviation in the Kinect reported volume either side of the mean.
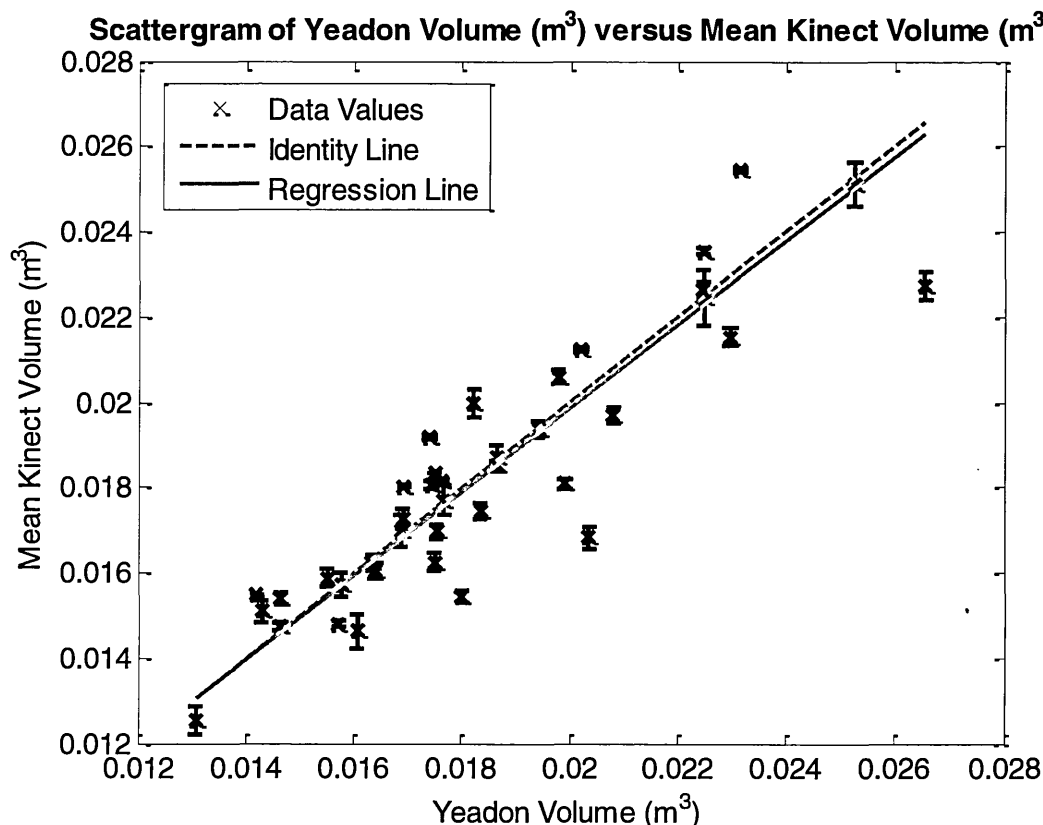
Figure 7.7 – Ordinary least products regression (OLP) plot showing the agreement between scan-derived volume and Yeadon-derived volume

Fixed bias (a) = 1.63 x 10^{-4}, 95% Confidence interval = -12.9 x 10^{-4} – 15.1 x 10^{-4}

Proportional bias (b) = 0.98, 95% Confidence interval = 0.91 - 1.06

Random differences (r^2) = 0.82

Results of the OLP analysis show a to be approximately equal to zero, combined with small 95% confidence intervals either side of zero. This agrees with the Bland-Altman plot, suggesting a lack of systematic bias in the scan-derived measurements. Similarly,

b was shown to be close to one, with small 95% confidence intervals either side of one, suggesting the lack of any proportional bias in the data.

The r^2 value was shown to be less than one. This also agrees with the previous analysis, suggesting presence of random differences between the two measurement techniques which cannot be explained by a proportional or fixed measurement bias. The variable differences between techniques may be owing to the ability of Yeadon's geometric model to represent some physiques better than others (Wicke & Dumas 2010), possibly masking the prior identified systematic overestimation in scan-derived volumes.

7.6.2 Reliability and comparison to acceptable tolerances

Reliability of the scan-derived and Yeadon-derived volume measurements was assessed using a variety of techniques.

Firstly, the average volume across all participants for each repeated measurement (scan-derived and Yeadon-derived) was calculated in order to determine differences in the three repeats, shown below in table 7.1.

Table 7.1 – Mean and standard deviations of Yeadon-derived and scan-derived volume data in the three repeated scans

Repeat	Mean scan-derived volume (m^3)	Mean Yeadon-derived volume (m^3)	Standard deviation scan-derived volume (m^3)	Standard deviation Yeadon-derived volume (m^3)
1	1.83×10^{-2}	1.83×10^{-2}	3.10×10^{-3}	3.00×10^{-3}
2	1.83×10^{-2}	1.84×10^{-2}	3.10×10^{-3}	3.20×10^{-3}
3	1.83×10^{-2}	1.85×10^{-2}	3.00×10^{-3}	3.20×10^{-3}

Results show almost equivalent means across the three repeated scans, suggesting a lack of differences in the three repeated scans. Similarly, approximately equal means were shown across the three repeated volumes calculated using Yeadon's model and the anthropometric measurements. Low standard deviations were shown across the

181

repeated measurements from both techniques, suggesting high reliability of scan-derived and Yeadon-derived volume.

Differences in the repeated scans were further assessed by calculating the change in the mean scan-derived and Yeadon-derived volume across all participants for each of the two possible trial pairs, table 7.2. Results are expressed in absolute measurements, with the associated 95% confidence interval shown adjacent in brackets.

Table 7.2 – Change in mean of scan-derived and Yeadon-derived volume measurements across different scan pairs

Scan Pair	Scan-derived change in mean (m^3)	Scan-derived 95% confidence interval (m^3)	Yeadon-derived change in mean (m^3)	Yeadon-derived 95% confidence interval (m^3)
2-1	2.07×10^{-5}	$\pm 7.57 \times 10^{-4}$	9.76×10^{-5}	$\pm 2.46 \times 10^{-3}$
3-2	-9.41×10^{-6}	$\pm 5.98 \times 10^{-4}$	-8.67×10^{-5}	$\pm 2.04 \times 10^{-3}$

These results further support those previously obtained, again suggesting a lack of order bias across the repeated scans, with only small noticeable differences in the mean scan-derived and Yeadon-derived measurements. The magnitude of the changes in the mean between repeats is comparable for the Yeadon-derived and scan-derived measurements, suggesting comparable reliability. Importantly, these results suggest the previously observed variable difference between techniques is not attributable to poor reliability of the scan-derived or Yeadon-derived measurements.

Reliability was also assessed by calculating intra-class correlation coefficients (ICC) for the Yeadon-derived and scan-derived volume data, table 7.3. ICCs were calculated for each technique, using the three repeated measurements across all thirty nine participants in conjunction with a two way random effects model with single measures accuracy (Shrout & Fleiss 1979).

Table 7.3 – Intra-class correlation coefficients for scan-derived and Yeadon-derived volume measurements

Measurement Technique	ICC(2,1)	95% confidence interval
Scan-derived	0.998	± 0.001
Yeadon-derived	0.978	± 0.015

The calculated scan-derived ICC suggests a lack of noticeable differences owing to different participants and repeated scans, suggesting high reliability of the system. Similarly the Yeadon-derived ICC suggests a lack of noticeable differences in the repeated manual measurements, also suggesting high reliability.

Technical error of measurement (TEM) (Stewart et al. 2011; Perini et al. 2005) was also calculated across the thirty nine participants and three repeated scans in order to assess relative accuracy. Average relative intra-tester TEM was calculated for both the scan-derived and Yeadon-derived volume measurements and is presented with their associated 95% confidence intervals, table 7.4. Average relative inter-tester TEM was also calculated for the scan-derived and Yeadon-derived volume measurements, providing a secondary measure of agreement between the two techniques. Further details regarding TEM and the calculation process can be found in section 6.6.3.

Table 7.4 – Intra-tester TEM for scan-derived and Yeadon-derived volume measurements

Measurement Technique	Average relative intra-tester TEM (%)	95% confidence interval (%)
Scan-derived	0.66	± 0.44
Yeadon-derived	2.21	± 1.18

Results show the average relative intra-tester TEM for the scan-derived volume measures to be lower than that of the Yeadon-derived volume measures. This suggests greater intra-participant reliability of scan-derived measures, possibly attributable to the lack of errors which are inherent in manual anthropometric measures (O'Haire &

Gibbons 2000). Inter tester TEM is understandably high, reflecting the observed random differences between the volume measures reported from the two techniques

The calculated intra-tester TEMs and ICCs are later compared to limits of clinical and sector acceptability, contextualising performance of the system, and indicating likely acceptance of the system by practitioners.

7.7 Discussion

7.7.1 Agreement between techniques

Results suggested a lack of systematic and proportional differences between the two techniques, instead, random differences were identified. Importantly, further analysis suggested the random differences were not attributable to poor reliability of the scan derived measurements, instead being attributable to other influential factors: discussed in more detail below. In addition, the previously identified systematic overestimation of scan-derived volumes was not apparent in the results, possibly being masked by the random differences between techniques.

It was expected the random differences could be attributed to the problems of modelling different physiques using Yeadon's geometric model (Wicke & Dumas 2010), leading to different volumes being reported by the two techniques and hence the observed random differences. It was therefore decided to conduct further investigations to greater understand the source of the observed random differences.

Twelve participants were chosen to be investigated further, four with a larger scan-derived volume than Yeadon-derived volume, four with close agreement between the scan-derived volume and Yeadon-derived volume, and four with a smaller scan-derived volume than Yeadon-derived volume. It was expected all the participants within each group would portray similar anatomical characteristics, explaining some of the trends apparent in the results.

Cross sectional point cloud slices (with a height of 2 mm) were extracted from the 3D scans at the four segmentation levels shown in figure 7.2. The corresponding manual

measurements taken at the four segmentation levels were used to produce the stadia used as part of Yeadon's model, representing the bounds of the stadium solids forming the trunk segment. The 3D cross sectional slices were 'collapsed' so the points lie in the same plane, enabling the stadia to be overlaid onto the points to determine agreement. This process was repeated for all twelve participants identified above, however, only the data corresponding to one participant for each category is shown in figure 7.8, figure 7.9, and figure 7.10. Importantly, equivalent analysis for all participants within each group demonstrated the same characteristics. The black points represent the surface geometry derived from the 3D scans, and the grey shapes represent the stadia used in conjunction with Yeadon's model.
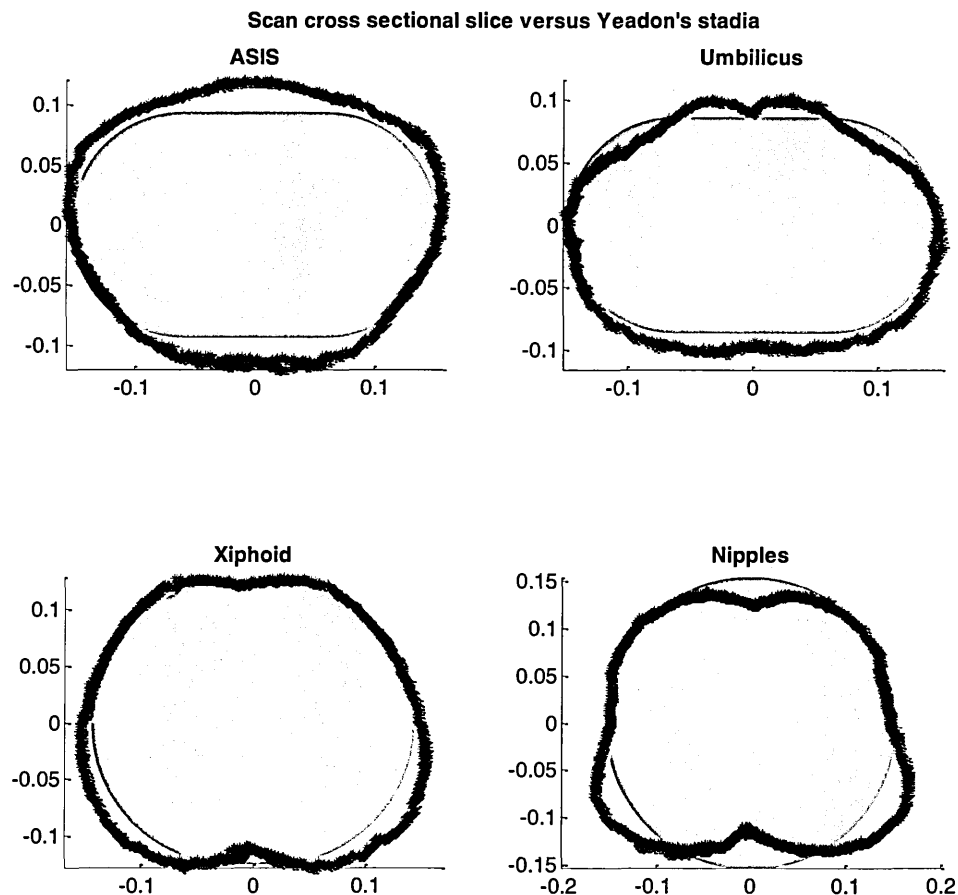
Figure 7.8 – Cross sectional point cloud slices and Yeadon's stadia for a participant having a greater scan-derived volume than Yeadon-derived volume

Figure 7.8 shows the agreement between scan-derived geometry and Yeadon's stadia for a participant having a greater scan-derived volume than Yeadon-derived volume. Inspection of figure 7.8 shows the area of the stadia to be smaller than that enclosed by

the scan-derived surface geometry at all four segmentation levels. With this in mind, it is immediately apparent why this participant's scan-derived volume is greater than the Yeadon-derived volume. It is important to note the cross sectional point slices shown in figure 7.8 are expected to include the previously identified systematic overestimation in scan-derived geometry. Were this not the case, then it is expected there would be greater agreement between the two techniques for this particular group of participants.

Figure 7.9 shows the equivalent overlaid stadia and scan-derived surface geometry for a participant having a greater Yeadon-derived volume than scan-derived torso volume.
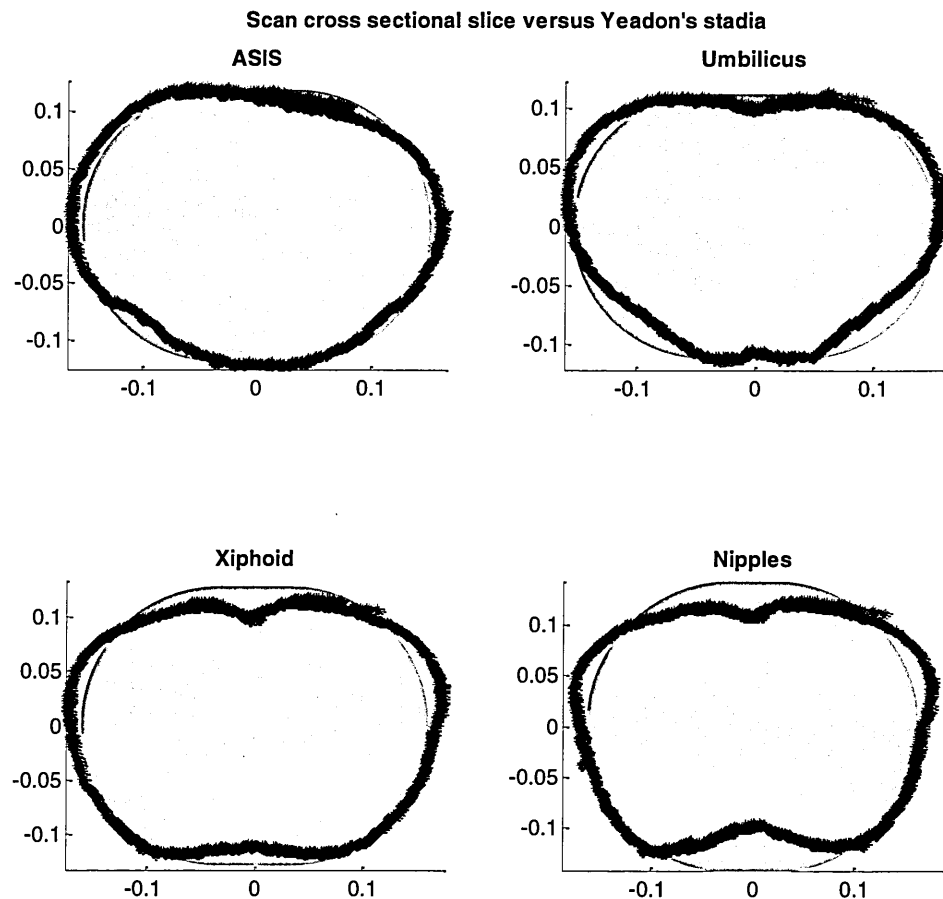
Figure 7.9 - Cross sectional point cloud slices and Yeadon's stadia for a participant having a greater Yeadon-derived volume than scan-derived volume

Inspection of figure 7.9 shows the stadia to have a greater surface area than that enclosed by the scan-derived outer surface at three out of the four segmentation levels. With this in mind, it suggests that two of the three stadium solids representing the torso will be greater in size than the scan-derived shape, resulting in the observed greater

Yeadon-derived volume. As the scanner is expected to overestimate surface geometry, it is expected that were the systematic measurement error corrected, there would be a greater difference between the two estimation techniques for this particular group of participants.

Figure 7.10 shows the equivalent overlaid stadia and scan-derived surface geometry for a participant having good agreement between Yeadon-derived and scan-derived torso volume.
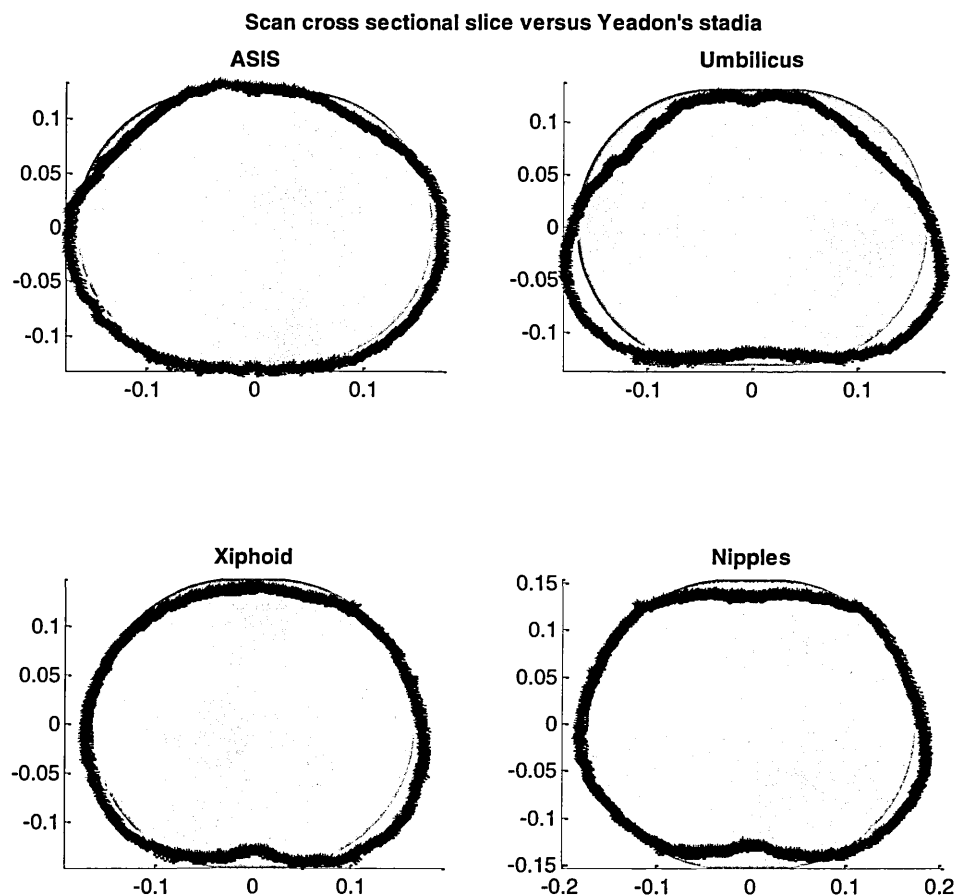
Figure 7.10 - Cross sectional point cloud slices and Yeadon's stadia for a participant having good agreement between Yeadon-derived volume and scan-derived volume

Inspection of figure 7.10 shows the surface areas of the stadia and that enclosed by the scan-derived surface to have close agreement at all four segmentation levels. This will result in stadium solids which closely match the 3D scans, and hence close numerical agreement between scan-derived volume and Yeadon-derived volume. It is important to note that whilst

figure 7.10 shows the umbilicus segmentation level to have some of the stadia lay outside the area enclosed by the scan-derived surface geometry there are some areas enclosed by the scan which are not covered by the stadia. These two areas are approximately equal in size, hence producing close numerical agreement.

Whilst there is closer numerical and geometric agreement in these four participants, were the scanning system's systematic measurement error corrected it is likely that Yeadon's technique would consistently produce greater volumes for this particular group of participants.

This analysis suggests the observed random differences between techniques can likely be explained by the differences in volume calculated by the two techniques. Figure 7.8, figure 7.9, and figure 7.10 suggest the differences in calculated volume are likely owing to the different geometries represented by the two techniques. Geometric agreement - and hence volumetric agreement – was found to differ depending upon the physique of the participant, and is likely due to the ability of Yeadon's model to represent some physiques better than others. Yeadon's stadia and the scanner appear to offer poor agreement in participants who have particularly pronounced anatomical features, such as pronounced muscles in the shoulder area, or a relatively large stomach. On the other hand, close agreement was found in participants who had no particularly prominent anatomical features and were of an 'average form', being close to the anatomical model upon which Yeadon's stadia were derived.

However, the lack of gold standard volumetric data means the expected intrinsic error of Yeadon's model cannot be quantified, nor can firm conclusions be formed regarding which technique is nearer the 'true' value for each of the three groups of participants. Furthermore, without gold standard volumetric data of living human participants it is impossible to further investigate the seemingly systematic overestimation of scan-derived volume, and determine why this was not apparent in the results above.

Further work is therefore concentrated on obtaining gold standard measurements of trunk segment volume using a highly accurate and reliable scanning system. Scan-derived volumes and surface geometries will be compared to those obtained from the

188

developed scanning system and Yeadon's model, providing greater understanding of the developed scanning system's operating characteristics.

7.7.2 Reliability and comparison to acceptable tolerances

Results suggested high reliability of scan-derived volume, and a lack of differences in the three repeated scans. Importantly, these results also suggested the observed random differences between techniques were not attributable to poor reliability of the scan-derived measurements.

TEM was used to assess reliability, being a measure of reliability commonly accepted by practitioners within the area of sports (Stewart et al. 2011). ISAK (Stewart et al. 2011) define acceptable limits for TEMs, based upon practitioners trained to take anthropometric measurements under ISAK protocols. The anthropometric measurements used within this study to form the stadium solid shapes, nor volume measurements are part of the ISAK measurement protocol, however, comparison of the results to ISAK standards provides context for the findings of this study. Further information regarding TEM and the limits of acceptance defined by ISAK can be found in section 6.6.3.

Intra-tester scan-derived volume measurements were found to have a relatively low TEM (0.66 (± 0.44)), meeting the intra-tester requirements of an ISAK level 2-4 assessor (TEM ≤ 1%), representing the highest level of qualification. On the other hand, the intra-tester Yeadon-derived volume measures were found to have a relatively high intra-tester TEM (2.21 (± 1.18)), greater than the TEM requirements for the lowest level of ISAK accreditation (level 1 assessor, TEM ≤ 1.5%)). Adherence to the intra-tester requirements of an ISAK level 2-4 assessor suggests high reliability of the scan-derived volume measures, and likely acceptance of the system by practitioners within the area of sports. The Yeadon-derived intra-tester TEM was found to be higher than expected (2.21 (± 1.18)), suggesting poor reliability. However, the absolute measures of reliability shown in table 7.1 suggest the reliability is not as poor as may be expected from simply examining the calculated TEM. The greater than expected TEM is likely due to the anthropometric measurements and their locations (used to form the Yeadon stadium solids) not being part of the ISAK measurement pro-forma (Stewart et al.

2011). It is possible these measurements have an inherently lower level of reliability when compared to other measurements in the ISAK pro-forma, meaning it is not possible to reach the ISAK TEM requirements when taking such measurements. ISAK defines full adherence to a particular level of accreditation based upon an inter-tester and intra-tester TEM score. The inter-tester TEM is calculated by comparing equivalent measurements from a person being examined and an ISAK assessor of level three or higher (Stewart et al. 2011), considered a gold standard measurement. However, given the magnitude of the differences between scan-derived and Yeadon-derived volume measurements, and the problems associated with Yeadon's geometric volume calculation technique (Wicke & Dumas 2010), it is unfair to consider the Yeadon-derived volume measurements to be gold standard and use them to infer inter-tester agreement of the scanning system. Instead, further testing should take place using a gold standard method of calculating volume, with results from this technique used to determine absolute inter-tester agreement, and hence adherence to ISAK standards.

ICC was also used to assess reliability, deemed appropriate for use in clinical environments (Munro 2000). Research suggests that measurement systems having an ICC greater than 0.7 (de Vet et al. 2006) are acceptable for use within such environments. The scanning system was found to have an ICC of 0.998, suggesting high reliability of scan-derived volume measurements and likely acceptance of the system by practitioners within clinical and healthcare fields. Similarly, the Yeadon-derived volume measurements were found to have an ICC of 0.978, again suggesting high reliability. In comparison, Outram et al (2012) investigated the reliability of torso volume measurements derived by two examiners with the use of Yeadon's BSP model. Results showed the two examiners to deliver average ICCs of 0.87 and 0.88 across the torso segment.

Throughout the design and execution of this study, care was taken to limit the influence of breathing, involuntary movement, and postural sway upon the scans, otherwise having potential to impact upon accuracy and reliability. The reliability metrics evaluated above suggest this was successful, meeting reliability targets typically accepted by practitioners in this area. However, it is inevitable that some participants are able to stand more stationary than others, and are able to repeatedly adopt the same scanning pose.

Further analysis of the results showed six participants to have a lower reliability than others. Visual inspection of the 3D scans showed the reduced reliability likely occurred due to involuntary movement throughout the scanning duration. A typical cross sectional point cloud slice showing such artefacts is shown in figure 7.11.

Figure 7.11 – A cross sectional point cloud slice showing participant movement during the scanning duration

Inspection of figure 7.11 without merit may suggest misalignment of the four viewpoints forming the complete 360° scan is simply due to poor calibration of the scanning system. However, the quality of the calibration was verified prior to each scan by scanning a torso manikin, showing good alignment between the four viewpoints. Instead, the misalignment is expected to be due to participant movement during the scanning operation. For example, in figure 7.11 the first collected scan was the back right of the torso (bottom left of figure 7.11), subsequent scans were collected around the participant's torso in a clockwise direction, meaning the last scan to be collected was the front right of the torso (bottom right of figure 7.11). Therefore, it is apparent the participant may have moved between the collection of the first and last scan, resulting in the misalignment highlighted above.

Given the interference of adjacent Kinects and the need to turn the IR projectors on and off, there is limited scope to make the scanning time shorter and further prevent the

influence of involuntary movements during the scanning operation. However, the scanning duration could be approximately halved by concurrently collecting scans from the two front and back Kinects as they are obscured from one another by the participant's body (meaning the projector of each Kinect will not interfere with the one opposite). This is something which should be investigated as part of the on-going future system improvements and development.

Although it is hard to quantify from visual inspection of the 3D scans, it is possible that reliability – and indeed fundamental accuracy – may have been affected by the adopted anatomical landmarking and segmentation protocol. Care was taken as part of the scanning protocol to ensure the participant was stood completely vertical and parallel to the scanning system, as this could otherwise have an influence upon the segmentation and volume calculation technique (discussed in more detail in chapter 5). Figure 7.12 and figure 7.13 show the effect upon the cross sectional slices used to calculate volume which may be caused by a participant leaning back or forth instead of being stood completely parallel to the scanning system.
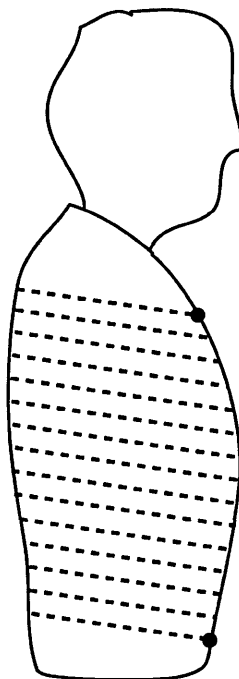
Figure 7.12 – Cross sectional slices of a participant's torso that was out of alignment with the global coordinate system during capture of the scan
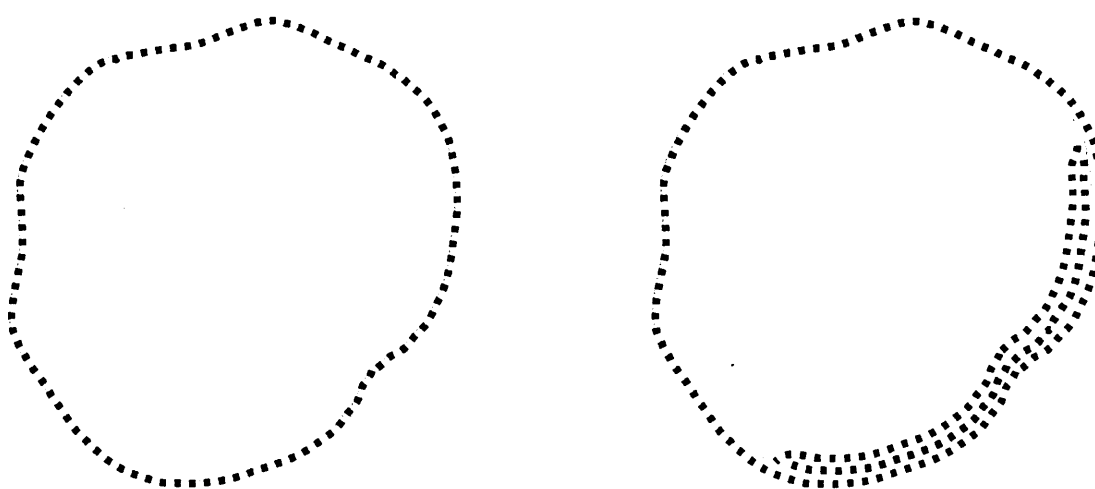
Figure 7.13 – Cross sectional torso slices of a torso in alignment with the global coordinate system (left) and out of alignment (right)

Figure 7.13 shows that in this case, the cross sectional slices become distorted (rightmost image in figure 7.13) and not truly cross-sectional (leftmost image in figure 7.13). The change in shape of the cross sectional slices leads to different points being used to represent the outer surface (figure 7.13), subsequently leading to errors in volume. Future work should seek to develop an improved anatomical landmarking protocol which is independent of participant orientation within the scanning area.

7.7.3 Relevance of the results and findings

Results and conclusions from this study relate to a sample of thirty nine recreationally active male participants with a BMI of 23 ± 2. When using the scanning system with participants falling outside of this demographic it is important to accept that different performance of the scanning system may be observed.

Relevance of the participants with respect to the UK population, and hence relevance of the findings of this study with respect to the UK population can be assessed by comparing participant BMIs to those of the UK population. Figure 7.14 shows a histogram of participant BMIs, with the UK BMI curve from 2011 overlaid onto the data (National Obesity Observatory 2011).

193

Histogram of Participant BMI vs UK Population
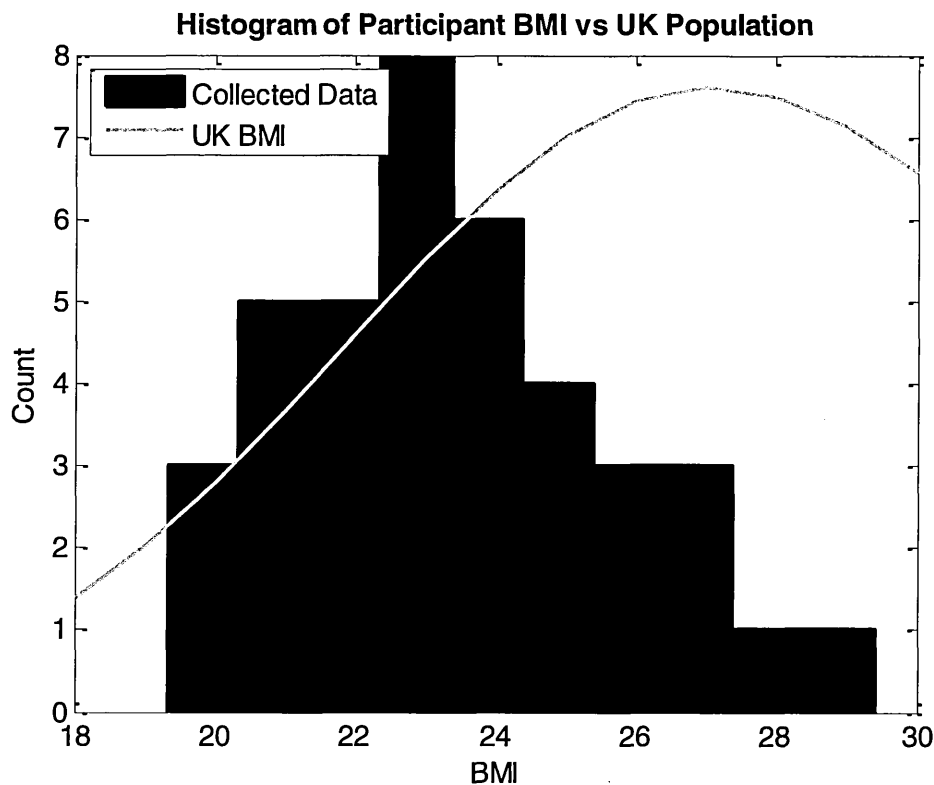
Figure 7.14 – Histogram of participant BMIs against the UK BMI curve (National Obesity Observatory 2011)

Figure 7.14 shows the participants to occupy a broad range of BMIs, representing those in the healthy and overweight BMI categories (NHS 2014). Owing to the demographic of the participants recruited for the study, the majority fall at the lower end of the UK BMI curve, meaning there are fewer people towards the upper end of the BMI range (> 25) where the majority of the UK population appear to fall.

However, the findings of this study suitably represent BMIs between around nineteen and twenty five, encompassing the healthy BMI category (NHS 2014). The majority of people that would be assessed within sports environments likely fall in this category, meaning the findings of this study are likely directly applicable to such participants. However, people likely to be assessed within healthcare communities are likely to be located at the extreme ranges of the tested BMI range, given the applicability of the system for profiling eating disorders or obesity for example. Before forming an assessment of system suitability for profiling the torso volume of the UK population as a whole, further testing is required using participants towards the upper end of the BMI range.

Importantly, this study only focussed upon the torso segment, cited to be the hardest segment to estimate and model with current BSP estimation techniques (Huijbregts 2002). Closer agreement between the scanning system and current techniques would be expected when investigating other body segments, as they contain less fat (Huijbregts 2002) and provide closer agreement with the stadium solid shape for example (Yeadon 1990a). Therefore, future studies should seek to characterise the accuracy and reliability of scan-derived measurements across all of the bodies segments.

Furthermore, this study only concentrated upon scan derived volume, representing only a small part of the range of measurements (volume, mass, COM, and MOI) encapsulated within body segment parameters. Therefore, future studies should seek to obtain the full array of body segment parameters from all of the bodies segments, with accuracy and reliability likely to differ between the different measurements. For example, MOI is cited to be the hardest to calculate accurately with conventional techniques (Wicke & Dumas 2010).

7.8   Chapter summary

This chapter documents an investigation into the agreement and reliability of human torso volume measurements obtained from 3D scans and Yeadon's BSP model.

Results show good agreement between scan-derived volume measurements and those derived using Yeadon's BSP model, with a lack of proportional or systematic measurement bias. However, agreement between techniques varied across the thirty nine participants, with some showing greater, smaller, and comparable volumes. Further analysis suggested this variable agreement is possibly due to Yeadon's BSP model being able to model some body physiques and shapes better than others. The previously identified systematic overestimation of scan-derived volume was not apparent in the results, possibly being masked by the variable accuracy of Yeadon's BSP model.

However, the lack of gold standard data prevented further analysis of the results. Future work should therefore involve obtaining gold standard 3D scans of the human body with which the scanning system and Yeadon's model can be compared.

Good reliability of scan-derived volume measurements was reported across the thirty nine participants and three repeated scans per participant. ICC values for clinically acceptable measurement systems were met, whilst intra-tester TEM requirements of an ISAK level 2-4 anthropometric practitioner were also met by the system. With this in mind, it is expected the scanning system would be well received by practitioners within the fields of sports and healthcare. Furthermore, the adherence to sector requirements for reliability meets one of the original aims of the project: do develop a scanning system able to deliver person-specific volume estimates with a reliability that is equal to or in excess of values for sector acceptability.

Care was taken to control postural sway and breathing during the scanning process, which would otherwise have likely reduced the accuracy and reliability of scan-derived volume measurements. These steps appeared to be effective for most participants, with the scans of only six participants displaying artefacts arising from movement during the scanning operation. Future work is focussed upon further reducing the scanning duration in order to limit the effects of such movement as much as possible.

Participants within the study represent those located towards the lower end of the UK BMI scale, possibly meaning reported performance of the scanning system may not be directly applicable to participants with a higher BMI. Whilst participants within a sports environment are likely to fall within the BMI range represented by this study, it is likely that were the scanning system used in a healthcare environment would be located at the extreme ends of the tested BMI range. For this reason, future work should focus upon participants of greater and smaller BMIs, being more representative of the UK population as a whole.

Future studies should also focus on characterising performance of the system across more of the bodies segments, as conventional BSP estimation techniques are cited to offer greater accuracy when scanning segments better obeying the stadium solid shape. These studies should also focus upon obtaining the full range of BSPs (volume, COM, and MOI), as accuracy and reliability is likely to differ between measurements.

Good agreement with a commonly used BSP estimation technique, combined with high reliability (in excess of sector guidelines) suggests the developed scanning system

would likely be well received by both healthcare and sports environments. The developed system offers many advantages and additional features when compared with conventional techniques, whilst inherently reducing the errors associated with manual anthropometric measurements (O'Haire & Gibbons 2000). Future work is focussed on identified improvements and further characterising the scanning system, assessing its suitability for replacing current BSP estimation techniques. With this information, practitioners will be able to select the most appropriate BSP estimation technique for a particular assessment activity.

8 FUTURE WORK

8.1 Introduction

The continued research, development, and validation process has resulted in three main areas requiring future work:

- Further development of the 3D scanning system
- Improvements to the data collection protocol and analysis techniques
- Further investigations into the influence of the work upon current BSP estimation techniques

These areas are discussed in greater detail below. Recommendations for future work include small modifications to the data collection system and protocol to improve accuracy and reliability. Larger changes are focussed on making the system more user friendly and quicker, whilst the largest changes focus on further studies in order to answer longer term research questions.

8.2 Further development of the 3D scanning system

Since the development of the 3D scanning system discussed throughout this program of study, the area of consumer level RGB-D sensors has changed considerably. For example, the Primesense 3D vision technology which forms the basis for a number of popular RGB-D sensors has been acquired by Apple (BBC 2013). This has led to the imminent discontinuation of a number of popular RGB-D sensors (Asus Xtion, Fotonic P70), including the Microsoft Kinect for Windows as used in this study (Microsoft Corporation 2014). The Microsoft Kinect for Windows has recently been succeeded by the Kinect v2 (Microsoft 2014d), however little information is currently available regarding the devices accuracy and reliability when used for 3D scanning. Rather than interest being focussed upon the large standalone RGB-D sensors discussed chapter 2, attention appears to be turning towards the development of much smaller OEM RGB-D sensors (Intel 2014). RGB-D sensors are becoming more ubiquitous, being integrated within consumer televisions for natural user interaction (Samsung 2013), and integrated within tablet computers and mobile phones for both natural user interaction and 3D

scanning of environments (Occipital 2014; Cubify 2014a). With this in mind, a thorough updated review of available RGB-D sensors is required in order to identify a suitable future replacement for the Kinect for Windows sensor currently used in the scanning system. Assessment of suitability would likely follow the procedures and techniques identified in chapter 6, in order that performance could be compared to the Kinect for Windows. In light of the findings of this study, attention may turn towards the development of a scanning system taking a very different form. This may include changes such as being much smaller, more portable, and simpler, meaning it could easily be used within home environments.

The current system was designed so the scanning process is as simple and quick as possible, however the system is designed to only scan a single segment of the body at once. Whilst this is sufficient for studies interested in a single segment of the body, other studies focussed upon full body BSPs would have to collect multiple scans of the body in order that all segments could be geometrically dimensioned. In this case the system would still offer many benefits over current BSP estimation techniques, but would lead to a longer data collection process: however, still likely to be less than the time requirement of current BSP estimation techniques, and under the target maximum data collection time of 10 minutes per person. Further work is however required to improve the functionality and user friendliness of the system by expanding the system to collect full body 3D scans in a single scanning operation. This would likely involve adding additional sensors in order that the scanning FOV is increased.

Although the performance, accuracy, and reliability improvements of the device specific calibration process was investigated as part of the development process, the stability of the calibration technique over a long period of time was not investigated. Whilst no changes in accuracy or reliability were detected whilst using the calibrated sensors over the course of one year, it is important this is quantified numerically. A number of factors may have an effect upon the stability of the calibration, including changes in the material and camera properties over time due to prolonged heating of the device as part of general use, and general knocks and component movement owing to day to day use of the sensors. Calibration stability should be numerically quantified over the period of a year for example, subjecting calibrated sensors to a controlled set of representative

day to day usage, and evaluating the impact upon accuracy and reliability (by scanning planes and cylindrical objects for example).

A significant limitation of the current device-specific calibration technique is the amount of time required to obtain the calibration data used to build the distortion correction model. Currently a large number of calibration images and scans are required (dependent upon the calibrated range), as discussed in chapter 4. Possible improvements lie in a number of areas, including the calibration model and data collection protocol. Improvements to the calibration model involve reducing the number of calibration images required to reliably constrain the distortion correction model, possibly involving the addition of a smoothing parameter to the input sensor data. This would make the data collection process quicker, increasing its practicality. Improvements to the data collection protocol lie around the automation of the data collection process. For example, the sensor being calibrated could be mounted onto a linear rail with the location of the sensor on the rail controlled via a stepper motor. In this case the calibration process could be started and left to run on its own accord. If the planar surface was large enough, multiple sensors could be mounted onto the rail and calibrated simultaneously: making the calibration process much quicker.

There are a number of possible improvements which could be made to the extrinsic calibration technique in order to improve its user friendliness, accuracy, and reliability. The current calibration technique assumes the sensors remain in the same place between calibration and data collection, with their relative position to one another determined by an initial calibration technique. Whilst this technique proved effective, it is possible the sensors and supporting framework may be knocked during general use – such as when a person enters or leaves the scanning area – meaning the system must be re-calibrated before further use. In a similar manner, a person being scanned may move slightly between the capture of the 3D surface profiles from the four different viewpoints. This would lead to one of the viewpoints becoming out of alignment with respect to the other three, but in other cases it would be a usable scan. Both of these problems could be circumvented with use of a post data capture alignment refinement process. The relative position of the cameras with respect to one another could first be approximated using a similar calibration process to that used currently, but using fewer (around four) calibration scans. This would be sufficient to obtain approximate inter-camera

200

transformations, whilst making the calibration data collection process quicker. After data collection, the scans and initial calibration parameters could be input to a scan alignment optimisation algorithm, such as ICP. The initial calibration parameters would be sufficient to constrain the ICP model in order that it converges to the correct solution, whilst also considerably reducing the execution time. The output of the ICP alignment process would be the same geometric scan, but with the four surface viewpoints better aligned with one another. The drawback of this approach would be the increase in scan post processing time before numeric parameters could be calculated, however, benefits would lie in terms of accuracy and reliability. Future work is therefore required to numerically quantify the accuracy and reliability benefits offered by such a technique against the trade-off of an increased data processing time.

8.3   Improvements to the data collection and analysis protocol

Despite the developed technique offering many time savings over many current BSP estimation techniques, a significant drawback is the need to palpate and mark the body and digitise the 3D scans, in order to define the bounds of the bodies segments. Depending upon the segments being scanned this process can be lengthy, reducing the benefits offered by the system, reducing its practicality for use in typical sports training environments, and therefore likely reducing its adoption by practitioners. With this in mind, alternative techniques of identifying anatomical landmarks in the 3D scans need to be investigated.

The need to digitise the 3D scans after collection could be eradicated by palpating and marking the anatomical landmarks using uniquely coloured markers. Image processing techniques and the Kinect's colour camera could be used to identify the markers in 2D which are later converted to 3D coordinates using the Kinects integral coordinate mapping functions. Whilst this offers time savings by eradicating the need to digitise the 3D scans, it still requires the body to be palpated and marked before collecting the 3D scans: likely to be the biggest time requirement. With that in mind, a better technique may be to eradicate any manual intervention in terms of palpation, marking, and digitisation by developing an automatic anatomical landmark detection algorithm (Werghi 2007). The algorithm should work on the raw 3D scans, identifying anatomical landmarks based upon anatomical definitions such as shape, size, and

relative location to other features of the body. If this were implemented, the time requirements for the current data collection and post processing workflow would be considerably reduced, offering further advantages and likely increasing adoption of the system by practitioners. Potential problems with this approach would likely arise when scanning participants with a high BMI, meaning bony or prominent anatomical points may be obscured by excess fat or skin.

A potential source of error in the current system is focussed upon the protocol for defining the bodies segments. Currently, it is assumed the person being scanned is stood parallel to one of the Kinect sensors from which the scanning system's global coordinate system is derived. For the torso segment this technique is practical, however, for other segments such as the arms it is impractical to request participants hold their arm completely parallel to the scanning heads. Similarly, if the system is adapted to scan the whole body at once, then it is likely only the torso segment would be in this orientation.

It is possible that were a person being scanned not stood completely parallel to the global coordinate system then the cross sectional slices would be formed out of alignment with the torso, leading to errors in calculated volume. Similarly, if a person is suffering from asymmetry of the body, then anatomical points on the left of the body are likely to be located at distinctly different positions to the right of the body for example. If only a single marker is used to define proximal and distal limits of the segment (as in chapter 7), then it would likely lead to errors in volume.

An alternative technique may be to define three anatomical landmarks at the proximal and distal bounds of the segment of interest, defining a local segmental coordinate system. This would eradicate any errors owing to misalignment with the global coordinate system, and would similarly eradicate any errors owing to asymmetry of the body. Importantly, this technique would also be transferrable to any of the bodies segments. However, when land marking the torso for example this would require a marker to be placed on the back of the torso to ensure the segmentation plane is in the correct orientation. As there are no anatomical landmarks in an equivalent position on the front and back of the body then this would be very hard, also likely leading to errors in volume.

One of the significant advantages offered by the scanning system is the ability to collect and archive 3D scans over a long period of time. In sports training environment for example, this would be useful for examining change in the body's physique over time in response to training interventions or changes in training programs. Rather than simply quantifying change of this nature in terms of anatomical measurements, a useful feature would be to overlay scans collected over periods of time with one another, allowing change in physique to be viewed in the form of an animated 3D body model. This feature is something which will later be revisited as part of continued system development.

8.4 Further investigations into the implication of the results upon current BSP estimation techniques

The developed scanning system offers many advantages in terms of simplicity, time, features, accuracy, and reliability when compared with current BSP estimation techniques. However, there is a range of future research and development which must be followed prior to the likely adoption of the system by practitioners in the area. For example, the first-step study conducted in chapter 7 suggested high reliability of scan-derived volume measures and good agreement between scan-derived volume measures and those derived using conventional BSP estimation techniques. However, volumetric measures are only a small part of the range of measurements (volume, mass, CoM, MOI) which are encompassed within BSPs. For this reason, future development must focus upon extending the data processing algorithms to provide such parameters. Future research should be focussed upon a study similar to that conducted in chapter 7, comparing the whole range of BSPs to current techniques rather than just volume. Based upon current research, it is expected the scanning system to also demonstrate high accuracy and reliability across the whole range of BSPs, particularly when calculating MOI and CoM which have been shown to be poorly estimated with current BSP estimation techniques (Wicke & Dumas 2010). With this information, practitioners interested in the system will be able to compare its performance to the techniques they currently use in order to determine the relative advantages offered by the system.

The initial system validation study conducted in chapter 6 showed the developed system to repeatedly overestimate volume across the four scanned cylinders. However, this trend was not apparent in the results in chapter 7, showing the scanning system to randomly produce smaller and larger volume measures than the geometric BSP model. However, analysis of the results conducted in chapter 7 suggested this was likely due to Yeadon's geometric BSP model randomly underestimating and overestimating volume, masking the consistent overestimation of scan-derived volume measures. In order that this seemingly systematic volume overestimation can be further investigated, quantified, and corrected, scan-derived volume measures of living participants need to be compared to gold standard measures of volume. A similar study to that conducted in chapter 7 should be repeated, instead replacing Yeadon's geometric model with an accurate and reliable gold standard body measurement technique, such as the 3D rapid structured light scanning system by 3dMD (3dMD 2014).

Participants used for the final system validation study discussed in chapter 7 had BMIs that were towards the lower end of the UK population, likely being representative of people who are regularly engaged with sport. Whilst it is likely that people of this physique would be scanned with the system were it used within a sports environment, if the system was used within a healthcare environment then it is likely the system would be used to scan people with a range of BMIs, likely including a number located towards the upper end of the BMI range. It is likely that the system would perform differently when scanning people with a high BMI, for example, having areas of excess or overhanging skin and fat, possibly creating occlusions or shadows, leading to degradation of the 3D scan data and hence inaccurate and unreliable calculated BSPs. For this reason, the system needs to be tested using a sample of participants that is representative of the UK population in terms of BMI, quantifying accuracy and reliability with respect to BMI. Qualitative analysis of the 3D scans is likely to lead to modifications to the scanning system, scanning protocol, and data processing algorithms, whilst quantitative analysis of the results will provide a greater understanding of the systems performance and hence its suitability for a range of applications.

8.5 Chapter summary

Considering the areas identified as requiring further work, four main future research projects have been identified:

- A fundamental system accuracy study, comparing BSPs derived using the scanning system to those derived using a gold standard body measurement technique, such as the rapid structured light scanning system by 3dMD (3dMD 2014). The study should involve further system development, enabling the whole range of BSPs to be compared rather than just volume. Participants for the study should be representative of the whole range of UK BMIs, in order that system performance can be quantified using a range of participants.
- Research and development of a technique to automatically identify anatomical landmarks from captured 3D scans. This process should also involve the development of an improved technique for defining the bounds of a body segment of interest and governing the ensuing 'slicing' process.
- Development of a technique to refine the alignment of collected 3D scans, including making the initial global system calibration process quicker than at present.
- Research and development of alternative scanner configurations in order that full body BSPs can be obtained in a single scanning operation. This will also involve the research into a suitable future replacement sensors for the Microsoft Kinect for Windows

These projects encompass the long term goal of the research project: to develop a low cost 3D scanning system which can be used to estimate person specific BSPs, improving current reliability and accuracy whilst making the whole process simpler, quicker, and adding additional functionality. The work conducted thus far represents significant progress towards this goal, with the four projects highlighted above required to fully achieve this long term goal, increasing the likelihood of adoption by practitioners in this area and hence changing current practice.

205

9 OVERALL DISCUSSION

The aim of this project was to develop a technique for obtaining estimates of person-specific segmental volume, with accuracy exceeding current techniques, and reliability in excess of requirements for clinical acceptability. The research was motivated by the apparent inability of current BSP estimation techniques to meet the needs of practitioners in the areas of sports and healthcare, with current techniques often leading to large errors in calculated BSPs, and resulting in large errors in subsequent calculations. Current techniques are also very time consuming, meaning many analysis activities that would be beneficial to a sports programme are not conducted as they would be prohibitive to normal activities. Therefore, development of a system with a short scanning time and capable of producing measurements in less than 10 minutes was a priority.

Background research suggested the developed technique should be based upon 3D scanning, whilst costing less than £2000. With this in mind, it was apparent that the use of low-cost consumer 3D sensors would be the most applicable technique. Therefore, the research aimed to answer the question: "Can low cost consumer 3D surface scanning techniques be used to obtain estimates of person-specific segmental volume with accuracy greater than current techniques and a reliability meeting clinical requirements."

9.1   Development of the 3D scanning system

One of the first steps in the project was to develop a 3D scanning system in accordance with criteria identified during the background research and defined in the project's aims and objectives (section 1.2). The scanning system comprises four Microsoft Kinect RGB-D sensors, providing 3D scan data and RGB colour images. The use of consumer level scanning hardware results in a low cost system, costing less than £1,200 including a PC, Microsoft Kinects, and all required hardware: meeting one of the original aims of the project.

The system is capable of obtaining 3D scans in around 0.8 seconds, however the scanning field of view is only capable of scanning a single segment at once. However,

it would only require 5 separate scans to obtain full coverage of the human body in the current system configuration. This would likely take significantly less than the 10 minute per-person target time set out in the original aim of the project, whilst also including any time for palpation, landmarking, and data post-processing. Future work will seek to reduce the overall time commitment per person even further.

An extrinsic rigid-body calibration process was developed, identifying the position of the four sensors with respect to one another, and enabling 3D scans of the human torso segment to be obtained in less than one second. Investigations with the Microsoft Kinect revealed significant distortion of the depth data, leading to the development of a device-specific initial calibration procedure applicable to a wide range of RGB-D sensors.

Various artefacts were apparent in the Kinect's 3D data, including point 'fringes' in areas of high curvature or high incidence between the Kinect and object. A fringe removal filter was therefore developed to remove these areas. Algorithms to calculate circumference and volume of the 3D scans were developed, enabling investigation of system accuracy and comparison to current techniques.

System calibration, data capture, post processing, and data analysis was integrated into a range of bespoke software applications, relying upon no external commercial software applications. This simplifies use of the system, whilst also leading to the development of a completely standalone low cost segmental volume estimation technique.

The developed scanning system is comparable to that by Kilner (2012) and Boehm (2012), however, this system offers a smaller footprint and greater capture area, a simple calibration technique, further processing of the Kinects raw 3D data in order to provide greater accuracy and reliability, and bespoke data collection and analysis software for the specific application of calculating BSPs.

9.2 Empirical findings

Accuracy and reliability of the scanning system was first assessed by scanning cylindrical objects representative of typical body segments. Results showed high

reliability of volume and circumference measurements derived from 3D scans provided by the system, in excess of sector requirements for acceptable ICCs and TEMs, and therefore meeting one of the original aims of the project. A systematic measurement overestimation was identified in scan-derived volumes and circumferences, which could be easily corrected with a simple offset factor.

Agreement between scan-derived torso segment volume and that obtained from a typical geometric BSP model (Yeadon's model) was also investigated. Results demonstrated high reliability of scan-derived volume measurements, in excess of sector requirements for TEM and ICC and hence meeting one of the original aims of the project. Results showed good agreement between scan-derived values and those derived using Yeadon's model across the range of participants included in the study. However, large differences were found between techniques in some participants. Background research and the initial system validation study, suggested this was due to the ability of Yeadon's model to reliably model the torso of some participants better than others. However, further studies involving gold standard person-specific BSPs and more body segments (as outlined in chapter 8) are required before forming final conclusions regarding accuracy and reliability of measures provided by the system.

9.3   Practical implication of the findings

Considering the developed system and findings from the studies conducted, it is expected the scanning system would be widely accepted by practitioners, offering many advantages over current BSP estimation techniques in terms of reliability, time, and the range of available data. However, further work – as outlined in chapter 8 – is required to characterise the system over a wider variety of participants, additional body segments, and the full range of BSPs before widespread use could be recommended.

Problems with current BSP estimation techniques are widely accepted across the sports sector. For example, practitioners investigating the dynamics of high acceleration movements would likely benefit from the greater reliability and accuracy offered by the scanning system when compared to current techniques. Importantly, the low cost and simple nature of the developed system likely means it could be easily integrated into current protocols. Practitioners within training and competition environments would

likely benefit from the time savings offered by the system, allowing assessment of BSPs in situations where it is currently infeasible owing to the time demands of current techniques. Other applications lie within the expanding area of talent identification and assessment. The ability to overlay and compare collected 3D scans allows changes in physique or anthropometric characteristics to be tracked over time, or enables scans of potential athletes to be overlaid onto those of a known 'good' athlete. This would offer a far deeper level of analysis and comparison than is possible with current anthropometric assessment techniques. Following on from this, many applications lie within 'consumer gyms'. Gym members may be allowed one scan per month for example, which compares their physique and anthropometric measurements to those from the previous month, whilst also tracking their progress towards the characteristics of a known 'good' athlete in their chosen sport.

A number of applications also lie within the medical and healthcare sector. For example, body volume index (BVI) has gained interest in recent years as a replacement for the currently used body mass index (BMI). However, BVI relies on the capture of accurate and reliable 3D scans of the body, from which volume can be calculated. The system developed as part of this project offers the ability to provide such data in a cost effective manner, enabling widespread measurement of BVI. The ability to overlay and compare scans would likely also be of benefit to those working in the areas of obesity and weight loss for example. Scans could be overlaid on top of one another and used to assess change in weight and size over time, providing far richer data than is possible with anthropometric measurements alone, whilst also acting as a visual weight loss motivational tool that could be used by patients. In a similar manner, there are many other applications of anthropometric and volumetric measures within the healthcare sector, including child growth monitoring and tracking, tracking the growth of a baby during pregnancy, and more widely, the use of 3D scanning for diagnosing and tracking the onset or development of skeletal conditions (including conditions such as curvature of the spine and body asymmetry). Other applications include the treatment of lymphedema patients who suffer from localised fluid retention throughout the body, and are often subject to lengthy and invasive anthropometric measurement processes (Armer & Stewart 2005). The patient experience would be immensely improved with a quick and accurate non-contact surface scanning system, which is able to obtain measures and distribution of segmental volume without touching sensitive areas of the patient's body.

Similarly, breast surgeons are currently seeking low cost, accurate, and non-invasive techniques for measuring breast volume (Cardoso et al. 2012), enabling a more objective method for selecting the correct implant to use for an operation (Cardoso et al. 2012).

Considering further applications of the developed system, many applications also lie in sectors such as retail and apparel. The ability to quickly capture accurate and reliable scans of the human body, and later use these scans to calculate anthropometric parameters has many potential applications. For example, the scanning system could be used to capture scans of the UK population as part of a sizing study. Given the quick scanning time it could be placed in places such as shopping centres, and scans collected of passing shoppers without significant intrusion on their time. Other applications lie around integrating such a system into retail outlets for virtual try on of clothes, recommendation of clothes size within a changing room, or for capture of anthropometric measurements to enable the manufacture of custom clothing or apparel.

Finally, considering broader implications of this research, the study as a whole represents a significant contribution to the advancement of knowledge in the areas of 3D scanning, the use of consumer RGB-D sensors for 3D scanning, and biomechanics as a whole. For example, the work on device specific calibration offers a significant advancement in the understanding of consumer RGB-D sensors and their inherent problems. The developed calibration technique improves the accuracy of such devices, extending the possibilities for their use. In a similar manner, the developed extrinsic calibration process extends the possible capture area of RGB-D sensors, hence increasing the range of applications in which RGB-D sensors could be used. The work on characterising the accuracy and reliability of the developed system also extends knowledge in this area, providing knowledge of the likely possible accuracy and reliability of a scanning system based upon consumer RGB-D sensors. With this information, researchers and developers would be able to determine the likely suitability of RGB-D sensors for a particular application. Finally, the developed algorithms and data processing software have many applications within biomechanics, healthcare, and retail for 3D scan analysis, post-processing of existing 3D scan data, and extracting measurements of interest.

9.4 Limitations of the research

An important acknowledgement in the work conducted this far is that comparisons of scan-derived volumes and those derived using current techniques (chapter 7) focus only upon the torso segment, cited to be the hardest segment to estimate and model with current BSP estimation techniques (Huijbregts 2002). However, in order to fully characterise performance of the developed system, it is important it is tested upon all segments of the body.

Furthermore, only scan-derived volume measurements (and circumferences) have been considered thus far, forming only a small part of the range of measurements (volume, mass, COM, and MOI) encapsulated within body segment parameters. Accuracy and reliability is likely to differ between the range of parameters (Wicke & Dumas 2010), meaning the system's ability to accurately and reliably calculate each parameter should be assessed.

Whilst participants from the study comparing scan-derived and Yeadon-derived volume measures (chapter 7) represented a wide range of BMIs (19-29), the majority were located towards the lower end of the examined BMI range. In comparison, the majority of the UK population would be located at the upper end of the examined BMI range. Scanning participants located towards the upper end of the BMI range may result in different accuracy and reliability of scan-derived volumes, owing to the amount of fat within the bodies segments (Huijbregts 2002) and the possibility for shadows and occlusions. Future studies should therefore characterise the accuracy and reliability of scan-derived measures from participants with a wide range of BMIs.

A significant limitation in the comparison between scan-derived and Yeadon-derived volume measurements is the lack of gold standard volume measurements, with which the scan-derived measurements can be compared in order to determine accuracy of the system. Whilst accuracy was inferred from the results of the validation study conducted in chapter 6, it is likely that human influential factors (breathing, involuntary movement) have an effect upon accuracy which must be quantified.

9.5 Future research

Four main future research projects were identified in chapter 8, encompassing the long term goal of the research project, whilst further characterising the system and making incremental improvements:

- A study to compare scan-derived BSPs of a variety of participant physiques to those derived using a gold standard measurement technique.
- Research and development of a technique to automatically identify anatomical landmarks from captured 3D scans.
- Development of a technique to refine the alignment of collected 3D scans, including further simplifying the initial global system calibration process.
- Research and development of alternative scanner configurations and data collection techniques, given the future discontinuation of the Microsoft Kinect for Windows v1 and the increased availability of RGB-D sensors.

Other research will focus on exploring specific applications of the scanning system within sports and healthcare communities.

9.6 Conclusion

A low cost 3D body scanning system capable of obtaining 3D scans of a single body segment and calculating segmental volume has been developed. The system offers many advantages and improvements over current estimation techniques in terms of simplicity, speed, and availability of data. Validation studies using cylindrical objects representative of common body segments showed a systematic overestimation in scan-derived volume and circumference which could be easily corrected with a simple offset factor. Studies comparing volume measures of living participants derived from 3D scans and a commonly used geometric model demonstrated close numerical agreement and reliability in excess of sector requirements, whilst highlighting commonly accepted problems with current estimation techniques.

Further work is needed to characterise the system's performance over more measures and more body segments before widespread use could be recommended. However, the

findings so far suggest the system would be well received by practitioners, having many applications within the sports and healthcare sectors.

10 REFERENCES

3dMD, 2014. 3dMD Torso Scanning System. Available at: http://www.3dmd.com/3dMDtorso/.

Adafruit, 2010. WE HAVE A WINNER – Open Kinect driver(s) released. *Adafruit.* Available at: http://www.adafruit.com/blog/2010/11/10/we-have-a-winner-open-kinect-drivers-released-winner-will-use-3k-for-more-hacking-plus-an-additional-2k-goes-to-the-eff/ [Accessed September 20, 2013].

ALGLIB, 2014. Clustering. Available at: http://www.alglib.net/dataanalysis/clustering.php.

Amazon, 2012. Kinect & Kinect Adventures for XBox 360. Available at: http://www.amazon.co.uk/Kinect-Sensor-Adventures-Xbox-360/dp/B0036DDW2G/ref=sr_1_1?ie=UTF8&qid=1353500423&sr=8-1 [Accessed October 9, 2013].

Anderson, M.R., Jensen, T. & Lisouski, P., 2012. *Kinect Depth Sensor Evaluation for Computer Vision Applications.* Aarhus University.

Arampatzis, A., Gao, J. & Brüggemann, G., 1997. Influences of inertial properties on joint resultants. In *XVIth Congress of the International society of Biomechanics.* Tokyo, Japan.

Artec, 2014. Artec L Scanner. Available at: http://www.artec3d.com/hardware/artec-l/specifications/ [Accessed December 2, 2014].

ASUSTeK Computer Inc, 2012. Asus Xtion Pro. Available at: http://www.asus.com/Multimedia/Motion_Sensor/Xtion_PRO/ [Accessed August 6, 2012].

Ballard, D., 1981. Generalizing the Hough Transform to Detect Arbitrary Shapes. *Pattern Recognition,* 13(2), pp.111–122.

Barter, J., 1957. *Estimation of the Mass of Body Segments,* Ohio.

Bauer, J.J. et al., 2007. MRI-derived body segment parameters of children differ from age-based estimates derived using photogrammetry. *Journal of Biomechanics,* 40(13), pp.2904–10. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17462656 [Accessed January 3, 2013].

BBC, 2013. Apple buys motion sensor maker PrimeSense. Available at: http://www.bbc.co.uk/news/technology-25083914 [Accessed January 24, 2015].

Behnke, A.R., 1959. Comment on the Determination of Whole Body Density and a Resume of Body Composition Data. In J. Brozek, A. Henschel, & Lehigh University, eds. *Techniques for Measuring Body Composition.* Natik, Massachusetts: QuartermasterResearch and Engineering Center, pp. 118–133.

Belhedi, A. et al., 2012. Depth Correction for Depth Cameras From Planarity. In *Proceedings British Machine Vision Conference*. Surrey, UK, pp. 1–10.

Belhedi, A. et al., 2012. Non-Parametric Depth Calibration of a TOF Camera. In *19th IEEE International Conference on Image Processing*. Orlando, Florida: IEEE, pp. 549–552.

Binney, D. & Boehm, J., 2011. Performance Evaluation of the PrimeSense IR Projected Pattern Depth Sensor. , (m), p.10.

Bjornstrup, J., 1995. *Estimation of Human Body Segment Parameters Historical Background*, Institute of Electronic Systems, Aalborg University.

Bland, J.M. & Altman, D.G., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1(8476), pp.307–10. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2868172.

Boehm, J., 2011. Accuracy Investigation for Natural User Interface Sensors. In *Low Cost 3D Sensors, Algorithms and Applications*. Berlin, Germany.

Boehm, J., 2012. Natural User Interface Sensors for Human Body Measurement. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIX(B3), pp.531–536.

Bouguet, J.-Y., 2010. Camera Calibration Toolbox for Matlab. Available at: http://www.vision.caltech.edu/bouguetj/calib_doc/index.html [Accessed October 9, 2012].

Braune, W. & Fischer, O., 1889. *On the Centre of Gravity of the Human Body Free Preview On the Centre of Gravity of the Human Body as Related to the Equipment of the German Infantry Soldier* 1st ed., Leipzig.

Bullock, N. et al., 2009. Talent identification and deliberate programming in skeleton: ice novice to Winter Olympian in 14 months. *Journal of sports sciences*, 27(4), pp.397–404. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19191166 [Accessed March 18, 2014].

Burrus, N., 2011a. Kinect Calibration. Available at: http://nicolas.burrus.name/index.php/Research/KinectCalibration [Accessed October 4, 2011].

Burrus, N., 2011b. Kinect Calibration. Available at: http://nicolas.burrus.name/index.php/Research/KinectCalibration [Accessed October 4, 2011].

Cambridge in Colour, 2013. Digital Camera Sensors. Available at: http://www.cambridgeincolour.com/tutorials/camera-sensors.htm [Accessed May 20, 2013].

Camplani, M., Salgado, L. & Polit, U., 2012. Efficient Spatio-Temporal Hole Filling Strategy for Kinect Depth Maps. *SPIE*.

Challis, J.H., 1995. A procedure for determining rigid body transformation parameters. *Journal of Biomechanics*, 28(6), pp.733–737.

Challis, J.H., 1996. Accuracy of Human Limb Moment of Inertia Estimations and Their Influence on Resultant Joint Moments. , 12, pp.517–530.

Challis, J.H., 1999. Precision of the Estimation of Human Limb Inertial Parameters. *Journal of Applied Biomechanics*, 15, pp.418–428.

Chandler, R.. et al., 1975. *Investigation of Inertial Properties of the Human Body*, Washington.

Chen, L., Lin, H. & Li, S., 2012. Depth Image Enhancement for Kinect Using Region Growing and Bilateral Filter. , (Icpr), pp.3070–3073.

Chen, S.-C. et al., 2011. A method for estimating subject-specific body segment inertial parameters in human movement analysis. *Gait & posture*, 33(4), pp.695–700. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21458993 [Accessed October 8, 2012].

Cheng, C.K. et al., 2000. Segment inertial properties of Chinese adults determined from magnetic resonance imaging. *Clinical Biomechanics*, 15(8), pp.559–66. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10936427.

Chiu, L. & Salem, G., 2005. Net joint moment calculation errors during weightlifting: Dempster versus DEXA. *Clinical Kinesiology*.

Choppin, S.B. et al., 2013. Breast Volume Calculation Using a Low-Cost Scanning System. In Hometrica Consulting, ed. *3D Body Scanning Technologies*. Long Beach, California: Hometrica Consulting, pp. 12–20.

Chow, J.C.K. et al., 2012. PERFORMANCE ANALYSIS OF A LOW-COST TRIANGULATION-BASED 3D CAMERA : MICROSOFT KINECT SYSTEM. , XXXIX(September), pp.175–180.

Clarkson, S. et al., 2012. Calculating Body Segment Inertia Parameters from a Single Rapid Scan Using the Microsoft Kinect. In Hometrica Consulting, ed. *3D Body Scanning Technologies*. Lugano, Switzerland: Hometrica Consulting, pp. 153–163. Available at: http://www.3dbodyscanning.org/2012/conference.html#program.

Clarkson, S. et al., 2013. Distortion Correction of Depth Data from Consumer Depth Cameras. In Hometrica Consulting, ed. *3D Body Scanning Technologies*. Long Beach, California: Hometrica Consulting, pp. 426–437. Available at: http://www.3dbodyscanning.org/2013/conference.html#program.

Clauser, C., McConville, J. & Young, J., 1969. *Weight, Volume and Center of Mass of Segments of the Human Body.*, Wright Patterson Air Force Base, Ohio: AMRL Technical Report.

Cohen, D. et al., 2011. Error Compensation in Three-Dimensional Mapping.

216

Computer Vision Group- Technische Universitat Mumchen, 2013. Intrinsic Camera Calibration of the Kinect. Available at: http://vision.in.tum.de/data/datasets/rgbd-dataset/intrinsic_calibration [Accessed September 25, 2013].

Crisco, J.J. & McGovern, R.D., 1998. Efficient calculation of mass moments of inertia for segmented homogenous three-dimensional objects. , 31, pp.97–101.

Cubify, 2014a. iSense 3D Scanner. Available at: http://cubify.com/en/Products/iSense [Accessed December 3, 2014].

Cubify, 2014b. Sense 3D Scanner. Available at: http://cubify.com/en/Products/Sense [Accessed December 3, 2014].

Cui, C. & Ngan, K.N., 2010. Plane-based external camera calibration with accuracy measured by relative deflection angle. *Signal Processing: Image Communication*, 25(3), pp.224–234. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0923596509001350 [Accessed March 26, 2012].

Cui, Y. et al., 2012. Algorithms for 3D Shape Scanning with a Depth Camera. *IEEE transactions on pattern analysis and machine intelligence*, pp.1–13. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22962071.

Cui, Y. & Stricker, D., 2011. 3D Body Scanning With One Kinect. In *3D Body Scanning Technologies*. Lugano, Switzerland, pp. 121–129.

Daanen, H., Brunsman, M. & Robinette, K., 1997. Reducing movement artifacts in whole body scanning. In *International Conference on Recent Advances in 3-D Digital Imaging and Modeling*. IEEE Comput. Soc. Press, pp. 262–265. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=603874.

Damavandi, M., Farahpour, N. & Allard, P., 2009. Determination of body segment masses and centers of mass using a force plate method in individuals of different morphology. *Medical Engineering & Physics*, 31(9), pp.1187–94. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19683955 [Accessed November 29, 2012].

Dellen, B., Rojas, I. & Rob, I. De, 2013. Volume measurement with a consumer depth camera based on structured infrared light. In *Catalan Conference on Artificial Intelligence*. Victoria, Canada.

Dempster, W., 1955. *Space Requirements of the Seated Operator*, Michigan: Wright Air Development Center.

Dempster, W.T. & Gaughran, G.R.L., 1889. Properties of Body Segments Based on Size and Weight. *American Journal of Anatomy*, 120, pp.33–54.

Derrer, D., 2013. DEXA Scan to Measure Bone Health. *WebMD Medical Reference*. Available at: http://www.webmd.com/osteoporosis/guide/dexa-scan.

Draelos, M., 2012. *The Kinect Up Close: Modifications for Short-Range Depth Imaging*. North Carolina State University.

217

Drillis, R., Contini, R. & Bluestein, M., 1966. *Body Segment Parameters; a Survey of Measurement Techniques.*, Available at: http://www.ncbi.nlm.nih.gov/pubmed/14208177.

Durkin, J.L. & Dowling, J.J., 2003. Analysis of Body Segment Parameter Differences Between Four Human Populations and the Estimation Errors of Four Popular Mathematical Models. *Journal of Biomechanical Engineering*, 125(4), pp.515–522. Available at: http://link.aip.org/link/JBENDY/v125/i4/p515/s1&Agg=doi [Accessed August 5, 2011].

Durkin, J.L., Dowling, J.J. & Andrews, D.M., 2002. The measurement of body segment inertial parameters using dual energy X-ray absorptiometry. *Journal of Biomechanics*, 35(12), pp.1575–80. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12445610.

Fiedler, D. & Heinrich, M., 2013. Impact of Thermal and Environmental Conditions on the Kinect Sensor. *Advances in Depth Image Analysis and Applications*, 7854, pp.21–31.

Fotonic, 2012. *Fotonic P70,*

Freedman, B., Shpunt, A., Machline, M., et al., 2010. Depth Mapping Using Projected Patterns.

Freedman, B., Shpunt, A. & Arieli, Y., 2010. Distance-Varying Illumination and Imaging Techniques for Depth Mapping.

Fuschillo, V.L. et al., 2012. Accelerometry-based prediction of movement dynamics for balance monitoring. *Medical & biological engineering & computing*, 50(9), pp.925–36. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22802142 [Accessed October 17, 2012].

Future Picture, Xbox Kinect IR Patterns. Available at: http://www.futurepicture.org/?p=116.

Gittoes, M.J.R., Bezodis, I.N. & Wilson, C., 2009. An Image-Based Approach to Obtaining Anthropometric Measurements for Inertia Modeling. *Journal of Applied Biomechanics*, 25(3), pp.265–270.

Gittoes, M.J.R. & Kerwin, D.G., 2006. Component inertia modelling of segmental wobbling and rigid masses. *Journal of Applied Biomechanics*, 22(2), pp.148–54. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16871005.

Gonzalez-Jorge, H. et al., 2013. Metrological evaluation of Microsoft Kinect and Asus Xtion sensors. *Measurement*, 46(6), pp.1800–1806. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0263224113000262 [Accessed July 12, 2013].

Gould, T. & Edmonds, M., 2010. How MRI Works. *How Stuff Works*. Available at: http://science.howstuffworks.com/mri.htm [Accessed January 20, 2015].

Gregory, D., 2006. The General Theory of Small Oscillations. In *Classical Mechanics*. Cambridge University Press, pp. 421–456.

Hanavan, E.P., 1964. *A Mathematical Model of the Human Body*. Ohio: USAF Institute of Technology.

Hancock, J. et al., 1997. High-performance laser range scanner. In *SPIE Proceedings of Intelligent Transportation Systems*.

Harless, E., 1860. Die statischen Momente der menschlichen Gliedmassen. *Abhandl Mathematische-Physikalischen Classe Konigl Bayerischen Akad Wissenschaft*, 8(69-96), pp.257–294.

Harris, T., 2002. How CAT Scans Work. *How Stuff Works*. Available at: http://science.howstuffworks.com/cat-scan.htm [Accessed January 20, 2015].

Hatze, H., 1980. A mathematical model for the computational determination of parameter values of anthropomorphic segments. *Journal of biomechanics*, 13(10), pp.833–43. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7462257.

Hatze, H., 1975. A New Method for the Simultaneous Measurement of the Moment of Inertia , the Damping Coefficient and the Location of the Centre of Mass of a Body Segment in situ. *Journal of Applied Physiology*, 34, pp.217–226.

Hencken, C. & White, C., 2006. Anthropometric assessment of Premiership soccer players in relation to playing position. *European Journal of Sport Science*, 6(4), pp.205–211. Available at: http://www.tandfonline.com/doi/abs/10.1080/17461390601012553 [Accessed March 8, 2014].

Henderson, R. & Schulmeister, K., 2003. *Laser Safety*, Taylor and Francis.

Henry, P. et al., 2012. RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments. *The International Journal of Robotics Research*, 31(5), pp.647–663. Available at: http://ijr.sagepub.com/cgi/doi/10.1177/0278364911434148 [Accessed October 5, 2012].

Herrera, D.C., Kannala, J. & Heikkila, J., 2012. Joint Depth and Color Camera Calibration with Distortion Correction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(10), pp.2058–2064.

Hinrichs, R., 1985. Regression Equations to Predict Segmental Moments of Inertia From Anthropometric Measurements. *Journal of Biomechanics*, 18(8), pp.621–624.

Hoare, D.G. & Warr, C.R., 2000. Talent identification and women's soccer: an Australian experience. *Journal of sports sciences*, 18(9), pp.751–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11043900 [Accessed March 18, 2014].

Huber, D. et al., 2010. Using laser scanners for modeling and analysis in architecture, engineering, and construction. In *44th Annual Conference on Information Sciences and*

Systems (CISS). Ieee, pp. 1–6. Available at:
http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5464818.

Hughes, S. & Lau, J., 2008. A technique for fast and accurate measurement of hand volumes using Archimedes principle. *Australasian physical & engineering sciences in medicine / supported by the Australasian College of Physical Scientists in Medicine and the Australasian Association of Physical Sciences in Medicine*, 31(1), pp.56–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18488965.

Hughes, S.W., 2005. Archimedes revisited: a faster, better, cheaper method of accurately measuring the volume of small objects. *Physics Education*, 40(5), pp.468–474. Available at: http://stacks.iop.org/0031-9120/40/i=5/a=008?key=crossref.e317a221e5ac5e5a9a099ee8e403dd3f.

Huijbregts, P. a., 2002. Spinal Motion Palpation: A Review of Reliability Studies. *Journal of Manual & Manipulative Therapy*, 10(1), pp.24–39. Available at: http://www.maneyonline.com/doi/abs/10.1179/106698102792209585.

Intel, 2014. Intel RealSense 3D Camera. Available at: http://www.intel.com/content/www/us/en/architecture-and-technology/realsense-3d-camera.html [Accessed January 25, 2015].

International Standards Office, 2010. *ISO 20685, 3-D Scanning Methodologies for Internationally Compatible Anthropometric Databases*,

Izadi, S. et al., 2011. KinectFusion : Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera. In *UIST Symposium*. Santa, Barbara, CA, pp. 559–568.

Jensen, R., 1978. Estimation of the Biomechanical Properties of Three Body Types Using a Photogrammetric Method. *Journal of Biomechanics*, 11(8-9), pp.349–358.

Kerr, D.A., Ackland, T.R. & Schreiner, A.B., 1995. The Elite Athlete- Assessing body shape, size, proportion and composition. *Asia Pacific Journal of Clinical Nutrition*, 4, pp.25–29.

Khoshelham, K., 2010. Accuracy analysis of Kinect depth data. In D. Lichti & A. Habib, eds. *ISPRS Workshop Laser Scanning*. Calgary, pp. 1–6.

Khoshelham, K. & Elberink, S.O., 2012. Accuracy and resolution of Kinect depth data for indoor mapping applications. *Sensors*, 12(2), pp.1437–54. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3304120&tool=pmcentrez&rendertype=abstract [Accessed October 29, 2012].

Kilner, J., Neophytou, A. & Hilton, A., 2012. 3D Scanning with Multiple Depth Sensors. In *3rd International Conference on 3D Body Scanning Technologies*. Lugano, Switzerland: Hometrica Consulting, pp. 295–301. Available at: http://www.3dbodyscanning.org/cap/abstracts/2012/295kilner.html.

Kingma, I. et al., 1995. Optimizing the Determination of the Body Center of Mass. *Journal of Biomechanics*, 28(9), pp.1137–1142.

Kingma, I. et al., 1996. Segment Inertial Parameter Evaluation in Two Anthropometric Models by Appliction of a Dynamic Linked Segment Model. *Journal of Biomechanics*, 29(5), pp.693–704.

Kirby, R., Price, N. & MacLeod, D., 1987. The Influence of Foot Position on Standing Balance. *Journal of Biomechanics*, 20(4), pp.423–427.

Klug, B., 2010. Taking apart the Kinect. Available at: http://www.anandtech.com/Show/Index/4057?cPage=6&all=False&sort=0&page=2&slug=microsoft-kinect-the-anandtech-review [Accessed October 5, 2011].

Kouzaki, M. & Masani, K., 2008. Reduced postural sway during quiet standing by light touch is due to finger tactile feedback but not mechanical support. *Experimental brain research*, 188(1), pp.153–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18506433 [Accessed March 7, 2014].

Kwon, Y., 1996. Effects of the Method of Body Segment Parameter Estimation on Airborne Angular Momentum. , 12, pp.413–430.

Labelle, K., 2011. *Evaluation of Kinect joint tracking for clinical and in home stroke rehabilitation tools*. Notre Dame, Indiana.

Lackner, J., Rabin, E. & DiZio, P., 2001. Stabilization of posture by precision touch of the index finger with rigid and flexible filaments. *Experimental Brain Research*, 139(4), pp.454–464. Available at: http://link.springer.com/10.1007/s002210100775 [Accessed February 28, 2014].

Lerch, T., MacGillivray, M. & Domina, T., 2006. 3D Laser Scanning: A Model of Multidisciplinary Research. *Journal of Textile and Apparel Technology and Management*, 5(3), pp.1–22.

Leva, P. De, 1996. Adjustments to Zatsiorsky-Seluyanov's Segment Inertia Parameters. *Journal of Biomechanics*, 29(9), pp.1223–1230.

Lindner, M. & Kolb, A., 2006. Lateral and Depth Calibration of PMD-Distance Sensors. In *Second international conference on Advances in Visual Computing*. Lake Tahoe, Navada, pp. 524–533.

Ludbrook, J., 1997. Comparing Methods of Measurement. *Clinical and Experimental Pharmacology & Physiology*, 24, pp.193–203.

Maccormick, J., 2011. *How does the Kinect work?*, Carlisle, United States.

Martin, P.E., Mungiole, M. & Longhill, J.M., 1989. The use of magnetic resonance imaging for measuring segment inertial properties. *Journal of Biomechanics*, 22(4), pp.367–376.

Matyunin, S., Vatolin, D. & Berdnikov, Y., 2011. Temporal Filtering for Depth Maps Generated by Kinect Depth Camera. *IEEE*.

Menna, F. et al., 2011. Geometric investigation of a gaming active device. *Proceedings of the SPIE*, 8085(XI), p.80850G(1)–80850G(15).

Mesa Imaging, 2011. *SwissRanger SR4000*, Zurich, Switzerland.

Microsoft, 2012a. Kinect for Windows. Available at: http://www.microsoft.com/en-us/kinectforwindows/discover/features.aspx [Accessed October 10, 2012].

Microsoft, 2014a. Kinect for Windows Sensor Components and Specifications. Available at: http://msdn.microsoft.com/en-us/library/jj131033.aspx.

Microsoft, 2014b. Kinect Sensor.

Microsoft, 2014c. Kinect Sensor. Available at: http://msdn.microsoft.com/en-gb/library/hh438998.aspx.

Microsoft, 2014d. Kinect v2. Available at: http://www.microsoft.com/en-us/kinectforwindows/meetkinect/features.aspx.

Microsoft, 2012b. Microsoft Kinect SDK 1.6. Available at: http://www.microsoft.com/en-us/kinectforwindows/develop/new.aspx [Accessed October 24, 2013].

Microsoft, 2012c. Near Mode: What it is (and isn't). Available at: http://blogs.msdn.com/b/kinectforwindows/archive/2012/01/20/near-mode-what-it-is-and-isn-t.aspx.

Microsoft Corporation, 2014. Original Kinect for Windows sensor sales to end in 2015. Available at: http://blogs.msdn.com/b/kinectforwindows/archive/2014/12/30/original-kinect-for-windows-sensor-sales-to-end-in-2015.aspx [Accessed January 24, 2015].

Mihelich, P., Accuracy / resolution of Kinect depth data. Available at: http://groups.google.com/group/openkinect/browse_thread/thread/1ff9b4501db3de05/49a578fce0da02ee [Accessed October 17, 2011].

Mirtich, B.V., 1996. *Impulse-based Dynamic Simulation of Rigid Body Systems*. University of California.

Mohamed, H. et al., 2009. Anthropometric and performance measures for the development of a talent detection and identification model in youth handball. *Journal of sports sciences*, 27(3), pp.257–66. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19153859 [Accessed March 4, 2014].

Molnar, B., Toth, C.K. & Detrekoi, A., 2012. Accuracy test of Microsoft Kinect for Human Morphologic Measurements. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIX(September), pp.543–547.

Mungiole, M. & Martin, P.E., 1990. Estimating segment inertial properties: comparison of magnetic resonance imaging with existing methods. *Journal of biomechanics*, 23(10), pp.1039–46. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2229087.

Munro, B., 2000. *Statistical Methods for Health Care Research* 4th ed., Philadelphia: Lippincott Williams and Wilkins.

National Obesity Observatory, 2011. *NOO Data Briefing: Adult Weight,*

Newcombe, R.A. et al., 2011. KinectFusion : Real-Time Dense Surface Mapping and Tracking. In *ISMAR*. Basel, Switzerland, pp. 127–136.

NHS, 2014. What's your BMI. *NHS Choices*. Available at: http://www.nhs.uk/Livewell/loseweight/Pages/BodyMassIndex.aspx#people [Accessed February 3, 2014].

Nigg, B.M., 1999. *Inertial properties of the human or animal body,*

Nikon Corp, 2011. Metris D100 Laser Scanner. Available at: http://www.nikonmetrology.com/en_EU/Products/Laser-Scanning/Handheld-scanning/ModelMaker-MMDx [Accessed September 24, 2012].

Norton, J., Donaldson, N. & Dekker, L., 2002. 3D whole body scanning to determine mass properties of legs. *Journal of biomechanics*, 35(1), pp.81–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11747886.

Norton, K. & Olds, T., 1996. *Anthropometrica*, Sydney, Australia: NewSouth Publishing.

O'Haire, C. & Gibbons, P., 2000. Inter-examiner and intra-examiner agreement for assessing sacroiliac anatomical landmarks using palpation and observation: pilot study. *Manual therapy*, 5(1), pp.13–20. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10688955 [Accessed December 12, 2014].

Occipital, 2014. Structure 3D Sensor. Available at: http://structure.io [Accessed December 3, 2014].

OpenNI, 2011. OpenNI NUI Documentation. Available at: http://openni.org/Documentation/home.html [Accessed September 27, 2012].

Outram, T. et al., 2011. The use of geometric shapes in estimating the geometry of body segments. *Journal of Sports Sciences*, 29(sup2), pp.24–25.

Outram, T., Domone, S. & Wheat, J., 2012. The Reliability of Trunk Segment Inertial Parameters Made From Geometric Models. In *30th Annual Conference of Biomechanics in Sports*. Melbourne, pp. 47–50.

Pataky, T.C., Zatsiorsky, V.M. & Challis, J.H., 2003. A simple method to determine body segment masses in vivo: reliability, accuracy and sensitivity analysis. *Clinical Biomechanics*, 18(4), pp.364–368. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0268003303000159 [Accessed April 30, 2014].

Paul, S.M. et al., 2009. Reliability, validity, and precision of an active stereophotogrammetry system for three-dimensional evaluation of the human torso.

Medical engineering & physics, 31(10), pp.1337–42. Available at:
http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3502070&tool=pmcentrez&
rendertype=abstract [Accessed November 23, 2012].

Paulus, W.M., Straube, a. & Brandt, T., 1984. Visual Stabilization of Posture. *Brain*,
107(4), pp.1143–1163. Available at:
http://brain.oxfordjournals.org/cgi/doi/10.1093/brain/107.4.1143.

Pearsall, D. & Reid, J., 1994. The Study of Human Body Segment Parameters in
Biomechanics: A Historical Review and Current Status Report. *Sports Medicine*, 18(5),
pp.126–140.

Pearsall, D.J. & Costigan, P.A., 1999. The effect of segment parameter error on gait
analysis results. *Gait & Posture*, 9(3), pp.173–83. Available at:
http://www.ncbi.nlm.nih.gov/pubmed/10575078.

Pearsall, D.J., Reid, J.G. & Livingston, L. a, 1996. Segmental inertial parameters of the
human trunk as determined from computed tomography. *Annals of Biomedical
Engineering*, 24(2), pp.198–210. Available at:
http://www.ncbi.nlm.nih.gov/pubmed/8678352.

Perini, T.A. et al., 2005. Technical error of measurement in anthropometry. *Revista
Brasileira de Medicina do Esporte*, 11(1), pp.81–85. Available at:
http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1517-
86922005000100009&lng=pt&nrm=iso&tlng=pt [Accessed September 15, 2014].

Peyer, K.E., Morris, M. & Sellers, W.I., 2014. Subject-specific body segment parameter
estimation using 3D photogrammtery with multiple cameras. *PeerJ PrePrints*, 2(573),
pp.1–20.

Pienaar, a E., Spamer, M.J. & Steyn, H.S., 1998. Identifying and developing rugby
talent among 10-year-old boys: a practical model. *Journal of sports sciences*, 16(8),
pp.691–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10189074 [Accessed
March 11, 2014].

Piovesan, D. et al., 2011. Comparative analysis of methods for estimating arm segment
parameters and joint torques from inverse dynamics. *Journal of Biomechanical
Engineering*, 133(3), p.31003(1)–31003(15). Available at:
http://www.ncbi.nlm.nih.gov/pubmed/21303179 [Accessed January 3, 2013].

PMD Technologies, 2010. *PMD Camcube 3.0*, Siegen, Germany.

Point Grey Research, 2014. Point Grey Bumblebee XB3. Available at:
http://www.ptgrey.com/bumblebee-xb3-1394b-stereo-vision-camera-systems-2.

Primesense, 2011. PrimeSense 3D Camera Solution. Available at:
http://www.primesense.com/en/technology/115-the-primesense-3d-sensing-solution
[Accessed October 5, 2011].

Primesense Ltd, 2012. *The PrimeSense 3D Awareness Sensor*, Tel-Aviv, Israel.

Rao, G. et al., 2006. Influence of body segments' parameters estimation models on inverse dynamics solutions during gait. *Journal of Biomechanics*, 39(8), pp.1531–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15970198 [Accessed December 6, 2012].

Raposo, C., Barreto, J.P. & Nunes, U., 2013. Fast and Accurate Calibration of a Kinect Sensor. In *International Conference on 3D Vision*. IEEE, pp. 342–349. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6599095.

Reichinger, A., 2011. Kinect IR Detailed Analysis. Available at: http://azttm.wordpress.com/tag/kinect/ [Accessed October 18, 2011].

Reid, J.G. & Jensen, R.K., 1990. Human Body Segment Parameters: A Survey and Status Report. *Exercise and Sport Sciences Reviews*, 18, pp.225–241.

Robinson, A. et al., 2012. Improving the Quality of Measurements through the Implementation of Customised Reference Artefacts. In *3D Body Scanning Technologies*. Lugano, Switzerland, pp. 235–246.

Rossi, M. et al., 2013. Body segment inertial parameters of elite swimmers using DXA and indirect methods. , (February), pp.761–775.

Rusinkiewicz, S. & Levoy, M., 2001. Efficient variants of the ICP algorithm. In *3-D Digital Imaging and Modeling*. Quebec: IEEE Comput. Soc, pp. 145–152. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=924423.

Samsung, 2013. Samsung Smart TV : Smart Interaction. Available at: http://www.samsung.com/global/microsite/tv/2013_vi/mobile/html/motion_control.html [Accessed January 24, 2015].

Schranz, N. et al., 2010. Three-dimensional anthropometric analysis: differences between elite Australian rowers and the general population. *Journal of sports sciences*, 28(5), pp.459–69. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20419589 [Accessed March 14, 2014].

Sheets, A.L., Corazza, S. & Andriacchi, T.P., 2010. An automated image-based method of 3D subject-specific body segment parameter estimation for kinetic analyses of rapid movements. *Journal of biomechanical engineering*, 132(1), p.011004(1) – 011004(10). Available at: http://www.ncbi.nlm.nih.gov/pubmed/20524742 [Accessed November 13, 2012].

Shotton, J. et al., 2011. Real-Time Human Pose Recognition in Parts from Single Depth Images. *IEEE CVPR*, 3, pp.1297–1304.

Shpunt, A., 2009. Optical Designs for Zero Order Reduction.

Shpunt, A. & Pesach, B., 2010. Optical Pattern Projection.

Shpunt, A., Rais, D. & Galezer, N., 2010. Reference Image Techniques for Three-Dimensional Scanning.

Shpunt, A. & Zalevsky, Z., 2011. Depth-Varying Light Fields For Three Dimensional Sensing.

Shpunt, A. & Zalevsky, Z., 2009. Three Dimensional Scanning using Speckle patterns.

Shrout, P.E. & Fleiss, J.L., 1979. Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), pp.420–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18839484.

Sizestream, 2014. Size Stream Body Scanner Assembly and Operation Manual. , (November).

Smisek, J., Jancosek, M. & Pajdla, T., 2011. 3D with Kinect. In *IEEE Workshop on Consumer Depth Cameras for Computer Vision*. Barcelona, Spain: IEEE.

Smith, T.B. & Hopkins, W.G., 2011. Variability and predictability of finals times of elite rowers. *Medicine and science in sports and exercise*, 43(11), pp.2155–60. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21502896 [Accessed December 12, 2014].

Soft Kinetic, 2011. Soft Kinetic TOF 3D Cameras. Available at: http://www.softkinetic.com/Solutions/DepthSensecameras.aspx [Accessed October 13, 2011].

Soudarissanane, S. et al., 2011. Scanning geometry: Influencing factor on the quality of terrestrial laser scanning points. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(4), pp.389–399. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0924271611000098 [Accessed October 20, 2014].

Stampfli, P. et al., 2012. SizeITALY - The Actual Italian Measurement Survey. In *3D Body Scanning Technologies*. Lugano, Switzerland, pp. 261–268.

Stančić, I., Musić, J. & Zanchi, V., 2013. Improved structured light 3D scanner with application to anthropometric parameter estimation. *Measurement*, 46(1), pp.716–726. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0263224112003594 [Accessed May 2, 2014].

Stern, B., IR output from the Kinect. Available at: http://blog.makezine.com/archive/2010/11/how-to-modded-camera-looks-at-kinec.html [Accessed October 5, 2011].

Stewart, A. et al., 2011. *International Standards for Anthropometric Assessment*, International Society for the Advancement of Kinanthropometry.

Stewart, A. & Sutton, L., 2012. *Body Composition in Sport, Exercise and Health* 1st ed. A. Stewart & L. Sutton, eds., Oxon, UK: Routledge.

TC2, 2011. *TC2 KX-16 3D Body Scanner*, Cary, USA.

The Association of German Engineers (VDI), 2008. *Optical 3D measuring systems - Multiple view systems based on area scanning*, Germany: VDI/VDE 2634 Part 3.

De Vet, H.C.W. et al., 2006. When to use agreement versus reliability measures. *Journal of clinical epidemiology*, 59(10), pp.1033–1039. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16980142 [Accessed May 26, 2014].

Vitronic, 2011. *Vitus 3D Body Scanner*, Germany.

Voosen, K., 2011. 3D Vision and the Kinect. In *NIDays Graphical System Design Conference*. London: kkkk, pp. 1–41.

Vuillerme, N. & Nafati, G., 2007. How attentional focus on body sway affects postural control during quiet standing. *Psychological research*, 71(2), pp.192–200. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16215747 [Accessed March 5, 2014].

Waldhausl, P. & Ogleby, C., 1994. 3x3 Rules for Simple Photogrammetric Documentation of Architecture. *International Archives of Photogrammetry and Remote Sensing*, XXX(5), pp.426–429.

Wan, Y., Hu, G. & Dong, G., 2013. Anthropometric techniques based on kinect 3D scanning. *Proceedings 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC)*, pp.1292–1296. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6885266.

Weiss, A., Hirshberg, D. & Black, M.J., 2011. Home 3D Body Scans from Noisy Image and Range Data. In *13th International Conference on Computer Vision*. Barcelona, Spain.

Werghi, N., 2007. Segmentation and Modeling of Full Human Body Shape From 3-D Scan Data: A Survey. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, 37(6), pp.1122–1136. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4344000.

Wheat, J.S. et al., 2011. Obtaining body segment inertia parameters using structured light scanning with Microsoft Kinect. *Journal of Sports Sciences*, 29(sup2), pp.S1–S132.

Wicke, J., Dumas, G. a & Costigan, P. a, 2008. Trunk density profile estimates from dual X-ray absorptiometry. *Journal of Biomechanics*, 41(4), pp.861–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18082166 [Accessed July 16, 2012].

Wicke, J. & Dumas, G.A., 2010. Influence of the Volume and Density Functions Within Geometric Models for Estimating Trunk Inertial Parameters. *Journal of Applied Biomechanics*, 26, pp.26–31.

Williams, a M. & Reilly, T., 2000. Talent identification and development in soccer. *Journal of sports sciences*, 18(9), pp.657–67. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11043892 [Accessed January 27, 2014].

WillowGarage, 2012. Open CV Documentation. Available at: http://docs.opencv.org/ [Accessed October 20, 2012].

Wrede, R., 2010. *Schaum's Outline of Advanced Calculus* 3rd ed., New York: Schaum's Outlines.

Yamazoe, H. et al., 2012. Easy Depth Sensor Calibration. In *21st International Conference on Pattern Recognition*. Tsukuba, Japan, pp. 465–468.

Yeadon, M.R., 1990a. The simulation of aerial movement--I. The determination of orientation angles from film data. *Journal of Biomechanics*, 23(1), pp.59–66. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2307692.

Yeadon, M.R., 1990b. The simulation of aerial movement--II. A mathematical inertia model of the human body. *Journal of Biomechanics*, 23(1), pp.67–74. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2307693.

Zatsiorsky, V., 1983. The Mass and Inertia Characteristics of the Main Segments of the Human Body. *Biomechanics*, VIII(B), pp.1152–1159.

Zatsiorsky, V.N. & Seluyanov, V.N., 1985. Estimation of the mass and inertia characteristics of the human body by means of the best predictive regression equations. *Biomechanics*, 9(B), pp.233–239.

Zhang, Z., 2000. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11), pp.1330–1334. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=888718.

RESEARCH ETHICS CHECKLIST (SHUREC1)

This form is designed to help staff and students to complete an ethical scrutiny of proposed research. The SHU Research Ethics Policy should be consulted before completing the form.

Answering the questions below will help you decide whether your research proposal requires ethical review by a Faculty Research Ethics Committee (FREC). In cases of uncertainty, members of the FREC can be approached for advice.

Please note: staff based in University central departments should submit to the University Ethics Committee (SHUREC) for review and advice.

The final responsibility for ensuring that ethical research practices are followed rests with the supervisor for student research and with the principal investigator for staff research projects.

Note that students and staff are responsible for making suitable arrangements for keeping data secure and, if relevant, for keeping the identity of participants anonymous. They are also responsible for following SHU guidelines about data encryption.

The form also enables the University and Faculty to keep a record confirming that research conducted has been subjected to ethical scrutiny.

– For student projects, the form may be completed by the student and the supervisor and/or module leader (as applicable). In all cases, it should be counter-signed by the supervisor and/or module leader, and kept as a record showing that ethical scrutiny has occurred. Students should retain a copy for inclusion in their research projects, and staff should keep a copy in the student file.

– For staff research, the form should be completed and kept by the principal investigator.

Please note if it may be necessary to conduct a health and safety risk assessment for the proposed research. Further information can be obtained from the Faculty Safety Co-ordinator.

General Details

(Table cells will expand as you type)

Name of principal investigator or student	Sean Clarkson
email address	s.clarkson@shu.ac.uk
Course or qualification (student)	PhD
Name of supervisor (if applicable)	Simon Choppin
email address	s.choppin@shu.ac.uk
Title of research proposal	An investigation into the agreement of first order body segment parameter estimates obtained from a 3D body scanning system and a geometric modelling technique.
Brief outline of research to include, rationale & aims (50 words). In addition for research with human, participants,	Many biomechanical and medical analyses rely on the availability of reliable body segment parameter (BSP) estimates. However, existing techniques have been criticised for their inherently generic nature, and the time required to complete the

| include recruitment method, participant details & proposed methodology (250 words). | necessary measurements. Body scanning systems such as DEXA, MRI and CT have previously been used to obtain subject person 3D scans, from which BSIPs were calculated. However, their cost prevents use within an ecologically valid laboratory environment. Recently, Wicke and Dumas (2010) suggested structured light scanning would be a suitable lower cost alternative.

The recent interest in natural user interaction (NUI) has led to the development of low cost (in the region of £200) sensors, typically using a combination of structured light and computer vision techniques, and able to capture 3D scans at a rate of 30Hz.

A non invasive body scanning system comprising four Microsoft Kinect NUI sensors has been developed. The system is able to collect 3D body scans in <2 seconds, and capable of calculating subject specific first order BSPs.

The aim of this study is to compare the scanning system to currently accepted geometric BSP estimation techniques to determine agreement. If the two methods agree with one another, then the newly developed system could be used within an ecologically valid laboratory environment: potentially offering significant advantages in terms of time and reliability.

Male participants aged 18+ will be recruited for the study. There are no other inclusion criteria, other than participants must be able to stand upright and unaided for periods of up to 20 minutes to allow the necessary body palpation and measurement.

Participants will be recruited via direct communication with friends, colleagues, and communication with SHU sports teams. A letter of collaboration to this effect is included as an attachment to this document.

Each data collection will take place on a single day, in a session lasting 25-45 minutes. Participants will be required to wear close fitting lycra shorts, and no top. Medical screens and blinds will be used within the room to provide appropriate privacy. Upon arrival, the height and weight of each participant will be recorded to allow characterisation based upon body mass index (BMI). Each participant will be palpated once, and measured 3 times by a female ISAK (International Society for the Advancement of Kinanthropometry) trained examiner. The participants will then be scanned three times in two different positions, with a one minute break between scans to allow the participant to leave the scanning booth and relax their posture.

After collection, the 3D scans will be digitised to segment the scan into the area of interest and allow calculation of the relevant BSPs and anthropometric measurements. Statistical analysis, such as the methods by Bland Altman (1986) shall be used to determine the agreement between the BSPs calculated from the 3D scans and the geometric models. A subset of the anthropometric data - collected in accordance with ISAK |

	procedures - shall be sent to ISAK as part of an accreditation process. The 3D scan data may also be used as part of additional preliminary investigations to validate algorithms and methods. Participants will not be individually identifiable in any of the data sets.
Will the research be conducted with partners & subcontractors?	**Yes/No** No (If **YES**, outline how you will ensure that their ethical policies are consistent with university policy.)

1. Research Involving the NHS or Social Care / Community Care

Question	Yes/No
1. Does the research involve the NHS or Social Care/Community Care (SC) as defined below? • Patients recruited because of their past or present use of the NHS or SC • Relatives/carers of patients recruited because of their past or present use of the NHS or SC • Access to data, organs or other bodily material of past or present NHS patients • Foetal material and IVF involving NHS patients • The recently dead in NHS premises • Prisoners recruited for health-related research • Participants who are unable to provide informed consent due to their incapacity	No
2. Is this a research project as opposed to service evaluation or audit? *For NHS definitions please see the following website* http://www.nres.nhs.uk/applications/is-your-project-research/	N/A

If you have answered **YES** to questions **1 & 2** then you **must** seek approval from the NHS or Social Care under their Research Governance schemes.

NHS https://www.myresearchproject.org.uk/Signin.aspx

If you are undertaking Social Care research in Sheffield you will require a favourable ethical review from a Faculty Committee but must use the Sheffield Council form for this. Full details from http://www.sheffield.gov.uk/caresupport/us/research. For other areas contact the relevant social services department directly for advice on procedures.

NB FRECs provide Independent Scientific Review for NHS or SC research and initial scrutiny for ethics applications as required for university sponsorship of the research. Applicants can use the NHS or SC proforma and submit this initially to the FREC.

2. Research with Human Participants

Question	Yes/No
1. Does the research involve human participants? This includes surveys, questionnaires, observing behaviour etc. *Note If YES, then please answer questions 2 to 10*	Yes

Question	Yes/No
If NO, please go to Section 3	
2. Will any of the participants be vulnerable? **Note** *'Vulnerable' people include young people under 18, people with learning disabilities, people who may be limited by age or sickness or disability from understanding the research, etc.*	No
3 Are drugs, placebos or other substances (e.g. food substances, vitamins) to be administered to the study participants or will the study involve invasive, intrusive or potentially harmful procedures of any kind?	No
4 Will tissue samples (including blood) be obtained from participants?	No
5 Is pain or more than mild discomfort likely to result from the study?	No
6 Will the study involve prolonged or repetitive testing?	No
7 Is there any reasonable and foreseeable risk of physical or emotional harm to any of the participants? **Note** *Harm may be caused by distressing or intrusive interview questions, uncomfortable procedures involving the participant, invasion of privacy, topics relating to highly personal information, topics relating to illegal activity, etc.*	No
8 Will anyone be taking part without giving their informed consent?	No
9 Is it covert research? **Note** *'Covert research' refers to research that is conducted without the knowledge of participants.*	No
10 Will the research output allow identification of any individual who has not given their express consent to be identified?	No

If you answered **YES only** to question 1, you **must** submit the signed form to the FREC for registration and scrutiny by the Chair. If you have answered **YES** to any of the other questions you are **required** to submit a SHUREC2A (or 2B) to the FREC.

3. Research in Organisations

Question	Yes/No
1 Will the research involve working with/within an organisation (e.g. school, business, charity, museum, government department, international agency, etc)?	No
2 If you answered YES to question 1, do you have granted access to conduct the research? *If YES, students please show evidence to your supervisor. PI should retain safely.*	N/A
3 If you answered NO to question 2, is it because: A. you have not yet asked B. you have asked and not yet received an answer C. you have asked and been refused access. **Note** *You will only be able to start the research when you have been granted access.*	N/A

4. Research with Products and Artefacts

Question		Yes/No
1.	Will the research involve working with copyrighted documents, films, broadcasts, photographs, artworks, designs, products, programmes, databases, networks, processes or secure data?	No
2.	If you answered YES to question 1, are the materials you intend to use in the public domain?	N/A

Notes	*'In the public domain' does not mean the same thing as 'publicly accessible'.*
	
	  − *Information which is 'in the public domain' is no longer protected by copyright (i.e. copyright has either expired or been waived) and can be used without permission.*
	  − *Information which is 'publicly accessible' (e.g. TV broadcasts, websites, artworks, newspapers) is available for anyone to consult/view. It is still protected by copyright even if there is no copyright notice. In UK law, copyright protection is automatic and does not require a copyright statement, although it is always good practice to provide one. It is necessary to check the terms and conditions of use to find out exactly how the material may be reused etc.*
	
	If you answered YES to question 1, be aware that you may need to consider other ethics codes. For example, when conducting Internet research, consult the code of the Association of Internet Researchers; for educational research, consult the Code of Ethics of the British Educational Research Association.	

3.	If you answered NO to question 2, do you have explicit permission to use these materials as data? *If YES, please show evidence to your supervisor. PI should retain permission.*	N/A
4.	If you answered NO to question 3, is it because: A. you have not yet asked permission B. you have asked and not yet received and answer C. you have asked and been refused access.	**A/B/C** N/A
Note	*You will only be able to start the research when you have been granted permission to use the specified material.*	

Adherence to SHU policy and procedures

Personal statement	
I can confirm that: – I have read the Sheffield Hallam University Research Ethics Policy and Procedures – I agree to abide by its principles.	
Student / Researcher/ Principal Investigator (as applicable)	
Name: Sean Clarkson	Date: 04/03/2014
Signature:	
Supervisor or other person giving ethical sign-off	
I can confirm that completion of this form has not identified the need for ethical approval by the FREC or an NHS, Social Care or other external REC. The research will not commence until any approvals required under Sections 3 & 4 have been received.	
Name: Simon Choppin	Date: 04/03/2014
Signature:	

Sheffield Hallam University

Faculty of Health and Wellbeing Research Ethics Committee
Sport and Exercise Research Ethics Review Group

Risk Assessment Pro Forma

****Please ensure that you read the accompanying
Risk Assessment Risk Ranking document before completing this form****

Title of research	An investigation into the agreement of first order body segment parameter estimates obtained from a 3D body scanning system and a geometric modelling technique.

Date Assessed	06/03/2014

Assessed by (Principal Investigator)	Sean Clarkson

Signed	Position
	Principal Investigator

Activity	Risks	Control Measures
Put in this box the activity which may cause harm.	**Risk** of [*place in here the harm that may be caused*] caused by [*put in the hazard (source of danger) here*]. Risk = consequence x likelihood. Identify risk category **Low Medium or High**	Place here what you would do to minimise the risk
Risk of slips and trips over wires connecting together the system components	2 x 1 = 2 (Low)	Ensure wires are appropriately secured, and not left trailing. Any wires crossing areas where participants walk should be secured to the floor and clearly marked with warning tape.
Discomfort from standing still for long periods of time whilst being palpated and measured	1 x 1 = 1 (Low)	Ensure participants are aware they can take rest or comfort breaks whenever they feel necessary.
Emotional harm from the participants seeing their body shape depicted in the 3D scan	2 x 1 = 2 (Low)	Ensure the screen operating the scanning system is occluded from view, and only show participants the 3D scans should they ask.

Risk Evaluation (Overall)

Low

General Control Measures

Is a pre-screen medical questionnaire required? Yes [ ]   No [x]

Emergency Procedures

Follow general university emergency procedures, and call ex 888 via an internal phone to request assistance from the emergency services if required.

Monitoring Procedures

Review Period	1 year

Reviewed By (Supervisor)	Date

Hallamunion

your **union.** your **choice**

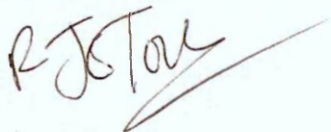|		
Paternoster Row	**F:**	**0114 225 4251**
Sheffield	**E:**	**info@hallamunion.org**
S1 2QQ	**W:**	**www.hallamunion.org**

Direct Line: (0114) 225 3450
Email: r.j.otoole@shu.ac.uk
Fax: 0114 225 4251

Re: Letter of collaboration

Further to your correspondence, I am happy to confirm you have permission to work with the participants that you have highlighted in relation to research for your dissertation at Sheffield Hallam University.

May I take this opportunity to wish you well for the successful completion of your study.

Yours Sincerely,

Rick O'Toole
Sports Union Manager
Sheffield Hallam Students' Union