Calculating Body Segment Inertia Parameters from a Single Rapid Scan Using the Microsoft Kinect

Sean CLARKSON*, Simon CHOPPIN, John HART, Ben HELLER, Jon WHEAT
Centre for Sports Engineering Research, Sheffield Hallam University, Sheffield, UK

Abstract

Many biomechanical analyses rely on the availability of reliable body segment inertia parameter (BSIP) estimates. Current processes to obtain these estimates involve many time consuming manual measurements of the human body, used in conjunction with models or equations. While such methods have become the accepted standard they contain many inherent errors arising from manual measurement and significant assumptions made in the underlying data used to form the models and equations. Presented here is an alternative approach to obtaining reliable estimates of body segment inertia parameters through the use of the Microsoft Kinect sensor. A 3D scanning system was developed, comprising four Kinects aligned to a single global coordinate system using rigid body calibration and random sample consensus (RANSAC) optimisation. The system offers the advantage of obtaining BSIP estimates in a single scanning operation of around three seconds, much quicker than the circa thirty minutes of manual measurements required for existing BSIP estimation methods. The results obtained with the system show a mean error of 0.04% and a standard deviation of 2.11% in volumetric measurements of a torso manikin, suggesting comparable and in many cases, greater accuracy volumetric estimates than a commonly used geometric BSIP model. Further work is needed to extend this study to include a full range of BSIP measurements across more of the bodies segments and to include scanning of living human subjects. However, this initial study suggests great potential for a low cost system that can provide quick and accurate subject specific BSIP estimates.

Keywords: Anthropometrics, body measurement, 3D body scanning, body segment inertia parameters, rigid transformation, geometric models, Kinect, RANSAC

1. Introduction

Within the field of biomechanics, body segment inertia parameters (BSIPs) are a very important measure due to their use in many biomechanical analyses. BSIPs commonly of interest are; mass, centre of mass and principal moments of inertia. There currently exists a number of methods that can be used to obtain such parameters, with common methods being those based on cadaver studies [1], mass scanning methods [2] and geometric models [3].

Early attempts to obtain BSIP estimates relate as far back as 1860 [4], with the most significant advancement being the work by Dempster [1]. Dempster's study collected measurement data from eight complete cadavers by dividing up the body into key segments to allow the calculation of segmental volume, mass, centre of gravity and moment of inertia. With this information, Dempster [1] created tables that enable segment mass to be calculated from total body mass and centre of mass and moment of inertia calculated from segment length. A number of questions regarding the validity of the work by Dempster have arisen in the years following due to the small sample size, meaning the model produced by Dempster may not be a true depiction of the population. Similarly, the model may be invalid when applied to female subjects or different ages.

Another approach for obtaining BSIP estimates adopts a medical imaging approach, using scanners such as MRI or CT to determine the mass and volume of the bodies segments, from which BSIPs can be estimated. Commonly accepted work in this field is by Zatsiorsky [2] who used a sample of 100 male and 15 female living subjects. The subjects were scanned using gamma mass scanning to estimate mass, centre of mass and principal moments of inertia for 15 body segments. Importantly, the subjects used by Zatsiorsky were from mixed genders and included subjects younger than the cadavers used in studies by Dempster et al. Zatsiorsky [2] subsequently created regression equations to allow the model to be used by other people in their own other studies.

*s.clarkson@shu.ac.uk; 0114 225 5590;
www.shu.ac.uk/research/cser; www.engineeringsport.co.uk; www.depthbiomechanics.co.uk

Whilst the approaches adopted by Dempster [1] and Zatsiorsky [2] have their own unique advantages, as highlighted by [5], many of the subjects used in the studies were heavily biased towards a specific gender or population type, possibly leading to errors when the models are applied to subjects from different populations.

The work by Yeadon [3] and [6] was one of the first studies to recognise this problem and highlighted the need for a method to obtain accurate subject specific BSIP measurements. To address this issue, Yeadon devised a geometric modelling approach which has become one of the most commonly used approaches to estimating BSIPs.

The geometric model produced by Yeadon [6] splits the human body into 40 segments, assuming a uniform density in each segment. Each segment is defined as a stadium solid from which volume and hence mass can be estimated. Stadium solids are defined by segment end point measurements of diameter, width and height, manually taken from the body using anatomical callipers and anthropometric measuring tapes.

Geometric methods of obtaining BSIPs can be very time consuming, for example, all 95 measurements required for Yeadon's model can take around 40 minutes of the subjects time. Methods based on regression equations and mathematical models require fewer, simpler measurements, but as highlighted by Durkin and Dowling [7], such models are highly simplified representations of the human body so will always be subject to a degree of error.

A study by Durkin and Dowling [7] investigated the accuracy of commonly used BSIP models by comparing the measurements with those from dual energy x-ray scans (DEXA). Table 1 highlights the percentage difference in mass of the thigh segment between the models and DEXA scans. The results below are an excerpt from the paper by Durkin and Dowling [7] and relate to their sample group of 19-30 year old males.

Table 1. Percentage differences in common BSIP models when compared to DEXA scans.

	Durkin [1]	Winter	Zatsiorsky (Regression) [2]	Zatsiorsky (Geometric) [2]	Hanavan
% Error in Mass of Thigh	7.91	21.30	16.08	8.51	21.21

The results suggest considerable differences when compared to the DEXA scan, but also a considerable inter-model variance. As would be expected, this percentage error varies significantly depending upon what segment was measured, but on a large segment like the thigh the difference is highly noticeable. A recent study by Outram et al [8] conducted a similar investigation, but focussed on assessing the accuracy of Yeadon's geometric model. The study compared the shank and trunk segment mass estimates of three healthy males obtained with the Yeadon model to a high accuracy laser scan. Again the results showed different errors depending upon the segment, with a range of +3.2% to -12.8% for the abdomen and upper arm segments respectively.

To improve on the inherent inaccuracies in the methods discussed above, attention is turning to methods of obtaining accurate subject specific BSIP measurements using 3D scanning approaches [9]. By offering subject specific measurements, such methods are able to overcome the shortfalls in existing BSIP models discussed above and provide BSIP estimates to a greater degree of accuracy.

Conventional 3D scanning systems such as laser scanners would appear to offer the best level of accuracy, their use is however inhibited by their expense and the time taken to produce a scan as errors may be introduced in the 3D model due to breathing and other involuntary movements of the human body [10].

Full body scanners are rapidly emerging, offering rapid 3D scans of the full human body and utilising existing technology such as electromagnetic wave and laser scanning to name but two. Whilst overcoming the problems of movement, the use of such systems is again inhibited by their expense and required space [10].

As highlighted by Wicke and Dumas [9], structured light scanners would seem suitable for this purpose as they are able to produce very quick 3D profiles of an object. However, such systems also currently have a low adoption rate due to their expense and complexity.

The recent introduction of the Microsoft Kinect games controller and associated software developer tools provides the potential for 3D scanning at a low price point, achieved with the integral pseudo structured light technology designed to track the motion of a person for animation of an avatar within a virtual gaming environment. The low price point, coupled with the ability to offer full 3D scans at a rate of 30 scans a second means the Kinect is able to overcome the shortfalls apparent in existing 3D scanning systems highlighted above [11]. Much work has been conducted in the field of 3D scanning with the Kinect, including notable papers who focus on scanning full human bodies and indoor environments respectively [10] and [12].

Given the original application of the Kinect sensor and its low price point the accuracy of the 3D data returned by the Kinect is not of paramount importance, posing potential problems for its use in this application. Work by Khoshelham [11] has however shown the Kinect to provide mean and standard deviation errors in 3D data of around 4mm and 10mm respectively at a distance of 0.5m from the object to be scanned when compared to a laser scan. Khoshelham [11] also reports the standard deviation to increase as much as 45mm when around 5m from the Kinect sensor.

Therefore, providing the subject to be scanned is able to get close enough to the Kinect sensor to achieve the smaller error values reported by Khoshelham [11] then it would seem that coupled with relevant data processing, the Kinect would be ideally suited to providing low cost, quick and accurate 3D scans of the human body, from which subject specific BSIPs can be estimated.

The purpose of the investigation was therefore to determine the Kinects suitability as a 3D scanner to obtain subject specific BSIP estimates by comparing the Kinects volumetric BSIP estimates of a torso manikin to those obtained with a high accuracy laser scanner and Yeadon's commonly accepted geometric model [6].

2. Method

2.1. The Microsoft Kinect

Custom software was written to access to the Microsoft Kinect device using the application programming interface (API) by OpenNI. The software allows access to the Kinects IR camera and depth feed to allow calibration and capture of 3D scanning data (Figure 1).
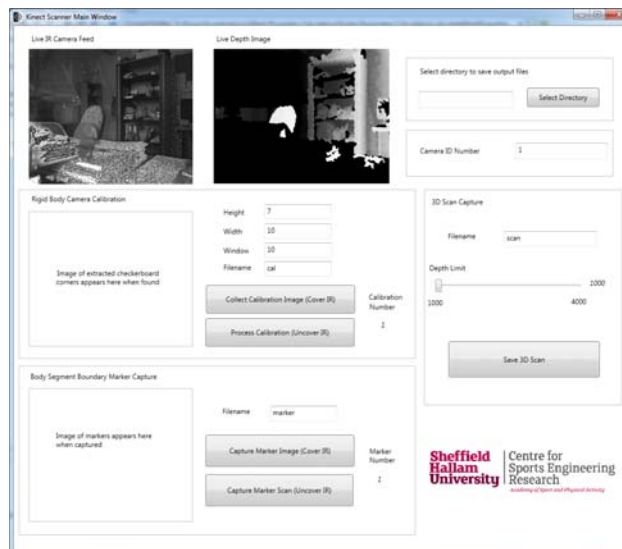

Fig 1. Software application to access the Microsoft Kinect.

2.2. Camera Calibration

Immediately apparent is that the field of view of a single stationary Kinect is unable to observe a full 360° view of the human body in one scanning operation. This issue has been apparent in existing studies with the Kinect, mainly focussing on producing 3D scans of large scale indoor environments. Efficient solutions have used one Kinect sensor which is moved around the area to be scanned [13] and [14]. As the Kinect moves only a small amount between successive frames it is possible to align the scans using algorithms such as iterative closest point (ICP) [12]. This process would require the subject to stand still for extended periods of time whilst the Kinect is moved around and as a primary objective of this project was to minimise the scanning duration to reduce the possible effects of breathing and involuntary movements, this approach would not seem viable.

An alternative approach is therefore proposed, based on multiple Kinect sensors arranged in a configuration so they are able to obtain a full 360^0 view of the subject to be scanned. Initial investigations showed that four Kinect sensors positioned diagonally opposite from one another provided the optimum compromise between the number of Kinects and the field of view. To allow a subject standing upright to be scanned, this equated to a scanning area of approximately $2m^2$.

Due to the arrangement of the Kinect sensors and hence their very different fields of view, their scans have dramatic differences. As these scans have a limited number of corresponding points it prohibits the use of conventional 3D stitching algorithms such as ICP. Extrinsic camera calibration is therefore required to determine the real world position of the Kinects in relation to one another in order that the 3D scans obtained from each Kinect can be aligned with a single global co-ordinate system.

The extrinsic calibration procedure proposed for use with this method is a novel approach, inspired by existing extrinsic camera calibration routines. [15]

The case of initial calibration is a relatively simple problem as the Kinects and object to be scanned remain stationary at all times, with the only difference between each Kinect being their point of view. Therefore, the co-ordinate systems between neighbouring Kinects are simply linked by a single rotation and translation which can be used in conjunction with a 3D rigid body transformation to align the scans with a common coordinate system.

Firstly, consider a rigid transformation between two points in non-homogenous coordinates.

$$p' = Rp + t$$

Where p' and p are 3D points, R is a 3x3 rotation matrix and t is a 3x1 translation vector

Therefore, given three corresponding 3D points obtained from two point clouds it is easy to calculate the rigid body transformation between them by calculating the rotation and then translation. With this knowledge, the transformation between camera pairs can be obtained and subsequently applied to the neighbouring point cloud to align it with the common coordinate system.

To calculate the rotation the mean position is first subtracted from each point cloud to make the rotation independent of the translation and to thereby simplify the problem.

$$m = \frac{1}{N}\sum\nolimits_{j=1}^{N} P_j$$

$$\overline{P_J} = P_j - m$$

Where m is the mean position and N is the number of points in the point cloud P_j

The resulting points are now only linked by a rotation. The points from each point cloud are then concatenated together in a 3 x N matrix with each row containing a single 3D point.

$$A = \overline{P}(\overline{P}')^T$$

Where A is the product of multiplication and $\overline{P}$ and $\overline{P}'$ are the concatenated points

The rotation can then be extracted from the matrix A by decomposing it with the use of singular value decomposition (SVD).

$$UDV^T = A$$

Where U, D and V are the products of the SVD operation, A is the product of multiplication from above

The rotation can then be obtained as below.

$$R = VU^T$$

Where V and U are products of the SVD operation, R is the rotation matrix

With the rotation known it is now easy to calculate the translation as below.

$$t = m' - Rm$$

Where t is the translation matrix, m and m' are the point cloud mean positions and R is the rotation matrix

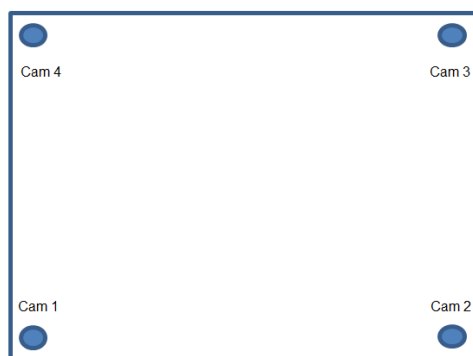To facilitate the scanning the Kinect sensors were laid out as shown in figure 2.

Fig 2. Layout of the Kinect sensors for the scanning system.

Kinect number 1 was used as the reference Kinect and therefore the coordinate system with which to align all other individual Kinect coordinate systems. Calibration is therefore required between Kinect 4 and 1, 2 and 1 and 3 and 2. Kinect number 3 also requires the transformation matrix from Kinect 2 to Kinect 1 to be applied to align its coordinate system with Kinect 1's reference coordinate system.

It is important to note that conventional intrinsic camera calibration is not required in this system as only the Kinects infra-red (IR) camera is used and investigations have shown this camera to be calibrated for intrinsic correction at manufacture.

2.3. Finding Corresponding 3D Point Pairs

The initial calibration procedure requires corresponding 3D points from each neighbouring Kinect, something which is not a trivial task. Many approaches have been considered, including finding the centre of spheres [16] and finding the corners of a planar board [17].

Conventional stereo camera calibration often makes use of a planar black and white checkerboard, from which the location of the checkerboard corners are found in each cameras frame. By knowing the 2D location of these points the cameras relative pose to one another can be determined to allow the 3D location of points within the camera pair's field of view to be obtained. [15]

The method adopted for finding 3D point pairs in the proposed system takes a similar form. A planar checkerboard is positioned in front of the Kinect pair and the open source Open CV checkerboard corner finding algorithm used to find the 2D projective coordinates of the checkerboard corners in the infra-red (IR) image from each Kinect (figure 3).
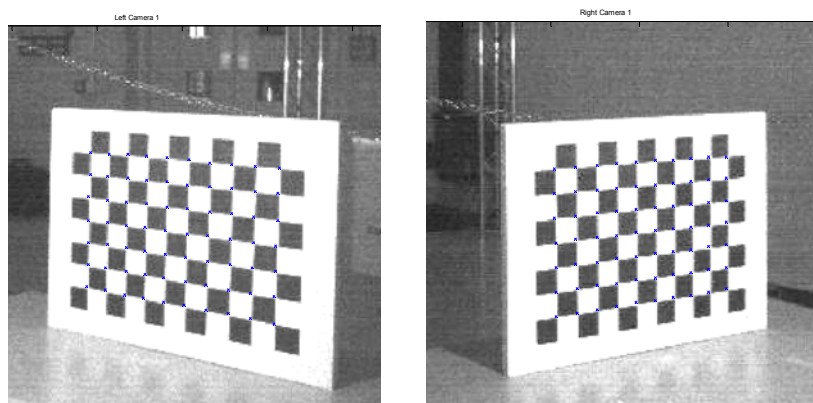

Fig 3. Checkerboard corners found from the IR cameras of two Kinects.

A 3D scan of the checkerboard is then taken by the Kinect which is later used to find the real world 3D coordinates of each checkerboard corner. As the Kinects 640x480 3D depth map directly relates to its 640x480 IR camera image it is a relatively simple process to convert from 2D projective coordinates to real world 3D coordinates.

A 10x7 planar checkerboard was used for calibration, giving 70 corner locations and hence 70 corresponding 3D points from one image capture that can be used for calibration.

It is important to note the IR projector of the Kinect must be covered up when using the Open CV checkerboard corner finding algorithm as otherwise the speckle pattern generated by the Kinect obscures the checkerboard corners, preventing the algorithm from working correctly. The scene observed by the Kinect must therefore be illuminated with an external infra-red light source. However, when performing the 3D scan of the checkerboard the IR projector must be uncovered and the external light source turned off to improve performance.

2.4. Calibration Optimisation

Due to the nature of the Kinect the returned depth information can be very noisy which may result in incorrect depth (Z) components of 3D coordinates. This poses a potential problem when trying to find the 3D location of a checkerboard corner using a single 2D pixel as the depth calculated by the Kinect at that point may be incorrect.

For this reason the proposed system adopts a depth windowing and averaging feature to improve upon the quality of the depth information. A search window decreasing in size from 15 pixels to 0 pixels in 1 pixel increments is employed to determine the Z value over the region surrounding the checkerboard corner. This concept is shown in figure 4 for a 4 pixel search window, with the central cross defining the checkerboard corner as found using the OpenCV algorithm and the outer four crosses defining the outer corners of the search square.
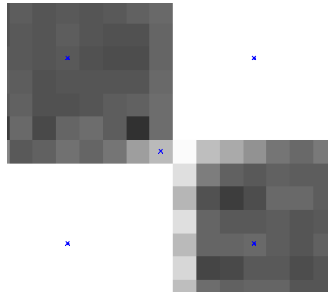
Fig 4. 4 Pixel Z component checkerboard corner search window.

The returned depth values across the search window are averaged and then used as the Z value for the real world checkerboard corner location.

This results in 16 sets of 140 point pairs (16 window sizes, 70 points per checkerboard and 2 Kinects) which may be used for the rigid body calibration procedure. With so many point pairs it is likely that some points will have errors associated with them which must be excluded from the calibration process. Furthermore, the optimum window size for each checkerboard position must be determined. To perform these computations the point pairs are input to a random sample consensus (RANSAC) style optimization routine to pick the best points to use for calibration, based upon the point pairs and window sizes which return the lowest root mean square (RMS) inter-camera calibration error.

Such rigid body algorithms require a minimum of 3 points to work effectively, with an optimum number being around 8 point pairs and little noticeable difference further increasing the number of point pairs up to around 35 points [18]. With this in mind and to improve the speed of the calibration the proposed RANSAC algorithm has a condition stating that point pair searches should include a minimum number of 8 point pairs.

Once the best point pairs have been picked the initial calibration routine is executed to obtain the transformation matrices for each Kinect pair. The transformation matrices obtained from this initial calibration procedure are later used to align the point cloud scans from each Kinect with the single global coordinate system.

2.5 Landmark Identification and Segmentation

3D scans obtained with the proposed system must be split up to allow BSIP's to be calculated for each of the bodies segments. The segment end points adopted for use by the proposed system are those defined by Yeadon [6].

Work has been done in this area to develop automatic anatomical landmark finding algorithms [19]. However, the physical condition of some subjects may mean the anatomical landmarks are hard to find and therefore a fully automatic system may lead to errors in determining the position of the landmarks.

To ensure the system works for all possible physical conditions of subjects and to obtain the greatest level of accuracy possible, the approach adopted by the proposed system adopts a conventional approach, involving the anatomical landmarks being manually located by an assistant and marked with black self-adhesive markers. This process, commonly known as palpation, defines the segment end points and hence where the 3D scan will be split up.

Figure 5 shows typical anatomical landmarks defined by Yeadon [6] that mark the trunk segment end points of stadium solids s2 and s3 as the umbilicus and nipples respectively.

Fig 5. Trunk manikin palpated at anatomical landmarks using black self-adhesive markers.

In a similar approach to the initial calibration procedure, the black circular markers can be clearly observed in the IR image. The Hough circle finding algorithm is used to locate the markers in the image and to find the 2D projective coordinates of their centre which can later be easily converted into 3D coordinates.

Although not apparent in figure 5, the upper end point of the trunk segment is defined with 4 markers and the lower end point defined with 3 markers. The placement of these markers is likely to have a degree of error due to asymmetry of the human body and human error in placement. Therefore the mean horizontal component of each segment end point is taken. These mean values are defined as horizontal planes in which to segment the point cloud scan, subsequently resulting in a point cloud of a given segment. Figure 6 shows a point cloud scan of the two stadium solids in figure 5.
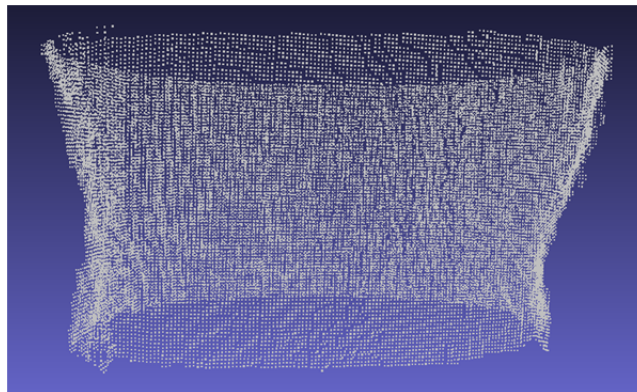
Fig 6. Trunk segment point cloud scan.

2.6. Watertight Model Generation

Immediately apparent from figure 6 are the discontinuities in the point cloud's top and bottom edges. This is due to the non-uniform scanning grid used by the Kinect, meaning there are variable gaps between the horizontal segmentation plane and the nearest points in the point cloud. These gaps result in a significant loss of volumetric data which would lead to considerable inaccuracy in BSIP calculations. To avoid this problem the point cloud must be fully extrapolated to the segmentation plane.

Firstly, the neighbouring points above and below the segmentation planes are found (figure 7). These two sets of points are then smoothed using a minimum least squares (MLS) smoothing algorithm to remove discontinuities in the depth information before being triangularly meshed using the open source Point Cloud Library's (PCL) greedy triangulation meshing algorithm.
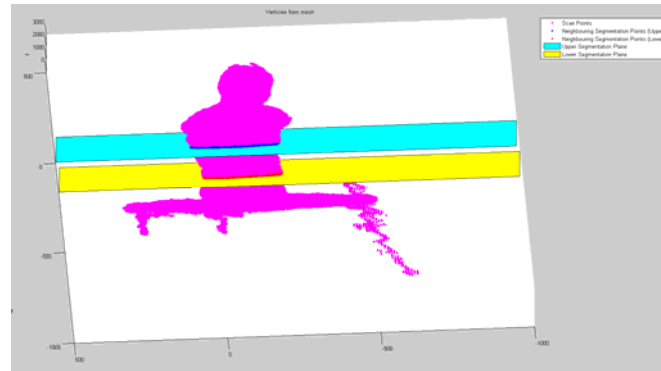
Fig 7. Segmentation process of the 3D point cloud scan.

The triangular meshing process creates a complete surface that passes through the points where the segmentation plane intersects. The triangles forming the mesh are then examined to determine if any intersect the segmentation plane, and therefore have 1 or 2 vertices above or below the segmentation plane. The vertices of the triangles matching this condition are then used to construct 3D geometric lines with one point above the plane and one of the triangles other vertices acting as a reference point below the plane. The intersection point of the segmentation plane and this line are then found which returns a 3D point that lies exactly on the plane where the point cloud was segmented. This new point is combined with the existing point cloud data and the point above the segmentation plane that was used to construct the line removed. This process is repeated for all mesh triangles matching the condition for both segmentation planes. Figure 8 shows the point cloud previously shown in figure 6 fully extrapolated to both segmentation planes.
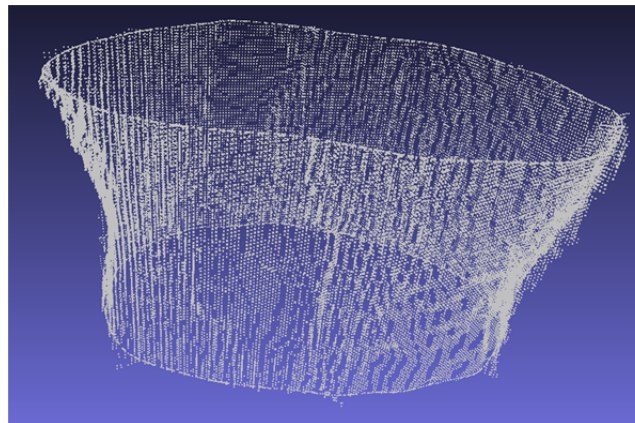

Fig 8. Trunk segment point cloud scan extrapolated to segmentation planes.

The 3D point cloud scan is then imported into GeoMagic's 'GeoMagic Studio' software to undergo a 5mm uniform sub sampling procedure, smoothing to remove discontinuities in the depth information and final triangular meshing. This results in a complete surface, but with a hole in the top and bottom of the mesh about the segmentation plane, subsequently filled using GeoMagic Studio's integral hole filling tools to result in a complete watertight mesh.

2.7. BSIP Calculation

To calculate BSIPs of the segment the complete mesh produced above is imported into PTC's 'Pro-E' CAD software and the integral analysis tools used to obtain the required BSIPs.

The complete process from 3D scan to BSIP estimation follows the workflow shown in figure 9.

Fig 9. Workflow for scanning process.

Due to the variation in chest and stomach shape it is hard to produce a geometric shape which is able to represent the body's trunk segment accurately, consequently, this segment exhibits the greatest level of error in BSIP estimation when using geometric based models. With this in mind the trunk segment was used to assess the accuracy of the system.

The results obtained with the proposed system were compared to results obtained from a gold standard laser scan and Yeadon's commonly used geometric model, with the same segment end points used across all methods.

A 'Choking Charlie' torso manikin, manufactured by Medical Plastics was used as the subject to be scanned by the proposed system. Firstly, the trunk segment of the manikin was scanned using a Modelmaker D100 non-contact laser scanner to obtain a gold standard measure of volume. Post processing with the use of GeoMagic Studio allowed segmentation of the point cloud, sub sampling and triangular meshing to result in a watertight mesh (figure 10). Further post processing as detailed above enabled calculation of the manikins BSIPs.
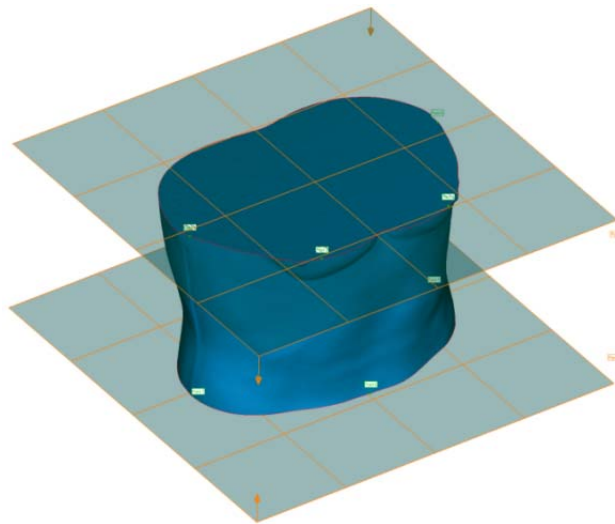

Fig 10. Laser scan of the manikin and the segmentation planes.

Next, the manikin was measured using anatomical tape and callipers to obtain the measurements necessary for the Yeadon model [6]. The measurement process was repeated twice with the measurements averaged before using Yeadon's formulae to calculate the segments BSIPs [6].

Finally, the initial calibration and scanning procedure as detailed above was adopted with six scans of the manikin being completed using the proposed system to assess accuracy and repeatability. To also assess the robustness of the calibration algorithm, each different scan used a new set of camera positions and hence new calibration.

As the Kinects IR projectors can interfere with one another in multiple device environments the scanning process was done in two stages with the IR projectors of neighbouring Kinects being covered up and the Kinects diagonally opposite from one another taking a scan together as they are obscured from one another by the manikin.

4. Results

Table 2 shows the mean and standard deviation of the RMS error in calibration across the six different calibrations

Table 2. Rigid body calibration RMS error values.

	Kinect 2 to Kinect 1 (mm)	Kinect 4 to Kinect 1 (mm)	Kinect 3 to Kinect 2 + Kinect 2 to Kinect 1 (mm)
Mean	2.76	4.03	6.91
Standard Deviation	0.59	0.71	0.97

Immediately apparent from table 2 is the large mean RMS error in the calibration for Kinect 3. This is purely caused by the need for two transformations to get back to the reference coordinate system of Kinect 1 and therefore the compounding of the error in the two transformations.

Table 3 and 4 shows the percentage errors apparent in the volume measurements of the manikin's trunk segment when compared to the gold standard laser scan.

Table 3. Percentage error in the volume measurements with Yeadon's geometric model [6].

	Volume
Error (%)	-1.17

Table 4. Mean percentage error and standard deviation in the volume measurements with the proposed system.

	Volume
Mean Error (%)	0.04
Standard Deviation	2.11

5. Discussion and Conclusions

The results in table 4 show a very small average difference in the volume measurements obtained with the Kinect scanning system, when compared to the gold standard laser scan. In the main the results obtained with the system also show smaller errors in volume estimation than is apparent with Yeadon's commonly used geometric model. However, the standard deviation in the error means that on some occasions the accuracy is slightly less than that of the Yeadon model.

The volume measurements obtained with the scanning system can be simply converted to mass measures by assuming uniform segment density in the same way as the investigations by Durkin and Dowling [7]. Directly comparing the results from these investigations suggests the system proposed here is able to offer better accuracy than other common BSIP estimation models, due to the subject specific nature of the measurements used to form BSIP estimations. Further work is however required in this area to compare the accuracy of the system when working with other BSIP measures and to determine if the uniform density assumption introduces unnecessary error. Something however shown by Wicke and Dumas [9] to have little effect.

As previously discussed, it is notoriously difficult to obtain accurate BSIPs of the trunk segment with the use of geometric models, such as that developed by Yeadon [6]. The trunk manikin used here is in fact relatively geometric in nature so fits the stadium solids of the geometric model very well. This has resulted in the geometric model providing a better than expected error, with normal errors in the region of ±6% to ±12% [9]. With this in mind it would be expected for the Kinect scanning system to consistently far outperform the geometric model when used with living subjects whose trunk segments differ greatly from the stadium solid shape due to excess fat and skin.

The standard deviation in volumetric data from the scans is also better than those apparent with existing BSIP models, as reported by Wicke and Dumas [9], thereby suggesting great potential for reliable usage of the system in the long term.

As well as offering accuracy advantages over existing BSIP estimation methods, the proposed system also offers a considerable advantage in terms of speed as a single scan of the whole body takes around 3 seconds, plus the time taken to initially palpate the body In contrast, the manual measurements required of Yeadon's geometric model can take around 40 minutes of the subject's time.

The greatest drawback of the system is the need to manually palpate the human body. The scanning system could be considerably improved if an accurate method could be devised to automatically determine the location of key anatomical landmarks from the 3D scans and to then use these to segment the body. Current methods however do not appear to have the required level of accuracy.

Part of the data processing procedure in this study involved the use of commercial software costing tens of thousands of pounds. The approach was adopted for this initial exploratory study to ensure the highest possible level of accuracy when processing the point cloud data to allow the accuracy of the scanning system to be truly examined without any possible interference from issues external to the scanning procedure. However, this rather defeats one of the initial objectives of the project to create low cost scanning solution. Further work therefore includes the production of fully standalone software that is able to mesh the point cloud scan and subsequently calculate the required BSIPs to a level of accuracy comparable with the commercial software used in this study.

The use of a static manikin in this investigation means that when required, it was possible to manually cover up the IR projectors to prevent interference. When scanning living subjects this process would be unsuitable due to the possibility of movement over the time taken to manually cover and uncover the IR projectors and hence the possible introduction of error. Further work therefore involves the development of an electronic system controlled from the scanning software to automatically cover and uncover the IR projectors with shutters when required.

Although clearly effective, the standard deviation in volume suggests the calibration approach adopted here is likely the main cause of error in volumetric estimations. The point cloud scan portrays visible misalignment between the four scans, particularly apparent with the scan from Kinect number 3 which must undergo two transformations. Further work is therefore required to improve the calibration method to reduce the inter-Kinect calibration RMS error and hence improve the overall accuracy of the system.

The study conducted here is only limited initial investigative work, purely focused on examining the volume and hence mass of the trunk segment of a static manikin. Further work includes scaling up the study to include the scanning of living subjects and to calculate the BSIPs of all the bodies segments.

The system proposed here will never be able to replicate the accuracy of a laser scanner when scanning static objects such as the torso manikin used here, but the short scanning duration is able to reduce the possibility of error arising due to involuntary movement of the body. The system presented here is a small financial investment and offers considerable advantages over current BSIP estimation techniques. It is hoped the principal demonstrated here can be further developed to produce a full system that is able to produce fast and accurate estimates of the body's segmental inertia parameters.

References

[1]   W. Dempster, "Space Requirements of the Seated Operator," 1955.

[2]   V. Zatsiorsky, "The Mass and Inertia Characteristics of the Main Segments of the Human Body," Biomechanics, vol. V, pp. 1182–1159, 1983.

[3]   M. R. Yeadon, "The simulation of aerial movement--I. The determination of orientation angles from film data.," Journal of biomechanics, vol. 23, no. 1, pp. 59–66, Jan. 1990.

[4]   E. Harless, "Die statischen Momente der menschlichen Gliedmassen," Abhandl Mathematische-Physikalischen Classe Konigl Bayerischen Akad Wissenschaft, vol. 8, no. 69–96, pp. 257–294, 1860.

[5]   M. J. R. Gittoes and D. G. Kerwin, "Component inertia modelling of segmental wobbling and rigid masses.," Journal of applied biomechanics, vol. 22, no. 2, pp. 148–54, May 2006.

[6]   M. R. Yeadon, "The simulation of aerial movement--II. A mathematical inertia model of the human body.," Journal of biomechanics, vol. 23, no. 1, pp. 67–74, Jan. 1990.

[7]   J. L. Durkin and J. J. Dowling, "Analysis of Body Segment Parameter Differences Between Four Human Populations and the Estimation Errors of Four Popular Mathematical Models," Journal of Biomechanical Engineering, vol. 125, no. 4, p. 515, 2003.

[8]   T. Outram, S. Domone, J. Hart, and J. Wheat, "The use of geometric shapes in estimating the geometry of body segments," Journal of Sports Sciences, no. May 2012, pp. 24–25, 2011.

[9]   J. Wicke and G. A. Dumas, "Influence of the Volume and Density Functions Within Geometric Models for Estimating Trunk Inertial Parameters," Journal of applied biomechanics, pp. 26–31, 2010.

[10]   A. Weiss, D. Hirshberg, and M. J. Black, "Home 3D Body Scans from Noisy Image and Range Data," in 13th International Conference on Computer Vision, 2011.

[11]   K. Khoshelham, "Accuracy analysis of Kinect depth data", 2010.

[12]   R. A. Newcombe, D. Molyneaux, D. Kim, A. J. Davison, J. Shotton, S. Hodges, and A. Fitzgibbon, "KinectFusion : Real-Time Dense Surface Mapping and Tracking."

[13]   S. Kean, J. Hall, P. Perry, and N. Burrus, Hacking The Kinect, New Editio. APRESS, 2012.

[14]   S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon, "KinectFusion : Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera," 2011.