Subject-specific body segment parameter estimation using 3D photogrammetry with multiple cameras

Kathrin E Peyer, Mark Morris, William I Sellers

Inertial properties of body segments, such as mass, centre of mass or moments of inertia, are important parameters when studying movements of the human body. These quantities are, however, not directly measurable. Current approaches include using regression models which have limited accuracy; geometric models with lengthy measuring procedures; or acquiring and post-processing MRI scans of participants. We propose a geometric methodology based on 3D photogrammetry using multiple cameras to provide subject-specific body segment parameters while minimizing the interaction time with the participants. A low-cost body scanner was built using multiple cameras and 3D point cloud data generated using structure from motion photogrammetric reconstruction algorithms. The point cloud was manually separated into body segments and convex hulling applied to each segment to produce the required geometric outlines. The accuracy of the method can be adjusted by choosing the number of subdivisions of the body segments. The body segment parameters of six participants (four male and two female) are presented using the proposed method. The multi-camera photogrammetric approach is expected to be particularly suited for studies including populations for which regression models are not available in literature and where other geometric techniques or MRI scanning are not applicable due to time or ethical constraints.

Authors

Kathrin Eva Peyer[1], Mark Morris[1], William Irvin Sellers[1]

[1]Faculty of Life Sciences, University of Manchester, Manchester, United Kingdom

1. Introduction

Inertial body segment parameters (BSP), such as mass, centre of mass (CoM) or moment of inertia, are used in motion analysis in research as well as in clinical settings. Accurate values are essential for techniques such as inverse dynamic analysis to allow the calculation of joint torques based on measured segmental accelerations (Winter, 1979). It is, however, not straightforward to measure these quantities from subjects directly. One approach is to use mathematical models of the body segments and rely on anthropometric measurements to determine the dimensions of the modelled segments. This type of methods requires a multitude of anthropometric measurements of the participants and is limited by the accuracy of the mathematical model of the body segments. The first mathematical model suggested by Hanavan in 1964 represented 15 body segments as cylinders and spheres and required 25 anthropometric measurements (Hanavan, 1964). More detailed models presented by Hatze or Yeadon required a total of 95 or 242 measurements respectively rendering these methods inefficient for studies with a large number of participants because of the time and discomfort experienced by the participant to acquire all the measurements needed (Hatze, 1980; Yeadon, 1990). Other types of approaches rely on X-ray or MRI based tomography to extract subject-specific BSP from participants. Unlike other methods, CT or MRI scans provide information about internal structures such as tissue composition which should improve the reconstruction accuracy (Martin et al., 1989; Mungiole & Martin, 1990; Pearsall, Reid & Livingston, 1996; Bauer et al., 2007). These approaches are, however, also difficult to implement in large-scale studies due to cost and ethical constraints. Alternatively, it is possible to approximate inertial BSP by adjusting previously reported average values or using regression models that require only a very few subject-specific measurements (commonly subject height and weight). Such average values and regression models were derived from cadavers or participants in a number of famous studies, such as the ones by Clauser, Dempster or Zatsiorsky (via de Leva) (Dempster, 1955; Clauser, McConville & Young, 1969; McConville, Clauser & Churchill, 1980; Leva, 1996). The reliability of such regression models is, however, rather low and the models are only applicable to a population similar to the one used to derive the regression equations.

Recently, other methods have been explored to obtain volumetric data of body segments that, in combination with body density assumptions, can provide subject-specific inertial BSP. Sheets et al. used a laser to scan the body surface of participants and morphing a generic model, which

42   contained joint location information, to the scanned surface (Sheets, Corazza & Andriacchi,

43   2010). Bonnechere et al used a Kinect sensor to estimate body segment lengths but not the

44   volumetric data required to estimate inertial properties (Bonnechère et al., 2014). Clarkson

45   evaluated the Kinect sensor as a surface scanner using a mannequin, but found the scanning

46   resolution to be quite low (Clarkson et al., 2012). Another approach to gain surface data is to use

47   photogrammetry. In 1978, Jensen proposed the use of stereophotogrammetry to estimate BSP

48   parameters (Jensen, 1978). In his model, the human body was divided into elliptical disks with a

49   thickness of 20 mm and the radii of the elliptical disks were estimated using images from the

50   front and side. The drawback of this approach lies in the simplifying assumptions of representing

51   body segments as the elliptical disks. It is, however, possible to reconstruct the surface of a 3D

52   object from multiple uncalibrated 2D images taken from different positions without requiring

53   any assumptions to the geometry of the body. This principle is referred to as "structure from

54   motion" and was initially used for producing 3D models of static objects and landscapes.

55   Perhaps the most striking example to date is the "Building Rome in a Day" project which used

56   images from the Flikr web site (http://www.flickr.com) to generate a 3D model of the whole city

57   (Agarwal et al., 2009). The reconstruction of a 3D surface from multiple cameras is two-stage

58   process. In stage one, the position, orientation and the parameters of the camera optics are

59   estimated. This is achieved by the bundle adjustment algorithm (Triggs et al., 2000) that

60   minimizes the error between the re-projected feature points using estimated camera pose and

61   parameters with the actual feature points in the images. In theory, feature points could be chosen

62   manually but this would be cumbersome and potentially not very accurate. Instead, Scale

63   Invariant Feature Transform (SIFT) algorithms are employed which automate this process by

64   identifying possible common points between multiple images (Lowe, 1999). Stage two uses the

65   calibrated views to produce a dense point cloud model of the 3D object. There are a number of

66   possible approaches to achieve this (for review see (Seitz et al., 2006)) but probably the most

67   widespread current approach is patch-based multi-view stereo reconstruction (Furukawa &

68   Ponce, 2010). This photogrammetric approach has gained wide acceptance for producing 3D

69   models in areas such as archaeology (McCarthy, 2014) and palaeontology (Falkingham, 2012),

70   and is even used for markerless motion capture (Sellers & Hirasaki, 2014).

71   The aim of this paper is to investigate whether an approach based on structure form motion

72   photogrammetric reconstruction can provide person-specific body segment parameters and to

73   identify the strength and weaknesses of such an approach with regards to ease of

74   implementation, cost-effectiveness, subject comfort and processing time. A low-cost body

75 scanner was built using multiple cameras and the body segment parameters of six participants

76 (four male and two female) are presented using the proposed method.

77

78 **2. Methods**

79

80 Photogrammetry relies on obtaining multiple photographs taken from different locations. These

81 photographs can be taken with any suitable device and for objects that do not move, the most

82 cost effective option is to take 50+ photographs with a single camera that is moved around the

83 object. This has the additional advantage that a single intrinsic calibration can be used since the

84 camera optics can be considered identical for multiple images. However for subjects that can

85 move, all the photographs must be taken simultaneously so that the subject is in exactly the same

86 position for all the images. Simultaneous photographs can be achieved in several different ways

87 including multiple still cameras with synchronised remote controls, multiple USB web cameras,

88 or multiple networked cameras. There is probably little to choose between these methods but

89 initial experimentation found that network/IP cameras provided a cost effective solution that

90 scaled well. The camera resolution should be as high as reasonably possible since higher

91 resolution images provide more information for the feature extraction algorithms and higher

92 point density in the eventual reconstruction. This means that low resolution cameras such as low

93 cost web cameras and standard resolution video cameras may not be suitable.

94 Most applications that employ photogrammetry aim to capture surface data in great detail with

95 the emphasis on creating almost true-to-live 3D models and thus maximizing the point cloud

96 density. Some applications require only the information available from the point cloud directly

97 (such as feature point locations) and do not require a surface mesh. In fact, meshing algorithms

98 tend to decrease the accuracy of the model (Falkingham, 2012). In applications where the

99 reconstructed object is to be 3D-printed (Garsthagen; Hobson; Straub & Kerlin, 2014) or where

100 volumetric data is required (such as for body segment estimations presented in this paper), a

101 (closed) surface mesh needs to be created from the point cloud. A high-resolution mesh is

102 commonly desired in 3D printing (e.g. for aesthetic or functional reasons), which requires a large

103 number of photographs and sophisticated algorithms to convert the point cloud to a mesh. In this

104 paper, we propose the use of convex hulling to generate simplified geometric outlines of the

105 body segments. Convex hulling is robust to low-density surface point clouds (and even potential

106 gaps in the point cloud) and can thus be implemented with ease and run automatically without

107 requiring user input. Furthermore, being able to generate surface meshes from a low-density

108   point cloud lowers the number of cameras required to build the 3D scanner (as opposed to

109   needing a large number of cameras to achieve densely packed point clouds).

110

111   ***2.1. 3D body scanner design***

112   Photogrammetric reconstruction can work well with as few as 4 cameras (Sellers & Hirasaki,

113   2014)but more cameras are necessary to provide a relatively gap free reconstruction. To estimate

114   the minimal number of cameras necessary to achieve a 360° reconstruction, we positioned a

115   single camera on a circle of radius 1.6 m and placed a stationary skeletal dummy as a test

116   subject in the centre.. Images were taken every 5° and the point cloud reconstructions using 72,

117   36, 24, 18, 12 and 9 images, corresponding to angular resolutions ranging from 5° to 40°, were

118   compared (see Fig. 1A). Acceptable reconstructions for the purpose of this paper, i.e. no loss of

119   body segment features, were found with 18 or more cameras although using larger numbers of

120   cameras certainly improved the point cloud density. After initial testing, the setup design was

121   adjusted by increasing the radius of the camera placements (to increase the field of view to

122   accommodate outstretched arms), placing the cameras above head-hight and angling the camera

123   views downwards (as opposed to placing the cameras at the bottom or at hip-height) and using

124   asymmetric patterns on the floor in the shared field of view of all cameras. The latter greatly

125   aided the reconstruction reliability as the camera calibration algorithm relies on shared features.[1]

126   The network camera was implemented using Raspberry Pi (RPi) modules, type A, each equipped

127   with an 8GB SD card and a Pi camera (http://www.raspberrypi.org). These modules run the

128   Linux operating system (Raspbian) and provide a flexible and cost-effective 5 megapixel

129   network camera platform. The 18 RPi modules (each with a camera) were attached to a 4.8 m

130   diameter hexagonal frame elevated to height of 2.3m by six support poles (see Fig. 1B). Each

131   RPi module was provided with a USB WiFi receiver (Dynamode WL-700-RX) and power was

132   provided using the standard RPi power adapter plugged into a multi-socket attached to each

133   support pole. Four 500 W Halogen floodlights were mounted to provide additional lighting to

134   increase the image quality.

135

136   RPi cameras can record either still images or movie files. For this application we needed to

137   trigger all the cameras to record a single image at the same instant. This was achieved using the

138   open source "Compound Pi" application (http://compoundpi.readthedocs.org), which uses the

---

1   Without the patterns on the floor, the camera calibration relied on shared features found on the subject, whereas
     the patterned floor provided a large (or even completely sufficient) number of features to run the camera
     calibration algorithm.

139 UDP broadcast protocol to control multiple cameras synchronously from a single server. Once
140 the individual images have been recorded, the application provides an interface to download all
141 the images obtained to the server in a straightforward manner. Since UDP broadcast is a one-to-
142 many protocol, all the clients will receive the same network packet at the same time and the
143 timing consistency for the images will be of the order of milliseconds which is adequate for a
144 human subject who is trying to stand still. Higher precision synchronisation can be achieved
145 using a separate synchronisation trigger but this was unnecessary in this application.
146

147 *2.2. Data acquisition*
148 Full body scans using the RPi setup were obtained from six voluntary participants. Additionally,
149 their body weight and height was measured (Table 1). The male visible human was used as an
150 additional data set for validation (National Library of Medicine's Visual Human Project (Spitzer
151 et al., 1996)). The experimental protocol (reference number 13310) was approved by the
152 University of Manchester ethics panel. In accordance with the experimental protocol, written
153 consent was obtained from all participants.
154

155 The reconstruction algorithms rely on finding matching points across multiple images so they do
156 not work well on images that contain no textural variation. We therefore experimented with
157 using different types of clothing in the scanner, such as sports clothing, leisure clothing, and a
158 black motion capture suit equipped with Velcro strips to aid feature detection. Clothing was
159 either body-tight or tightened using Velcro strips if they were loose since loose clothing would
160 lead to an overestimation of the body volume. The participants stood in the centre of the RPi
161 setup with their hands lifted above their head (see Fig. 2) and the 18 images were then acquired.
162

163

164

165 *2.3. Data processing*
166 The 3D point cloud reconstruction was initially done using open source application VisualSFM
167 (http://ccwu.me/vsfm/) which performed adequately, but we then switched to using Agisoft
168 PhotoScan Standard Edition v1.0.4 (http://www.agisoft.com) which proved to be much easier to
169 install and use. Agisoft PhotoScan also achieved a better reconstruction quality with fewer holes

170   in the point cloud and smoother surfaces.[2] The parameters used in the reconstructions are

171   reported in the supporting information. Agisoft PhotoScan runs identically on Windows, Mac or

172   Linux. The full 3D reconstruction with 18 images took an average of 30 minutes using an 8 core

173   3GHz Xeon MacPro with 12GB RAM. The actual time taken was variable depending on the

174   image file size and the reconstruction parameters. The output of the Agisoft PhotoScan is an

175   unscaled 3D point cloud of the participants and surrounding scenery (see Fig. 2), which requires

176   further post-processing to calculate BSP values. First, the point cloud was scaled and oriented

177   using CloudDigitizer (Sellers & Hirasaki, 2014), the oriented point clouds were then divided into

178   anatomical segments using Geomagic (http://geomagic.com), and the convex hulls computed in

179   Matlab® (http://www.mathworks.com, see supporting information). The reference points for the

180   body segmentation are listed in the supporting information Table S1. The body segments were all

181   oriented into the standard anatomical pose before the volume, centre of mass and inertial tensor

182   were calculated based on the hull shape and segment density using a custom function

183   implemented in Matlab® (see supporting information). The choice of body density is an

184   interesting issue. Different tissues within segments have different densities and tissue

185   composition is moderately variable between individuals. Indeed variations in density are

186   commonly used to estimate body fat percentage (Siri, 1961; Brožek et al., 1963). MRI and CT

187   based techniques can allow individual tissue identification and can compensate for this but

188   surface volumetric techniques need to use an appropriate mean value. Segment specific densities

189   are available (e.g. (Winter, 1979)) but the quoted trunk density is after subtraction of the lung

190   volume. For a surface scan model, we need to use a lower value trunk density that incorporates

191   the volume taken up by the air within the lungs. Therefore for the purpose of this paper a trunk

192   density value of 940 kg/m³ was chosen, while a uniform density of 1000 kg/m³ was assumed for

193   all other body segments (Weinbach, 1938; Pearsall, Reid & Ross, 1994). The body mass

194   calculated from the volume was never exactly the same as the recorded body mass so the density

195   values were adjusted pro-rata to produce a consistent value for total mass.

196    $$s = \frac{m_{Participant}}{\sum m_{SegmHull,i}} \qquad (1)$$

197   The factor s effectively scales the body densities and is thus also applied the moments and

198   products of inertia obtained from the convex hull segments (see supporting information).

199

---

2   This is based on the comparison of the best reconstruction result achieved with each software after testing an
extensive, but not complete, combination of reconstruction parameters.

3. Results

Six participants were scanned using the RPi photogrammetry setup and their point cloud segmented. In order to be able to calculate the inertial properties, the point cloud needs to be converted into a closed surface mesh. To calculate the volume of an arbitrary shape defined by a surface mesh, the mesh needs to be well defined, i.e., it should be two-manifold, contain no holes in the mesh, and have coherent face orientations. The processing of converting a point cloud to a well defined mesh is known as hulling and there are several possible methods available. The simplest is the minimum convex hull where the minimum volume convex shape is derived mathematically from the point cloud (www.qhull.org). This approach has the advantage of being extremely quick and easy to perform and it is very tolerant of point clouds that may contain holes where the reconstruction algorithm has partially failed. However it will always overestimate the volume unless the shape is convex. There are also a number of concave hulling approaches. Some are mathematically defined such as AlphaShapes (Edelsbrunner & Mücke, 1994) and Ball Pivoting (Bernardini et al., 1999) and require additional parameters defining the maximum level of permitted convexity. Others are proprietary and can require considerable manual intervention such as the built in hole-filling algorithms in Geomagic. This latter group provides the highest quality reconstructions but at the expense of considerable operator time. For this paper we concentrated on convex hulls under the assumption that the level of concavity in individual body segments was likely to be relatively small. The relative segment mass of all participants are reported in Fig. 3 (the segmented convex hulls are shown in Fig. S1 in the supporting information). Figure 3 also displays average values from literature. As the participants were imaged wearing shoes, the foot volume was overestimated significantly. It is possible to adjust the value using a foot-specific scaling factor that accounts for this overestimation although of course if the subsequent use of the BSP parameters is in experiments with participants wearing shoes then the shoe mass becomes an important part of the segment. For the purpose of this paper, a scaling factor was derived based on a single participant (P5), by comparing the convex hull volume of the of the foot imaged in socks versus the convex hull volume wearing shoes, and this factor (of 0.51) applied to all participants' inertial values of the feet. The moments of inertia are shown in in Fig. 4 together with average values from literature. Geometric methods also allow us to calculate the products of inertia which are otherwise simply assumed to be zero. The average products of inertia are depicted in Fig. 5 (absolute values shown only, signed values reported in the supporting information Table S2-S4). Some segments, e.g. the

233 thigh or trunk, have products of inertia that are of a similar order of magnitude as their moments

234 of inertia, which is indicative of a noticeable difference between the inertial principal axes and

235 the anatomical principal axes. The majority of the products of inertia are however significantly

236 smaller than the moments of inertia (of the same segment) by one to two orders of magnitude.

237 Figure 6 contains the relative centre of mass in the longitudinal segment direction, i.e. along the

238 z-axis with the exception of the foot whose longitudinal axis corresponds to the x-axis (see Fig.

239 2). Figure 7 shows the shift of CoM from the longitudinal axis in the transverse plane (x-y

240 plane). The CoM values in literature assume a zero shift from the principal anatomical

241 (longitudinal) axis. The shift values we found with our geometric method are generally unequal

242 to zero, but they have be to viewed with caution as the placement of the reference anatomical

243 axis itself has uncertainties associated with it. The numerical values presented in Fig. 3-7 and the

244 segment lengths are reported in the supporting information (Tables S2-S13)

245

246 To estimate the effect of the convex hull approximation on the mass estimation versus the

247 original body segment shape, the volumes of a high resolution 3D body scan and of their convex

248 hull approximation were calculated and compared. A detailed surface mesh was obtained from

249 the National Library of Medicine's Visible Human Project (Spitzer et al., 1996) by isosurfacing

250 the optical slices using the VTK toolkit (http://www.vtk.org) and cleaning up the resultant mesh

251 using Geomagic. The surface mesh of the 3D body scan was separated into body segments and

252 the volume calculated following the same methodology as used for the point cloud data. A

253 convex hull was applied to each body segment and the volume calculated again (see Fig. 8). The

254 volume overestimations for each body segment (averaged between left and right) are shown Fig.

255 9 (column CH). Several body segments showed a large relative volume overestimation (using

256 10% error as a cutoff, although the choice would depend on the required accuracy): foot (26%),

257 shank (31%), hand (47%) and forearm (16%). This is due to the relatively strong curvatures in

258 these segments. To minimize the effect, these body segments were subdivided (see Fig. 10) and

259 the convex hulls recalculated. The results of the divided segments are also shown in Fig. 9

260 (column CHD), and the decrease in volume overestimation is apparent. The volume

261 overestimation of the subdivided foot (11%), shank (11%) and forearm (5%) are at a similar

262 level to the other body segments and would probably be acceptable in many cases. The hands

263 show the largest relative mass overestimation still (25%), which is due to its curved position and

264 slightly open fingers. The convex hull error of the hand is, however, expected to be significantly

265 smaller if the hand is imaged while being held in a straight position with no gaps between the

266 digits.

267

268 Figure 11 contains the relative mass estimations of the original surface mesh, the convex hulls

269 with and without subdivision, and the average and regression model values found in literature.

270 With a BMI value of almost 28, the male visible human is not well represented by the average or

271 regression model values found in literature, where the majority of the studies involve relatively

272 athletic people (BMI average of around 24) or obese individuals (BMI over 30). The convex

273 hulls of the subdivided segments (CHD in Fig. 11) give the closest approximation to the original

274 mesh and, with the exception of the hands, are within a relative error of less than 5%. The

275 relative error of the convex hull of the whole segments (CH in Fig. 11) is larger but closer to the

276 original mesh than average and regression values given in literature. The moments of inertia are

277 overestimated as well as they are a product of the mass of the segment. Their overestimation

278 follows the same trend as the mass overestimation, i.e. the largest overestimation occurs for the

279 hands, followed by the shanks and feet (see Fig. S2 in supporting information), and the

280 subdivided segments produce more accurate values with an average relative error below 10%.

281

282

283 **4. Discussion**

284

285 We can see from the results that the proposed methodology produces values that are very similar

286 to those derived using regression equations. There are no consistent problems although it is

287 clearly important that the hand is held in a suitable flat position but with fingers adducted so that

288 the hulling can provide an accurate volume estimation. We would expect that the

289 photogrammetric process will work as well as any of the published geometrical approaches

290 (Hanavan, 1964; Hatze, 1980) since it is simply an automated process for achieving the same

291 outcome. The procedure is currently moderately time consuming in total but the interaction time

292 with the participant is extremely short and involves no contact which can be very beneficial for

293 certain experimental protocols or with specific vulnerable participants. Since most of the time is

294 spent post-processing the data, we expect that this post-processing could be streamlined

295 considerably by writing dedicated software rather than the current requirement of passing the

296 data through multiple software packages.

297 In general regression equations work well for applicable populations and are probably more

298 suitable if body mass distribution is not a major focal point of the research, particularly given

299 that in some cases it can be shown that experimental outcomes are not especially sensitive to the

300 BSP parameters chosen (Yokoi et al., 1998). The values generated in our sample are relatively

301 close to those generated by using regression equations but BSP values are highly variable

302 between individuals and current regression equations are only suitable for a very limited range of

303 body shapes. This is particularly the case when we are dealing with non-standard groupings such

304 as children, the elderly or people with particularly high or low BMI values.

305

306 However there are some specific issues with this technique that could to be improved for a more

307 streamlined and potentially more accurate workflow (see Fig.12, which summarises the steps

308 involved in estimating body segment parameters using photogrammetry).

309 Convex hulling of the point cloud is a robust and fast way to produce surface meshes. The fact

310 that it systematically overestimates the volume of concave features can be improved by

311 subdividing body segments into smaller parts and the decision then becomes what level of

312 subdivision is appropriate for an acceptable level of accuracy (see Fig. 12 C). For example, with

313 only one subdivision of the shank and forearm the relative error of their volume overestimation

314 was reduced by a factor of three, and the end result was within 10% of the true value which is

315 probably sufficient in most cases, especially given the level of uncertainty in other parameters

316 such as segment specific density. It is important to note that the scaling factor used in our method

317 minimises the segment mass estimation errors introduced by the volume overestimation

318 significantly. In fact, if all hulled segment masses (i.e. the product of segment volume and

319 segment-specific density, see Fig. 12 D) were overestimated by 10% , the final body segment

320 mass would be calculated correctly due to the scaling factor applied to each segment. Using a

321 pro-rata scaling factor therefore performs best when the relative errors of the volume estimation

322 of each segment are within a small range of each other.

323

324 The adoption of one of the concave hulling techniques is likely to lead to a similar level of

325 improvement again with a minimum (but not zero) level of additional work. The level of

326 subdivision required not only depends on the body segment, but also the population studied so it

327 may well be appropriate that the segmentation level is adjusted according to the type of study

328 and its sensitivity to inaccuracies in the BSP (i.e. multiple segment subdivisions increase

329 accuracy of volume estimation). In this work, a uniform scaling factor and constant body density

330 (apart from the trunk) was assumed. It is well known that the density varies among body

331   segments as well as among populations due to different percentages of fat and muscle tissue
332   (Drillis, Contini & Bluestein, 1964; Durnin & Womersley, 1974; Zatsiorsky, 2002). Thus, using
333   segment and population specific densities (and scaling factors) may improve the accuracy of the
334   presented methodology if such values are available or derived. Similarly important contributions
335   to segmental mass distribution such as the presence of the lungs within the torso can be modelled
336   explicitly which may lead to small but important shifts in the centre of mass (Bates et al., 2010).
337
338   In terms of technology, the current arrangement of using 18 Raspberry Pi cameras is reasonably
339   straightforward and relatively inexpensive. It requires no calibration before use, and the process
340   of moving the subject into the target area is extremely quick. However it does take up a great
341   deal of room in the laboratory and the current software is reliant on clothing contrast for the
342   reconstructions which limits the flexibility of the technique. One area where this could be
343   improved is by projecting a structured light pattern onto the subject so that areas with minimal
344   contrast can be reconstructed accurately (Casey, Hassebrook & Lau, 2008). Our results show that
345   18 cameras is currently the minimum needed for full body reconstruction and a system with 36
346   or more cameras would produce better point cloud reconstruction results by minimizing areas of
347   potential occlusions (such as between the legs or between the arms and trunk) and increasing the
348   point cloud density. To what degree a more densely packed point cloud would significantly
349   improve the accuracy of the estimated inertial parameters based on convex hulls would be an
350   interesting aspect to investigate further. We would expect a denser point cloud to facilitate the
351   use of more complex meshing methods instead of convex hulling.
352
353   One future use of this technology is clearly the use of such systems and algorithms for complete
354   motion capture (Sellers & Hirasaki, 2014). The limitation currently is that these cameras would
355   need to be closely synchronised and whilst the proposed system is adequate for producing a
356   single still image, it is currently not able to adequately synchronise video. In addition the video
357   resolution is much lower and this makes the reconstruction more difficult. However we predict
358   that markerless, multiple video camera structure from motion systems will become a much more
359   common mainstream tool for experimental motion capture in the near future. Ideally we could
360   imagine that such a system would both do the motion capture and also the body segment
361   parameter reconstruction since much of the computational technology would be shared.
362
363   **Conclusion**

364

365 A methodology based on structure form motion photogrammetric reconstruction has been

366 presented that provides subject-specific body segment parameters. The method relies on the

367 surface depth information extracted from multiple photographs of a participant, taken

368 simultaneously from multiple different view points. The brief interaction time with the

369 participants (taking all required photos simultaneously, and measuring the height and weight

370 only) makes this a promising method in studies with vulnerable subjects or where cost or ethical

371 constraints do not allow the use of other imaging methods such as CT or MRI scans. Unlike

372 regression models that are valid only for a small population sample, we expect the proposed

373 methodology to be able to perform equally well for a wide range of population samples. The

374 post-processing time is lengthy compared to using regression models or average values from

375 literature but not compared to processing MRI or CT data. The 3D scanner presented in this

376 paper was able to produce a sufficient 3D data points to estimate body segment volumes with

377 only 18 RPi cameras, which kept the hardware cost to a minimum. Depending on the accuracy

378 required for the project, we would expect both more cameras and higher resolution cameras to

379 improve the robustness of the 3D point cloud reconstruction.

380 While the results presented in this work were derived using commercial software, such as

381 AgiSoft, Geomagic and Matlab®, we were able to to achieve similar results using open-source

382 software only (such as VisualFMS (http://ccwu.me/vsfm/) for calculating 3D point clouds and

383 MeshLab (http://meshlab.sourceforge.net/) for point cloud segmentation, hulling and BSP

384 calculation). This makes our proposed methodology, in combination with the low hardware

385 costs, particularly promising for small-budget projects.

386

387 **Acknowledgements**

388

389 The authors would like to thank Dave Jones for the development of the Compound Pi

390 programme and his generous help with the network setup of the Raspberry Pi scanner.

391

392

393 **References**

394 Agarwal S, Snavely N, Simon I, Seitz SM, Szeliski R. 2009. Building Rome in a day. IEEE Int.

395   Conf. on Computer Vision:72–79.

396 Bates KT, Manning PL, Margetts L, Sellers WI. 2010. Sensitivity analysis in evolutionary robotic
397   simulations of bipedal dinosaur running. Journal of Vertebrate Paleontology 30:458–466.

398 Bauer JJ, Pavol MJ, Snow CM, Hayes WC. 2007. MRI-derived body segment parameters of
399   children differ from age-based estimates derived using photogrammetry. Journal of
400   biomechanics 40:2904–10.

401 Bernardini F, Mittleman J, Rushmeier H, Silva C, Taubin G. 1999. The Ball-Pivoting Algorithm
402   for Surface Reconstruction. IEEE Transactions on Visualization and Computer Graphics
403   5:349–359.

404 Bonnechère B, Jansen B, Salvia P, Bouzahouene H, Sholukha V, Cornelis J, Rooze M, Van Sint
405   Jan S. 2014. Determination of the precision and accuracy of morphological measurements
406   using the KinectTM sensor: comparison with standard stereophotogrammetry. Ergonomics
407   57:622–31.

408 Brožek J, Grande F, Anderson JT, Keys A. 1963. Densitometric Analysis of Body Composition:
409   Revision of some Quantitative Assumptions. Annals of the New York Academy of Sciences
410   110:113–140.

411 Casey CJ, Hassebrook LG, Lau DL. 2008. Structured light illumination methods for continuous
412   motion hand and face-computer interaction. In: Asai K ed. Human Computer Interaction:
413   New Developments. InTech, 297–308.

414 Clarkson S, Choppin S, Hart J, Heller B, Wheat J. 2012. Calculating Body Segment Inertia
415   Parameters from a Single Rapid Scan Using the Microsoft Kinect. Proceedings of the 3rd
416   International Conference on 3D Body Scanning Technologies, Lugano, Switzerland, 16-17
417   October 2012:153–163.

418 Clauser CEC, McConville JJT, Young JWJ. 1969. Weight, volume, and center of mass of
419   segments of the human body. Antioch Coll Yellow Springs OH.

420 Dempster W. 1955. Space requirements of the seated operator: geometrical, kinematic, and
421   mechanical aspects of the body, with special reference to the limbs. Wright-Patterson Air
422   Force Base Ohio.

423 Drillis R, Contini R, Bluestein M. 1964. Body segment parameters. Artificial limbs 8:44–66.
424   Durnin J, Womersley J. 1974. Body fat assessed from total body density and its estimation
425   from skinfold thickness: measurements on 481 men and woman aged from 16 to 72 years.
426   British Journal of Nutrition 32:77–97.

427 Edelsbrunner H, Mücke EP. 1994. Three-dimensional alpha shapes. ACM Transactions on
428   Graphics 13:43–72.

429 Falkingham PL. 2012. Acquisition of high resolution three-dimensional models using free, open-
430   source, photogrammetric software. Palaeontologia Electronica 15:1T:15p.

431 Furukawa Y, Ponce J. 2010. Accurate, dense, and robust multiview stereopsis. IEEE transactions
432   on pattern analysis and machine intelligence 32:1362–76.

433 Garsthagen R. www.pi3dscan.com.

434 Hanavan EP. 1964. A Mathematical Model of the human body. Air force aerospace medical
435   research lab Wright-Patterson AFB Ohio.

436 Hatze H. 1980. A mathematical model for the computational determination of parameter values of
437   anthropomorphic segments. Journal of Biomechanics 13:833–843.

438 Hobson J. http://hackaday.com/2014/03/07/an-affordable-full-body-studio-grade-3d-scanner/.
439   Jensen R. 1978. Estimation of the biomechanical properties of three body types using a
440   photogrammetric method. Journal of Biomechanics 11:349–358.

441 Leva P De. 1996. Adjustment to Zatsiorsky-Seluyanov's segment inertia parameters. J. Biomech.
442   29:1223–30.

443 Lowe DG. 1999. Object recognition from local scale-invariant features. Proceedings of the
444   Seventh IEEE International Conference on Computer Vision 2:1150–1157.

445 Martin P, Mungiole M, Marzke M, Longhill J. 1989. The use of magnetic resonance imaging for
446   measuring segment inertial properties. Journal of Biomechanics 22:367–376.

447 McCarthy J. 2014. Multi-image photogrammetry as a practical tool for cultural heritage survey
448   and community engagement. Journal of Archaeological Science 43:175–185.

449 McConville J, Clauser C, Churchill T. 1980. Anthropometric relationships of body and body
450   segment moments of inertia. Anthropology research project inc Yellow Springs Ohio.
451   Mungiole M, Martin PE. 1990. Estimating segment inertial properties: comparison of
452   magnetic resonance imaging with existing methods. Journal of biomechanics 23:1039–46.

453 Pearsall DJ, Reid JG, Livingston L a. 1996. Segmental inertial parameters of the human trunk as
454   determined from computed tomography. Annals of biomedical engineering 24:198–210.

455 Pearsall D, Reid J, Ross R. 1994. Inertial properties of the human trunk of males determined from
456   magnetic resonance imaging. Annals of biomedical engineering 22:692–706.

457 Seitz SM, Curless B, Diebel J, Scharstein D, Szeliski R. 2006. A Comparison and Evaluation of
458   Multi-View Stereo Reconstruction Algorithms. 2006 IEEE Computer Society Conference on
459   Computer Vision and Pattern Recognition - Volume 1 (CVPR'06) 1:519–528.

460 Sellers WI, Hirasaki E. 2014. Markerless 3D motion capture for animal locomotion studies.
461   Biology open 3:656–668.

462 Sheets AL, Corazza S, Andriacchi TP. 2010. An automated image-based method of 3D subject-
463   specific body segment parameter estimation for kinetic analyses of rapid movements.
464   Journal of biomechanical engineering 132:011004.

465 Siri WE. 1961. Body composition from fluid spaces and density: analysis of methods. In: Brožek
466   J, Henschel A eds. Techniques for measuring body composition. Washington DC: National
467   Academy of Sciences - National Research Council, 223–44.

468 Spitzer V, Ackerman MJ, Scherzinger AL, Whitlock D. 1996. The visible human male: a
469   technical report. Journal of the American Medical Informatics Association 3:118–30.
470   Straub J, Kerlin S. 2014. Development of a Large, Low-Cost, Instant 3D Scanner.
471   Technologies 2:76–95.

472 Triggs B, McLauchlan P, Hartley RI, Fitzgibbon AW. 2000. Bundle adjustment—a modern
473   synthesis. Lecture Notes in Computer Science 1883:298–372.

474 Weinbach A. 1938. Contour maps, center of gravity, moment of inertia and surface area of the
475   human body. Human Biology 10:356–371.

476 Winter DA. 1979. Biomechanics of Human Movement. New York: Wiley and Sons.
477   Yeadon M. 1990. The simulation of aerial movement—II. A mathematical inertia model of
478   the human body. Journal of Biomechanics 23:67–74.

479 Yokoi T, Takahashi A, Okada H, Ohyama KB, Muraoka M. 1998. Is the Selection of Body
480   Segment Interia Parameters Critical to the Results of the Kinematic and Kinetic Analysis of
481   Human Movement? Anthropological Science 106:371–383.

482 Zatsiorsky VM. 2002. Kinetics of HumanMotion. Human Kinetics.

1

Body Scanner Design.

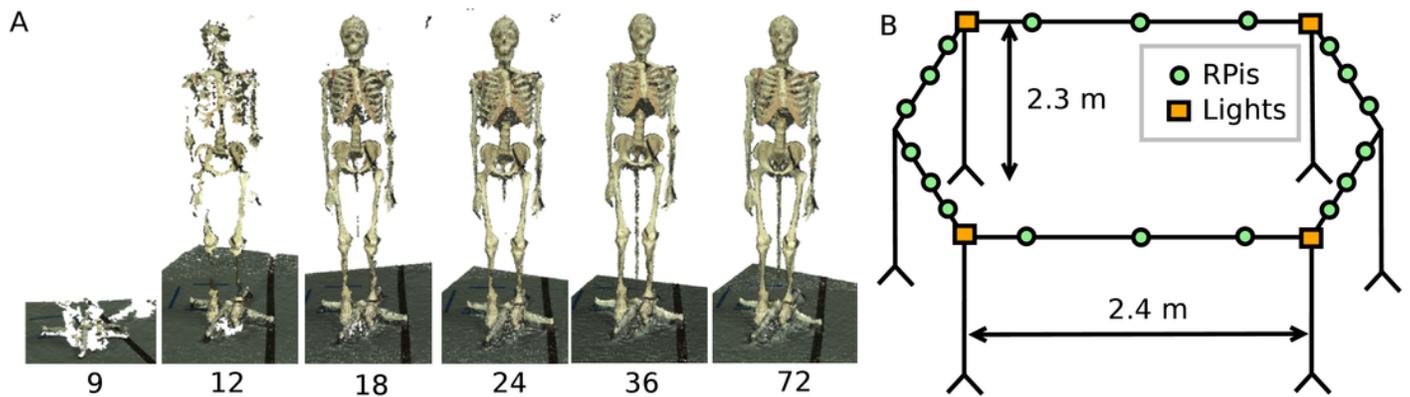
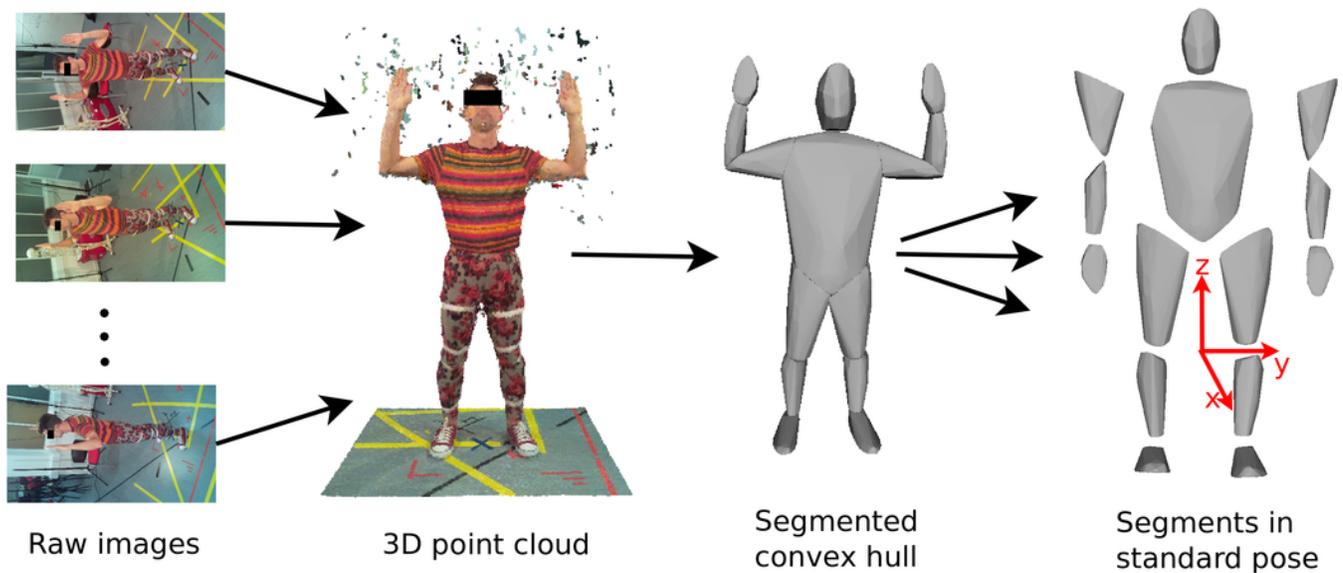
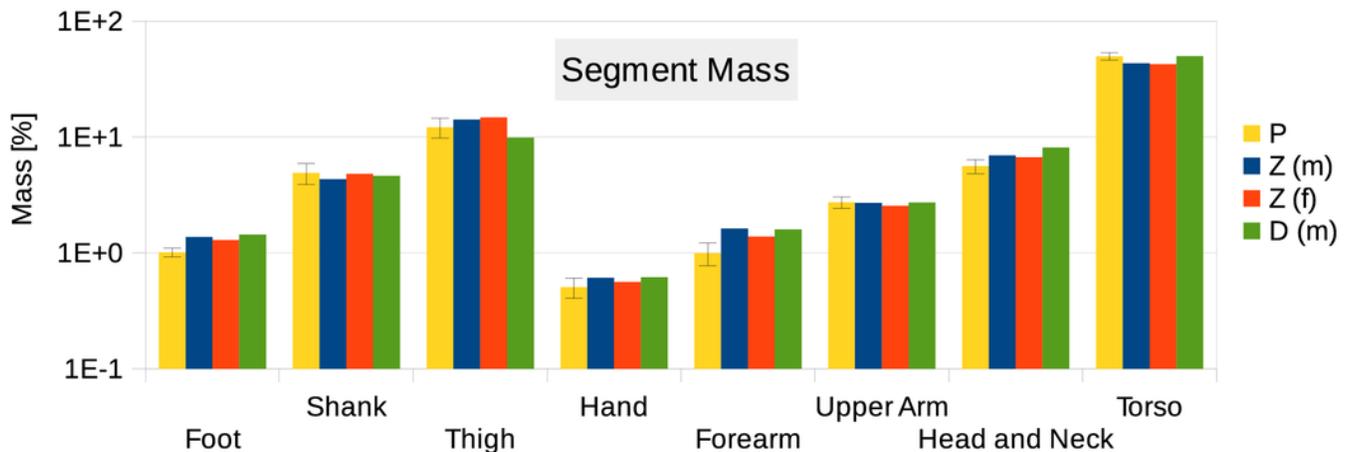
A: Point cloud reconstruction with varying number of cameras. B: Schematic representation of the RPi scanner design.

2

Image processing work flow.

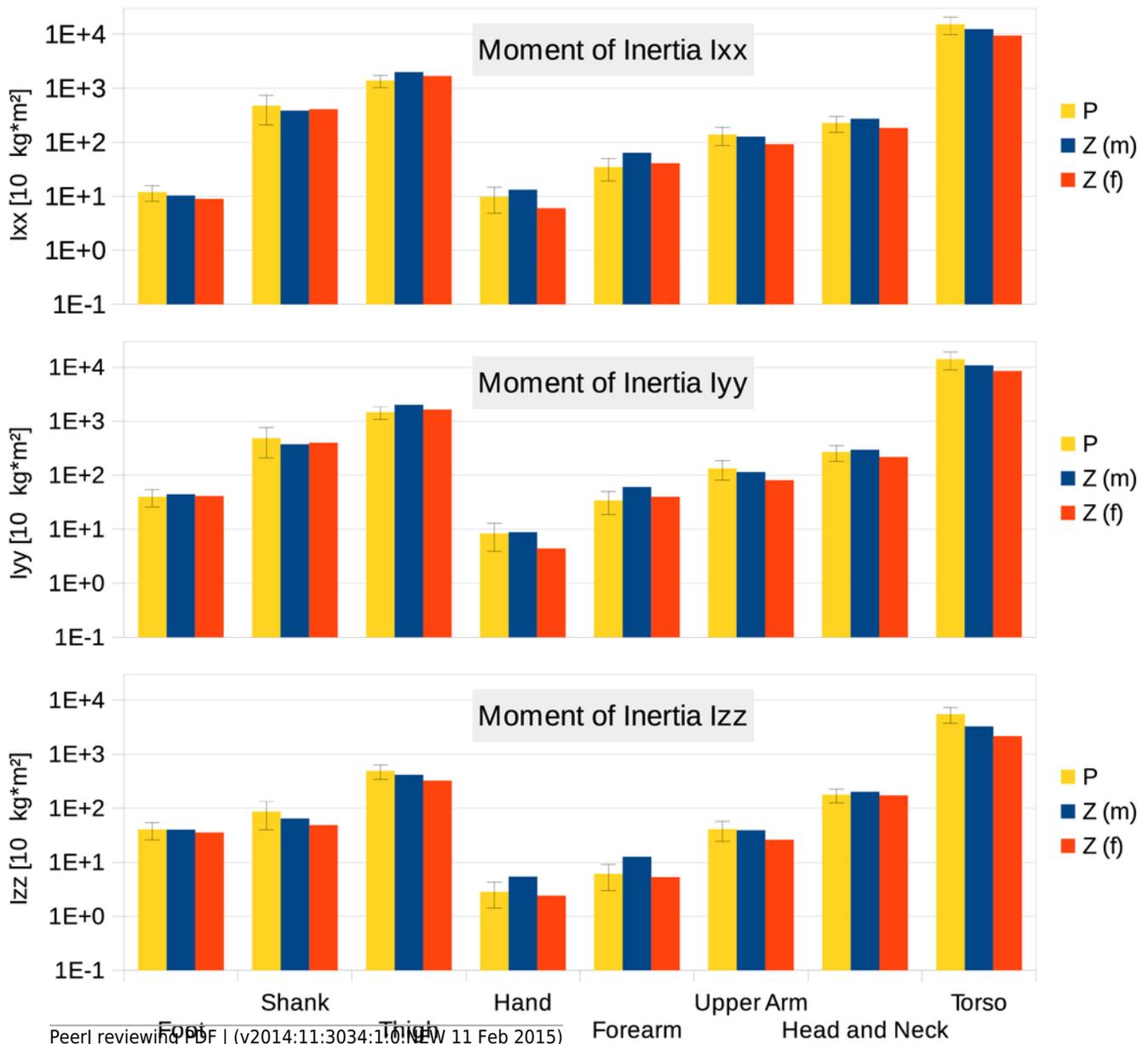
Images from the RPI scanner are converted to 3D point clouds which are then scaled and segmented manually. Subsequently, convex hulling is used to produce a surface mesh around each body segment.

Raw images   3D point cloud   Segmented   Segments in
                                            convex hull   standard pose

3

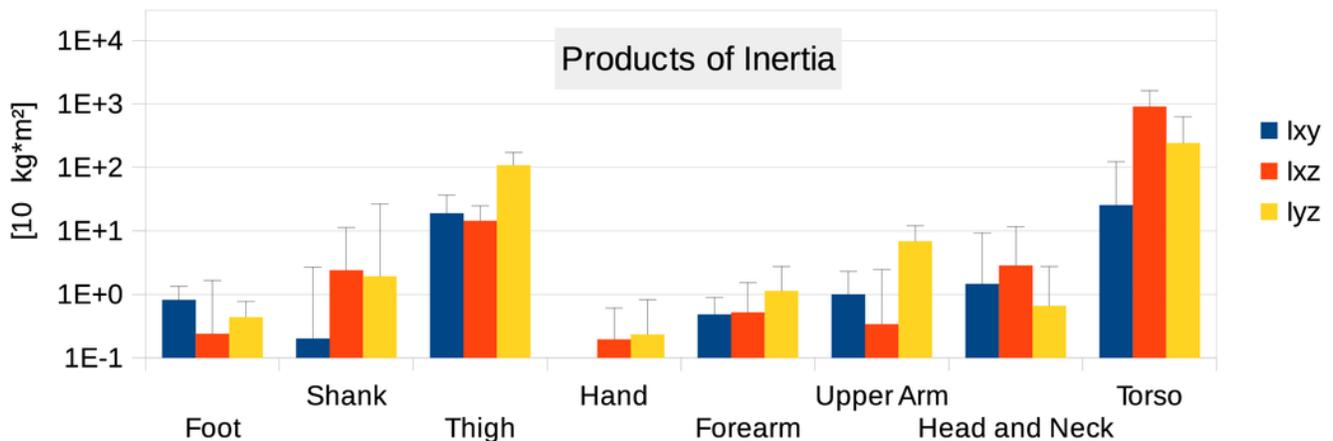
Segment mass (as % of body mass).

P: Average value of all six participants (error bars show standard deviation). Foot mass adjusted by a factor of 0.51 to compensate for volume overestimation due to wearing shoes. Z(m): Male average values reported by Zatsiorsky. Z(f): Female average values reported by Zatsiorsky (Leva, 1996; Zatsiorsky, 2002). D(m): Male average values by Dempster (via Zatsiorsky) (Dempster, 1955; Zatsiorsky, 2002).

4

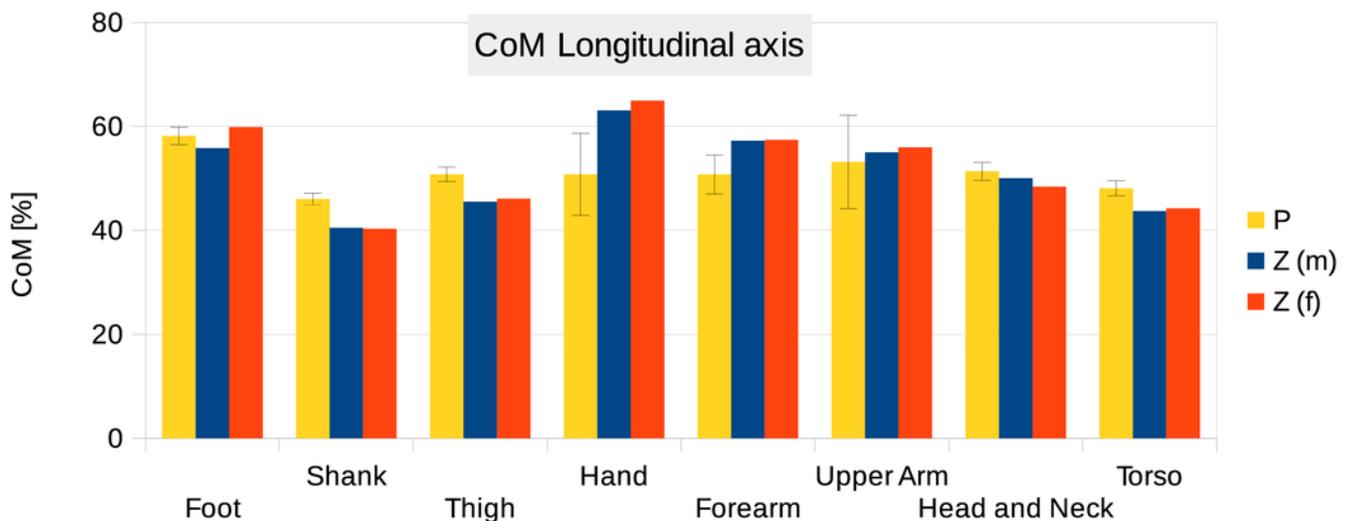
Moment of inertia in $[10^4 \text{ kg*m}^2]$.

P: Average value of all six participants (error bars show standard deviation). Foot moment of inertia adjusted by a factor of 0.51 to compensate for volume overestimation due to wearing shoes. Z(m): Male average values reported by Zatsiorsky. Z(f): Female average values reported by Zatsiorsky (Leva, 1996; Zatsiorsky, 2002). The definition of the coordinate system is shown in Fig. 2.

5

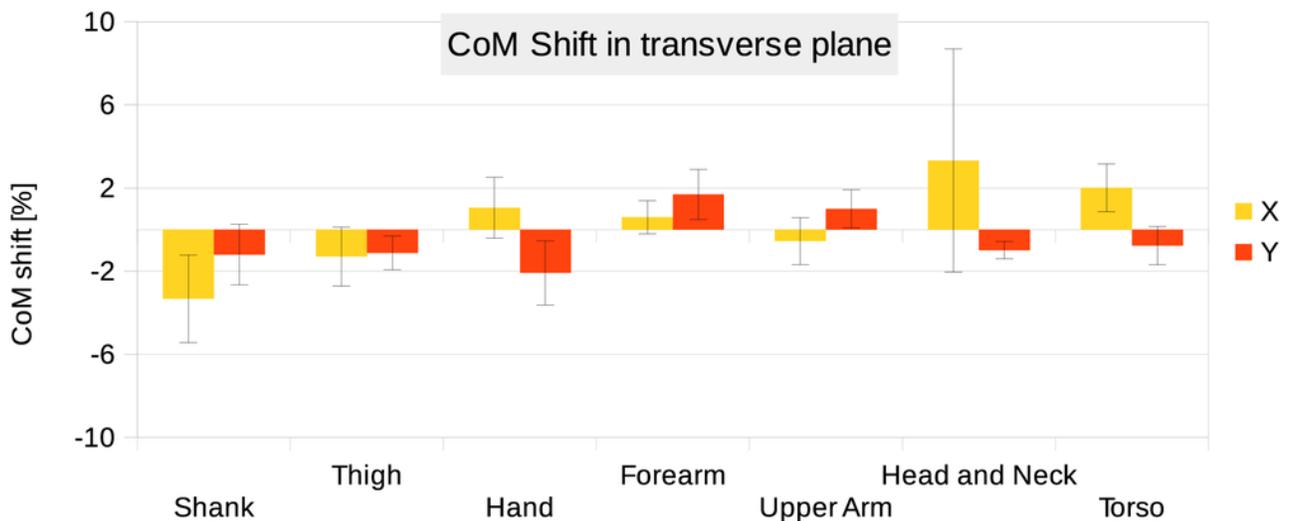
Absolute values of products of inertia in $[10^4 \text{ kg}*\text{m}^2]$.

he absolute values of Ixy, Ixz and Iyz are shown together with a positive error bar (negative error bar is symmetrical) equal to one standard deviation. The signed values are reported in the supporting information in Tables S2-S4. The Ixy value of the hand is smaller than 10^3 kg*m² and is not displayed. Foot products of inertia adjusted by a factor of 0.51 to compensate for volume overestimation due to wearing shoes.

6

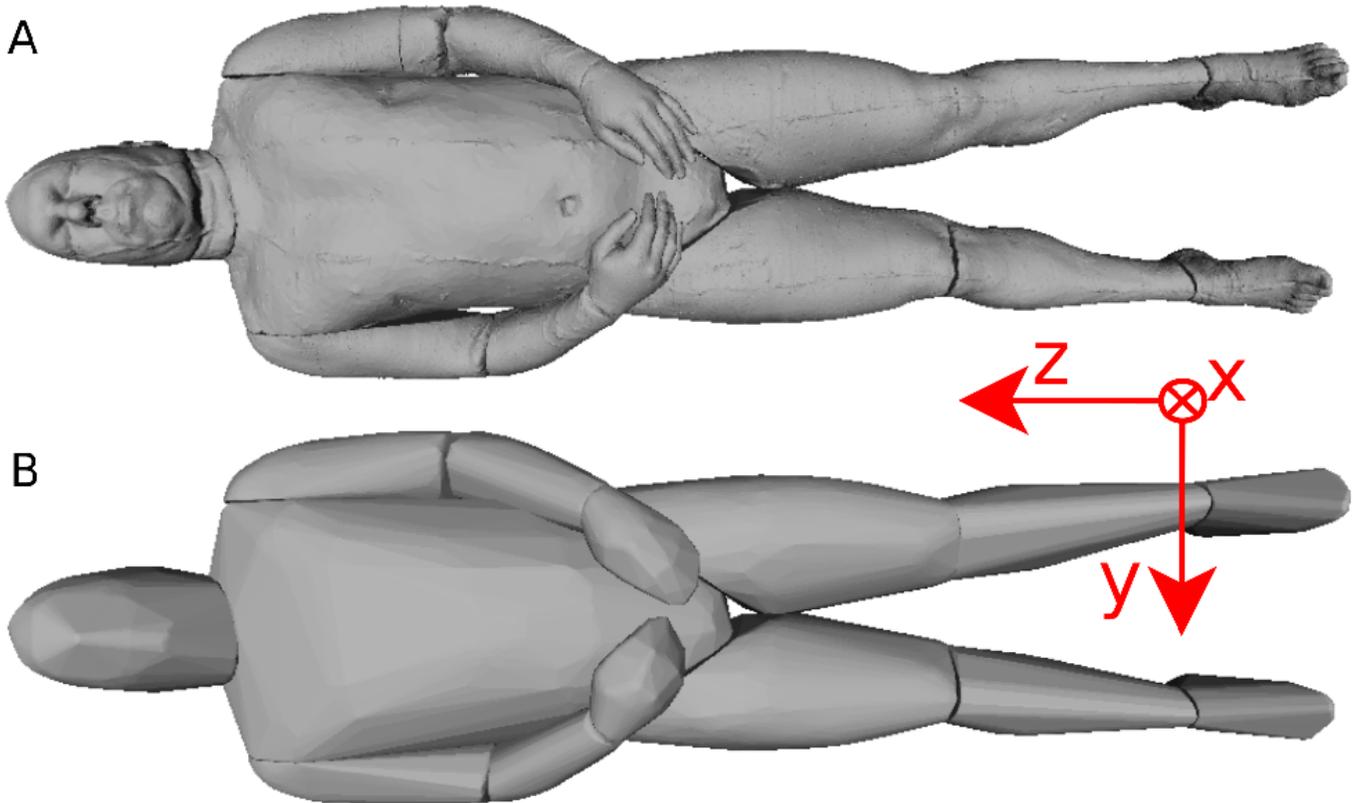
Centre of mass along the longitudinal axis.

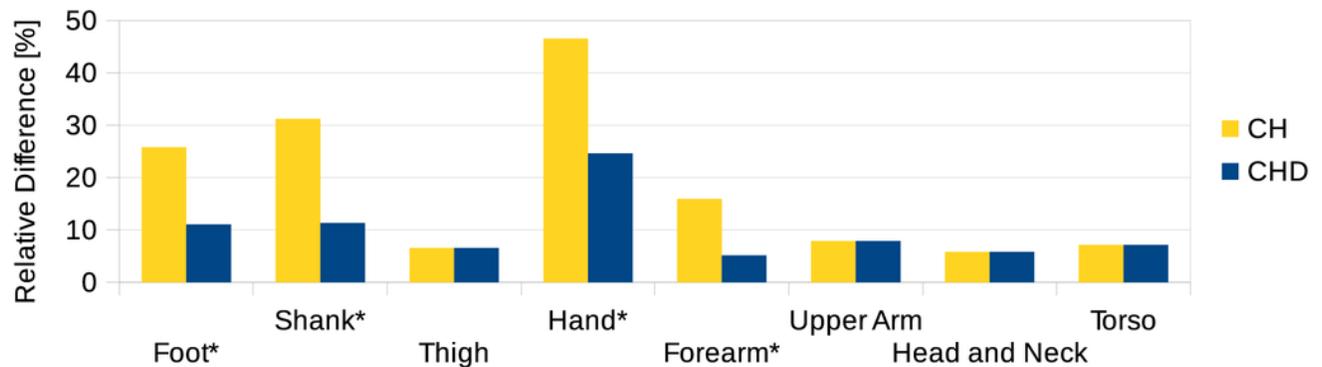
P: Average value of all six participants (error bars show standard deviation). Z(m: male, f: female): Average values by Zatsiorsky, adjusted by de Leva . The CoM is given as % of the segment length. The definition of the segments and reference points are given in the supporting information Table S1 - Exceptions: * Foot of participants: Heel and toe end point of participant's shoes instead of foot. ** Forearm and Upper Arm of Z: Elbow reference point is the elbow joint centre instead of the Olecranon (Leva, 1996; Zatsiorsky, 2002).

7

CoM shift from the anatomical longitudinal axis in the transverse (x-y) plane.

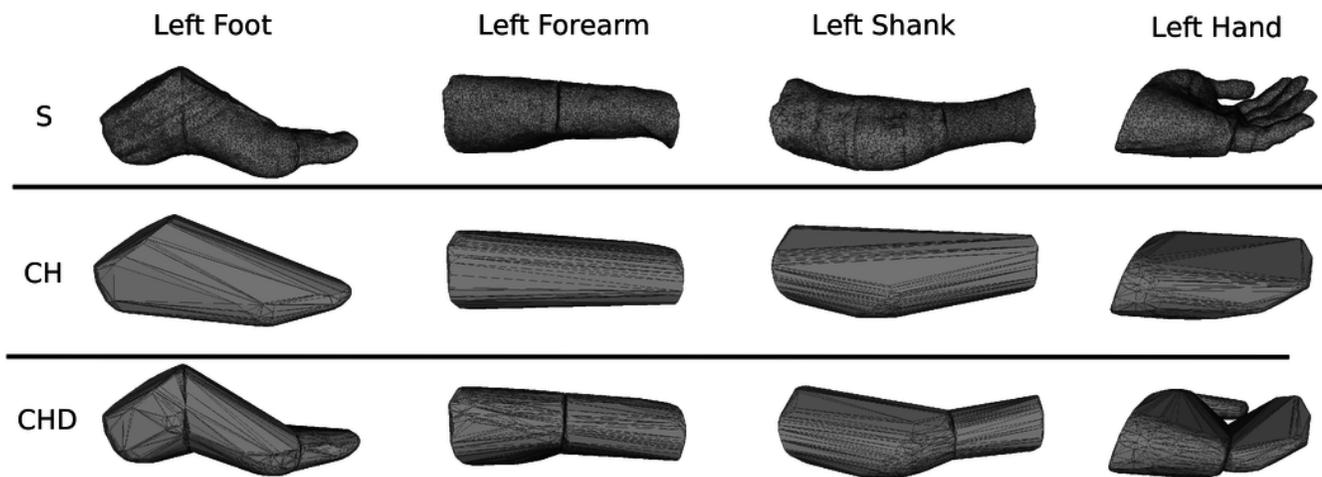
Average values of all six participants are shown (error bars show standard deviation). Due to mirror-symmetry, the y-values of the segments on the left- and right-hand side have opposite signs. To calculate the average, the sign of the segments on the left-hand side was inverted. The CoM is given as % of the segment length. The data of the foot is not included due to the participants wearing shoes.

8

Visible Human surface mesh.

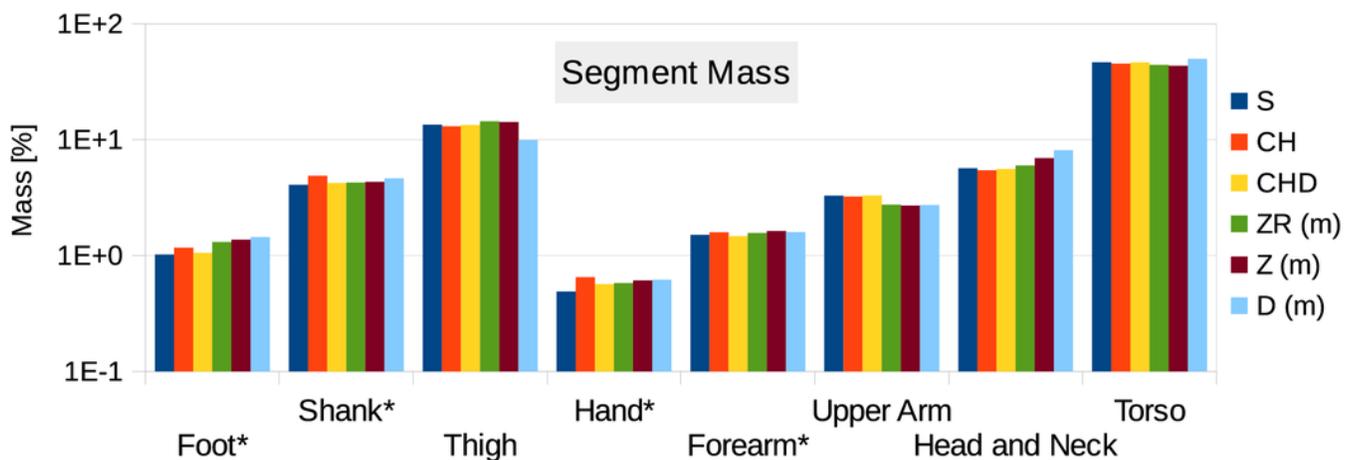
A: High-resolution surface mesh. B: Convex hull mesh.

Segment volume overestimation of the hulled mesh versus the high-resolution surface mesh of the Visible Human.

Data shown as the relative difference of the hull with respect to the original mesh. CH: Convex hull of body segment. CHD: Convex hull of divided body segments (only segments indicated with an * were subdivided, see Fig. 10).

10

Subdivision of the body segments with large curvature.

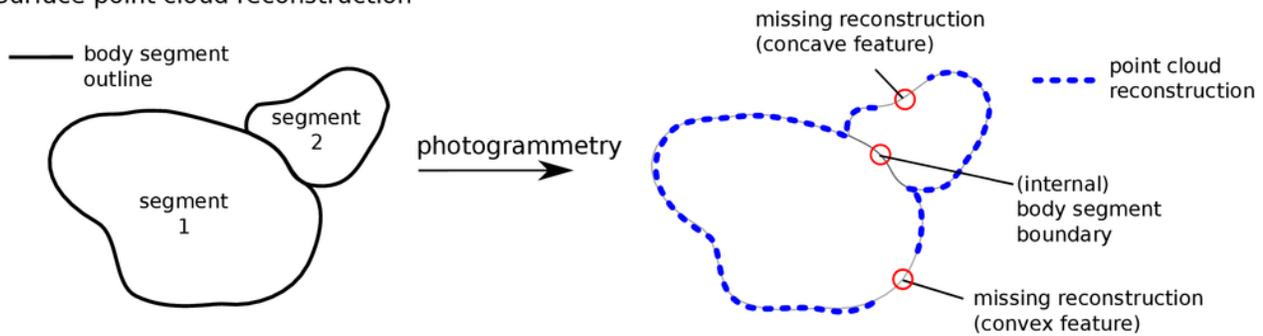
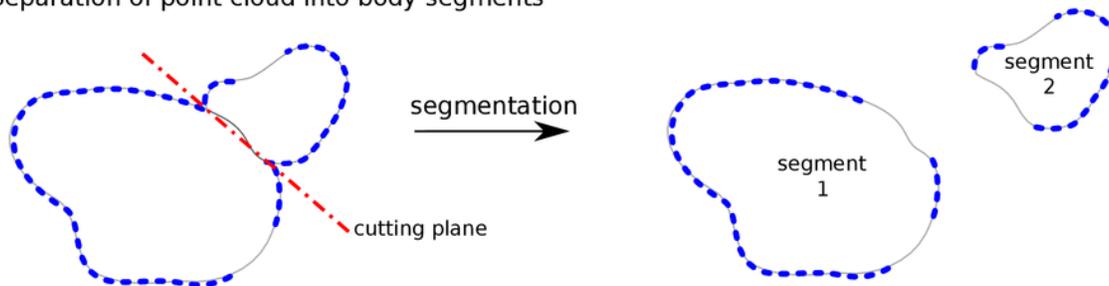
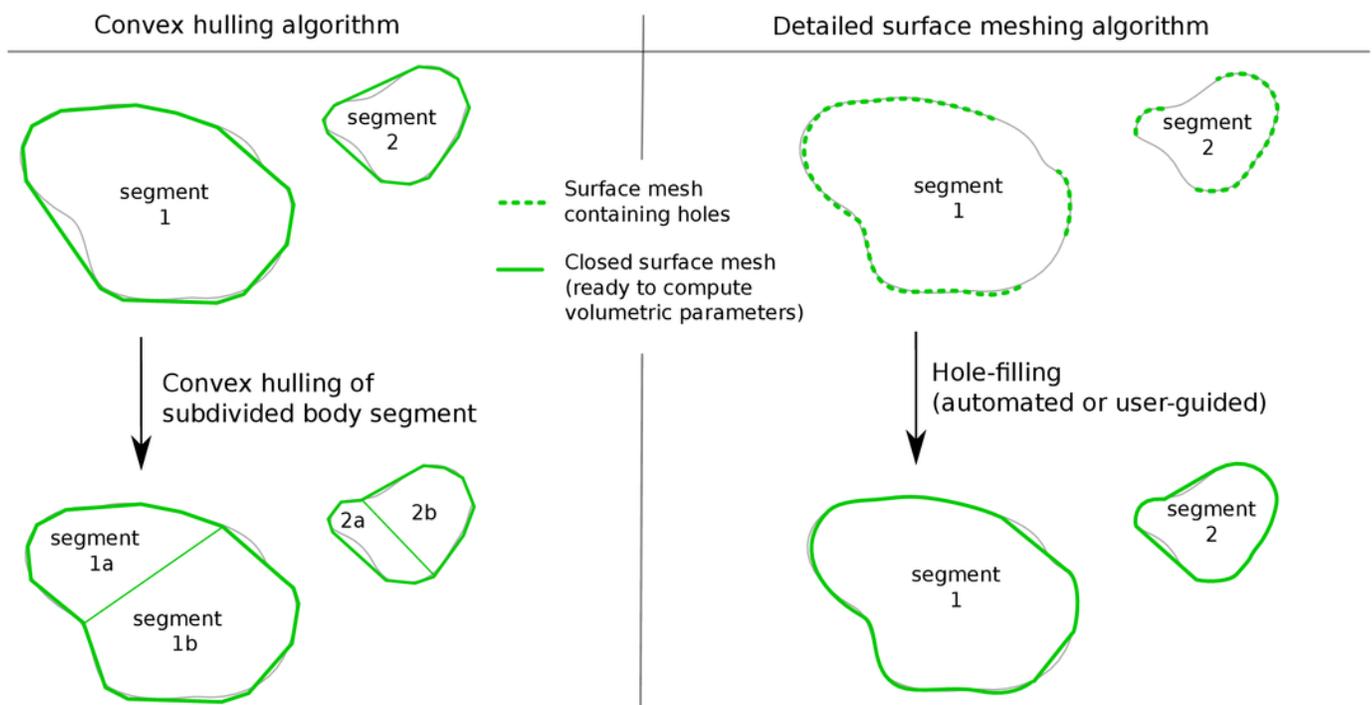
The first row (S) shows the high-resolution surface mesh, the second row (CH) the convex hull of the whole body segment, and the bottom row (CHD) the convex hulls of the subdivided body segments.

11

Male Visible Human segment mass (as % of body mass) of the high-resolution mesh, convex hull, regression model and average values.

S: High-resolution surface mesh. CH: Convex Hull of whole body segments. CHD: Convex Hull with subdivided body segments (only segments indicated with an * were subdivided as shown in Fig. 10). ZR: Values predicted using Zatsiosrky's linear regression model (using weight and height). Z: Male average values reported by Zatsiorsky. D: Male average values reported by Dempster (Dempster, 1955; Leva, 1996; Zatsiorsky, 2002).

12

Methodology to estimate subject-specific body segment parameters using photogrammetry.

A. Photogrammetry; B. Body segmentation; C. Segment hulling; D. Inertial parameter estimation.

A. Photogrammetry:
Surface point cloud reconstruction

B. Body segmentation:
Separation of point cloud into body segments

C. Segment hulling:
Creation of closed surface mesh using hulling algorithms

D. Inertial parameter estimation
Estimation of subject inertial segment parameters using body segment densities and scaling factor

(1) Hull mass = Hull volume x Segment density

(2) Scaling factor = $\dfrac{\text{Subject mass}}{\text{Sum of all hull masses}}$

(3) Segment mass = Hull mass x Scaling factor

(4) Segment inertial tensor =
Hull inertial tensor x Segment density x Scaling factor

Participant mass and weight.

P1 – P6: Participants (m: male, f: female). VH: Male Visible Human.

	P1 (m)	P2 (m)	P3 (m)	P4 (m)	P5 (f)	P6 (f)	VH (m)
Mass [kg]	73.4	77.0	88.2	87.8	65.4	55.2	90.3
Height [m]	1.81	1.83	1.85	1.83	1.65	1.58	1.80