IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received April 4, 2021, accepted April 18, 2021, date of publication April 29, 2021, date of current version May 12, 2021.

*Digital Object Identifier 10.1109/ACCESS.2021.3076595*

# A Review of Body Measurement Using 3D Scanning

**KRISTIJAN BARTOL, (Graduate Student Member, IEEE), DAVID BOJANIĆ, (Graduate Student Member, IEEE), TOMISLAV PETKOVIĆ, (Member, IEEE), AND TOMISLAV PRIBANIĆ, (Senior Member, IEEE)**

Department of Electronic Systems and Information Processing, Faculty of Electrical Engineering and Computing, University of Zagreb, 10000 Zagreb, Croatia

Corresponding author: Kristijan Bartol (kristijan.bartol@fer.hr)

This work was supported by the Croatian Science Foundation under Project IP-2018-01-8118.

**ABSTRACT** The understanding of body measurements and body shapes in and between populations is important and has many applications in medicine, surveying, the fashion industry, fitness, and entertainment. Body measurement using 3D surface scanning technologies is faster and more convenient than measurement with more traditional methods and at the same time provides much more data, which requires automatic processing. A multitude of 3D scanning methods and processing pipelines have been described in the literature, and the advent of deep learning-based processing methods has generated an increased interest in the topic. Also, over the last decade, larger public 3D human scanning datasets have been released. This paper gives a comprehensive survey of body measurement techniques, with an emphasis on 3D scanning technologies and automatic data processing pipelines. An introduction to the three most common 3D scanning technologies for body measurement, passive stereo, structured light, and time-of-flight, is provided, and their merits w.r.t. body measurement are discussed. Methods described in the literature are discussed within the newly proposed framework of five common processing stages: preparation, scanning, feature extraction, model fitting, and measurement extraction. Synthesizing the analyzed prior works, recommendations on specific 3D body scanning technologies and the accompanying processing pipelines for the most common applications are given. Finally, an overview of about 80 currently available 3D scanners manufactured by about 50 companies, as well as their taxonomy regarding several key characteristics, is provided in the Appendix.

**INDEX TERMS** Body measurement, 3D surface scanning, body shape, anthropometry, deep learning.

## I. INTRODUCTION

Anthropometry, a subfield of applied metrology, is the study of how to measure humans. General anthropometry includes the complete process of data collection, documentation, summarization, and analysis [174]. In a narrower sense, anthropometry can be defined as the science of *body measurement*, where lengths, breadths, heights, and circumferences are used to numerically describe body segments and the overall body shape [11]. Body measurement is essential in quantifying the variations in and between populations of different countries, ethnicities, cultures, and ages [26], [137], and it strongly impacts medicine [46], [68], surveying [56], [174], the fashion industry [174], fitness [32], and entertainment [38].

Body dimensions may be measured in various ways, e.g., they can be obtained manually using traditional tools such as calipers and tape measures [174] or automatically using 3D scanners where the measurements are extracted from the obtained data. To ensure both comparability and repeatability, body measurements are standardized via the definition of measurement postures and body landmarks [74], [75]. Although manual measurement is the gold standard, several reports suggested that human expert measurers and 3D scanners achieve comparable accuracies and that the repeatability is generally better for 3D scanners [82], [85], [109]. Another advantage of using 3D scanners over expert measurers is the measurement speed [91]: the duration of an automatic scan is

The associate editor coordinating the review of this manuscript and approving it for publication was Mohamad Forouzanfar.

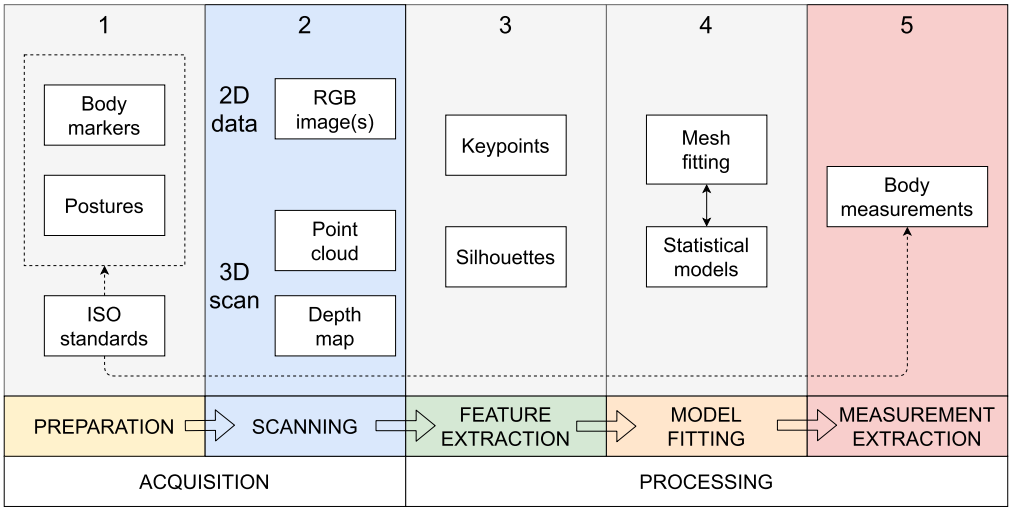



**FIGURE 1.** **The proposed body measurement framework. All measurement methods include the scanning and measurement extraction stages and one or more of the remaining stages.**

often under a few seconds and may go up to 30 seconds for high-quality scans.[1] Therefore, although the first commercial 3D body scanners appeared in the 1990s [39] and were expensive, requiring trained personnel and extensive manual post-processing [114], the scanning technology is currently mature and is comparable in performance to human measurers [24], [64], [175].

There are three commonly used scanning technologies for human body data acquisition: (a) passive stereo (PS); (b) structured light (SL); and (c) time-of-flight imaging (ToF). PS uses images from multiple viewpoints to reconstruct the 3D body surface using the triangulation principle [63]; it fails in the case of low or no texture. SL extends the PS approach by projecting known light patterns, which mitigates the main drawback of PS. In SL, the 3D body surface is reconstructed from the deformations of the projected light pattern [54]. Regarding SL, we distinguish projector- and laser-based methods. In ToF modulated light is projected onto a person and the 3D body surface is directly obtained by measuring the travel time of the modulated light [70]. Considering the multitude of data acquisition and processing methods that have been described in the literature, there exists an increased interest in the topic, which is also substantiated by many large public 3D human body datasets [3], [8], [24], [64], [175] released over the last decade.

In this work we provide a comprehensive review of body measurement based on 3D scanning, starting from a review of 3D scanning technologies and ending by describing the most recent advances in pose and shape estimation. We propose to divide the body measurement processing pipeline into five stages: (1) preparation, (2) scanning, (3) feature extraction, (4) model fitting, and (5) measurement extraction (Fig. 1). In the preparation stage (stage 1), markers that identify standard body landmarks may be placed on the body [3], [24], [166]. The person is asked to take a pre-defined pose [75] and to hold still until the scan ends. Scanning (stage 2) produces a 3D point cloud or depth map(s), along with a set of images, if RGB cameras are used. In stage 3, features such as keypoints and silhouettes are extracted from a 3D scan and images. Based on the features or raw image data [84], in stage 4, the optimal human 3D template mesh[2] is estimated. The primary advantage of fitting the template mesh (a model) to the 3D scan is that any measurement may be easily and conveniently determined from the semantics of the model. Mesh fitting techniques enable the creation of statistical body models, as described in Sec. IV-D. The statistical models enable template mesh regression directly from images and image features. Finally, body measurements are extracted from the processed data (3D scan, images, features, and template mesh) in stage 5. Note that stages 2 and 5 are mandatory, while stages 1, 3, and 4 are optional.

The remainder of this paper is structured as follows: Prior review works on body measurements and on 3D scanning are briefly listed in Sec. II. The three most common scanning technologies (PS, SL, and ToF) are described in Sec. III. The proposed body measurement framework and the five processing stages are introduced and discussed in more detail in Sec. IV. In Sec. V, a methodology for the comparison of the reviewed methods is described, and the methods are discussed w.r.t. their limitations and introduced measurement errors. We also recommend specific scanning technology and the most suitable measurement pipeline for selected anthropometric applications. Finally, Appendix A provides an overview of currently available commercial body scanners, and Appendix B lists currently available mobile applications for body scanning.

[1] See Appendix A for more details.

[2] A template mesh is a graphical model, comprised of vertices and surfaces, which depicts a standard human, usually in a T-pose, with a known number of parameters that control the appearance of the mesh (sex included).

## II. PRIOR REVIEWS

We briefly describe prior reviews in a chronological order. The review covers 3D scanning technology and body measurement.

**3D scanning technology.** One of the first reviews on 3D scanning technology for anthropometry was done by Daanen and Van de Water [39] in 1998, covering 8 commercially available full-body scanners. The most developed scanning technology at that time was laser line-based scanners with vertically moving scanning heads, projecting a horizontal line over the human body. A review on 3D body scanners for the apparel industry [76] (2001) distinguishes laser, LED SL, and white-light SL scanners. Based on their analysis, the scanning time of laser scanners is usually higher than the latter two, but SL scanners have longer data processing times. Olds and Honey [114] claimed in 2005 that structured light 3D scanners using white light are generally cheaper and faster than their laser counterparts, but they produce lower quality scans. A review by D'Apuzzo from 2007 [43] focuses on 3D body scanning technology and its application in the fashion and apparel industry. The paper distinguishes SL and the photogrammetry (passive stereo) approach. ToF sensors were still not commercially used for 3D scanning in 2007. Even though scanning systems were becoming smaller in size, there were no commercial handheld or mobile scanners dedicated to anthropometry. Another review from 2007, by Treleaven and Wells [156], analyzes the 3D scanning technology and methodology for various medical applications, like skin analysis and burn treatment, deformity detection, and obesity treatment.

The updated review by Daanen and Van de Water, from 2013 [40], points out that 3D scanning technology improved in terms of transportation (mobility), speed, price, and accuracy, especially regarding SL scanners. Around that time, ToF scanners appeared on the market. The review focuses on stationary 3D scanners. A book on 3D cameras [56], from 2018, describes and provides in-depth comparisons of ToF, SL and photogrammetry-based (PS) 3D cameras. Finally, a survey by Haleem and Javaid [61] from 2020, similar to the one by Treleaven and Wells [156], is focused on 3D scanning technology in medicine. The difference is that they also take into account X-ray, CT, MRI, and ultrasound, analyze strengths and limitations, and discuss the specific applications of each technology.

**Body measurement.** A body measurement review by Wang *et al.* [164] from 2000 is focused on the measurement and analysis of body length, width, circumference, and skinfold thickness to predict body fat percentage. The main issue in their survey that still has not been completely solved is the lack of standardization in body measurement. A review by Lescay *et al.* [91] compares different anthropometric measurement techniques, including traditional anthropometry, structured light, photogrammetry, and mobile applications, based on precision, the number of measurements, speed, and price. Another review by Heymsfield *et al.* from 2019 [68] describes the process of acquiring 3D human body scans, creating and processing meshes, validating the acquired data, and the applications of the obtained data in anthropometry and medicine. It also distinguishes between SL and ToF scanners in terms of data acquisition techniques and mentions several stationary scanner models. A review by Dianat *et al.* [44] focuses on the methodology and applications of anthropometry in ergonomics. Their paper mostly mentions measurement methods in terms of traditional anthropometry and covers the existing 3D scanning technology on a high level only.

Taking into account prior work on 3D scanning technology, we detect in the reviews a lack of existing handheld and mobile scanners, as well as a lack of reviews of existing mobile applications for 3D scanning and especially anthropometry. To the best of our knowledge, we are the first to provide a complete and modern overview of body measurement based on 3D scans and RGB images.

## III. 3D SCANNING TECHNOLOGIES

Several 3D scanning technologies have been proposed over the years. As mentioned in the Introduction, we distinguish between three common approaches, passive stereo, structured light, and time-of-flight imaging, which we now describe in more detail.

### A. PASSIVE STEREO

Passive stereo is a measuring technique for 3D reconstruction from multiple camera views. Photogrammetry is the science of measuring objects from photographs. Passive stereo and photogrammetry are sometimes used interchangeably in the context of 3D scanning [41], [48], [135]. For clarity, we use the term passive stereo in the remainder of the paper. PS-based 3D scanners use RGB cameras to obtain color images. PS assumes that multiple cameras are pointing to a person. Under passive stereo, in this section, we describe stereo and monocular reconstruction principles, as well as motion capture systems.

**Stereo reconstruction.** The simplest PS configuration is a binocular stereo, a configuration of two horizontally or vertically aligned RGB cameras (see Fig. 2). The reconstruction is based on the correspondences found on the images and triangulation [63]. The point $P$ in the 3D scene projects to pixels $p_1$ in the first image and $p_2$ in the second image (for example, as in Fig. 2). However, for a fixed pixel location $p_1$, the corresponding pixel location $p_2$ is not known a priori. The location $p_2$ is determined by matching an image block around $p_1$ with the most similar block along the epipolar line $l$ [63]. The difference between the corresponding pixel coordinates[3] $|p_1 - p_2|$ (the disparity), is used to triangulate the depth of a point $P$ [63]. The stereo approach can be extended to more than two cameras by coupling pairs of cameras [147] or by using multi-view-stereo techniques [52].

**Monocular reconstruction.** A monocular moving-camera-based 3D reconstruction is a special case of stereo

[3]Note that the images first need to be rectified [101].

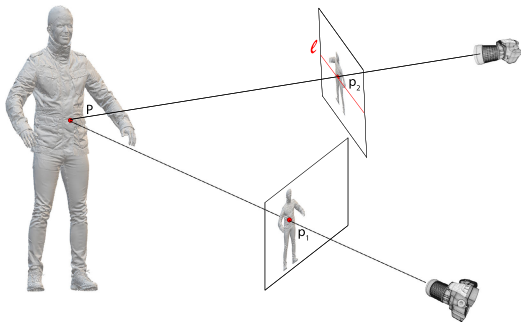



**FIGURE 2.** Passive stereo approach. Point $p_2$ is the most similar image pixel to point $p_1$ along the epipolar line $l$, as described in Sec. III-A.

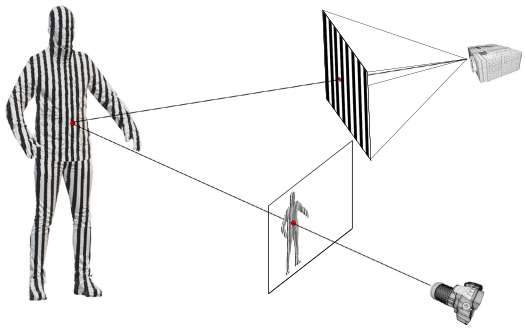


**FIGURE 3.** Structured light (projector based) approach.

reconstruction, where each viewpoint (frame) is considered a separate camera [52]. The general monocular approaches [21], [146], [179] jointly reconstruct a 3D scene and estimate camera locations in every frame. First, the keypoints are detected [10], [104], [138], [139] and matched between the images [36] to find the correspondences. The correspondences are then used for the initial 3D reconstruction and camera parameter estimation, usually followed by bundle adjustment (BA) refinement [157]. Human 3D scanning is usually simpler, as camera locations can be obtained prior to the reconstruction. This is implemented in a way such that either the camera is rotating around the person or the person is standing on a rotating platform, mimicking camera rotation. Note that the person needs to stay still during the quasi-static scanning. The relative camera positions with respect to the subject are extracted based on timestamps. To acquire a dense 3D reconstruction, the principles of stereo reconstruction described above can be used.

**Motion capture.** MoCap is a (semi-) passive stereo technique that uses body markers visible under standard or near-infra-red light. The MoCap markers are usually small, round objects with reflective surfaces. MoCap produces sparse 3D reconstructions and is usually used for motion tracking. The number of body markers is between 30 [72] and 300 [121]. Multiple markers are often used to estimate the location of a single keypoint (joint), as markers can only be placed on the surface of the body.

**Human body scanners using PS.** Commercial 3D scanners use either a rotating monocular system or multiple fixed cameras. For example, Texel Portal MX, Fit3D, and BodyGee Orbiter rotate a person that is standing on a platform, while Texel Portal BX circles around a static body. A few examples of fixed-camera scanning systems are Bootscan Neo, TC$^2$-21B, and 3IOSK by Mantis Vision, which uses from several to more than 50 RGB cameras to obtain the reconstruction. There are several advantages of fixed multi-camera over single-camera scanners. The first advantage is reduced scanning time, because neither the cameras nor the person need to move. The second advantage is the ability to scan people in motion over a period of time, also called 4D scanning (Move4D scanner by IBV). Thirdly, it is possible to reconstruct multiple people at once, if the scanning area is large enough to avoid occlusions, for example, as in Panoptic Studio [78].

Based on the images and the reconstruction described in this section, a mobile device camera can be considered a special case of a monocular PS-based scanner, where a camera is moved around a person to record a video or take individual images. For a comprehensive overview of the commercial 3D scanners, please refer to Appendices A and B.

### B. STRUCTURED LIGHT

To address the poor 3D reconstruction quality of PS in the case of low or repeating texture, the usual approach is to project a textured pattern over the scene. Active stereo (AS) [59], [60], [71], [100] upgrades PS by projecting a light pattern onto the body to improve the correspondence search between views. Structured light approaches [17], [89], [142], [160], on the other hand, search for the camera-to-light-pattern correspondences. In the remainder of this section, we focus on SL technology and methods.

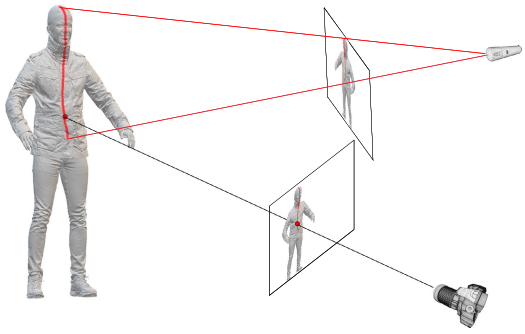
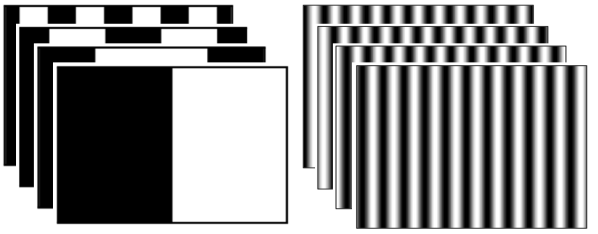


**FIGURE 4.** Structured light (laser based) approach.

**Technology.** We distinguish two scanner types based on SL technology — laser and projector scanners. Laser scanners [9], [47], [100], seen in Fig. 4, use a laser to project dot or stripe patterns over the scene. Laser scanners present sub-millimeter accuracy [22], [161], [174] and a simpler decoding procedure with respect to projector-based scanners [39]. However, laser scanners usually suffer from a slow scanning time, since the laser line needs to sweep the whole body [43]. Projector-based scanners are usually faster than

laser scanners [160], since more complex 2D patterns can be projected and the whole body can be scanned at once from one view. Additionally, projector-based scanners present less safety constraints compared to laser scanners [136]. Even though projector-based scanners are not as accurate as laser scanners, their accuracy range ($\mu$m-mm) is sufficient for high-quality body measurement (see Sec. V).

In general, many classifications of the projected light patterns have been proposed: they may be based on the number of projected patterns (single- or multi-shot), color (achromatic or colored), transitions (discrete or continuous), or structured form (stripes, grids, dot arrays, gradients, etc.) [54], [118], [143], [144], [160], as seen in Fig. 5. For (quasi-) static human 3D scanning, short-duration achromatic multi-shot patterns are usually used, presenting a trade-off between acquisition speed and reconstruction accuracy [160]. For dynamic scenes where fast acquisition is needed (see Sec. IV-B), single-shot patterns are more suitable [81].



(a) Discrete achromatic multi-shot stripe pattern - Gray Code [145]

(b) Continuous achromatic multi-shot stripe pattern - Phase Shift [134]

(c) Discrete colored single-shot grid pattern - M-array [117]

(d) Discrete colored single-shot stripe pattern - De Brujin code [119]

**FIGURE 5.** Structured light pattern examples.

**Reconstruction.** The camera-to-light-source correspondences are found depending on the light and pattern projected. Laser-based approaches mostly use pattern detection algorithms to find the (monochromatic) light projections in the image [50], [53]. Visible-light scanners, on the other hand, have more complex pattern decoding mechanisms [115], [129], especially in the case of multiple projectors and light interference [115], [154]. For more details, we refer readers to the relevant survey papers [144]. After the correspondences have been obtained, ray-to-ray or ray-to-plane triangulation can be applied [54], [56], [89], [100] to reconstruct the 3D human body.

**Human body scanners using SL.** Commercial SL scanners either rotate around a person or have a fixed multi-sensor configuration that surrounds them. Stationary scanners, such as the HP Pro S3, 4DDynamics EOS, TC$^2$-105, or Hexagon Aicon Primescan, rotate around the body to obtain a whole 3D scan. Another way to move around the body is to use hand-held scanners, such as the Artec Eva, TechMed3D BodyScan Scanner, Mantis Vision F6 Smart or ScanTech Axe B17. Stationary scanners with fixed sensor positions, such as the Artec Shapify Booth, botscan Neo, botscan OptaONE+, TC$^2$-105, and 4D Dynamics IIID Body Scan, showcase a booth filled with cameras and projectors in fixed positions that surround the scanned subject. Solutions to avoid light interference [163] from multiple projectors have been proposed, but in practice, every projector illuminates the subject in its designated time interval. Hence, the acquisition time is prolonged and proportional to the number of scanners. For a comprehensive overview of the commercial 3D scanners, please refer to Appendix A.

### C. TIME-OF-FLIGHT

ToF scanners, shown in Fig. 6, measure the time needed for an emitted light signal to travel from the illumination source to the 3D scene and back to the sensor. The distance information is directly proportional to the time of flight of the light signal [51], [56], [70], [92].

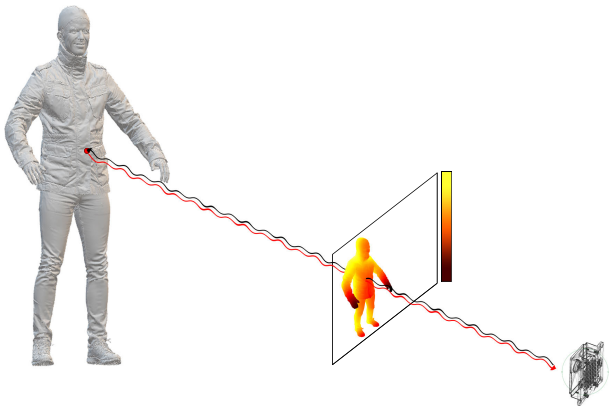


**FIGURE 6.** Time-of-flight approach. The black arrow indicates the emanated light signal path. The red arrow indicates the received light signal path.

**Technology.** The main components of a ToF scanner are the light emitter and the photodetector [56]. The light emitter uses a laser or an LED to send a modulated beam of light, typically in the NIR range [92]. The lens is used to spread the light from the emitter over the whole scene. The photodetector usually uses a matrix of point-wise sensors [70]. For human 3D scanning, CCD/CMOS matrix sensors are usually used.

**Reconstruction.** Two reconstruction methods can be distinguished: pulsed-light (direct) and continuous-wave (indirect) [56], [70]. Continuous-wave (CW) methods indirectly measure the round-trip time of an emitted light pulse and collect the time-dependent intensity information of the signal [51], [126]. The distance of a point is then

**TABLE 1.** Main properties of the three 3D scanning technologies with respect to human body scanning.

| | Passive stereo | Structured light | Time-of-flight |
|---|---|---|---|
| Method | triangulation | triangulation | time-to-distance conversion |
| Illumination | passive (ambient) | active (visible, IR) | |
| Scanning range | several meters | $< 5$ m (illumination source limited) | |
| Dynamic scanning | yes | yes (slower movement only) | yes |
| Accuracy range | mm - cm | $\mu$m - cm | mm - cm |
| Resolution range | mm | $\mu$m - mm | mm |
| Main issues | textureless body parts | light interference | lower resolution, multi-camera interference |

retrieved (demodulated) from the phase shift of the emitted and received light signals by their cross-correlation [70], [130]. The emitted illumination signal amplitude is usually modulated using a sine or square wave [18]. The periodicity of the waves implies a maximum scanning range at half of the modulation wavelength, after which an ambiguity problem arises [65]. Increasing the modulation frequency increases the measurement accuracy but shortens the maximum range [70]. The range can be extended using multiple modulation frequencies [57], [125]. Fortunately, this does not usually present a problem in the anthropometry application, since human bodies are scanned from close range. Pulsed-light (PL) methods directly measure the round-trip time of an emitted light pulse using time-to-digital (TDC) or time-to-amplitude (TAC) circuitry [70], [126]. Since the speed of light is very fast, PL methods require extremely precise timing information, on the order of picoseconds, to obtain a millimetric distance range [56], [92], [126]. Hence, PL is not usually used for 3D body scanning.

ToF cameras present low-cost, compact-size, accurate, and reliable sensors with lower power consumption [51], [65], [67]. Compared to SL, ToF does not have a spatially separate light source and camera, avoiding occlusion problems between views. Additionally, it is texture-independent, with a minimal post-processing time and lower-light capabilities [151]. Even though fast frame rates that are suitable for dynamic scanning can be achieved [51], [151], the biggest problem of single ToF camera scanners is a low scanning resolution [51]. It is possible to increase the resolution by using multiple ToF cameras [167], but complex light interference issues then need to be addressed [130]. Therefore, ToF is still less applicable for (quasi-) static scanning and body measurement.

**Human body scanners using ToF.** Most of the commercial human body scanners, such as the SizeStream SS20, Styku S100, and TC$^2$-30R, are based on indirect ToF methods. In general, ToF as a standalone solution is unable to provide high-quality 3D human body scans due to its lower resolution. Hence, it is usually used in combination with RGB cameras. Noticeably, a bigger percentage of stationary scanners, such as the TC$^2$-19R, Naked scanner, and BodyGee Orbiter, come with a turntable on which subjects take a standard scanning position. This alleviates the problem of light interference caused by having multiple cameras. Note that all mini scanners are ToF-based and therefore used for 3D data acquisition in mobile applications (see Appendix B). For a comprehensive overview of the commercial 3D scanners, please refer to Appendix A.

### D. SUMMARY OF SCANNING TECHNOLOGIES

A comparison of the three scanning technologies is provided in Table 1. Regarding methodology, PS and SL rely on finding the correspondences between the views to triangulate 3D points in space, while ToF uses time-to-distance conversion and thus avoids the correspondence search problems. The common challenge for the triangulation approaches is the potential (self-) occlusions between the views, which might result in holes in the 3D point cloud [52]. A way to cope with these occlusions is to use more cameras or viewpoints (achieved by rotating the subject or the scanner) and to use the T-pose where self-occlusions are mitigated.

SL and ToF use light sources. In one way, this helps SL in low-textured body areas, but it also limits its applications to specific indoor lightning conditions. For multi-ToF scanners, light causes interference problems. Regarding the scanning ranges, SL and ToF are limited by the illumination source. The PS scanning range is, in theory, only limited by the optics, but in practice it is several meters. All the scanning ranges are suitable for human body scanning.

With regard to the scanning of moving subjects (dynamic scanning), PS is the most suitable because of its fast acquisition time, good overall reconstruction performance, and lack of light interference issues. ToF has a high reconstruction frame rate, making it applicable for dynamic applications [62], [152]. SL can also be used for dynamic scanning with single-shot patterns, but for scanning slower movement only. Moreover, single-shot patterns offer a lower reconstruction accuracy compared to multi-shot patterns.

Finally, SL offers the best accuracy[4] and resolution,[5] making it the method of choice for quasi-static scanning and body measurement. This can also be seen in the number of commercial SL scanners.[6] PS and ToF have similar accuracies and resolution ranges (see Table 1), but ToF generally has a lower resolution.

[4]Accuracy is the distance between the reconstructed location and true location of a 3D point in space.

[5]Resolution is the minimal distance of two points in space that can be differentiated and reconstructed.

[6]See Appendix A for more details.

## IV. BODY MEASUREMENT

We describe our proposed body measurement framework (Fig. 1), dedicating subsections to each of the five processing stages: preparation, scanning, feature extraction, model fitting, and measurement extraction. The first two stages are the acquisition stages, and the latter three are the processing stages (see Fig. 1). In the acquisition stage, the subjects are prepared and the data in the form of 3D point clouds, depth maps, or 2D images are obtained. There are two acquisition protocols — static and dynamic. In the processing stage, the collected input is used for body measurement. Body measurement can be done directly on the given inputs, but usually the features are extracted first, and the body model is fitted based on these features or the inputs.

### A. PREPARATION

The standardization of body landmarks, measurements, and postures is the first step to ensure the comparability of measurements between the body measurement surveys [174] and to compare the scientific results. Body landmarks represent the same semantics for every measured subject (Fig. 7A) and some of the body measurements can be directly derived from landmarks (see Sec. IV-E). The landmarks are defined on the skin to reduce the ambiguity in their locations between the subjects. In practice, markers that represent the landmarks are manually placed on the human body. The markers are useful in the feature extraction (stage 3); however, marker placing is a tedious and error-prone process, so successful marker-less systems have been proposed [79], [106].

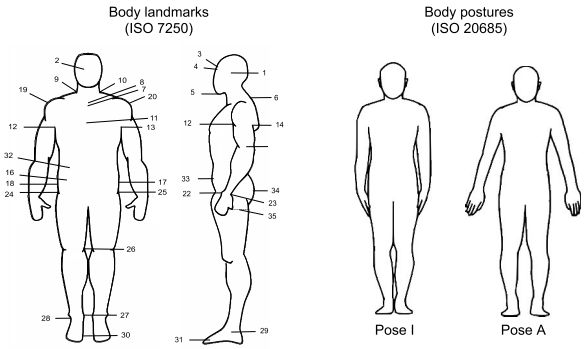



**FIGURE 7.** Body landmarks according to ISO 7250-1:2017 [75] standard and two standing postures according to 20685-1:2018 [74]. Left axilla point posterior (#12) is not shown.

**Standardization.** ISO standard 7250-1:2017 [75] specifies a list of body landmarks and measurements. The complete list of body landmarks is given in Table 2 and the corresponding points are shown in Fig. 7 (left). The complete list of body measurements is given in Table 3 and the corresponding Fig. 11. There are two standard standing poses recommended by the ISO standard 20685-1:2018 [74] (Fig. 7). The person is asked to take one of the two poses, hold their breath during the scanning, and try to keep as calm as possible [106]. In the first pose (I-pose), the subject stands upright with the shoulders relaxed and the arms hanging down naturally and the feet together. In the second pose (A-pose), the feet are 20 cm apart, the arms form a 20° angle with the torso, the elbows are straight, and the palms face backward [174]. Using the standard or fixed body postures is not always required for body measurement, but usually it is when creating datasets that capture shape variations [3], [8], [12], [24], [64], [72], [175]. Another pose that is also often used for scanning is the T-pose, as seen on a neutral template mesh in Fig. 9.

**TABLE 2.** The list of human body landmarks according to ISO 7250-1:2017 standard [75]. The numbers correspond to the numbers in Fig. 7 (left). The letters *R* and *L* abbreviate right and left.

| Human body landmarks (ISO 7250) | | |
|---|---|---|
| 1 tragion | 12 axilla pnt. post. RL | 23 supratarsale fib. R |
| 2 orbitale | 13 iliocristale RL | 24 metatarsale tib. R |
| 3 glabella | 14 iliospinale ant. R | 25 metatarsale fib. R |
| 4 sellion | 15 acromiale RL | 26 waist level |
| 5 gnathion | 16 radiale R | 27 stylion ulnare R |
| 6 cervicale | 17 stylion R | 28 stylion ulnare R |
| 7 suprasternale | 19 stylion ulnare R | 29 trochanterion RL |
| 8 front neck | 20 trochanterion RL | 30 abdom. ext. level |
| 9 side neck | 21 tibiale R | 31 buttock pnt. R level |
| 10 mesosternale | 22 sphyrion R | 32 gluteale R level |
| 11 axilla pnt. ant. RL | 22 sphyrion fib. R | |

### B. SCANNING

Regarding the acquisition protocol when using 3D scanners, the human body may be measured in a stationary position [3], [6], [8] or in motion [2], [108], [158]. In static scanning, a person is asked to take a pre-defined pose and to hold still until the scan ends. For 3D scanners that have longer acquisition times, e.g., scanners with rotating heads or handheld scanners, subjects may unintentionally move during acquisition, which introduces artifacts, so we may distinguish such situations as quasi-static scanning. Static scanning is the method of choice to obtain the most precise body measurements and is routinely used in the production of relatively large and diverse public 3D human body datasets [3], [8], [24], [64], [175]. Scanning in motion usually limits the technology to either PS or ToF. The most common systems are motion capture (MoCap) systems [12], [103], [107], [108], [158], which are PS-based and use markers attached to the body to track movement. Other dynamic 3D scanning systems [2] record a person in motion to analyze soft-tissue deformations over time [128].

Scanning usually produces a 3D point cloud, one or more depth maps, or a set of RGB images. In the case of dynamic scanning, so-called 4D scans are obtained [2]. In the processing stage, some or all of these data are used to extract the measurements.

### C. FEATURE EXTRACTION

Two types of features that are usually extracted from 3D scans and images are keypoints and silhouettes. The locations of keypoints[7] can be determined based on markers or can

[7]Note that keypoints are called landmarks if they refer to standardized body locations [75].

**TABLE 3. An example list of 44 standardized human body measurements [75]. The measurements consist of distances (lengths, breadths, depths, and heights), circumferences, and soft biometrics (weight, height, BMI).**

| Human body measurements | | | |
|---|---|---|---|
| 1 eye | 12 forearm circum. L | 23 weight | 34 bicep circum. R |
| 2 cervicale | 13 forearm circum. R | 24 height | 35 shoulder breadth |
| 3 shoulder-elbow L | 14 neckbase breadth | 25 BMI | 36 elbow circum. L |
| 4 shoulder-elbow R | 15 thigh clearance | 26 neck circum. | 37 elbow circum. R |
| 5 crotch height | 16 wall-acromion distance | 27 chest circum. | 38 knee circum. L |
| 6 tibial height | 17 grip and forward reach | 28 waist circum. | 39 knee circum. R |
| 7 chest depth | 18 elbow-wrist L | 29 thigh circum. L | 40 neck base circum. |
| 8 body depth | 19 elbow-wrist R | 30 thigh circum. R | 41 neck circum. |
| 9 thorax depth | 20 hip circum. | 31 calf circum. R | 42 head circum. |
| 10 chest breadth | 21 buttock-popliteal | 32 calf circum. R | 43 trouser waist circum. |
| 11 hip breadth | 22 buttock-knee | 33 bicep circum. L | 44 iliac spine breadth |

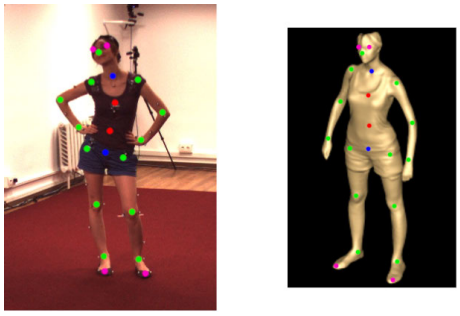


**FIGURE 8. An example of 2D keypoints and their corresponding 3D scan keypoints. The typical keypoint extraction algorithm finds between 13 (green) and 21 keypoints (other colors). The blue keypoints represent the neck and pelvis, the red ones two spinal points, and the pink ones details on the face and feet. The images are adapted from the Human3.6M dataset [72]. Note that the keypoints are manually annotated and do not necessarily reflect the H36M dataset ground truth locations.**

be estimated automatically from a 3D scan [79], [106]. The keypoints usually represent a selected subset of human joints (see Fig. 8). Silhouettes may represent 3D points or pixels of the whole human body or body segments.

**Keypoint extraction.** Most of the keypoint estimation algorithms detect human joints from single or multiple images. The joints can be represented by 2D pixel coordinates in an image or 3D points in a scene. If there is a moving person, the time component can be exploited, and temporal smoothness can be applied to improve the estimation accuracy [124]. Therefore, keypoint estimation methods can be divided into single-image [14], [30], [31], [111], [176], multi-frame [124], and multi-view methods [66], [73], [133] for 2D [14], [30], [31], [153], [176] or 3D [66], [73], [111], [124], [133] keypoint estimation. The keypoint estimation algorithms usually find between 14 and 21 keypoints, as shown in Fig. 8. Most of the state-of-the-art keypoint estimation methods are deep learning-based, due to the availability of large, annotated datasets [72], [110], [132], [148], [162]. In practice, the extracted 2D and 3D keypoints are used for mesh fitting [23], [83], [116] (see the next subsection) and are typically not combined with landmark extraction from 3D scans.

Motion capture [158] is a movement tracking technique that enables the direct acquisition of precise ground truth 2D and especially[8] 3D keypoint locations. Most of the previously mentioned keypoint estimation algorithms take advantage of the ground truth data obtained using motion capture. Examples of MoCap datasets are the Human3.6M [72] (Fig. 8), HumanEva [148] (Fig. 10), and TotalCapture [158] datasets. The disadvantage of motion capture systems is that they are impractical for in-the-wild scenarios.

Regarding keypoints from a 3D scan, Lu and Wang [106] proposed a system for marker-less 3D scan keypoint detection. A body scan is firstly cleaned by removing the outlier points, and then it is segmented into five parts: head and torso, left arm, right arm, left leg, and right leg. The initial keypoint locations are derived from the anthropometric database [165] and then refined using four algorithms: silhouette analysis, minimum circumference determination, grayscale detection, and human body contour plots. The results of the four algorithms are combined to determine the final keypoint locations and body measurements.

**Silhouette extraction.** Silhouette extraction methods separate pixels that represent an object of interest (the human body) from other pixels in an image [15]. There are three approaches to silhouette extraction: background subtraction [19], semantic segmentation [168], and multi-view segmentation (visual hull) [90]. In the work by Lin and Wang [99], two silhouettes are extracted using background subtraction, from front and side input images, and 60 feature points in total are detected on the edge of the silhouette, based on the curve distance between them [98]. The extracted feature points are directly used for approximate body measurement extraction (see Subsec. IV-E).

State-of-the-art semantic segmentation methods [33], [49], [93], [96], [171], similar to human pose estimation, are also deep learning-based. In addition to whole-body segmentation [97], [112], there are also body-part-segmented datasets [95], [178]. Both whole body and body part segmentation problems are particularly interesting in terms of silhouette and body measurement extraction, as they achieve a relatively high accuracy,[9] even on difficult examples. A visual hull is

[8]Another way to obtain 2D pose estimation data is to manually label human joints on a large number of images. However, this is impractical and unreliable in the case of 3D data.

[9]The accuracy is measured as a mean IoU (intersection over union).

reconstructed by applying background subtraction or semantic segmentation for multiple images of a fixed object from different views [52]. A visual hull can be used as an initial solution for mesh fitting.

### D. MODEL FITTING

Model fitting is a set of techniques for finding a 3D template mesh that best represents a given input. The given input can be a 3D scan, 2D or 3D keypoints, or silhouette(s). The advantage of using template meshes in the context of body measurement estimation is that the number of vertices is fixed, and corresponding vertices have the same semantics for all the registered meshes in the dataset. Once body measurements are obtained for a single mesh, they can be obtained in the same way for all the meshes. We distinguish two model fitting techniques — mesh fitting (registration or deformation) and mesh regression using statistical models. In this Subsection, we describe mesh fitting, statistical model creation, and mesh regression from 3D scans and images.

**Mesh fitting.** Mesh fitting is an optimization process deforming an initial, template mesh to the 3D scan.[10] Mesh fitting consists of pose and shape fitting [12], [69], [102], [116], [123]. Before the optimization, a 3D scan is usually subsampled so that the number of points is the same or larger than the number of vertices in the template mesh [12], [172]. First, the landmarks are used to roughly align the 3D scan and mesh [127]. Next, pose fitting is done by rigging [16] the body skeleton parts of the template mesh and then skinning the surface points, using linear blend (LBS) [12], [102] or dual quaternion skinning (DQS) [116]. Once the pose satisfies the convergence criterion, shape fitting is done using a non-rigid registration, minimizing a loss function that usually consists of three components: a landmark term, a smoothness term, and a data term. The landmark term minimizes the distance between the corresponding landmarks of the template mesh and the 3D scan. The smoothness term minimizes the difference between the spatial transformations of the neighboring vertices. Finally, the data term minimizes the distances between the corresponding vertices. Note that the correspondence is determined at the beginning of a shape fitting phase. Pose and shape fitting is usually done alternately multiple times, until the final convergence [127]. Some works also take texture into account [24], which improves fitting. The described fitting is a method of choice for almost-complete 3D scans, obtained using high-quality scanners. For partial 3D scan fitting, a method based on implicit functions [37] has shown promising results on the SHARP 2020 (SHApe Recovery from Partial textured 3D scans) challenge [141]. The result of fitting is a clean mesh that fills up the holes in the original, noisy 3D scan.

**Statistical models (SMs).** Statistical models represent the population of human bodies with respect to pose and shape variations, usually represented by the principal components (PCs). To create a statistical model, a mesh fitting procedure needs to be applied to each scan in a dataset. The work by Hirshberg *et al.* proposed simultaneously fitting meshes while creating the body model [69]. One of the advantages of the simultaneous fitting and creation of the model is that the occluded 3D scan regions are properly fitted based on the scans of different poses where these regions are not occluded. To describe pose and shape variations in the set of fitted template meshes, principal component analysis (PCA) is used. The purpose of PCA is to compress the dataset of registered meshes by finding pose and shape principal components that explain the maximal variance of the dataset. An important advantage of PCA is that the PCs can be used to generate novel template meshes [107], [122], [149], [177] from a pose-shape parameter space. The datasets commonly used for building SMs are CAESAR [3], Size-UK [8], ScanDB [64], and possibly other datasets containing 3D scans [12], [24], [72], [148], [175].

SCAPE [12] is the first SM for both pose and shape deformations, as well as pose-dependent shape changes (for example, muscle contractions in different poses). They use a set of initial, physical markers and the correlated correspondence algorithm [13] to generate around 150 additional markers. Then they apply non-rigid registration to obtain the articulated human model. One of the main disadvantages of SCAPE is that each body part is independently rotated, which introduces artifacts near joints. To that end, BlendSCAPE [69] smooths SCAPE body part segmentations across part boundaries, which solves the artifacts problem. A disadvantage of both BlendSCAPE and SCAPE is that they use triangle deformations for the PCA. One of the most popular statistical models, SMPL [102], showed that using vertex instead of triangle transformations improves the final SM. SMPL also enforces body symmetry to produce models that are visually more pleasing for animation. Enforcing the symmetry, however, sacrifices realism in particular poses. An improvement over SMPL is the STAR [116] model, which enforces spatially local and sparse pose corrective blend shapes and is independent of the symmetry optimization component. STAR is the most expressive SM, partly due to the fact that it is built using the largest database, a combination of the CAESAR [3] (4000 scans) and SizeUSA [6] (9000 scans) datasets.

**Mesh regression from 3D scans using SMs.** Once a statistical model is built, it can be used for mesh regression. The idea of mesh regression is to find the pose and shape parameters of the SM that best fit a given input. An example of such an approach is done by Kwok *et al.* [88], consisting of iteratively selecting the mesh from the statistical pose-shape space and fitting the clothes to match the input 3D scan. Prokudin *et al.* [131] propose a deep learning model for template fitting, supervised by SMPL templates fitted to the dataset before learning. The learning is based on the distances between the set of 3D scan features, called the basis point set, and the ground truth template mesh. The advantage of using the features to find the optimal parameters is that the (slow)

[10]For simplicity, we only describe mesh fitting on 3D scans, but similar techniques can be applied to features or images [84].

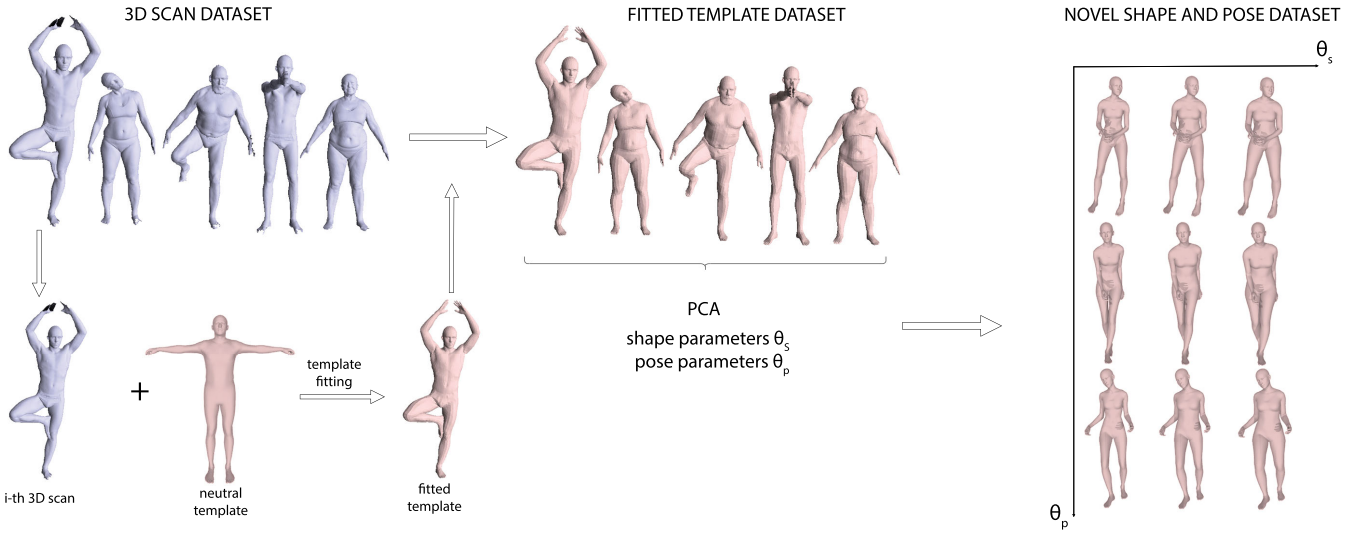



**FIGURE 9.** An overview of a mesh fitting process for the creation of statistical models. For every 3D scan in the scanning dataset, a neutral template mesh is registered to the scan, producing a dataset of registered template meshes. Based on the shape and pose variations of the registered templates, PCA can be applied to create the statistical model. The principal components can be used to generate novel 3D meshes from the pose-shape space. The 3D scans and template meshes are retrieved from the FAUST dataset [24]. The novel 3D meshes are generated using SMPL-X [123].

**FIGURE 10.** An example of a SMPL-X [23] mesh regression approach based on 2D keypoint estimation. Note that yellow keypoints represent the hands and face, which are usually modelled separately. The image is adapted from [123].

rendering step that is needed to verify the parameters is avoided.

**Mesh regression from images using SMs.** There is a group of methods that use extracted image features (body pose or silhouette) or RGB images directly and exploit the SMs for mesh regression. A large body of these methods are based on the SMPL statistical model [23], [84], [150]. For example, SMPLify [23] is a deep learning model for 3D shape and pose estimation from 2D keypoints. The keypoints are detected using a 2D pose estimation algorithm [31]. Using sex-specific SMPL models, SMPLify simultaneously estimates 3D pose and shape parameters and produces a template mesh (see Fig. 10). The main disadvantage of the SMPLify approach is that it does not exploit image information. A multi-task learning approach by Smith *et al.* [149] uses front- and side-view silhouettes and feeds them into a convolutional model to estimate 3D joints, mesh volume, shape parameters, and pose angles (the angles between the adjacent joints), using SMPL as an underlying statistical model. The results, as seen in Sec. V, show that silhouette-based approaches can be used to obtain accurate body measurements. However, the major issue of silhouette-based approaches is the clothed-people scenario, where it is difficult to estimate the underlying body shape. A recent method by Kolotouros *et al.* [84] uses raw pixels and deforms an initial mesh based on a graph CNN [169]. The most similar mesh from the SMPL pose-shape space can then be matched to the deformed mesh. Note that graph CNN approaches can also be interesting for mesh-from-3D-scan regression.

### *E. MEASUREMENT EXTRACTION*

Body measurements can be extracted from a 3D scan, template mesh, or image features. We focus on two measurement types — distances (lengths, breadths, depths, and heights) and circumferences. For other measurement types, such as surface area measurement, we refer readers to [155]. A subset of the standardized body measurements [75] is listed in Table 3.

**Measurements from the template mesh.** When the fitted or regressed template mesh is obtained, the number of vertices is known, and their semantics are the same across all samples [102]. To calculate distance measures, for example, elbow-wrist, hip breadth, or chest depth, the distance between the semantically corresponding vertices can be used. The circumferences, for example, the waist, thigh or calf circumference (see Fig. 11, left), can be calculated as the extent of an intersection between the mesh and a plane.

**Measurements from the 3D scan.** The measurements can also be extracted directly from a 3D scan. The landmarks can help to obtain distances and some of the circumferences [166]. In the work by Lu and Wang [106], the circumferences are calculated from a point cloud using a convex hull

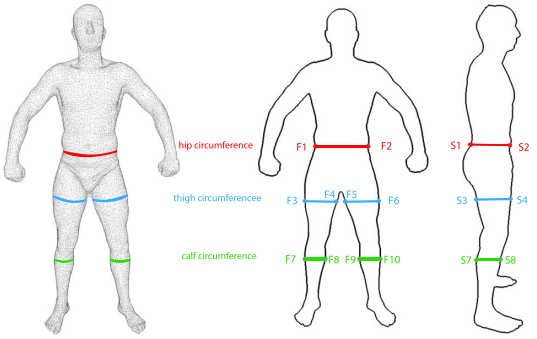



**FIGURE 11.** Body measurements on a 3D mesh (left) and corresponding feature points on front- and side-view silhouettes (right). The feature points can be used to approximate the measurements. The mesh is generated using the SCAPE model [12].

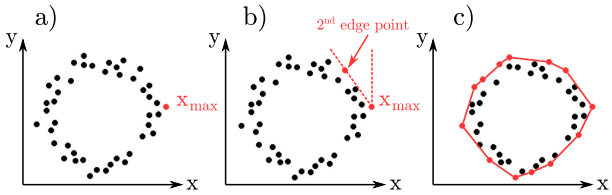



**FIGURE 12.** The convex hull polygonal approximation method.

polygonal approximation method. The circumference points are obtained by slicing the point cloud with a perpendicular plane. The algorithm starts at the point with the highest X coordinate (Fig. 12b). The next point is selected as the point with the minimal angle between the Y-axis and the line connecting the current point X and the next point, in the counterclockwise direction (Fig. 12b). The process is continued until the polygon is closed. The circumference is approximated as the sum of the line lengths between the selected points.

**Measurements from image features.** If the front- and side-view silhouettes are extracted [77], [98], [99], the measurements can be approximated using the distances between the feature points on the silhouette (see Fig. 11). The circumference can be approximated by a circle or an ellipse. For example, the waist breadth is the distance between F1 and F2, and the waist depth is the distance between S1 and S2. The distance between F1 and F2 and the distance between S1 and S2 can be used as a major and a minor axis, respectively, to approximate the hip circumference.

## V. DISCUSSION

We discuss the main limitations and issues of current scanning technology and the body measurement framework, as well as the gathered state-of-the-art results from Table 4. Based on the presented framework and the scanner types introduced in this section, we recommended pipelines for particular body measurement applications (see Fig. 15).

### A. LIMITATIONS AND CHALLENGES

**Absolute scale.** An important practical challenge for some body measurement approaches, in particular, the monocular and self-calibrated [63] PS methods, is the unknown absolute scale. The simplest way to obtain the scale is to use a calibrated 3D scanner data as input. Selected body measurement methods [7], [64], [159], [172], compared in Table 4, use 3D scans on the absolute scale as input. Another way to recover the scale is to place an object of known size (the calibration object) next to the subject, since, to recover absolute body measurements, it is sufficient to retrieve the scale of a single measurement. Usually, the body height is the most convenient body measure. Selected approaches presented in Table 4 use either the height [45], [149], [173] or camera parameters [25], [170] to scale images or silhouettes in order to extract anthropometric measurements on the absolute scale. Finally, some of the approaches [34], [80] presented in Table 4 do not know the height prior to body measurement. Hence, they estimate the camera parameters as a part of the learning procedure to infer the absolute scale. While [80] uses an encoder and regression approach, [34] uses a Gaussian process latent variable model to estimate the camera parameters.

**Evaluation measures.** There are multiple evaluation measures concerning accuracy, precision, and reliability [113] that are usually reported [29], [87], [172]. This lack of standardized evaluation measures complicates straightforward comparisons of different body measurement methods, since the different error measures cannot be converted from one to the other. To compare selected methods, we focus on a single reported measure, the mean absolute error (MAE), since it is mostly reported for the published body measurement methods. The MAE is a measure of accuracy, and it is calculated between the body measurement method estimation, $E_{est}$, and the ground truth, $E_{gt}$, as follows:

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N} \mid E_{est} - E_{gt} \mid \tag{1}$$

for every subject $i$ from the dataset. The ground truth is usually obtained by manual measurement, as noted in the ISO 20685-1:2018 [74] standard.

**Allowable error.** The allowable error (AE), based on the ANSUR study [58], defines an upper bound on the acceptable MAE for a measurement method. The study reports the median absolute deviation between measurements made by human experts, which are considered to be the gold standard in anthropometry, and the ground truth for the body measurements. This indicates the limitation of anthropometric measurement methods since the ground truth is never exactly known. Additionally, note that the AE measure represents the median, while the MAE represents the mean absolute deviation from the ground truth. Hence, the MAE is affected by possible outlier measurements and can present higher or lower values than the AE.

**Datasets.** While most methods are evaluated on the CAESAR dataset, there is a fair number of methods that evaluate their results using their own data [27], [35], [85], [94], [99], [105], [172]. Additionally, methods evaluated on the CAESAR dataset tend to use its random subsets

**TABLE 4.** MAE in millimeters for different measurement methods for measurements shown in Fig. 13. The measurements are grouped into circumferences, lengths, and breadths. All the methods are evaluated on some sample of the CAESAR [137] dataset with the exception of Yan *et al.* [172] (denoted with †). The results of each method were extracted from the corresponding paper listed in the "From" column. The table is split into three parts: 2D-based methods, 3D-based methods (further split into published and commercial methods), and the allowable error (AE) [58] for some of the body parts. The best results are bolded for both the 2D and 3D categories. The mean MAE for every method is provided.

| | | | Circumference | | | | | | | | | | | Length | | | Breadth | Height | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | From | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Mean |
| 2D | Smith et al. [149] | [149] | 14.2 | **11.4** | **16.2** | 25.0 | 15.2 | **5.5** | **10.4** | 7.9 | **11.1** | **10.4** | **6.3** | **11.0** | **6.0** | **8.0** | 8.4 | **7.9** | **10.9** |
| | Yan et al. [173] | [173] | 11.6 | 12.3 | 26.1 | 28.7 | 22.6 | 6.9 | 13.0 | **7.8** | 18.2 | 11.7 | 7.8 | 13.9 | 9.5 | 11.2 | 7.6 | 20.1 | 14.3 |
| | Dibra et al. 17 [45] | [173] | 10.8 | 13.1 | 28.3 | 38.6 | 26.0 | 6.5 | 13.4 | 8.0 | 18.5 | 11.8 | 7.9 | 13.4 | 6.9 | 8.7 | 7.7 | 11.8 | 14.5 |
| | Boisvert et al. [25] | [25] | **11.0** | 27.0 | 21.0 | **14.0** | 42.0 | 21.0 | 23.0 | 13.0 | 33.0 | 12.0 | 14.0 | 20.0 | 20.0 | 34.0 | 30.0 | 9.0 | 21.5 |
| | Chen et al. [34] | [149] | 23.0 | 27.0 | 18.0 | 37.0 | **15.0** | 24.0 | 59.0 | 76.0 | 19.0 | 16.0 | 28.0 | 52.0 | 53.0 | 9.0 | 12.0 | 21.0 | 30.6 |
| | Kanazawa et al. [80] | [173] | 16.3 | 27.2 | 68.3 | 85.3 | 62.8 | 14.3 | 35.6 | 16.7 | 39.3 | 21.4 | 13.6 | 28.6 | 45.3 | 37.2 | 21.8 | 96.5 | 39.4 |
| | Xi et al. [170] | [149] | 50.0 | 59.0 | 36.0 | 55.0 | 23.0 | 56.0 | 146.0 | 182.0 | 35.0 | 33.0 | 61.0 | 119.0 | 109.0 | 19.0 | 24.0 | 49.0 | 66.0 |
| | Bogo et al. [23]* | [173] | 28.1 | 24.4 | 74.5 | 72.8 | 99.1 | 11.9 | 28.4 | 25.9 | 51.3 | 28.4 | 28.8 | 57.8 | 150.2 | 219.1 | 51.9 | 398.5 | 84.4 |
| 3D | Yan et al. [172]† | [172] | - | **9.1** | 14.3 | **12.4** | **8.9** | **4.5** | **5.5** | - | **7.9** | **3.0** | 10.6 | - | 13.2 | - | - | - | **8.9** |
| | Tsoli et al. [159] | [159] | **5.9** | 15.8 | **12.7** | - | 12.4 | - | - | - | - | - | **6.2** | - | **10.1** | - | - | **7.5** | 10.1 |
| | Hasler et al. [64] | [159] | 7.5 | 17.0 | 13.0 | - | 16.2 | - | - | - | - | - | 6.6 | - | 10.4 | - | - | 10.2 | 11.5 |
| | Anthroscan [7] | [159] | 7.4 | 21.1 | 12.4 | - | 7.5 | - | - | - | - | - | 7.6 | - | 11.7 | - | - | 5.6 | 10.4 |
| | AE [58] | [58] | ± 5 | ± 11 | ± 15 | ± 12 | ± 12 | - | - | - | ± 6 | - | ± 4 | - | - | - | ± 8 | ± 10 | ± 9.2 |

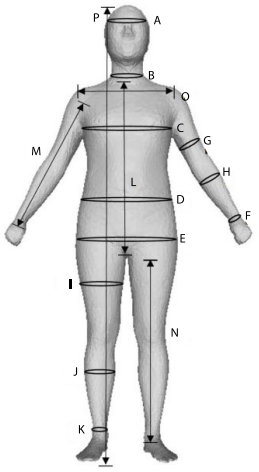
[149] or gender specific subsets [159]. This lack of a standardized benchmarking dataset presents problems for the direct comparison of measurement methods. We present and compare the results of different body measurement methods evaluated on different datasets in Sec. V-B, assuming that the quality and variability of each dataset are similar enough.

**Body measurements.** Different body measurement methods can be compared on standardized measurements defined in the ISO 7250-1:2017 [75] standard. However, different methods tend to report their evaluations for different measurements, which are not always equal, as can be seen by the missing values in Table 4. This hinders a comprehensive comparison of each method.

### B. METHODOLOGY COMPARISON

In Table 4, we gather the MAEs of different anthropometric measurement methods for measurements from Fig. 13, addressing the lack of literature on body measurement method evaluations and comparisons. We classify each method depending on its input (2D or 3D) and highlight the best obtained results. Additionally, we group the measurements into three categories, namely circumferences, lengths, and breadths, to compare their performance on different measurement tasks.

The 2D methods presented in Table 4 can be divided into two groups. One group of methods [23], [45], [80], [149], [173] uses images or silhouettes to estimate a 3D point cloud or learn the shape and pose parameters of a SMPL model. As can be seen, these methods outperform the second group [25], [34], [170], which tries to map a 2D PCA space into a 3D PCA space using a Gaussian process latent variable model [25], [34] or simply a linear regression [170]. Additionally, methods that try to estimate the absolute scale [34], [80], in parallel to estimating the body measurements, seem to perform worse in their appropriate groups.




| | Measurements |
|---|---|
| A | Head C |
| B | Neck base C |
| C | Chest C |
| D | Waist C |
| E | Hip C |
| F | Wrist C |
| G | Bicep C |
| H | Forearm C |
| I | Thigh C |
| J | Calf C |
| K | Ankle C |
| L | Shoulder-crotch L |
| M | Shoulder-wrist L |
| N | Inside leg L |
| O | Shoulder B |
| p | Height |

**FIGURE 13.** Body measurements reported in Table 4 abbreviated accordingly: C stands for circumference, L for length and B for breadth. The image is adapted from: [149].

The 3D methods presented in Table 4 are based on fitting the SCAPE model to a 3D scan and extracting body measurements from the fitted model. In general, there are fewer 3D methods than 2D methods, probably because image data is more accessible that the 3D scanning datasets.

Based on the presented results, 3D-based methods are generally better, but do not outperform 2D-based methods by a large margin. Intuitively, 3D scans hold more information about the shape of the human body than 2D images, and hence obtain better circumference measures. On the other hand, 2D methods slightly outperform 3D methods in length estimations, as seen from the shoulder-wrist measure (measure M in Table 4), which may be easier to estimate in 2D. Breadth measurements are unfortunately not comparable, since measurements from 3D methods are not provided, confirming the

limitations noted in Sec. V-A. The best 2D [149] and 3D [172] methods are both based on the SMPL [102] model. While Smith *et al.* [149] (2D) use a deep learning model to predict the shape and pose parameters of a SMPL model, Yan *et al.* [172] (3D) fit an initial SMPL template to a 3D scan using an iterative closest point (ICP) [20].

Compared to the commercially available anthropometry software Anthroscan [7], 3D methods present slightly better results. Anthroscan predicts measurements directly from a 3D scan in the standing pose and is frequently used as a body measurement approach [42], [87], [159]. It achieves an average MAE of 1.5mm, worse than the 3D method from Yan *et al.* [172].

In the third part of Table 4 we show the allowable error (AE) for measurements for which the AE was measured [58]. While we can observe that the MAE is decreasing with more recent measurement methods, none of the presented methods are within the allowable error, indicating that automatic body measurement methods are still lagging behind human anthropometers. However, this does not indicate that the assessment methods are insufficient for real-world applications [28]. Additionally, there are commercially available 3D scanners with 3D anthropometry software [86] that claim to obtain results lower than the AE and can hence be used in applications that require greater accuracies, such as medical and surveying applications.

### C. RECOMMENDATIONS

Based on the presented technologies, the proposed measurement framework, and the previous discussion, we finally provide practical recommendations for body measurement, as shown in Fig. 15. First, the scanner classification is introduced. Next, specific pipelines are proposed with respect to their input. Finally, the requirements for the applications are described along with the introduced scanner types and pipeline recommendations.

**Scanner types.** We classify scanners based on their mobility/size into following categories: (a) stationary; (b) handheld; (c) mini; and (d) mobile camera.[11] Stationary scanners (see Fig. 14a) are usually installed in a fixed location, e.g., a lab or a medical facility. They are usually SL or PS-based. Compared to other scanner types, they are the most accurate and reliable and are therefore typically used to obtain ground truth data, e.g., stationary scanners were used to create 3D body scanning datasets like CAESAR [3], SIZE-UK [8], Scan DB [64], and FAUST [24]. Handheld scanners (see Fig. 14b) are designed to be moved around the imaged body area by hand. Most of the existing handheld 3D scanners are SL-based. Mini-scanners (see Fig. 14c) are embedded in or attached to mobile devices like smartphones and tablets to enable 3D data acquisition. Most mini-scanners are ToF- or SL-based. Finally, we distinguish mobile RGB cameras as a separate scanner type, because they are widespread and convenient for non-demanding users, and usually rely simply on monocular measurement estimation techniques.[12] The four scanner types represent the data acquisition techniques for body measurement, as shown in Fig. 15.


**FIGURE 14.** Three types of 3D scanners in terms of mobility and size: stationary (a), handheld (b) and mini-scanners (c).

**Pipelines.** We propose and distinguish three possible pipelines for body measurement, as shown in the right part of Fig. 15. The first pipeline, sufficient for the majority of applications, consists of preparation, 3D scanning, and measurement extraction. The second pipeline is more flexible and consists of 3D scanning (without prior subject preparation), feature extraction with or without mesh fitting, and measurement extraction. In both pipelines, 2D images acquired using RGB cameras are often useful for improving the reconstruction [120]. Finally, the third and usually the least precise pipeline only takes 2D RGB images as input. These images are then used for feature extraction, mesh fitting, and measurement extraction.

**Applications.** We recommend specific measurement pipelines and scanner types for different anthropometric applications: medicine, surveying, the fashion industry, fitness, and entertainment.

For medical applications [46], [68], it is usually desirable that high-quality body measurements are obtained. Therefore, 3D scanning using stationary or handheld scanners, along with the preparation stage (marker placement), is recommended (see the first pipeline in Fig. 15). The measurements can then be directly extracted from the 3D scan (see Sec. IV).

The second application is surveying, a systematical measurement of a population sample for the purpose of analyzing and tracking the properties of human bodies over time [56], [174]. High-quality surveys sometimes release their data publicly [3], [8], which allows for the creation and improvement of statistical models [12], [69], [79], [102], [116], [127]. Surveying is usually done using stationary scanners, and markers are sometimes placed on the body to improve and simplify the measurement [56].

For fashion industry applications (garment and clothing design), all four of the data acquisition techniques are used. For individually designed garments, stationary scanners are preferable [174]. For less reliable measurements and

[11] For more details on the currently available scanners on the market see Appendix A.

[12] For more details on mobile devices and applications for body measurement assessment, see Appendix B.

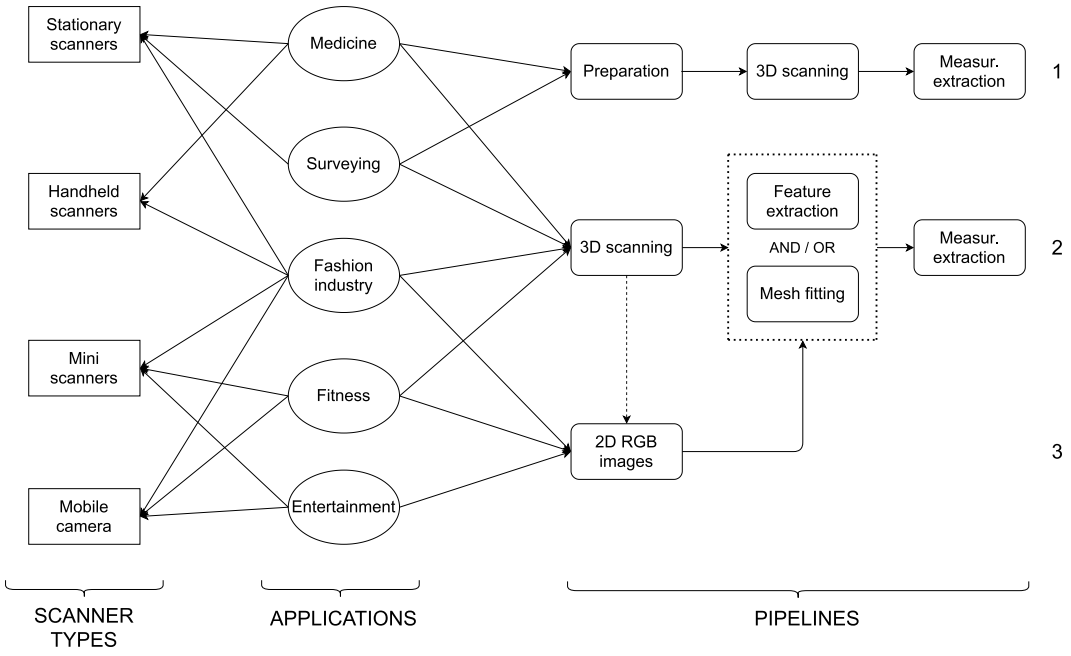



**FIGURE 15.** The diagram of practical body measurement recommendations.

mass-produced clothes, other data acquisition techniques are sufficient.

For fitness and entertainment applications (gaming, AR, VR, etc.), low-budget solutions using mini scanners and mobile cameras are ideal for individual users. For fitness applications, body measurements are used for tracking physical progress over time. As seen in Appendix B, there are a few fitness-based mobile applications that estimate body measurements. Most of them use one or two RGB images from different views. For gyms or fitness centers, stationary 3D scanners might be more convenient. Regarding entertainment, a 3D human pose [73], [124] in an AR setup allows the creation of a rigged character [16]; therefore, only a rough estimate of body measurements is needed.

## VI. CONCLUSION

Anthropometry is a very important, interdisciplinary area of research, still strongly entangled with 3D scanning technology for the purpose of body measurements. We have concisely reviewed the fast developing and improving scanning technologies, which are therefore becoming more applicable for automatic body measurements. As a consequence of this development, larger and more diverse body scanning datasets became publicly available. This work has also proposed and discussed different processing stages of the body measurement framework. It was pointed out that a particularly important processing stage is model fitting, which includes mesh fitting and mesh regression, since it allows the development of the expressive statistical body models that describe the pose and shape variances of a human population sample. The 3D and 2D measurement methodologies and published works have been compared and the main challenges and limitations have been identified; based on this, several measurement pipelines have been proposed for various applications. Reflecting on the future, we recall that pose and shape estimation from images is increasingly becoming a very active area of research. Consequently, it is now possible to estimate human pose and shape from an RGB image only, to a large extent due to the advances in deep learning research and optimization. Combining those advances with improvements in scanning technology, primarily that scanners are becoming smaller and more convenient while maintaining a high reconstruction accuracy, we conclude that the accuracy and reliability of body measurements from 3D as well as from 2D data will be significantly improved in the near future.

## APPENDIX A - 3D SCANNERS

Table 5 presents an overview of the commercial 3D scanners that have the ability to scan human bodies, excluding scanners that are not fit for the task, such as the Revopoint Tanso S1 [4], used to reconstruct smaller objects. We provide more than 80 currently available 3D scanners manufactured by more than 50 companies, as well as their taxonomy regarding several key characteristics: their mobility, method of reconstruction, price, resolution, accuracy, number of sensors, dimensions, provided texture, scanning time and provided anthropometric software. Additionally, we comment on their effect on human body scanning.

We observe an equal amount of stationary (booth-like) and handheld scanners, whereas only a few mini scanners are on the market. While handheld scanners offer a quicker scanning setup time in new environments, stationary scanners are more ideal for fixed scenarios, omitting (almost) entirely the setup process. Naturally, the mobility of a scanner is correlated with its dimensions. Stationary scanners are large and bulky, while mini scanners are compact and portable. Hence, mini

**TABLE 5.** The existing 3D scanners on the market capable of scanning the human body in its entirety. Each scanner is described by its respective mobility gradation: mini, handheld, and stationary, with mini scanners being the most portable. "Method" describes the 3D reconstruction approach as discussed in Sec. III. Additionally, we denote laser-based SL approaches with "SL"* and projector-based SL approaches with "SL". "Acc". and "Res". define the reconstruction capabilities of the scanners in terms of accuracy and resolution. The measures are given in millimeters. "No. Sens". reports the number of cameras and lighting sensors in the given scanner. The "Dims". column reports the dimensions of the scanner in centimeters. The dimensions can be given as a product of the height, width, and depth, or as a product of the diameter and the height of the scanner. The "Scan. Time" column reports the scanning duration in seconds of one single working volume. "Anthropo". reports if anthropometric measurement extraction is available in the given software. The remaining "Link", "Price", and "Texture" columns are self-explanatory.

| Manufact. | Product | Link | Mobility / Size | Method | Price ($) | Res. (mm) | Acc. (mm) | No. Sens. | Dims. (cm) | Texture | Scan. Time | Anthropo. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SizeStream | SS20 | sizestream.com | Stationary | ToF | from 15k | - | - | 20 | 145x188x203 | Yes | - | - |
| Vitronic | Vitus | vitronic.com | Stationary | SL* | 5k-10k | - | 1 | 8 | - | Yes | - | Yes |
| Texel | Portal BX | texel.graphics | Stationary | SL* | 31k | 1 | 1 | 4 | 225×258 | Yes | 30 s | Yes |
| | Portal MX | | Stationary | SL* | 26k | 1 | 1 | 4 | 260×60×60 | Yes | 30 s | Yes |
| IBV | Move4D | ibv.org/move4d | Stationary | PS | - | 1 | - | 16 | 200x200x300 | Yes | - | Yes |
| Artec | ArtecLeo | artec3d.com | Handheld | SL | 29.8k | 0.2 | 0.1 | 3 | 23.1x16.2x23 | Yes | - | Yes |
| | Artec Eva | | Handheld | SL | 19.8k | 0.2 | 0.1 | 3 | 26.2x15.8x6.3 | Yes | - | Yes |
| | Artec Eva Lite | | Handheld | SL | 9800 | 0.5 | 0.1 | 2 | 26.2x15.8x6.3 | No | - | Yes |
| | Artec Space Spider | | Handheld | SL | 24.8k | 0.1 | 0.05 | 5 | 19x14x13 | Yes | - | Yes |
| | Shapify Booth | | Stationary | SL | 180k | 1.5 | 0.25 | 12 | 330x330x280 | Yes | 12 s | Yes |
| Thor3D | Calibry | thor3dscanner.com | Handheld | SL | 5790 | 0.6 | 0.1 | 5 | 16.5x8.5x27.3 | Yes | 60 s | - |
| Fit3D | Fit3D | fit3d.com | Stationary | PS | 10k | - | - | 3 | - | - | - | Yes |
| Styku | Styku S100 | styku.com | Stationary | ToF | 10k | - | - | - | 254x254x117 | No | - | Yes |
| Revopoint3D | Handysense | revopoint3d.com | Handheld | SL | 3000 | 0.3 | 0.1 | 3 | 21.5x12x33.6 | Yes | 300 s | - |
| | Acusense A1 | | Mini | SL | 1000 | - | 0.1-1 | 3 | 15x25x38 | Yes | - | No |
| Apple | iPhone 12 Pro Lidar | | Mini | ToF | 999 | - | - | | | | - | No |
| HP | HP Pro S3 | hp.com | Stationary | SL | 3400 | 0.05 | - | 2 | - | Yes | - | No |
| DexaFit | DexaFit | dexafit.com | Stationary | SL | - | - | - | - | - | No | - | Yes |
| botspot | botscan Neo | botspot.de | Stationary | SL | - | 0.1 | - | 120 | 305x246 | Yes | - | No |
| | botscan Pro S | | Stationary | SL | - | 0.2 | - | 300 | 355x2600 | Yes | - | No |
| | botscan Cargo | | Stationary | SL | - | 0.2 | - | 70 | 605x243x259 | Yes | - | No |
| | OptaOne+ | | Stationary | SL | 10k- 50k | 0.2 | - | 68 | 314x254 | Yes | - | No |
| 3dMD | 3dMDBody | 3dmd.com | Stationary | PS | From 100K | 0.5 | - | 30+ | 250x200x300 | Yes | 1.5 ms | Yes |
| | 3dMDflex | | Stationary | PS | From 50K | - | 0.2 | 9+ | 100x60x60 | Yes | 1.5 ms | Yes |
| TechMed3D | BodyScan Scanner | techmed3d.com | Handheld | SL | - | - | - | 2 | - | Yes | - | Yes |
| 4DDynamics | IIID Body | 4ddynamics.com | Stationary | SL | - | - | 0.5 | 10 | 170x170x210 | - | - | Yes |
| | IIID Trailer | | Stationary | SL | - | - | - | - | - | - | - | Yes |
| | IIID ScanBooth | | Stationary | SL | - | - | - | - | 170x170x170 | - | - | Yes |
| | Memphisto EOS | | Stationary | SL | - | - | - | 2 | - | - | - | - |
| | Memphisto EX | | Stationary | SL | - | - | - | 3 | - | - | - | - |
| | Pico | | Stationary | SL | - | - | - | 2 | - | - | - | - |
| | Pico Pro | | Stationary | SL | - | - | - | 2 | - | - | - | - |
| | Gotcha | | Handheld | SL | - | - | - | 3 | - | - | - | - |
| | Gotcha Pro | | Handheld | SL | - | - | - | 3 | - | - | - | - |
| Shape Labs | ShapeScale | shapescale.com | Stationary | SL | 499 | 3.1 | 1.58 | 3 | 120x145 | Yes | 60 s | Yes |
| Naked Labs | Naked | nakedlabs.com | Stationary | ToF | 1395 | - | 5 | 3 | 158x30x30 | No | 20 s | Yes |
| mPort Ltd. | mPod | mport.com | Stationary | SL | - | - | 10 | - | - | No | - | Yes |
| Telemat Industrie | Symcad II ST | telmat.fr | Stationary | SL | 44k | - | - | 4 | 402x160x235 | - | - | - |
| | Symcad II HD | | Stationary | SL | 31.3k | - | - | 4 | 358x134x230 | - | - | - |
| | Symcad III | | Stationary | SL | 15.9k | - | - | 16 | 190x173x210 | - | - | - |
| $[TC]^2$ Labs | TC2-105 | tc2.com | Stationary | SL | 100k- 250k | 0.7 | 0.1 | - | - | Yes | - | No |
| | TC2-30R | | Stationary | ToF | - | 2 | - | - | 177x102 | Yes | - | Yes |
| | TC2-19M | | Stationary | ToF | - | 1 | - | - | - | - | - | Yes |
| | TC2-21B | | Stationary | PS | 30k | - | - | - | - | - | - | - |
| | TC2-19R | | Stationary | ToF | 1k- 10k | - | - | - | - | - | - | - |
| | TC2-19B | | Stationary | ToF | - | - | - | - | - | - | - | - |
| Spacevision | SCUVEG4-Portable | spacevision.tokyo | Stationary | - | 50k- 100k | - | 2 | - | 205x60x80 | - | - | - |
| | SCUVEG4-Flex | | Stationary | - | 50k- 100k | - | 2 | - | 205x60x80 | - | - | - |
| QuantaCorp | Shapewatch | quantacorp.io | Stationary | PS | - | - | - | - | - | - | - | Yes |
| Mantis Vision | Studio 3iosk | mantis-vision.com | Stationary | PS | 30k | - | - | 15 | 250x247 | Yes | - | - |
| | Studio 3iosk XT | | Stationary | PS | - | - | - | 16 | 250x247 | Yes | - | - |
| | F6 SR | | Handheld | SL | 10k- 50k | 0.4 | 0.1 | 3 | - | Yes | - | No |
| | F6 Smart | | Handheld | SL | 10k- 50k | 0.4 | 0.5 | 3 | - | Yes | - | No |
| TG3D Studio | TG 2000-F | tg3ds.com | Stationary | SL* | from 15k | - | - | - | 152x132x202 | - | - | Yes |
| GOM | ATOS Q | gom.com | Handheld | SL | from 60k | 0.03 | - | 3 | 34x24x8.3 | No | - | - |

**TABLE 5.** *(Continued.)* **The existing 3D scanners on the market capable of scanning the human body in its entirety. Each scanner is described by its respective mobility gradation: mini, handheld, and stationary, with mini scanners being the most portable. "Method" describes the 3D reconstruction approach as discussed in Sec. III. Additionally, we denote laser-based SL approaches with "SL"* and projector-based SL approaches with "SL". "Acc". and "Res". define the reconstruction capabilities of the scanners in terms of accuracy and resolution. The measures are given in millimeters. "No. Sens". reports the number of cameras and lighting sensors in the given scanner. The "Dims". column reports the dimensions of the scanner in centimeters. The dimensions can be given as a product of the height, width, and depth, or as a product of the diameter and the height of the scanner. The "Scan. Time" column reports the scanning duration in seconds of one single working volume. "Anthropo". reports if anthropometric measurement extraction is available in the given software. The remaining "Link", "Price", and "Texture" columns are self-explanatory.**

| Manufact. | Product | Link | Mobility / Size | Method | Price ($) | Res. (mm) | Acc. (mm) | No. Sens. | Dims. (cm) | Texture | Scan. Time | Anthropo. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ScanTech | KScan 20 | 3d-scantech.com | Handheld | SL | from 40k | 0.01 | 0.02 | 6 | - | No | - | - |
| | KScan Magic II | | Handheld | SL | from 40k | 0.01 | 0.02 | - | - | No | - | - |
| | Prince 775 | | Handheld | SL | from 40k | 0.02 | 0.03 | - | 31x16x10 | No | - | - |
| | HScan 771 | | Handheld | SL | from 40k | 0.05 | 0.03 | - | 31x16x10 | No | - | - |
| | Axe B17 | | Handheld | SL | from 40k | 0.025 | 0.02 | - | - | No | - | - |
| | IReal 2E | | Handheld | SL | 4980 | 3 | 0.1 | 9 | 14x9.4x25.8 | Yes | - | No |
| Hexagon | Aicon Primescan | hexagonmi.com | Stationary | SL | 35k | 0.016 | 0.016 | 3 | 300x210x175 | No | - | No |
| EvixScan3D | Heavy Duty Basic | evixscan3d.com | Stationary | SL | 10k- 50k | - | 0.02 | 3 | 430x220x65 | Yes | - | - |
| | Heavy Duty Optima | | Stationary | SL | 10k- 50k | - | 0.0183 | 3 | 430x220x65 | Yes | - | - |
| | Heavy Duty Quadro | | Stationary | SL | 10k- 50k | - | 0.013 | 5 | 520x280x95 | Yes | - | - |
| Polyga | Carbon | polyga.com | Stationary | SL | 14.9k | 0.362 | 0.05 | 2 | 13x41x12 | Yes | 1.2 s | - |
| | Compact L6 | | Stationary | SL | 11.9k | 0.18 | 0.08 | 2 | 5.5x12.9x40 | Yes | 1.2 s | No |
| | Polyga H3 | | Handheld | SL | 9990 | 0.5 | 0.08 | 2 | 28x20x6 | Yes | - | No |
| Shining 3D | Freescan X7 | shining3d.com | Handheld | SL* | - | 0.05 | 0.03 | 2 | 13x9x31 | No | - | - |
| | EinscanPro 2X | | Handheld | SL | 5499 | 0.2 | 0.04 | 3 | 37x36.5x13.5 | Yes | - | - |
| | EinscanPro 2X Plus | | Handheld | SL | 6899 | 0.2 | 0.04 | 3 | 37x36.5x13.5 | Yes | - | - |
| | Einscan H | | Handheld | SL | | 3 | 0.05 | - | 10.8x11x23.7 | Yes | - | - |
| Peel 3D | Peel 1 | peel-3d.com | Handheld | SL | 5990 | 0.5 | 0.25 | 2 | 9.6x14x25.8 | No | 90 s | Yes |
| | Peel 2 | | Handheld | SL | 7690 | 0.5 | 0.25 | 4 | 15x17.1x25.1 | Yes | 90 s | Yes |
| Faro | Freestyle3D | faro.com | Handheld | SL* | 10k- 20k | 0.2 | 1.5 | - | 260x310x105 | Yes | - | - |
| | Freestyle3D X | | Handheld | SL* | 10k- 50k | 0.2 | 1 | - | 26x31x10.5 | Yes | - | - |
| Creaform | GO!Scan Spark | creaform3d.com | Handheld | SL | 39.9k | 0.2 | 0.05 | - | 8.9x11.4x34.6 | Yes | - | - |
| | HandyScan Black | | Handheld | SL* | 50k- 100k | 0.1 | 0.035 | 9 | 14.2x7.9x28.8 | No | - | - |
| | HandyScan Black Elite | | Handheld | SL* | 50k- 100k | 0.025 | 0.025 | 11 | 14.2x7.9x28.8 | No | - | - |
| Occipital | Structure Sensor | structure.io | Mini | SL | 379 | - | - | 3 | - | Yes | - | No |
| BodyGee | Orbiter | bodygee.com | Stationary | ToF | - | - | - | - | - | Yes | 90 s | Yes |
| | Boxx | | Stationary | ToF | - | - | - | - | - | Yes | - | Yes |
| Intel | RealSense LiDAR Camera L515 | intelrealsense.com | Mini | ToF | 349 | - | 14 | 2 | 6.1x2.6 | - | - | No |
| | RealSense Depth Camera D455 | | Mini | PS | 239 | - | - | 3 | 12.4x2.6x2.9 | No | - | No |
| | RealSense Depth Camera D435i | | Mini | PS | 179 | - | - | 4 | 9x2.5x2.5 | No | - | No |
| | RealSense Depth Camera D415 | | Mini | PS | 149 | - | - | 3 | 9.9x2x2.3 | No | - | No |

and handheld scanners offer better applicability to the task of the distributed data collection process [174] since they present higher portability. On the other hand, stationary scanners offer faster scanning times, in the range of seconds, while handheld scanners offer scanning times in the range of minutes, presenting a trade-off between their dimensions and applicability. Since breathing and fidgeting causes human bodies to move during the scanning process, faster scanning times are more desirable. Nevertheless, the performance of handheld scanners does not seem to lag behind stationary 3D scanners, as seen by their accuracy.

The mobility and scanning time of a scanner seem to mostly drive its price. Smaller scanners tend to be cheaper, while scanners offering faster scanning times tend to be pricier, indicating that the market is still more appreciative towards stationary scanners. Most of the scanners use structured light (SL) to reconstruct the human body since it offers the best reconstruction accuracy within the methods presented in Sec. III. Additionally, they present the lowest resolution, followed by passive stereo (PS) and time-of-flight (ToF), respectively. Hence, they allow dense 3D human body reconstructions, appropriate for the anthropometric application. To this end, we additionally report if the scanner comes with an anthropometry software that can automatically extract body measurements from a 3D scan. While texture does not directly impact the scanning process, arguments have been made in favor of the greater usecase for textured 3D human body models [140].

The market is moving towards handheld and mini scanners. Mini scanners are particularly important for the future of tablet and smartphone scanning, because they can be attached to or even embedded into devices. For example,

**TABLE 6. The existing 3D body measurement mobile applications for iOS and Android. We distinguish between the apps based on the company, main application, scanner, and OS. Note that the main applications specified in the table are retrieved based on the app description, and the actual purpose might differ in reality. The applications that use 2D input data only are listed above the double line; those ones that also use 3D data are listed below the line. The scanner that is used for data acquisition is specified within the "Scanner" column. The values denoted with "-" are not available.**

| Company | App | Manufacturer link | Main application | Scanner | OS |
|---|---|---|---|---|---|
| Sony | 3DCreator | sony.com | Entertainment | RGB camera | Android |
| Standard Cyborg | Capture: 3D Scan Anything | standardcyborg.com | Entertainment | RGB camera | iOS |
| EyeCue | Qlone | qlone.pro | Entertainment | RGB camera | iOS+Android |
| SmartMobileVision | Scann3D | smartmobilevision.com | Entertainment | RGB camera | Android |
| IBV | 3DAvatarBody | ibv.org/3davatarbody | Fitness / Fashion | RGB camera | iOS+Android |
| 3DLook | Mobile Tailor | 3dlook.me | Fashion | RGB camera | - |
| | Your Fit | | Fashion | RGB camera | - |
| | Uniform Pro | | Fashion | RGB camera | - |
| QuantaCorp | Two Pictures 3D BODYSCAN | quantacorp.io | Fitness / Fashion | RGB camera | iOS+Android |
| Size Stream LLC | Mobile Scanner | www.sizestream.com | Fashion | RGB camera | iOS+Android* |
| | MeThreeSixty | | Fitness | RGB camera | iOS+Android |
| Fision AG | Meepl | meepl.com | Fashion | RGB camera | iOS+Android |
| NetVirta | NetVirta | netvirta.com | Fashion | - | - |
| SizeYou | SizeYou | sizeyou.it | Fashion | RGB camera | iOS+Android |
| Nettelo | Nettelo | nettelo.com | Fashion | RGB camera | Android |
| Xplorazzy Tech | 3D Scanner Pro | xplorazzi.com | Entertainment | RGB camera | Android |
| BodyGee | BodyGee Coach App | bodygee.com | Fitness | RGB camera | Android |
| Scandy | Scandy Pro | scandy.co | Entertainment | LIDAR | iOS |
| Itseez3D, Inc. | ItSeez3D | itseez3d.com | Entertainment | Occipital | iOS |
| TechMed3D | 3DSizeMe | techmed3d.com | Entertainment | Occipital | iOS |

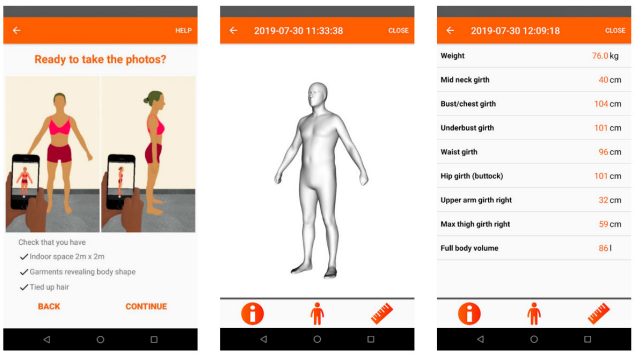



**FIGURE 16. A common mobile application body measurement pipeline is to take front and side images, estimate a 3D human mesh, and assess the body measurements from the mesh. Image credits: [1].**

the Occipital sensors can be attached to a smartphone device, while the Apple iPhone 12 has an embedded LIDAR sensor (see Appendix B). Mini scanners are usually ToF-based, which can be seen from Table 5. As the computing capabilities of mobile devices improve further and ToF-based mini scanners increase their resolution, we expect that mobile devices will become more reliable and accurate 3D scanners.

## APPENDIX B - MOBILE APPLICATIONS

Mobile phones have become an emerging market for making 3D scanners more approachable. Multiple cameras [55], new ToF sensors [5], and the general development of said phones have made the implementation of 3D scanning techniques easier. Additionally, stationary scanners are relatively bulky and pricey. Hence, mobile 3D scanning technology has become important, particularly for the distributed data collection process [174]. In this appendix, we present a comprehensive list of available applications for 3D human body measurement estimation (see Table 5).

The majority of the existing applications use a single RGB camera for computing body measurements. The most common approach (as seen in MeThreeSixty, Meepl, 3DAvatarBody, and many others) is to fit a template mesh to a front and side image of the subject (see Fig. 16). The measurements can then be extracted from the template mesh, as described in Sec. IV. Some of the applications extract the measurements from a single image (Nettelo), and some take multiple images from different angles and rely on photogrammetry for 3D reconstruction and measurements (3DCreator, Qlone, Scann3D, 3DAvatarBody, Two Pictures 3D BODYSCAN, Mobile Scanner, SizeYou, 3D Scanner Pro, BodyGee Coach App). A few applications use a 3D Occipital scanner attached to the smartphone device (ItSeez3D, TechMed3D, Occipital original app), while one (Scandy Pro) uses Apple's embedded LIDAR sensor to directly retrieve 3D human scans.

## REFERENCES

[1] *3D Avatar Body—Institute of Bomechanics Valencia*. Accessed: Feb. 9, 2020. [Online]. Available: https://www.ibv.org/en/3d-avatar-body-3/
[2] *4D Human Body Motion Scanning*. Accessed: Mar. 12, 2020. [Online]. Available: https://www.ibv.org/en/technologies/4d-human-body-motion-scanning/
[3] *Caesar I, the Most Comprehensive Source for Body Measurement Data*. Accessed: Nov. 20, 2020. [Online]. Available: http://store.sae.org/caesar/
[4] *Revopoint Tanso S1*. Accessed: Mar. 16, 2021. [Online]. Available: https://www.revopoint3d.com/portable-3d-scanner-tanso-s1/
[5] *Scandy Time-of-Flight Smartphone Technology*. Accessed: Jan. 1, 2021. [Online]. Available: https://www.scandy.co/
[6] *Sizeusa Dataset*. Accessed: Mar. 17, 2021. [Online]. Available: https://www.tc2.com/size-usa.html
[7] *Uhuman Solutions GmbH Anthroscan*. Accessed: Nov. 20, 2020. [Online]. Available: https://www.human-solutions.com/
[8] *UK National Sizing Survey*. Accessed: Nov. 20, 2020. [Online]. Available: http://www.size.org/
[9] A. Abdelhafiz, "Integrating digital photogrammetry and terrestrial laser scanning," Ph.D. dissertation, Inst. Geodesy Photogramm., Braunschweig, Germany, Feb. 2009.
[10] P. Alcantarilla, J. Nuevo, and A. Bartoli, "Fast explicit diffusion for accelerated features in nonlinear scale spaces," in *Proc. BMVC*, 2013, pp. 1–12.

[11] B. Allen, B. Curless, B. Curless, and Z. Popović, "The space of human body shapes: Reconstruction and parameterization from range scans," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 587–594, 2003.
[12] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, and J. Rodgers, "SCAPE: Shape completion and animation of people," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 408–416, Jul. 2005.
[13] D. Anguelov, P. Srinivasan, H.-C. Pang, D. Koller, S. Thrun, and J. Davis, "The correlated correspondence algorithm for unsupervised registration of nonrigid surfaces," in *Proc. 17th Int. Conf. Neural Inf. Process. Syst. (NIPS)*. Cambridge, MA, USA: MIT Press, 2004, pp. 33–40.
[14] B. Artacho and A. Savakis, "UniPose: Unified human pose estimation in single images and videos," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 7035–7044.
[15] G. Ascenso, M. H. Yap, T. Allen, S. S. Choppin, and C. Payton, "A review of silhouette extraction algorithms for use within visual hull pipelines," *Comput. Methods Biomech. Biomed. Eng., Imag. Vis.*, vol. 8, no. 6, pp. 649–670, Nov. 2020.
[16] I. Baran and J. Popović, "Automatic rigging and animation of 3D characters," *ACM Trans. Graph.*, vol. 26, no. 3, p. 72, Jul. 2007.
[17] T. Bell, B. Li, and S. Zhang, *Structured Light Techniques and Applications*. American Cancer Society, 2016, pp. 1–24. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1002/047134608X.W8298
[18] S. Bellisai, F. Villa, S. Tisa, D. Bronzi, and F. Zappa, "Indirect time-of-flight 3D ranging based on SPADs," *Proc. SPIE*, vol. 8268, pp. 282–289, Jan. 2012.
[19] P.-M. Jodoin, "Comparative study of background subtraction algorithms," *J. Electron. Imag.*, vol. 19, no. 3, Jul. 2010, Art. no. 033003.
[20] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, Feb. 1992.
[21] S. Bianco, G. Ciocca, and D. Marelli, "Evaluating the performance of structure from motion pipelines," *J. Imag.*, vol. 4, no. 8, p. 98, Aug. 2018.
[22] W. Boehler, M. Vicent, and A. Marbs, "Investigating laser scanner accuracy," in *Proc. 19th Int. Symp. CIPA*, vol. 34, Jan. 2003, pp. 1–19.
[23] F. Bogo, A. Kanazawa, C. Lassner, P. Gehler, J. Romero, and M. J. Black, "Keep it SMPL: Automatic estimation of 3D human pose and shape from a single image," in *Proc. ECCV*, 2016, pp. 561–578.
[24] F. Bogo, J. Romero, M. Loper, and M. J. Black, "FAUST: Dataset and evaluation for 3D mesh registration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Piscataway, NJ, USA, Jun. 2014, pp. 3794–3801.
[25] J. Boisvert, C. Shu, S. Wuhrer, and P. Xi, "Three-dimensional human shape inference from silhouettes: Reconstruction and validation," *Mach. Vis. Appl.*, vol. 24, no. 1, pp. 145–157, Jul. 2011.
[26] J. Bougourd and P. Treleaven, "UK national sizing survey—SizeUK," in *Proc. 1st Int. Conf. 3D Body Scanning Technol.*, Lugano, Switzerland, Oct. 2010, pp. 327–337. [Online]. Available: https://www.3dbody.tech/cap/abstracts/2010/327bougourd.html, doi: 10.15221/10.327.
[27] B. Bradtmiller and M. E. Gross, "3D whole-body scans: Measurement extraction software validation," *SAE Trans.*, vol. 108, pp. 394–400, Jan. 1999.
[28] B. Bradtmiller and M. E. Gross, "3D whole-body scans: Measurement extraction software validation," *SAE Trans.*, vol. 108, pp. 394–400, Jan. 1999.
[29] S. Bragança, P. Arezes, M. Carvalho, S. P. Ashdown, I. Castellucci, and C. Leão, "A comparison of manual anthropometric measurements with kinect-based scanned measurements in terms of precision and reliability," *Work*, vol. 59, no. 3, pp. 325–339, Apr. 2018.
[30] A. Bulat, J. Kossaifi, G. Tzimiropoulos, and M. Pantic, "Toward fast and accurate human pose estimation via soft-gated skip connections," 2020, *arXiv:2002.11098*. [Online]. Available: http://arxiv.org/abs/2002.11098
[31] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime multi-person 2D pose estimation using part affinity fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 1, pp. 172–186, Jan. 2021.
[32] K. Casadei and J. Kiel, *Anthropometric Measurement*. Treasure Island, FL, USA: StatPears Publishing, 2020.
[33] L.-C. Chen, D. M. Collins, Y. Zhu, G. Papandreou, B. Zoph, F. Schroff, H. Adam, and J. Shlens, "Searching for efficient multi-scale architectures for dense image prediction," in *Proc. NIPS*, 2018, pp. 1–14.
[34] Y. Chen, T.-K. Kim, and R. Cipolla, "Inferring 3D shapes and deformations from single views," in *Computer Vision*, K. Daniilidis, P. Maragos, and N. Paragios, Eds. Berlin, Germany: Springer, 2010, pp. 300–313.
[35] Y. Chen, D. Robertson, and R. Cipolla, "A practical system for modelling body shapes from single view measurements," in *Proc. Brit. Mach. Vis. Conf. (BMVC)*, Jan. 2011, pp. 1–11.
[36] J. Cheng, C. Leng, J. Wu, H. Cui, and H. Lu, "Fast and accurate image matching with cascade hashing for 3D reconstruction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1–8.
[37] J. Chibane and G. Pons-Moll, "Implicit feature networks for texture completion from partial 3D data," in *Proc. ECCV Workshops*, 2020, pp. 717–725.
[38] J. D. Camba, A. Leon, J. Cantero, J. Saorín, and M. Contero, "Application of low-cost 3D scanning technologies to the development of educational augmented reality content," in *Proc. IEEE Frontiers Educ. Conf. (FIE)*, Oct. 2016, pp. 1–6.
[39] H. M. Daanen and G. J. van de Water, "Whole body scanners," *Displays*, vol. 19, no. 3, pp. 111–120, Nov. 1998.
[40] H. A. M. Daanen and F. B. T. Haar, "3D whole body scanners revisited," *Displays*, vol. 34, no. 4, pp. 270–275, Oct. 2013.
[41] M. Daneshmand, A. Helmi, E. Avots, F. Noroozi, F. Alisinanoglu, H. S. Arslan, J. Gorbova, R. E. Haamer, C. Ozcinar, and G. Anbarjafari, "3D scanning: A comprehensive survey," 2018, *arXiv:1801.08863*. [Online]. Available: http://arxiv.org/abs/1801.08863
[42] N. D'Apuzzo, "Recent advances in 3D full body scanning with applications to fashion and apparel," in *Proc. Conf. Opt. 3-D Meas. Techn.*, N. D'Apuzzo, Ed. Vienna, Austria, 2007. [Online]. Available: https://hometrica.ch/publ/2009_optical3d.pdf
[43] N. D'Apuzzo, "3D body scanning technology for fashion and apparel industry," *Proc. SPIE*, vol. 6491, pp. 203–214, Jan. 2007.
[44] I. Dianat, J. Molenbroek, and H. I. Castellucci, "A review of the methodology and applications of anthropometry in ergonomics and product design," *Ergonomics*, vol. 61, no. 12, pp. 1696–1720, Dec. 2018.
[45] E. Dibra, H. Jain, C. Öztireli, R. Ziegler, and M. Gross, "Human shape from silhouettes using generative HKS descriptors and cross-modal neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 4826–4836.
[46] M. Đonlić, "Three-dimensional analysis of back surface under dynamic conditions in scoliosis diagnostics," Ph.D. dissertation, Fac. Elect. Eng. Comput., Dept. Electron. Syst. Inf. Process., Univ. Zagreb, Zagreb, Croatia, 2019.
[47] M. Ebrahim, *3D Laser Scanners: History, Applications, and Future*. LAP LAMBERT Academic Publishing, 2016. [Online]. Available: https://books.google.hr/books?id=vpszvgAACAAJ
[48] S. El-Hakim, J.-A. Beraldin, and F. Blais, "A comparative evaluation of the performance of passive and active 3-D vision systems," *Proc. SPIE*, vol. 2646, pp. 14–25, May 2003.
[49] H.-S. Fang, G. Lu, X. Fang, J. Xie, Y.-W. Tai, and C. Lu, "Weakly and semi supervised human body part parsing via pose-guided knowledge transfer," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 70–78.
[50] D. Fofi, T. Sliwa, and Y. Voisin, "A comparative survey on invisible structured light," *Proc. SPIE*, vol. 5303, pp. 90–97, May 2004.
[51] S. Foix, G. Alenya, and C. Torras, "Lock-in time-of-flight (ToF) cameras: A survey," *IEEE Sensors J.*, vol. 11, no. 9, pp. 1917–1926, Sep. 2011.
[52] Y. Furukawa and C. Hernández, "Multi-view stereo: A tutorial," *Found. Trends Comput. Graph. Vis.*, vol. 9, nos. 1–2, pp. 1–148, 2015.
[53] Z. Gan, *Visual Sensing and Its Applications Integration of Laser Sensors to Industrial Robots* (Advanced Topics in Science and Technology in China). Hangzhou, China: Zhejiang Univ. Press, 2011.
[54] J. Geng, "Structured-light 3D surface imaging: A tutorial," *Adv. Opt. Photon.*, vol. 3, no. 2, pp. 128–160, Jun. 2011.
[55] I. A. Ghaffar and M. N. H. Mohd, "Selection of Android smartphones with built-in dual lens camera for stereo vision Android app development," in *Proc. 7th Int. Conf. Comput. Commun. Eng. (ICCCE)*, Sep. 2018, pp. 74–78.
[56] S. Giancola, M. Valenti, and R. Sala, *A Survey on 3D Cameras: Metrological Comparison of Time-of-Flight, Structured-Light and Active Stereoscopy Technologies* (Springer Briefs in Computer Science). Cham, Switzerland: Springer, 2018.
[57] S. B. Gokturk, H. Yalcin, and C. Bamji, "A time-of-flight depth sensor–system description, issues and solutions," in *Proc. Conf. Comput. Vis. Pattern Recognit. Workshop*, 2004, p. 35.
[58] C. Gordon, T. Churchill, C. Clauser, B. Bradtmiller, J. McConville, I. Tebbetts, and R. Walker, "Anthropometric survey of U.S. Army personnel: Summary statistics, interim report for 1988," U.S. Army Natick Res., Develop. Eng. Center, Natick, MA, USA, Tech. Rep. NATICK/TR-89/027, Jan. 1989. [Online]. Available: https://apps.dtic.mil/dtic/tr/fulltext/u2/a209600.pdf
[59] J. Hafeez, A. Hamacher, S. Kwon, and S. Lee, "Performance evaluation of patterns for image-based 3D model reconstruction of textureless objects," in *Proc. Int. Conf. 3D Immersion (IC3D)*, Dec. 2017, pp. 1–5.

[60] J. Hafeez, S.-C. Kwon, S.-H. Lee, and A. Hamacher, "3D surface reconstruction of smooth and textureless objects," in *Proc. Int. Conf. Emerg. Trends Innov. ICT (ICEI)*, Feb. 2017, pp. 145–149.

[61] A. Haleem and M. Javaid, "3D scanning applications in medical field: A literature-based review," *Clin. Epidemiol. Global Health*, vol. 7, no. 2, pp. 199–210, Jun. 2019.

[62] M. Hansard, S. Lee, O. Choi, and R. Horaud, *Time-of-Flight Cameras*. London, U.K.: Springer, 2013.

[63] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[64] N. Hasler, C. Stoll, M. Sunkel, B. Rosenhahn, and H.-P. Seidel, "A statistical model of human pose and body shape," *Comput. Graph. Forum*, vol. 28, no. 2, pp. 337–346, Apr. 2009.

[65] Y. He and S. Chen, "Recent advances in 3D data acquisition and processing by time-of-flight camera," *IEEE Access*, vol. 7, pp. 12495–12510, 2019.

[66] Y. He, R. Yan, K. Fragkiadaki, and S.-I. Yu, "Epipolar transformers," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 7776–7785.

[67] Y. He, B. Liang, Y. Zou, J. He, and J. Yang, "Depth errors analysis and correction for time-of-flight (ToF) cameras," *Sensors*, vol. 17, no. 1, p. 92, Jan. 2017.

[68] S. B. Heymsfield, B. Bourgeois, B. K. Ng, M. J. Sommer, X. Li, and J. A. Shepherd, "Digital anthropometry: A critical review," *Eur. J. Clin. Nutrition*, vol. 72, no. 5, pp. 680–687, May 2018.

[69] D. Hirshberg, M. Loper, E. Rachlin, and M. J. Black, "Coregistration: Simultaneous alignment and modeling of articulated 3D shape," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, in Lecture Notes in Computer Science, vol. 7577. Berlin, Germany: Springer-Verlag, Oct. 2012, pp. 242–255.

[70] R. Horaud, M. Hansard, G. Evangelidis, and C. Ménier, "An overview of depth cameras and range scanners based on time-of-flight technologies," *Mach. Vis. Appl.*, vol. 27, no. 7, pp. 1005–1020, Oct. 2016.

[71] A. H. Ahmadabadian, A. Karami, and R. Yazdan, "An automatic 3D reconstruction system for texture-less objects," *Robot. Auton. Syst.*, vol. 117, pp. 29–39, Jul. 2019.

[72] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu, "Human3.6M: Large scale datasets and predictive methods for 3D human sensing in natural environments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 7, pp. 1325–1339, Jul. 2014.

[73] K. Iskakov, E. Burkov, V. Lempitsky, and Y. Malkov, "Learnable triangulation of human pose," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 7717–7726.

[74] *3-D Scanning Methodologies for Internationally Compatible Anthropometric Databases—Part 1: Evaluation Protocol for Body Dimensions Extracted From 3-D Body Scans*, Standard ISO 20685-1:2018, 2018.

[75] *Basic Human Body Measurements for Technological Design—Part 1: Body Measurement Definitions and Landmarks*, Standard ISO 7250-1:2017, 2017.

[76] C. Istook and S.-J. Shin, "3D scanning systems with application to the apparel industry," *J. Fashion Marketing Manage.*, vol. 5, pp. 120–132, Jun. 2001.

[77] L. Jiang, J. Yao, B. Li, F. Fang, Q. Zhang, and M. Q.-H. Meng, "Automatic body feature extraction from front and side images," *J. Softw. Eng. Appl.*, vol. 5, no. 12, pp. 94–100, 2012.

[78] H. Joo, T. Simon, X. Li, H. Liu, L. Tan, L. Gui, S. Banerjee, T. Godisart, B. Nabbe, I. Matthews, T. Kanade, S. Nobuhara, and Y. Sheikh, "Panoptic studio: A massively multiview system for social interaction capture," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 1, pp. 190–204, Jan. 2019.

[79] H. Joo, T. Simon, and Y. Sheikh, "Total capture: A 3D deformation model for tracking faces, hands, and bodies," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 8320–8329.

[80] A. Kanazawa, M. J. Black, D. W. Jacobs, and J. Malik, "End-to-end recovery of human shape and pose," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7122–7131.

[81] H. Kawasaki, R. Furukawa, R. Sagawa, and Y. Yagi, "Dynamic scene shape reconstruction using a single structured light pattern," in *Proc. IEEE Conf. Comput. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[82] N. Koepke, M. Zwahlen, J. C. Wells, N. Bender, M. Henneberg, F. J. Rühli, and K. Staub, "Comparison of 3D laser-based photonic scans and manual anthropometric measurements of body size and shape in a validation study of 123 young swiss men," *PeerJ*, vol. 5, p. e2980, Feb. 2017.

[83] N. Kolotouros, G. Pavlakos, M. J. Black, and K. Daniilidis, "Learning to reconstruct 3D human pose and shape via model-fitting in the loop," in *Proc. ICCV*, 2019, pp. 2252–2261.

[84] N. Kolotouros, G. Pavlakos, and K. Daniilidis, "Convolutional mesh regression for single-image human shape reconstruction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 4496–4505.

[85] M. Kouchi and M. Mochimaru, "Errors in landmarking and the evaluation of the accuracy of traditional and 3D anthropometry," *Appl. Ergonom.*, vol. 42, no. 3, pp. 518–527, Mar. 2011.

[86] A. Koval, "Quantitative comparison of manual vs. 3D scanner human body measurements," in *Proc. 11th Int. Conf. Exhib. 3D Body Scanning Process. Technol. Online/Virtual (3DBODY TECH)*, Nov. 2020. [Online]. Available: https://www.3dbody.tech/cap/abstracts/2020/2035koval.html

[87] A. Kuehnapfel, P. Ahnert, M. Loeffler, A. Broda, and M. Scholz, "Reliability of 3D laser-based anthropometry and comparison with classical anthropometry," *Sci. Rep.*, vol. 6, no. 1, pp. 1–11, May 2016.

[88] T.-H. Kwok, K.-Y. Yeung, and C. C. L. Wang, "Volumetric template fitting for human body reconstruction from incomplete data," *J. Manuf. Syst.*, vol. 33, no. 4, pp. 678–689, Oct. 2014.

[89] D. Lanman and G. Taubin, "Build your own 3D scanner: 3D photography for beginners," in *Proc. ACM SIGGRAPH Courses*, Jan. 2009, p. 8.

[90] A. Laurentini, "The visual hull concept for silhouette-based image understanding," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 2, pp. 150–162, Feb. 1994.

[91] R. N. Lescay, A. Becerra, and A. H. González, "Anthropometry. Comparative analysis of technologies for the capture of anthropometric dimensions," *Revista EIA*, pp. 47–59, Dec. 2016. [Online]. Available: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1794-12372016000200004&nrm=iso

[92] L. Li, "Time-of-flight camera—An introduction," May 2014. [Online]. Available: https://www.ti.com/lit/wp/sloa190b/sloa190b.pdf

[93] P. Li, Y. Xu, Y. Wei, and Y. Yang, "Self-correction for human parsing," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Dec. 29, 2020, doi: 10.1109/TPAMI.2020.3048039.

[94] Z. Li, W. Jia, Z.-H. Mao, J. Li, H.-C. Chen, W. Zuo, K. Wang, and M. Sun, "Anthropometric body measurements based on multi-view stereo image reconstruction," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2013, pp. 366–369.

[95] X. Liang, K. Gong, X. Shen, and L. Lin, "Look into person: Joint body parsing & pose estimation network and a new benchmark," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 4, pp. 871–885, Apr. 2019.

[96] K. Lin, L. Wang, K. Luo, Y. Chen, Z. Liu, and M.-T. Sun, "Cross-domain complementary learning using pose for multi-person part segmentation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 3, pp. 1066–1078, Mar. 2021.

[97] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, and P. Dollár, "Microsoft COCO: Common objects in context," 2014, *arXiv:1405.0312*. [Online]. Available: http://arxiv.org/abs/1405.0312

[98] Y.-L. Lin and M.-J.-J. Wang, "Automated body feature extraction from 2D images," *Expert Syst. Appl.*, vol. 38, no. 3, pp. 2585–2591, Mar. 2011.

[99] Y.-L. Lin and M.-J.-J. Wang, "Constructing 3D human model from front and side images," *Expert Syst. Appl.*, vol. 39, no. 5, pp. 5012–5018, Apr. 2012.

[100] Y. Liu, N. Pears, P. L. Rosin, and P. Huber, Eds., *3D Imaging, Analysis and Applications*. Cham, Switzerland: Springer, 2020.

[101] C. Loop and Z. Zhang, "Computing rectifying homographies for stereo vision," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1, Feb. 1999, p. 131.

[102] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, and M. J. Black, "SMPL: A skinned multi-person linear model," *ACM Trans. Graph.*, vol. 34, no. 6, pp. 248:1–248:16, Oct. 2015.

[103] M. M. Loper, N. Mahmood, and J. M. Black, "MoSh: Motion and shape capture from sparse markers," *ACM Trans. Graph.*, vol. 33, no. 6, pp. 220:1–220:13, Nov. 2014.

[104] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, vol. 2, Sep. 1999, pp. 1150–1157.

[105] J.-M. Lu and M.-J.-J. Wang, "The evaluation of scan-derived anthropometric measurements," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 8, pp. 2048–2054, Aug. 2010.

[106] J. Lu and M. Wang, "Automated anthropometric data collection using 3D whole body scanners," *Expert Syst. Appl.*, vol. 35, nos. 1–2, pp. 407–414, Jul. 2008.

[107] N. Mahmood, N. Ghorbani, N. F. Troje, G. Pons-Moll, and M. Black, "AMASS: Archive of motion capture as surface shapes," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 5442–5451.
[108] C. Malleson, A. Gilbert, M. Trumble, J. Collomosse, A. Hilton, and M. Volino, "Real-time full-body motion capture from video and IMUs," in *Proc. Int. Conf. 3D Vis. (3DV)*, Oct. 2017, pp. 449–457.
[109] J. Medina-Inojosa, V. Somers, T. Ngwa, L. Hinshaw, and F. López-Jiménez, "Reliability of a 3D body scanner for anthropometric measurements of central obesity," *Obesity, Open Access*, vol. 2, no. 3, p. 10, 2016.
[110] D. Mehta, H. Rhodin, D. Casas, P. Fua, O. Sotnychenko, W. Xu, and C. Theobalt, "Monocular 3D human pose estimation in the wild using improved CNN supervision," in *Proc. Int. Conf. 3D Vis. (3DV)*, Oct. 2017, pp. 506–516.
[111] D. Mehta, O. Sotnychenko, F. Mueller, W. Xu, S. Sridhar, G. Pons-Moll, and C. Theobalt, "Single-shot multi-person 3D pose estimation from monocular RGB," in *Proc. Int. Conf. 3D Vis. (3DV)*, Sep. 2018, pp. 120–130.
[112] R. Mottaghi, X. Chen, X. Liu, N.-G. Cho, S.-W. Lee, S. Fidler, R. Urtasun, and A. Yuille, "The role of context for object detection and semantic segmentation in the wild," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 891–898.
[113] K. Norton and T. Olds, *Anthropometrica: A Textbook of Body Measurement for Sports and Health Courses*. Sydney, NSW, Australia: UNSW Press, 1996.
[114] T. Olds and F. Honey, "The use of 3D whole-body scanners in anthropometry," in *Proc. 9th Int. Conf. Int. Soc. Adv. Kinanthropometry*, Jan. 2005, pp. 1–12.
[115] M. Đonlić, T. Petković, and T. Pribanić, "3D surface profilometry using phase shifting of de Bruijn pattern," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 963–971.
[116] A. A. A. Osman, T. Bolkart, and J. M. Black, "STAR: Sparse trained articulated human body regressor," in *Proc. ECCV*, 2020, pp. 1–17.
[117] J. Pages, C. Collewet, F. Chaumette, and J. Salvi, "An approach to visual servoing based on coded light," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2006, pp. 4118–4123.
[118] J. Pages, J. Salvi, R. Garcia, and C. Matabosch, "Overview of coded light projection techniques for automatic 3D profiling," in *Proc. IEEE Int. Conf. Robot. Automat.*, vol. 1, Sep. 2003, pp. 133–138.
[119] J. Pagès, J. Salvi, C. Collewet, and J. Forest, "Optimised de Bruijn patterns for one-shot shape acquisition," *Image Vis. Comput.*, vol. 23, no. 8, pp. 707–720, Aug. 2005.
[120] J. Park, Q.-Y. Zhou, and V. Koltun, "Colored point cloud registration revisited," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 143–152.
[121] S. I. Park and J. K. Hodgins, "Capturing and animating skin deformation in human motion," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 881–889, Jul. 2006.
[122] C. Patel, Z. Liao, and G. Pons-Moll, "TailorNet: Predicting clothing in 3D as a function of human pose, shape and garment style," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 7365–7375.
[123] G. Pavlakos, V. Choutas, N. Ghorbani, T. Bolkart, A. A. Osman, D. Tzionas, and M. J. Black, "Expressive body capture: 3D hands, face, and body from a single image," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 10967–10977.
[124] D. Pavllo, C. Feichtenhofer, D. Grangier, and M. Auli, "3D human pose estimation in video with temporal convolutions and semi-supervised training," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 7753–7762.
[125] D. A. Payne, P. P. A. Jongenelen, A. A. Dorrington, J. M. Cree, and A. D. Carnegie, "Multiple frequency range imaging to remove measurement ambiguity," Tech. Rep., Jul. 2009. [Online]. Available: https://researchcommons.waikato.ac.nz/handle/10289/4032
[126] F. Piron, D. Morrison, M. R. Yuce, and J.-M. Redoute, "A review of single-photon avalanche diode time-of-flight imaging sensor arrays," *IEEE Sensors J.*, early access, Nov. 19, 2020, doi: 10.1109/JSEN.2020.3039362.
[127] L. Pishchulin, S. Wuhrer, T. Helten, C. Theobalt, and B. Schiele, "Building statistical shape spaces for 3D human modeling," *Pattern Recognit.*, vol. 67, pp. 276–286, Jul. 2017.
[128] G. Pons-Moll, J. Romero, N. Mahmood, and J. M. Black, "Dyna: A model of dynamic human shape in motion," *ACM Trans. Graph.*, vol. 34, no. 4, pp. 120:1–120:14, Aug. 2015.
[129] T. Pribanic, T. Petkovic, D. Bojanic, K. Bartol, and M. Gupta, "Scene adaptive structured light 3D imaging," in *Proc. 15th Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, vol. 4, 2020, pp. 576–582.
[130] T. Pribanić, T. Petković, D. Bojanić, K. Bartol, and A. Gupta, "Smart time-multiplexing of quads solves the multicamera interference problem," in *Proc. 3DV*, 2020, pp. 811–819.
[131] S. Prokudin, C. Lassner, and J. Romero, "Efficient learning on point clouds with basis point sets," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019, pp. 4331–4340.
[132] A. Pumarola, J. Sanchez, G. P. T. Choi, A. Sanfeliu, and F. Moreno, "3DPeople: Modeling the geometry of dressed humans," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 2242–2251.
[133] H. Qiu, C. Wang, J. Wang, N. Wang, and W. Zeng, "Cross view fusion for 3D human pose estimation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 4341–4350.
[134] C. Quan, X. Y. He, C. F. Wang, C. J. Tay, and H. M. Shang, "Shape measurement of small objects using LCD fringe projection with phase shifting," *Opt. Commun.*, vol. 189, nos. 1–3, pp. 21–29, Mar. 2001.
[135] F. Remondino, A. Guarnieri, and A. Vettore, "3D modeling of close-range objects: Photogrammetry or laser scanning," *Proc. SPIE*, vol. 5665, pp. 216–225, Dec. 2004.
[136] M. Ribo and M. Brandner, "State of the art on vision-based structured light systems for 3D measurements," in *Proc. Int. Workshop Robot. Sensors, Robot. Sensor Environ.*, 2005, pp. 2–6.
[137] K. M. Robinette, H. Daanen, and E. Paquet, "The CAESAR project: A 3-D surface anthropometry survey," in *Proc. 2nd Int. Conf. 3-D Digit. Imag. Modeling*, 1999, pp. 380–386.
[138] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *Computer Vision*, A. Leonardis, H. Bischof, and A. Pinz, Eds. Berlin, Germany: Springer 2006, pp. 430–443.
[139] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2564–2571.
[140] A. Saint, E. Ahmed, A. E. R. Shabayek, K. Cherenkova, G. Gusev, D. Aouada, and B. Ottersten, "3DBodyTex: Textured 3D body dataset," in *Proc. Int. Conf. 3D Vis. (3DV)*, Sep. 2018, pp. 495–504.
[141] A. Saint, A. Kacem, K. Cherenkova, K. Papadopoulos, J. Chibane, G. Pons-Moll, G. Gusev, D. Fofi, D. Aouada, and B. Ottersten, "Sharp 2020: The 1st shape recovery from partial textured 3D scans challenge results," in *Proc. ECCV Workshops*, 2020, pp. 741–755.
[142] J. Salvi, E. Mouaddib, and J. Batile, "An overview of the advantages and constraints of coded pattern projection techniques for autonomous navigation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot Syst. Innov. Robot. Real-World Appl. (IROS)*, vol. 3, Sep. 1997, pp. 1264–1271.
[143] J. Salvi, J. Pagès, and J. Batlle, "Pattern codification strategies in structured light systems," *Pattern Recognit.*, vol. 37, no. 4, pp. 827–849, Apr. 2004.
[144] J. Salvi, S. Fernandez, T. Pribanić, and X. Llado, "A state of the art in structured light patterns for surface profilometry," *Pattern Recognit.*, vol. 43, no. 8, pp. 2666–2680, Aug. 2010.
[145] K. Sato and S. Inokuchi, "Three-dimensional surface measurement by space encoding range imaging," *J. Robot. Syst.*, vol. 2, no. 1, pp. 27–39, 1985.
[146] J. L. Schönberger and J.-M. Frahm, "Structure-from-motion revisited," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 4104–4113.
[147] T. Schöps, J. L. Schönberger, S. Galliani, T. Sattler, K. Schindler, M. Pollefeys, and A. Geiger, "A multi-view stereo benchmark with high-resolution images and multi-camera videos," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2538–2547.
[148] L. Sigal, A. O. Balan, and M. J. Black, "HumanEva: Synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion," *Int. J. Comput. Vis.*, vol. 87, nos. 1–2, pp. 4–27, Mar. 2010.
[149] B. M. Smith, V. Chari, A. Agrawal, J. M. Rehg, and R. Sever, "Towards accurate 3D human body reconstruction from silhouettes," in *Proc. Int. Conf. 3D Vis. (3DV)*, Sep. 2019, pp. 279–288.
[150] B. M. Smith, V. Chari, A. Agrawal, J. M. Rehg, and R. Sever, "Towards accurate 3D human body reconstruction from silhouettes," in *Proc. Int. Conf. 3D Vis. (3DV)*, Sep. 2019, pp. 279–288.
[151] E. Stoykova, A. A. Alatan, P. Benzie, N. Grammalidis, S. Malassiotis, J. Ostermann, S. Piekh, V. Sainov, C. Theobalt, T. Thevar, and X. Zabulis, "3-D time-varying scene capture technologies—A survey," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 11, pp. 1568–1586, Nov. 2007.

[152] L. Streeter and Y. C. Kuang, "Metrological aspects of time-of-flight range imaging," *IEEE Instrum. Meas. Mag.*, vol. 22, no. 2, pp. 21–26, Apr. 2019.
[153] K. Sun, B. Xiao, D. Liu, and J. Wang, "Deep high-resolution representation learning for human pose estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 5686–5696.
[154] Z. Sun, G. Qian, Z. Peng, W. Dai, D. Sun, G. Zhang, N. Zhang, J. Xu, R. Wang, and C. Li, "Orthogonal coded multi-view structured light for inter-view interference elimination," in *Proc. IEEE Int. Conf. Vis. Commun. Image Process. (VCIP)*, Dec. 2020, pp. 181–184.
[155] P. Tikuisis, P. Meunier, and C. E. Jubenville, "Human body surface area: Measurement and prediction using three dimensional body scans," *Eur. J. Appl. Physiol.*, vol. 85, nos. 3–4, pp. 264–271, Aug. 2001.
[156] P. Treleaven and J. Wells, "3D body scanning and healthcare applications," *Computer*, vol. 40, no. 7, pp. 28–34, Jul. 2007.
[157] B. Triggs, F. P. McLauchlan, I. R. Hartley, and W. A. Fitzgibbon, "Bundle adjustment—A modern synthesis," in *Vision Algorithms: Theory and Practice*, B. Triggs, A. Zisserman, and R. Szeliski, Eds. Berlin, Germany: Springer, 2000, pp. 298–372.
[158] M. Trumble, A. Gilbert, C. Malleson, A. Hilton, and J. Collomosse, "Total capture: 3D human pose estimation fusing video and inertial sensors," in *Proc. BMVC*, 2017, pp. 1–13.
[159] A. Tsoli, M. Loper, and M. J. Black, "Model-based anthropometry: Predicting measurements from 3D human scans in multiple poses," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Mar. 2014, pp. 83–90.
[160] S. Van der Jeught and J. J. J. Dirckx, "Real-time structured light profilometry: A review," *Opt. Lasers Eng.*, vol. 87, pp. 18–31, Dec. 2016.
[161] N. Van Gestel, S. Cuypers, P. Bleys, and J.-P. Kruth, "A performance evaluation test for laser line scanners on CMMs," *Opt. Lasers Eng.*, vol. 47, nos. 3–4, pp. 336–342, Mar. 2009.
[162] G. Varol, J. Romero, X. Martin, N. Mahmood, M. J. Black, I. Laptev, and C. Schmid, "Learning from synthetic humans," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1–9.
[163] J. Wang, C. Zhang, W. Zhu, Z. Zhang, Z. Xiong, and P. A. Chou, "3D scene reconstruction by multiple structured-light based commodity depth cameras," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2012, pp. 5429–5432.
[164] J. Wang, J. Thornton, S. Kolesnik, and R. Pierson, "Anthropometry in body composition: An overview," *Ann. New York Acad. Sci.*, vol. 904, pp. 317–326, Jun. 2000.
[165] M. Wang, M. Wang, and Y. Lin, "Hsinchu: The ergonomics society of Taiwan," Tech. Rep., 2000. [Online]. Available: https://www.worldcat.org/title/taiwan-diqu-ren-ti-ji-ce-zi-liao-ku-shou-ce-anthropometric-data-book-of-the-chinese-people-in-taiwan/oclc/53197435?referer=di&ht=edition
[166] M.-J.-J. Wang, W.-Y. Wu, K.-C. Lin, S.-N. Yang, and J.-M. Lu, "Automated anthropometric data collection from three-dimensional digital human models," *Int. J. Adv. Manuf. Technol.*, vol. 32, nos. 1–2, pp. 109–115, Feb. 2007.
[167] F. Wermke, T. Wübbenhorst, and B. Meffert, "Interference avoidance for two time-of-flight cameras using autonomous optical synchronization," in *Proc. 6th Int. Conf. Control, Autom. Robot. (ICCAR)*, Apr. 2020, pp. 586–595.
[168] Y. Wu, A. Kirillov, F. Massa, W.-Y. Lo, and R. Girshick. (2019). *Detectron2*. [Online]. Available: https://github.com/facebookresearch/detectron2
[169] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 4–24, Jan. 2021.
[170] P. Xi, W.-S. Lee, and C. Shu, "A data-driven approach to human-body cloning using a segmented body database," in *Proc. 15th Pacific Conf. Comput. Graph. Appl. (PG)*, Oct. 2007, pp. 139–147.
[171] F. Xia, P. Wang, X. Chen, and A. L. Yuille, "Joint multi-person pose estimation and semantic part segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 6080–6089.
[172] S. Yan, J. Wirta, and J.-K. Kämäräinen, "Anthropometric clothing measurements from 3D body scans," *Mach. Vis. Appl.*, vol. 31, nos. 1–2, pp. 1–11, Feb. 2020.
[173] S. Yan and J.-K. Kämäräinen, "Learning anthropometry from rendered humans," Jan. 2021, *arXiv:2101.02515*. [Online]. Available: https://arxiv.org/abs/2101.02515
[174] N. Zakaria and D. Gupta, *Anthropometry, Apparel Sizing and Design* (The Textile Institute Book Series). Amsterdam, The Netherlands: Elsevier, 2019.
[175] C. Zhang, S. Pujades, M. Black, and G. Pons-Moll, "Detailed, accurate, human shape estimation from clothed 3D scan sequences," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4191–4200.
[176] F. Zhang, X. Zhu, H. Dai, M. Ye, and C. Zhu, "Distribution-aware coordinate representation for human pose estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 7091–7100.
[177] Y. Zhang, M. Hassan, H. Neumann, M. J. Black, and S. Tang, "Generating 3D people in scenes without people," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 6193–6203.
[178] J. Zhao, J. Li, Y. Cheng, T. Sim, S. Yan, and J. Feng, "Understanding humans in crowded scenes: Deep nested adversarial learning and a new benchmark for multi-human parsing," in *Proc. 26th ACM Int. Conf. Multimedia (MM)*. New York, NY, USA: Association for Computing Machinery, 2018, pp. 792–800.
[179] O. Özyeşil, V. Voroninski, R. Basri, and A. Singer, "A survey of structure from motion," *Acta Numerica*, vol. 26, pp. 305–364, May 2017.

**KRISTIJAN BARTOL** (Graduate Student Member, IEEE) received the B.Sc. degree in computing, in 2016, and the M.Sc. degree in computer science, in 2019. He is currently pursuing the Ph.D. degree in computing with the Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia. His current research interests include computer vision, human pose and shape estimation, and deep learning.

**DAVID BOJANIĆ** (Graduate Student Member, IEEE) received the B.Sc. *(magna cum laude)* and M.Sc. degrees in mathematics from the University of Zagreb, Croatia, in 2016 and 2019, respectively, where he is currently pursuing the Ph.D. degree in computer science with the Faculty of Electrical Engineering and Computing. His current research interest includes computer vision, focusing on shape analysis and 3D reconstruction of human bodies.

**TOMISLAV PETKOVIĆ** (Member, IEEE) received the degree in engineering and the magister and Ph.D. degrees from the University of Zagreb, in 2002, 2006, and 2010, respectively, all in electrical engineering. He is currently an Assistant Professor with the Faculty of Electrical Engineering and Computing, University of Zagreb. He teaches several graduate courses in the field of digital signal and image processing. His main research interests include digital image processing and analysis, depth sensing, and 3D imaging. He is a member of ACM and HDBIMF.

**TOMISLAV PRIBANIĆ** (Senior Member, IEEE) was a Visiting Researcher with the INRIA Grenoble Rhône-Alpes, Grenoble, France, and the Fraunhofer IGD, Darmstadt, Germany. He was a Fulbright Visiting Scholar with the University of Wisconsin, Madison, USA. He is currently a Professor with the Faculty of Electrical Engineering and Computing, University of Zagreb. He teaches several undergraduate and graduate courses in the field of algorithms and data structures, image processing, sensors, and human motion analysis. He has led a number of domestic and international scientific projects, collaborating with researchers from within and outside the EU. His main research interests include computer vision, and biomedical signal measurement and analysis. The results of his research have been implemented in technological projects and received recognition for innovations. He is a Senior Member of IFMBE and a Collaborating Member of the Croatian Academy of Engineering.