IEEE SA
STANDARDS ASSOCIATION

INDUSTRY CONNECTIONS REPORT

3D BODY PROCESSING INDUSTRY CONNECTIONS

COMPARATIVE ANALYSIS OF ANTHROPOMETRIC METHODS: PAST, PRESENT, AND FUTURE

Authored by

Alfredo Ballester, Instituto de Biomecánica, Universitat Politècnica de València, València, Spain

Warren Wright, Size Stream LLC, Cary, NC, USA

Jorge Valero, Instituto de Biomecánica, Universitat Politècnica de València, València, Spain

Emma Scott, Fashion Should Empower, Vancouver Island, BC, Canada

Tim Devlin, Kalypso, Portland, OR, USA

Alice Bullas, Centre for Sports Engineering Research, Sheffield Hallam University, UK

Javier Silva, Instituto de Biomecánica, Universitat Politècnica de València, València, Spain

Carol McDonald, Gneiss Concept, Washougal, WA, USA

TRADEMARKS AND DISCLAIMERS

IEEE believes the information in this publication is accurate as of its publication date; such information is subject to change without notice. IEEE is not responsible for any inadvertent errors.

The ideas and proposals in this specification are the respective author's views and do not represent the views of the affiliated organization.

ACKNOWLEDGEMENTS

Special thanks are given to the following participants of this paper:

Julianne Harris (Independent, apparel industry)
Emma Scott (Fashion Should Empower)
Gina Patterson (Independent, apparel industry)
Ruth Dang (Independent, apparel industry)
Alice Bullas (ISAK Level 2 - Sheffield Hallam University, Sports Engineering)

Special thanks are given to the following participants of the study described in this paper:

Fred McDonald, Gneiss Concept
Tim Guenzel, Human Solutions of North America, Inc
Karen Davis, David Bruner, Richard Allen, and Mark Koivuniemi, Size Stream LLC
Eduardo Parrilla, Ana V. Ruescas, Paola Piqueras, Sandra Alemany, Álvar Gallego, Juan A. Solves, Sofía Ramos, and Romina Tracchia, Instituto de Biomecánica
Andrei Coval, Texel LLC
Alex Arapov and Anna Antoniuk, 3Dlook Inc.
Tracy Mok and Ouyang Cheng, PolyU

The Institute of Electrical and Electronics Engineers, Inc. 3 Park Avenue, New York, NY 10016-5997, USA

PDF: STDVA25397 978-1-5044-8682-8

NOTICE AND DISCLAIMER OF LIABILITY CONCERNING THE USE OF IEEE SA INDUSTRY CONNECTIONS DOCUMENTS

This IEEE Standards Association ("IEEE SA") Industry Connections publication ("Work") is not a consensus standard document. Specifically, this document is NOT AN IEEE STANDARD. Information contained in this Work has been created by, or obtained from, sources believed to be reliable, and reviewed by members of the IEEE SA Industry Connections activity that produced this Work. IEEE and the IEEE SA Industry Connections activity members expressly disclaim all warranties (express, implied, and statutory) related to this Work, including, but not limited to, the warranties of: merchantability; fitness for a particular purpose; non-infringement; quality, accuracy, effectiveness, currency, or completeness of the Work or content within the Work. In addition, IEEE and the IEEE SA Industry Connections activity members disclaim any and all conditions relating to: results; and workmanlike effort. This IEEE SA Industry Connections document is supplied "AS IS" and "WITH ALL FAULTS."

Although the IEEE SA Industry Connections activity members who have created this Work believe that the information and guidance given in this Work serve as an enhancement to users, all persons must rely upon their own skill and judgment when making use of it. IN NO EVENT SHALL IEEE OR IEEE SA INDUSTRY CONNECTIONS ACTIVITY MEMBERS BE LIABLE FOR ANY ERRORS OR OMISSIONS OR DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO: PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS WORK, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE AND REGARDLESS OF WHETHER SUCH DAMAGE WAS FORESEEABLE.

Further, information contained in this Work may be protected by intellectual property rights held by third parties or organizations, and the use of this information may require the user to negotiate with any such rights holders in order to legally acquire the rights to do so, and such rights holders may refuse to grant such rights. Attention is also called to the possibility that implementation of any or all of this Work may require use of subject matter covered by patent rights. By publication of this Work, no position is taken by the IEEE with respect to the existence or validity of any patent rights in connection therewith. The IEEE is not responsible for identifying patent rights for which a license may be required, or for conducting inquiries into the legal validity or scope of patents claims. Users are expressly advised that determination of the validity of any patent rights, and the risk of infringement of such rights, is entirely their own responsibility. No commitment to grant licenses under patent rights on a reasonable or non-discriminatory basis has been sought or received from any rights holder. The policies and procedures under which this document was created can be viewed at https://standards.ieee.org/about/bog/iccom/.

This Work is published with the understanding that IEEE and the IEEE SA Industry Connections activity members are supplying information through this Work, not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought. IEEE is not responsible for the statements and opinions advanced in this Work.

TABLE OF CONTENTS

COMPARATIVE ANALYSIS OF ANTHROPOMETRIC METHODS: PAST, PRESENT, AND FUTURE

ABSTRACT

The IEEE 3D Body Processing Industry Connections (3DBP IC) Group develops recommendations that foster interoperability for three-dimensional (3D) body processing, enabling an ecosystem for consumers and creators to use and interchange 3D human body anthropometric information. The 3DBP IC's Quality subgroup focuses on developing preliminary methods, tools, benchmarks, resources, and testing procedures to define and quantify the quality of 3D models as well as the quality of the critical metadata, such as body landmarks and measurements. To further this goal, the 3DBP IC Group completed a comparative analysis of anthropometric methods focused on body measurements and captured 3D body shapes.

The aim of this study is twofold: first, to propose a methodology to assess the compatibility and repeatability of different body measurement extraction methods (e.g., body scanners, smartphone apps, and traditional anthropometry) using consistent criteria, and second, to conduct data collection and the corresponding analyses for the first instance of the application of the proposed methodology.

The experimental study was conducted in two phases. Phase 1 took place in Portland, Oregon (USA) in 2018 and included two booth scanners and four traditional anthropometry measurers. Phase 2 took place in Valencia (Spain) in 2019 and included two booth scanners, four smartphone apps, and two traditional anthropometry measurers. A total of 133 subjects participated in the study. Participants were selected to represent a wide range of body heights and weights. Each subject had 11 body measurements collected twice by each measurer and measuring technology.

The proposed methodology synthesizes: varied, widely used mathematical tools (including statistical models); minimum sample size required for the analyses; and several estimators and plots appropriate for describing the data and quantifying the repeatability and compatibility of the different measurement extraction methods. It provides commonly used criteria to assess most of the statistics. In addition, the methodology includes a procedure to obtain comparable snapshots of the 3D outcomes from digital technologies, facilitating visual assessment by potential data consumers.

This study has gathered a unique dataset that can be used for future research work. The results of the analyses of this dataset constitute useful benchmarks and references for similar studies. All of the participating technologies exhibited a good repeatability for the 11 body measurements considered; however, despite using similar measurement definitions, the compatibility between technologies was low, even for the experts using traditional anthropometry. These results highlight the need for definitions better adapted to 3D imaging systems as well as new methods to assess compatibility. The visual assessment of the 3D objects also shows a positive performance for all digital methods and reveals the differences between the various body scanning solutions and each of the phone apps.

Data and additional resources can be accessed at the IEEE website (https://standards.ieee.org/industry-connections/3d/bodyprocessing/), specifically measurement data, 3D data, and data from the analyses (i.e., tables, plots, and images).

1. INTRODUCTION

The IEEE 3D Body Processing Industry Connections (IEEE 3DBP IC) Group is a collection of diverse stakeholders from across technology, retail, research, and standards development working together to streamline three-dimensional (3D) practices for widespread adoption around 3D body processing technology [1].[1]

Despite the rising number of body scanning solutions available to capture the human body shape and extract or estimate body measurements from it, there are no common practices among technology suppliers for assessing and reporting the performance of these systems. Therefore, it is not easy to compare technologies—and the data they produce—or to assess their suitability for a given use case.

The purpose of the work presented in this paper was to study different measurement methods that capture anthropometric data, using consistent populations, data cleaning processes, and analytical methods. Although there have been previous, similar studies that focus on the compatibility and reliability of different anthropometric methods, the unique aspects of this study are the number of anthropometric methods, the number of subjects, and the consistent methodologies to collect and analyze the data.

The high-level objectives of the study were to:

- Help body data "consumers" understand 3D body digitization technologies and make informed decisions when choosing the most suitable technology for their particular business needs,
- Create a common understanding among body scanning "providers" for reporting repeatability and compatibility, and
- Collect datasets enabling the development of new IEEE standards that complement existing ones—for example, International Standards Organisation (ISO) and American Society for the Testing of Material (ASTM) standards—and advancing research into new quality metrics for 3D bodies.

The low-level objectives of the study were to:

- Propose evaluation techniques (or methods) and consistent criteria to assess reliability and compatibility,
- Quantify and compare the repeatability and compatibility of a representative number of body measurement extraction methods (an example of the application of the proposed methodology is presented and reference values, or benchmarks, are provided), and
- Create and publicly release data resources, specifically, the 3D scans and the measurement data gathered in the study, making them available for other research works.

[1] Numbers in brackets reference the sources listed in Section 6.

As a result, the 3DBP IC Group expects to demonstrate cross-platform compatibility evaluation techniques sufficient to instill confidence in various use-case scenarios and to propose methodology options for future application.

1.1. BACKGROUND

Anthropometry is widely used and deemed central to research and practice in a variety of industry contexts, such as apparel, ergonomics, and health assessment. Traditionally, anthropometric data collection has followed international guidelines and standards for taking body measurements physically over the body surface using equipment like stadiometers, tapes, calipers, and anthropometers [2]. Standardization bodies such as the International Society for the Advancement of Kinanthropometry (ISAK), ISO [3], [4], and ASTM [5] as well as reference studies like the Anthropometric Survey of U.S. Army (ANSUR) [6] and the National Health and Nutrition Examination Survey (NHANES) have provided direction [7]. Traditional anthropometric methods are widely used in industry, sports, and clinical practice because both the equipment and methods are affordable and accessible. These methods, however, rely on the training and skills of the measurer and are plagued by several shortcomings, including susceptibility to human error, time constraints, reliance on predictive equations to estimate complex anthropometrics, and a lack of standardized procedures for such measurements [8]. The accuracy and repeatability errors of traditional anthropometry are extensively reported in literature with a wide variety of results ranging from error values of a few millimeters to a few centimeters, depending on the body measurement and on the study (see [9]– [17]). Systematic biases between measurers lead to larger interobserver errors in studies involving more than one measurer [11], [13], [18]. Measurements taken by untrained measurers are less reliable and less accurate [19].

Over the past few years, 3D surface imaging has gained popularity as an anthropometric acquisition method due to the fast speed of contactless measuring, the availability of a 3D image, and the ability to acquire measurements that are not attainable through manual means. Also known as *body scanning* or *optical scanning*, 3D surface imaging captures the geometry of the human body surface as data points in a 3D space. The approach is particularly suitable for studies with large sample sizes and atypical populations because the creation of a digital 3D image permits both the immediate and retrospective analysis of data, the collection of vast amounts of data within a short timeframe, and the ability to produce a digital representation of body changes over time. The 3D image also presents an opportunity for the analysis of human size and shape using metrics out of the reach of traditional anthropometry, such as volume, area, and other geometric measurements. Due to these advantages, 3D imaging has experienced rapid market growth, expected to reach a value of at least $5.345 billion by 2025 [20]. A wide variety of 3D body scanners are available now [21]– [23]. They differ from each other in many ways, depending on—among other

things—the 3D point cloud acquisition technology (e.g., photogrammetry, ultrasound, or projected light such as laser, white, or infrared); the number of sensors; and the type of embodiment (e.g., booths with static sensors, turntables for the body or for the sensors, and handheld scanners). Each body scanner incorporates its own software and computer vision algorithms to filter and clean the point cloud data and to create meshed surfaces out of it. These differences lead to very different specifications in terms of scanning volume, scanning resolution, and scanning time, as well as to variations in the quality of 3D outcomes. Different protocols have been proposed to evaluate the 3D accuracy of these systems using reference objects ([24]– [26]) but not with real subjects.

Most body scanners also incorporate software for the extraction or estimation of body measurements from the 3D content captured; in most cases, this software is fully automatic and does not require the interaction of a human. There is also standalone software for this purpose. The body measurements extracted from 3D scans can be as reliable as those obtained by trained measurers using traditional anthropometry ([16], [17], [27], [28]) and can be more reliable than those obtained by untrained users [29]. There is additional literature reporting accuracy and reliability errors for different 3D systems with human subjects ([30]– [39]) and with reference manikins [40]). The main sources of error for body measurements obtained by body scanning are: the resolution and 3D accuracy of the body surface obtained from scanning; body sway during scanning; slight posture changes between repeated scans; and the measurement extraction software [14].

The addition of high-quality cameras to smartphones and the increase in their computing power made it possible to create 3D body content with these ubiquitous and widely available devices [41]. There are two types of smartphone-based applications for body scanning. The first type uses photogrammetry principles for 3D surface reconstruction. Similar to a classic 3D imaging system, while the subject stands still in the scanning pose, the phone is used like a hand-held 3D scanner to capture a sequence of images from around the subject to obtain a large number of points of view [42]. The second type of phone-based technology uses sparse information—such as features extracted from one or two photographs (e.g., silhouettes or key points) along with other information that can be reported easily by a subject (e.g., gender, age, weight, or body height)—to estimate a 3D body avatar within a body shape model ([43]– [47]), typically developed from a dataset of thousands of 3D body scans ([48], [49]). Recently, these technologies have been improved by the development of deep learning techniques that facilitate the efficient and robust extraction of features from images [29]. There are other phone-based solutions that directly estimate body measurements from image features without producing any 3D body representation.

There are two fundamental differences between the model-based technologies and actual 3D imaging systems: first, the actual information captured from the real world is much lower (e.g., a 2D outline versus millions of 3D point coordinates), and second, the 3D bodies they can produce are constrained by the body model used. These

differences provide several advantages and limitations. Model-based solutions are very robust and always provide a complete, watertight body representation within the parameters of the model. However, the accuracy of the 3D representation is dependent on the representativeness of the model and on the method used to compute the parameters of the model. As with any model, the ability to produce different outputs depends on the number of parameters and on the richness of the input. In theory, for the same underlying data model, model-based solutions that use richer inputs (e.g., an actual outline versus a distance between key points) potentially could cope better with diversity in body shape, provided that the input information is not redundant. For example, a solution that uses only body height and weight to create a 3D body human shape will be extremely robust but will not produce different bodies for two subjects of similar weight and height but with different torso shapes or limb lengths.

As previously mentioned, each body scanner or phone-based technology uses proprietary algorithms for the computation of body measurements from the 3D objects, which often leads to different locations for anatomical references and measurement values. Therefore, it is critical to establish standards to ensure cross-platform compatibility and suitability of use for the appropriate interpretation of the extracted data [50]– [54].

The compatibility between measurements obtained by different body scanners and measuring software is still an open issue being addressed by different standardization bodies such as ISO and the IEEE Standards Association (SA), as noted in references [55] and [56]. Most prior studies that compare the performance between anthropometric methods include just one body scanner and traditional anthropometry measurements. There is a lack of comparable datasets obtained by more than one scanner and more than one traditional anthropometry measurer. The different sample sizes, methods, and indicators used in these studies make it difficult to use them as reference data in new studies. Moreover, criteria for assessing the performance and compatibility of body measurements within and between different methods are limited. As a result of the ANSUR study, Gordon et al. [6] proposed criteria to accept or reject an expert measurer based on its reliability when taking repeated measurements. These criteria were more or less severe depending on the higher or lower repeatability error observed in the experimental study. These ANSUR thresholds are widely used to assess the repeatability of any anthropometric method. ISO 20685-1:2018 [64] sets the criteria to determine the valid compatibility between body measurements obtained from 3D body scanning technologies and those obtained by an expert using traditional anthropometry methods. These ISO thresholds are intended to be generalized to assess compatibility between two anthropometric methods. Both criteria are used and discussed in this study.

Analyses in this study will focus on the repeatability of each individual method and on the pairwise compatibility between technologies. This study does not focus on assessing the accuracy of body measurements acquired, and none of the technologies used is being considered as a ground truth.

1.2. TERMINOLOGY

For the purposes of this document, the following terms, and definitions apply. Where possible, these are in concurrence with ISO 20685-1:2018 [64], ISO 13528:2015 [57], and ISO 3534-1:2006 [58].

- *3D imaging system:* Hardware and software system that creates a digital 3D image.
- *3D image:* Digital representation of a human form, or parts thereof, in three dimensions.
- *accuracy:* Extent to which the measured value approximates a true value [63].
- *analysis of variance (ANOVA):* A collection of statistical models and procedures used to analyze the differences among means in a sample.
- *anatomical landmark* or *anatomical reference:* Point clearly defined on the body that can be used for defining anthropometric measurements [4].
- *anthropometric:* As derived from individual body measurements (anthropometric data), for example, waist girth.
- *anthropometric database:* Collection of individual body measurements (anthropometric data) and background information (demographic data) recorded on a group of people (the sample) [65].
- *compatibility*: Extent to which the measured value of a system approximates the measurement value of another system.
- *demographic:* Background or descriptive information (demographic data), for example, age.
- *hardware:* Physical components of a 3D imaging system scanner and any associated computer(s).
- *kernel density estimation (KDE):* A non-parametric method of estimating the probability density function of a random variable.
- *linear mixed model (LMM):* Statistical linear model that allows for both fixed and random effects.
- *measuring station:* A station at which instruments are used for making measurements. For this study, it could be either manual measurements using a tape measure or a 3D imaging system.
- *morphometry:* Study of the size, shape, and relation between size and shape (allometry).

- *point cloud:* Collection of raw 3D points (to be translated to a human axis system) in space referenced by their coordinate values [64].
- *repeatability:* Extent to which the values of a variable measured several times on the same subject is the same. It can be estimated between methods (interobserver) or with the same measurers (intraobserver).
- *software:* Operating system, user interface, programs, algorithms, and instructions associated with a 3D imaging system.
- *three-dimensional (3D):* Pertaining to the use of three orthogonal scales on which the three coordinates, x, y, and z, can be measured to give the precise position of any relevant anatomical point in the considered space [64].
- *z scores:* A standardized measure of performance calculated using the participant result, assigned value, and the standard deviation for proficiency assessment [58].

2. MATERIALS AND METHODS

2.1. GENERAL RESEARCH DESIGN

The design of the study was observational with repeated measures and was conducted in two phases, during which participants had a series of anthropometric measurements collected using different measurement methods.

According to optimal experiment design methods proposed by Walter et al. [59], a minimum number of 13 subjects is required to conduct reliability analyses for a test reliability hypothesis at 5% significance level with 80% power, a minimum acceptable intraclass correlation coefficient (ICC) of 0.8, an expected ICC of 0.95, and two repetitions. The minimum total number of participants per phase was set at 52 (4×13) so that future analyses could be also conducted with data segmentations by gender, age, body mass index, and any combination thereof, provided that the analyzed dataset included 13 subjects.

The study consisted of two phases: Phase 1 was carried out in December 2018 in Portland, Oregon (USA), and Phase 2 in October 2019 in Valencia (Spain). Participants were selected to represent a wide range of body heights and weights. Each subject had 11 body measurements collected twice by different measuring stations (3D scanning booths, smartphone apps, and traditional anthropometry experts). Phase 1 included two 3D scanners—Vitus bodyscan (Human Solutions GmbH) and SS20 (Size Stream LLC)—and four experts. Phase 2 included two 3D scanners: MOVE4D (Instituto de Biomecánica) and Portal MX (Texel LLC); four smartphone apps: 3D avatar body (Instituto de Biomecánica), SS@Home (Size Stream LLC), 3DLOOK (3DLOOK Inc.), and eM+ (PolyU); and two experts.

Both phases were compliant with the applicable regulations at the date of the data collection for ethical human participant research and data protection in the United States[2] and Europe.[3] The two phases of the study were approved by the Institutional Review Board (IRB) of Kalypso and by the Comité de Ética en la Investigación of the Universitat Politècnica de València (CEI-UPV), respectively.

2.2. PARTICIPANTS

A total of 133 participants participated in the study (77 females and 56 males). Participants were proactively recruited for comprehensive body height and mass representation (see TABLE 1 and FIGURE 1). The recruitment in Phase 1 was coordinated by Kalypso and was announced via website, email, and flyers distributed in Portland. The recruitment in Phase 2 was coordinated by IBV and was communicated by email and phone using IBV's own database of participants. All participants were volunteers and consented to the use of their data in the study.

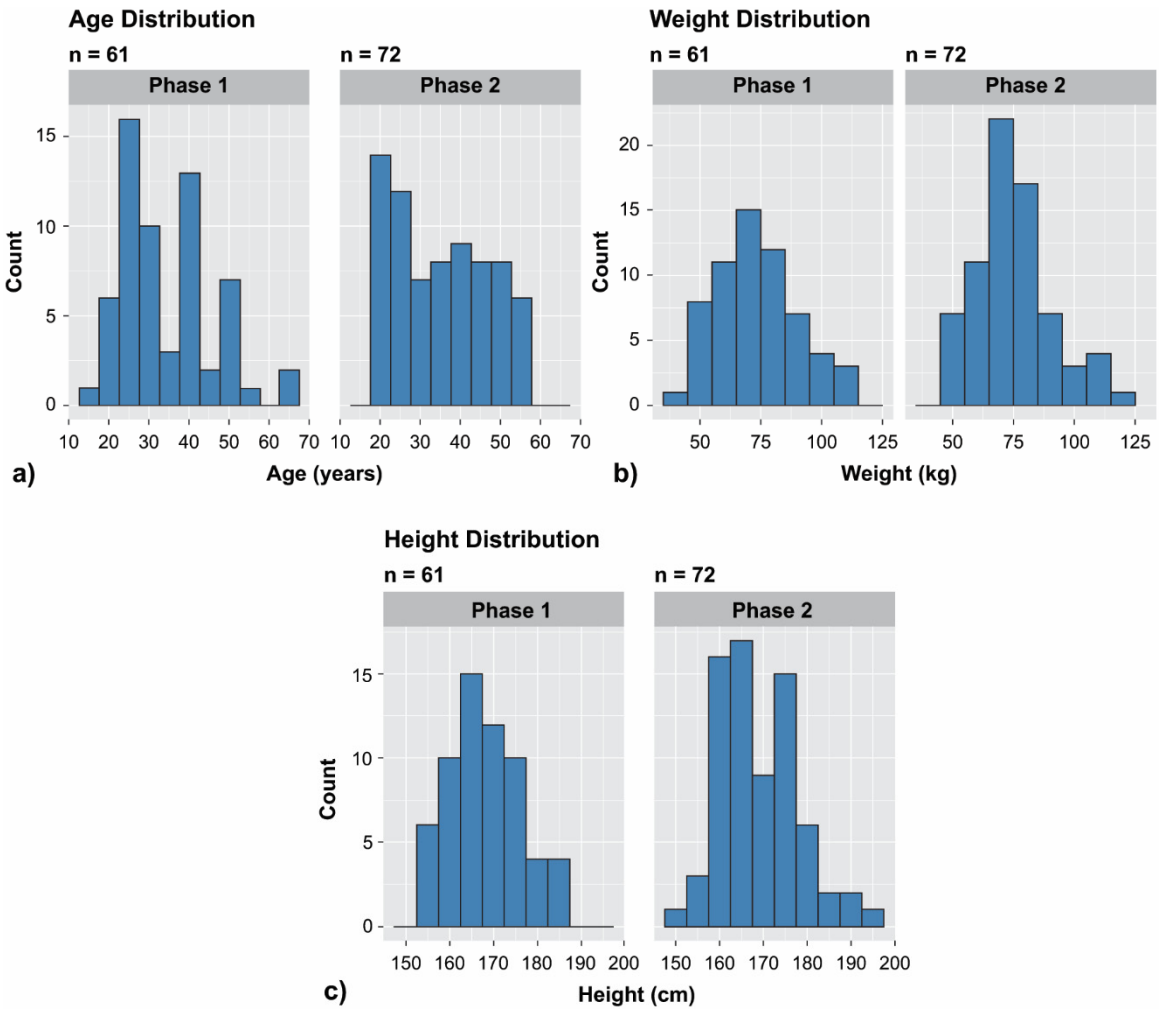
TABLE 1 Frequencies, mean ± standard deviation with range in parentheses, basic data for all participants, segmented by gender and phase

Phase	Demographic	Female	Male	Total
1	*N*	41	20	61
	Age (years)	32 ± 10 (16–57)	39 ± 13 (20–65)	34 ± 12 (16–65)
	Mass (kg)	70 ± 17 (46–114)	82 ± 16 (44–115)	74 ± 17 (44–115)
	Standing height (cm)	165 ± 6 (153–180)	174 ± 8 (156–185)	168 ± 8 (153–185)
2	*N*	36	36	72
	Age (years)	36 ± 13 (18–57)	33 ± 10 (18–55)	35 ± 12 (18–57)
	Mass (kg)	71 ± 16 (50–109)	81 ± 14 (62–122)	75± 16 (50–122)
	Standing height (cm)	163 ± 6 (150–181)	175 ± 8 (160–193)	169 ± 9 (150–193)
Total	*N*	77	56	133
	Age (years)	33.8 ± 11.9 (16–57)	35.5 ±11.3 (18–65)	34.5 ±11.7 (16–65)
	Mass (kg)	70.2 ± 16.2 (46.1–114.2)	80.9 ±14.6 (44.2–122)	74.7 ± 16.4 (44.2–122)
	Standing height (cm)	164.1 ± 6.1 (149.6–181)	174.5 ± 7.8 (156.3–193)	168.5 ± 8.6 (149.6–193)

[2] Social-Behavioral Education Foundations for Human Subjects Research, https://about.citiprogram.org/course/social-behavioral-educational-sbe-foundations/.

[3] General Data Protection Regulation (GDPR) is available from https://eur-lex.europa.eu/legal content/EN/TXT/PDF/?uri=CELEX:32016R0679.

FIGURE 1 **Histograms showing the distribution of a) age, b) body mass, and c) standing height for each phase**



Reprinted with permission from IBV.

2.3. DATA COLLECTION

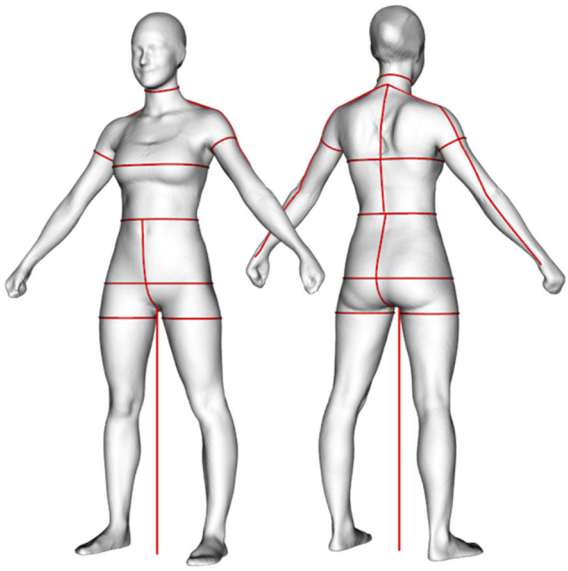
Eleven body measurements were selected for inclusion in this study (TABLE 2 and FIGURE 2). Measurements were selected for industry relevance (apparel), body segment coverage, and type variety (i.e., lengths, girths, and heights).

TABLE 2 **List of the designation and reference to ISO definitions of each of the measurements used in the study**

Number	Measurements	ISO Reference [3]
1	Neck girth	ISO 8559-1:2017, 5.3.2
2	Back neck point to waist	ISO 8559-1:2017, 5.4.5
3	Upper arm girth	ISO 8559-1:2017, 5.3.16

Number	Measurements	ISO Reference [3]
4	Back neck point to wrist	ISO 8559-1:2017, 5.4.17
5	Across back shoulder width	ISO 8559-1:2017, 5.4.3
6	Bust girth	ISO 8559-1:2017, 5.3.4
7	Waist girth	ISO 8559-1:2017, 5.3.10
8	Hip girth	ISO 8559-1:2017, 5.3.13
9	Thigh girth	ISO 8559-1:2017, 5.3.20
10	Total crotch length	ISO 8559-1:2017, 5.4.18
11	Inside leg height	ISO 8559-1:2017, 5.1.15

FIGURE 2 Illustration of the 11 body measurements selected over the surface of a 3D avatar



Fourteen measuring stations—each using a different measuring method or technique—were set up for the study: six in Phase 1 and eight in Phase 2. They included both digital technologies and traditional methods (TABLE 3). The maximum number of stations at each phase was determined by the space and resources available at each location. IEEE IC 3DBP members were prioritized for participation. For Phase 2, the call for full body scanning booths and phone apps was open to any company complying with the minimum requirements, namely eligible candidates were able to: a) set up and operate the scanning station during the five days of the study; b) deliver a 3D file in standard format of the outer body surface (e.g., mesh or point cloud); and c) deliver the set of 11 body measurements (TABLE 2). The public call was posted on LinkedIn[4] and the IEEE website. Moreover, more than 30 companies and startups identified by the IEEE IC 3DBP Group were contacted directly by email.

[4]The call for participation was posted at https://www.linkedin.com/feed/update/urn:li:activity:6557973921222664192/.

TABLE 3 Measuring stations participating in each phase by type of technology

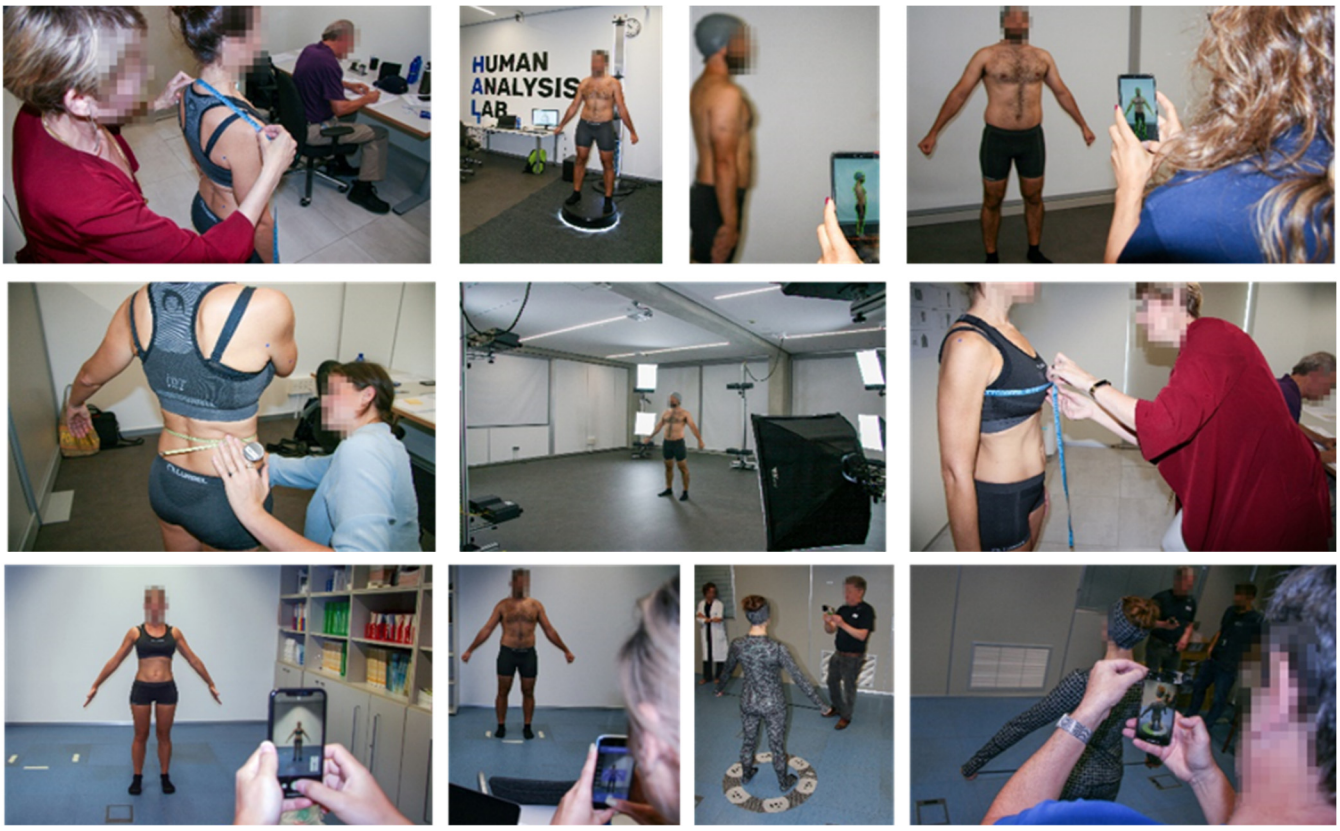
Type	Phase 1	Phase 2
Traditional anthropometry experts	▪ Julianne Harris (Independent, apparel industry) ▪ Emma Scott (Fashion Should Empower) ▪ Gina Patterson (Independent, apparel industry) ▪ Ruth Dang (Independent, apparel industry)	▪ Emma Scott (Fashion Should Empower) ▪ Alice Bullas (ISAK Level 2; Sheffield Hallam University, Sports Engineering)
Full body scanner	▪ VITUS bodyscan, Human Solutions GmbH ▪ SS20, Size Stream LLC	▪ MOVE4D, Instituto de Biomecánica ▪ Portal MX, Texel LLC
Smartphone apps		▪ 3D avatar body, Instituto de Biomecánica ▪ Size Stream @Home, Size Stream LLC ▪ 3DLOOK, 3Dlook Inc. ▪ eM+, PolyU

In order to reduce the impact of clothing on the accuracy and compatibility of the measurement method, all participants were provided with the same type of form-fitting attire, which was compatible with all the measuring stations All participants wore close-fitting shorts on top of their own underwear. Female participants also wore tight-fitting tops over their own bras. In Phase 2, one station required a specific tight-fitting two-piece, full-body suit, which was worn on top of the other garments. Participants were allowed to keep the garments provided.

Upon arrival, all participants provided basic demographic data (i.e., age and gender) and were measured with digital scales (Tanita BF-684W and SECA) to acquire mass and stadiometers (SECA) to acquire standing height. In Phase 2, these data were used as input for the phone apps requiring it.

All participants were measured twice at each measuring station (FIGURE 3). In the case of digital stations, the 3D images or photographs were taken with repositioning between the two repetitions (i.e., participants went in and out at each repetition). At each station, the participants adopted the specified pose for that station; therefore, participants' poses were slightly different from one station to another. All digital stations were operated by personnel designated by each company (e.g., trained users took the images). Each measuring station was responsible for providing to the data manager the data gathered and processed (i.e., 3D images plus anthropometrics). For traditional anthropometry stations, to prevent both order and report effects and to ensure the measurers' independence, experts were not able to share methods; the measurers did not take repetition measurements consecutively (two to three other participants were measured between a participant's repetitions); stickers were allowed to mark body features but were refreshed at each repetition; and the measurement values were annotated on paper by an independent recorder. An independent annotator digitized both the traditional measurements and the basic characteristics gathered at intake.

FIGURE 3 **Images of the data gathering during Phase 2**



**Permission granted by IBV*

The data manager, IBV, conducted additional post processing of the 3D images from actual 3D scanners (Vitus, SS20, SS@Home, MOVE4D, and PortalMX) in order to remove all facial details. Upon completion of the data collection, the data acquired from each participant included the following:

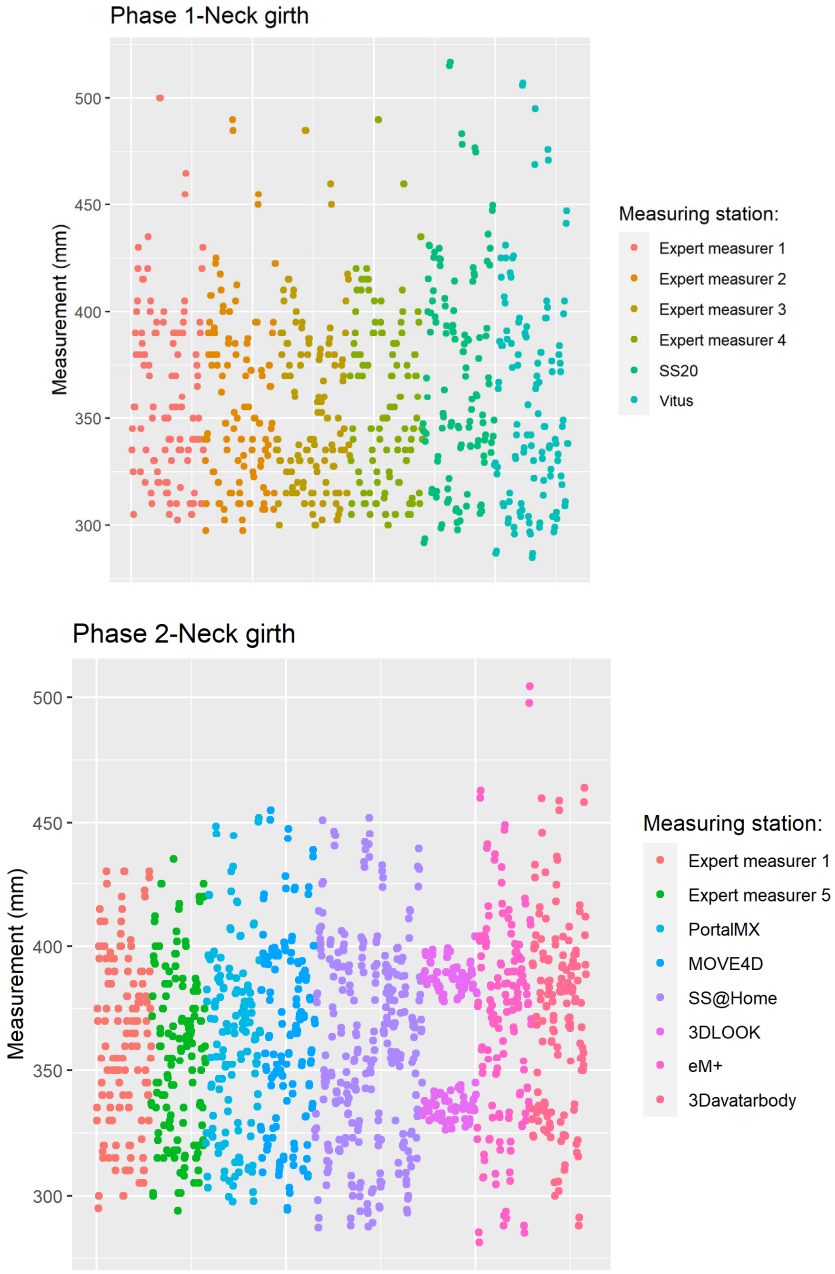
- Basic demographic and anthropometric characteristics: standing height (cm), body mass (kg), gender (female or male), and age (years)
- 3D image: a 3D point cloud or mesh representing a person's external body surface
- Faceless 3D image: a 3D point cloud or mesh representing a person's external body surface, excluding facial traits
- 11 body measurements: see TABLE 2 for a comprehensive list

There was a significant time delay between the data collection and the publication of results; therefore, four measuring stations (SS20, SS@Home, 3DLOOK, and Texel) asked to update their datasets (measurements and/or 3D scans) and resubmit data because they were able to reprocess it with improved algorithms or technology. Data resubmissions were accepted if they complied with the following conditions:

- Automatic processing was applied to all originally gathered data, which did not include any human interaction.
- No special processing was applied to a single data point or subset of data.
- Data from the study were not used to train, improve, or modify the processing algorithms used.

FIGURE 4 provides two descriptive plots of measurement data, one for each phase. As an illustrative example, only one measurement, neck girth, is chosen for the plots. The descriptive plots for all measurements and the full measurement data collected are available at the IEEE 3DBP website. The colors correspond to the measuring stations and are distributed along the x-axis. The y-axis is the raw neck girth measurement value. These plots include both genders.

FIGURE 4 Neck girth measurement data example from Phase 1 and Phase 2



2.4. QUANTITATIVE DATA ANALYSIS

After data gathering and prior to data analyses, the data were digitized and screened for outliers (due to, for example, large discrepancies or transcription errors) using a bespoke algorithm based on kernel density estimation (KDE) between pairs of measurements (e.g., neck girth versus waist girth). The cut off probability percentage was made below a given threshold (i.e., 98%). A single value was considered an outlier if its scaled two-dimensional Kernel Density Estimation (2D KDE) was below 98% for four or more pairs. Therefore, grossly inaccurate values were

identified as outliers and, where there was no clear evidence of transcription error, were removed and declared missing. Apparent transcription errors for manual measurements were corrected.

Although there are several appropriate methods to analyze repeatability and compatibility, a linear mixed model (LMM) is proposed for these analyses due to its flexibility and its capacity to handle missing data [60]. Other approaches like analysis of variance (ANOVA) models are equally suitable and would produce similar results.

Given *m* repeated measurements made by each of the *t* measuring stations on a random sample of *n* participants, an LMM can be defined as Equation (1):

$$y_{ijk} = \mu + \pi_i + \gamma_j + \epsilon_{ijk}$$
$$i = 1,\ldots,n;\ j = 1,\ldots,t;\ k = 1,\ldots,m \quad (1)$$

where y_{ijk} is the k^{th} measurement by the j^{th} measuring station on the i^{th} participant, μ is the overall population mean of the measurements, π_i is the participant effect, and γ_j is the measuring station effect; ϵ_{ijk} represents the random error. The independent components π_i and ϵ_{ik} are assumed to vary normally with means of zero and variances of σ_S^2 and σ_e^2, respectively.

Note that the number of measuring stations *t*, listed in TABLE 3, is 14. However, for repeatability calculations, the number of stations reduces to 13. This is because one expert human measurer participated in both phases; therefore, repeatability performance—but not compatibility—can be computed across both phases.

If the measuring station effects are fixed, then the component γ_j is constrained to satisfy Equation (2):

$$\sum \gamma_j = 0 \quad (2)$$

The essential repeatability estimators proposed are the standard error of measurement (SEM), intra-class correlation coefficient (ICC), coefficient of variation (CV), and the mean absolute difference (MAD), which is also known as the Gini mean difference (GMD) [61]. To simplify the process and avoid potential overfitting, it can be necessary to constrain the model to each measuring station *j*, that is, build *t* different models y_{ik} using data only from station *j* as in Equation (3):

$$y_{ik} = \mu + \pi_i + \epsilon_{ik} \quad (3)$$

where μ is now the population mean of station *j*. The specific repeatability estimators of station *j* can be easily obtained as follows:

- The SEM estimates how repeated measures on the station *j* tend to be distributed around its "true" score and it is defined as the standard deviation of the error ϵ_{ik}, as given in Equation (4):

$$SEM_j = \sigma_e \tag{4}$$

- The ICC is a measure of the reliability of measurements. It describes how strongly measurements on the same participant resemble each other and is given as Equation (5):

$$ICC_j = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_e^2} \tag{5}$$

- The CV measures the variability of the error standardized to the overall mean as in Equation (6):

$$CV_j = \frac{\sigma_e}{\mu} \tag{6}$$

The CV is a unitless measure; therefore, it is possible to condense information by averaging the CV over all measurement types (neck girth, back neck point to waist, etc.).

- With repeated measurements, the MAD is calculated using all possible differences between repetitions for each participant. Thus, for any measuring station *j*, the MAD is given by Equation (7):

$$MAD_j = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\binom{m}{2}}\sum_{k=1}^{m-1}\sum_{s=k+1}^{m}\left|y_{ijk} - y_{ijs}\right|\right) \tag{7}$$

where *n* is the number of participants, and *m* is the number of repeated measurements.

The main compatibility estimators proposed are the mean signed difference (bias) and the standard deviation of the differences between pairs of measuring stations, or pairwise standard deviation (PSD).

- The bias gives a measure of the difference between a station mean and the overall population mean μ. It can be obtained from the global model shown in Equation (1). For this purpose, the measuring station effects must be defined as fixed; however, this choice is based on the type of analysis required and, therefore, on whether the set of measuring stations are taken as a random sample of an underlying population of interest or as a specific object of interest themselves. It can be expressed as Equation (8):

$$BIAS_j = \gamma_j \tag{8}$$

Consequently, any pairwise comparison between stations, for example, can be easily calculated by subtracting their respective biases as in Equation (9):

$$BIAS_{jl} = \gamma_j - \gamma_l \tag{9}$$

Moreover, for any stations *j* and *l*, this result may provide a useful and easy way to construct a linear function that enables the transformation of body measurement data from one station to another as shown

in Equation (10):

$$f_{jl}(x) = x + BIAS_{jl} \quad (10)$$

More complex and precise functions than the above can be used to determine the pairwise transformation between stations. The bias function outlined above is deemed the primary and simplest approach.

- The PSD of stations *j* and *l* provides a measure of the distribution of the difference between them and can be obtained as in Equation (11):

$$PSD_{jl} = \sqrt{\sigma^2_{y_j - y_l} + \frac{SEM_j^2}{m} + \frac{SEM_l^2}{m}} \quad (11)$$

where $\sigma^2_{y_j - y_l}$ is the variance of the subject differences between stations *j* and *l*

These two complementary compatibility indicators can be visualized for pairs of stations with Bland-Altman plots [62], which are very useful for visualizing the agreement between two measuring methods.

In addition to studying the performance of measuring stations and the paired compatibility between them, some use cases may require analyzing groups of interest for repeatability or compatibility. For example, it can be relevant to investigate how digital technologies perform compared to traditional experts or how smartphone apps work compared to full body scanners or even how a single measurer performs compared to a group of measurers. The flexibility of the LMM allows the analysis of differences by grouping the data in the desired classes and applying the same model described previously [Equation (1)]. In this paper, as an example, three groups are taken into consideration: Digital (all digital technologies), Expert (all traditional anthropometry experts), and Global (all stations). By grouping the data, it is assumed that all measurements come from a single source, regardless of the measuring methods employed. Companies often face different data from unknown origin, so this depicts a realistic situation.

In this study, the Global station is used as a reference point for the calculation of biases. In any case, it must be understood to be the benchmark. In Section 3, each station's bias is calculated [refer to Equation (8)] against the Global station.

Regarding assessments, ANSUR thresholds were used for the evaluation of repeatability (MAD only), and ISO 20685-1:2018 thresholds and methods were used for the evaluation and discussion of pairwise compatibility (TABLE 4).

TABLE 4 Acceptability thresholds from literature

Measurement	ANSUR acceptability thresholds for MAD (mm)	ISO compatibility thresholds for biases (mm)
Neck girth	6	4
Back neck point to waist	5	5
Upper arm girth, right	6	4
Back neck point to wrist, right	9	5
Across back shoulder width	8	4
Bust girth	14	9
Waist girth	12	9
Hip girth	12	9
Thigh girth, right	6	9
Total crotch length	18	9
Inside leg height	10	4

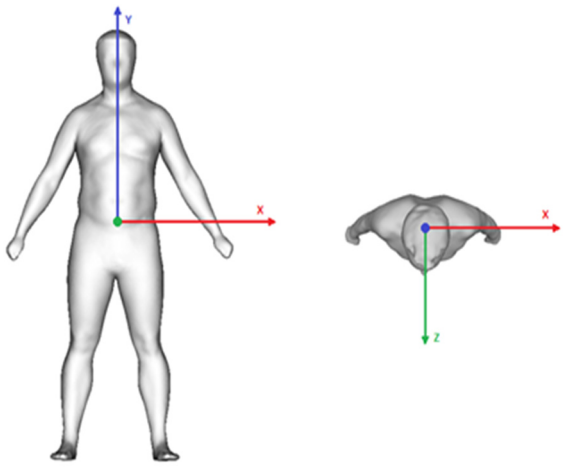
2.5. QUALITATIVE DATA ANALYSIS

With digital technologies, there is another critical aspect to consider when selecting a 3D imaging technology: the 3D digital object itself that represents the body surface of the subject. In some use cases, the 3D digital object is as important as the measurements.

To visually assess the accuracy of the 3D objects provided by each measuring station, image renders of the front, side, and back of each participant from each measuring station were acquired through the following framework:

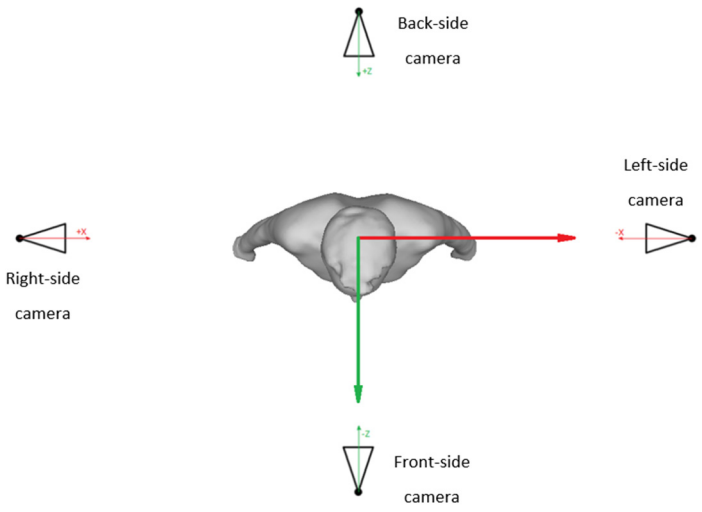
1. Object orientation and location:
    - Object (body model) orientation: Common axis convention is used for all objects. A reference system configuration is provided in FIGURE 5.
    - Object (body model) location: The object is located by matching the average centroid of the bounding box calculated for each participant (including all repetitions) with the origin of the reference system (pure coordinates).
    - Camera orientation: The camera orientation changes depending on the view: front camera points toward negative Z, back camera toward positive Z, left-side camera toward negative X, and right-side camera toward positive X (FIGURE 5).

FIGURE 5 Reference system configuration



2. Virtual camera settings:
    - Parallel projection (orthogonal view)
    - Camera location: The camera is placed using a proportional distance to the height of each participant (*p*) from the origin of the reference system in the pure direction of axis X and Z according to the view (FIGURE 6).
        - Front-side camera coordinates: (0, 0, *p*)
        - Back-side camera coordinates: (0, 0, -*p*)
        - Left-side camera coordinates: (*p*, 0, 0)
        - Right-side camera coordinates: (-*p*, 0, 0)
    - Light conditions: Same for all objects

FIGURE 6 Camera orientation



3. Image format:
    - Alpha channel for transparent background
    - High resolution

The translation of the object that can be made during the process is equivalent to the rotation of the camera. All of these parameters can be applied using the Visualization Toolkit (VTK) library (Python). The proportional distance used to place the camera is 2.2 times the height of each participant ($p=2.2\times h$). Light conditions are set by default. The dimensions of the images obtained are 5400 × 6366 pixels for front and back views and 3600 × 6366 pixels for left and right views.

3. RESULTS AND DISCUSSION

3.1. OUTLIERS AND MISSING DATA

The original algorithm identified a total of 34 outliers (out of 20 592 data points), 13 from the Expert measurement set and 21 from the Digital set. Details for each station are provided in TABLE 5. Expert measurement outliers were mainly due to transcription errors during digitization, and some of the clear typographic errors were amended, while digital measurement outliers were due to problems in the acquisition, reconstruction, or measurement. After removing these data points, the percentage of missing data was 0.7%.

TABLE 5 Measuring station and number of outliers

Measuring Station	Number of Outliers	Phase
Expert measurer 1	1	1
Expert measurer 2	0	
Expert measurer 3	0	
Expert measurer 4	5	
Vitus	2	
SS20	6	
Expert measurer 1	3	2
Expert measurer 5	4	
3DLOOK	2	
MOVE4D	0	
Texel	0	
3Davatar body	0	
eM+	1	
SS@Home	10	

3.2. REPEATABILITY

SEM, ICC, CV, and MAD results are presented, respectively, in TABLE 6 to TABLE 9. Each column corresponds to one body measurement, while each row represents a measuring station or a specific group of interest (Digital, Expert, or Global). FIGURE 6 to FIGURE 8 present, respectively, bar plots of the SEM, MAD, and CV of each measurement, where each colored bar represents one station. The red horizontal line in FIGURE 8 represents the ANSUR tolerances [6] for each body measurement. This tolerance can be considered the achievable performance of manual measurement obtained under the best conditions, with the best training, and with a full suite of anthropometric equipment. FIGURE 9 shows the average CV for each station using the same color code.

TABLE 6 SEM (mm) for each anthropometric, for each measuring station, all expert measurers, all digital stations, and all stations

Measuring station	Neck girth	Back neck point to waist	Upper arm girth (right)	Back neck point to wrist (right)	Across back shoulder width	Bust girth	Waist girth	Hip girth	Thigh girth (right)	Total crotch length	Inside leg height
Expert measurer 1	5.4	10.0	5.5	9.6	8.0	11.7	13.5	11.6	9.2	21.3	11.3
Expert measurer 2	5.3	12.5	6.0	8.2	9.4	12.7	10.2	7.6	9.4	19.6	6.1
Expert measurer 3	5.8	10.0	5.5	9.4	12.0	9.5	8.0	9.8	12.6	13.5	6.0
Expert measurer 4	5.3	9.4	6.3	9.3	7.7	13.0	13.1	9.7	8.8	19.2	7.1
Expert measurer 5	4.6	10.5	3.3	6.1	7.9	10.3	9.7	7.8	6.1	16.2	5.9
SS20	4.5	7.2	3.9	5.5	5.4	6.8	10.4	8.0	9.1	14.4	6.5
Vitus	4.0	6.2	5.6	5.3	9.8	8.2	8.6	5.0	3.7	10.1	4.2
PortalMX	3.9	3.6	4.7	7.5	8.7	7.3	4.1	3.7	5.5	8.4	4.6
MOVE4D	2.3	5.0	2.5	5.0	6.1	7.9	6.5	2.7	2.6	8.4	3.1
SS@Home	7.5	11.3	5.9	11.6	9.6	12.6	11.2	9.3	9.1	16.0	7.6
3DLOOK	1.2	2.4	3.9	3.8	3.5	10.6	16.4	7.6	7.0	7.0	3.0
eM+	2.4	2.5	1.9	4.6	2.8	5.4	5.1	6.8	3.9	4.5	4.5
3Davatar body	4.6	5.3	4.3	8.0	6.9	11.3	8.6	7.6	5.3	8.6	4.8
All expert	6.5	18.6	8.5	19.2	12.8	15.8	18.2	13.3	11.4	29.3	12.3
All digital	13.7	24.0	14.8	19.1	23.1	22.6	38.2	16.8	15.6	39.7	17.0
Global (all)	13.0	22.5	13.1	19.6	21.8	22.8	34.4	17.3	14.7	38.8	16.3

TABLE 7 ICC for each anthropometric, for each measuring station, all expert measurers, all digital stations, and all stations

Measuring station	Neck girth	Back neck point to waist	Upper arm girth (right)	Back neck point to wrist (right)	Across back shoulder width	Bust girth	Waist girth	Hip girth	Thigh girth (right)	Total crotch length	Inside leg height
Expert measurer 1	0.980	0.911	0.982	0.964	0.938	0.990	0.990	0.988	0.980	0.925	0.948
Expert measurer 2	0.983	0.769	0.980	0.965	0.919	0.990	0.995	0.995	0.982	0.929	0.984
Expert measurer 3	0.978	0.865	0.983	0.952	0.851	0.995	0.997	0.992	0.968	0.963	0.983
Expert measurer 4	0.984	0.912	0.979	0.958	0.951	0.990	0.991	0.992	0.983	0.911	0.975
Expert measurer 5	0.980	0.884	0.994	0.984	0.936	0.990	0.995	0.994	0.990	0.945	0.988
SS20	0.992	0.968	0.990	0.984	0.981	0.998	0.995	0.994	0.982	0.965	0.980
Vitus	0.994	0.943	0.976	0.987	0.924	0.997	0.997	0.998	0.997	0.983	0.991
PortalMX	0.988	0.986	0.987	0.974	0.895	0.995	0.999	0.999	0.993	0.977	0.991
MOVE4D	0.996	0.972	0.996	0.988	0.973	0.994	0.998	0.999	0.998	0.978	0.995
SS@Home	0.966	0.932	0.972	0.946	0.917	0.985	0.991	0.991	0.980	0.969	0.966
3DLOOK	0.998	0.992	0.986	0.991	0.985	0.990	0.985	0.993	0.984	0.987	0.997
eM+	0.997	0.995	0.997	0.993	0.993	0.997	0.998	0.996	0.996	0.995	0.994
3Davatar body	0.985	0.968	0.986	0.972	0.968	0.988	0.996	0.995	0.993	0.977	0.990
All expert	0.969	0.665	0.959	0.855	0.853	0.981	0.982	0.983	0.969	0.847	0.939
All digital	0.903	0.601	0.854	0.841	0.644	0.965	0.925	0.972	0.940	0.704	0.875
Global (all)	0.896	0.607	0.894	0.840	0.640	0.962	0.936	0.971	0.948	0.721	0.888

TABLE 8 CV (%) for each anthropometric, for each measuring station, all expert measurers, all digital stations, and all stations

Measuring station	Neck girth	Back neck point to waist	Upper arm girth (right)	Back neck point to wrist (right)	Across back shoulder width	Bust girth	Waist girth	Hip girth	Thigh girth (right)	Total crotch length	Inside leg height
Expert measurer 1	1.5	2.4	1.8	1.2	1.9	1.2	1.6	1.1	1.5	2.8	1.5
Expert measurer 2	1.5	3.1	2.0	1.1	2.3	1.3	1.2	0.7	1.5	2.6	0.8
Expert measurer 3	1.6	2.4	1.8	1.2	2.9	1.0	0.9	1.0	2.1	1.8	0.8
Expert measurer 4	1.5	2.2	2.1	1.2	1.8	1.3	1.5	0.9	1.5	2.7	0.9
Expert measurer 5	1.3	2.6	1.0	0.8	1.8	1.0	1.1	0.8	1.0	2.1	0.8
SS20	1.2	1.7	1.3	0.7	1.3	0.7	1.2	0.8	1.5	1.9	0.9
Vitus	1.1	1.5	1.9	0.7	2.3	0.8	1.0	0.5	0.6	1.2	0.6
PortalMX	1.1	0.9	1.4	1.0	2.0	0.7	0.5	0.4	0.9	1.1	0.6
MOVE4D	0.6	1.2	0.8	0.7	1.5	0.8	0.7	0.3	0.4	1.1	0.4

Measuring station	Neck girth	Back neck point to waist	Upper arm girth (right)	Back neck point to wrist (right)	Across back shoulder width	Bust girth	Waist girth	Hip girth	Thigh girth (right)	Total crotch length	Inside leg height
SS@Home	2.1	2.5	1.8	1.5	2.4	1.3	1.2	0.9	1.5	2.1	1.0
3DLOOK	0.3	0.5	1.3	0.5	0.8	1.1	1.9	0.7	1.1	0.9	0.4
eM+	0.6	0.6	0.6	0.6	0.6	0.5	0.6	0.7	0.6	0.6	0.6
3Davatar body	1.2	1.2	1.3	1.0	1.7	1.1	1.0	0.7	0.9	1.1	0.6
All expert	1.8	4.5	2.7	2.4	3.0	1.6	2.1	1.3	1.9	3.8	1.6
All digital	3.8	5.7	4.7	2.4	5.5	2.3	4.3	1.6	2.6	5.1	2.3
Global (all)	3.6	5.3	4.2	2.5	5.2	2.3	3.9	1.7	2.4	5.1	2.2

TABLE 9 MAD (mm) for each anthropometric, for each measuring station, all expert measurers, all digital stations, and all stations

Measuring station	Neck girth	Back neck point to waist	Upper arm girth (right)	Back neck point to wrist (right)	Across back shoulder width	Bust girth	Waist girth	Hip girth	Thigh girth (right)	Total crotch length	Inside leg height
Expert measurer 1	5.8	11.0	5.3	10.2	8.3	11.5	12.2	10.8	9.9	20.0	9.4
Expert measurer 2	4.8	11.4	5.5	7.8	9.8	10.3	11.2	8.0	10.1	20.0	6.0
Expert measurer 3	5.1	10.5	5.5	10.0	12.0	9.6	8.7	10.4	11.1	14.8	6.8
Expert measurer 4	4.5	10.5	6.7	10.0	8.0	13.2	13.4	10.7	8.0	17.1	6.5
Expert measurer 5	4.8	10.9	3.7	6.7	8.2	9.4	10.8	8.4	5.9	17.3	6.0
SS20	4.6	7.1	4.1	4.9	5.9	7.6	7.9	7.1	6.1	14.3	5.1
Vitus	3.8	6.3	5.4	6.0	9.3	8.6	8.4	5.3	3.7	9.8	4.2
PortalMX	4.2	3.9	4.9	9.2	9.8	8.5	4.5	4.1	6.5	8.7	5.0
MOVE4D	2.3	5.1	2.7	5.6	7.0	8.5	6.5	3.1	2.6	8.0	3.1
SS@Home	7.8	12.0	6.8	12.8	10.6	14.3	11.6	10.1	9.8	17.8	7.1
3DLOOK	1.2	2.7	3.3	3.9	3.2	10.1	12.7	6.9	6.5	6.0	3.3
eM+	3.0	3.1	2.2	5.7	3.7	6.4	6.0	8.4	5.0	5.6	5.3
3Davatar body	5.1	5.9	4.7	9.0	7.4	12.2	9.7	7.9	6.1	9.7	5.6
All expert	6.7	20.3	9.2	23.3	12.3	15.6	17.9	14.1	10.7	27.2	12.8
All digital	12.6	23.6	14.1	16.8	22.3	20.2	35.5	14.2	14.2	43.5	14.9
Global (all)	12.4	24.2	12.7	21.5	23.0	22.8	32.3	17.3	15.4	41.6	16.8

FIGURE 7 **SEM for each anthropometric, for each measuring station, all expert measurers, all digital stations, and all stations**

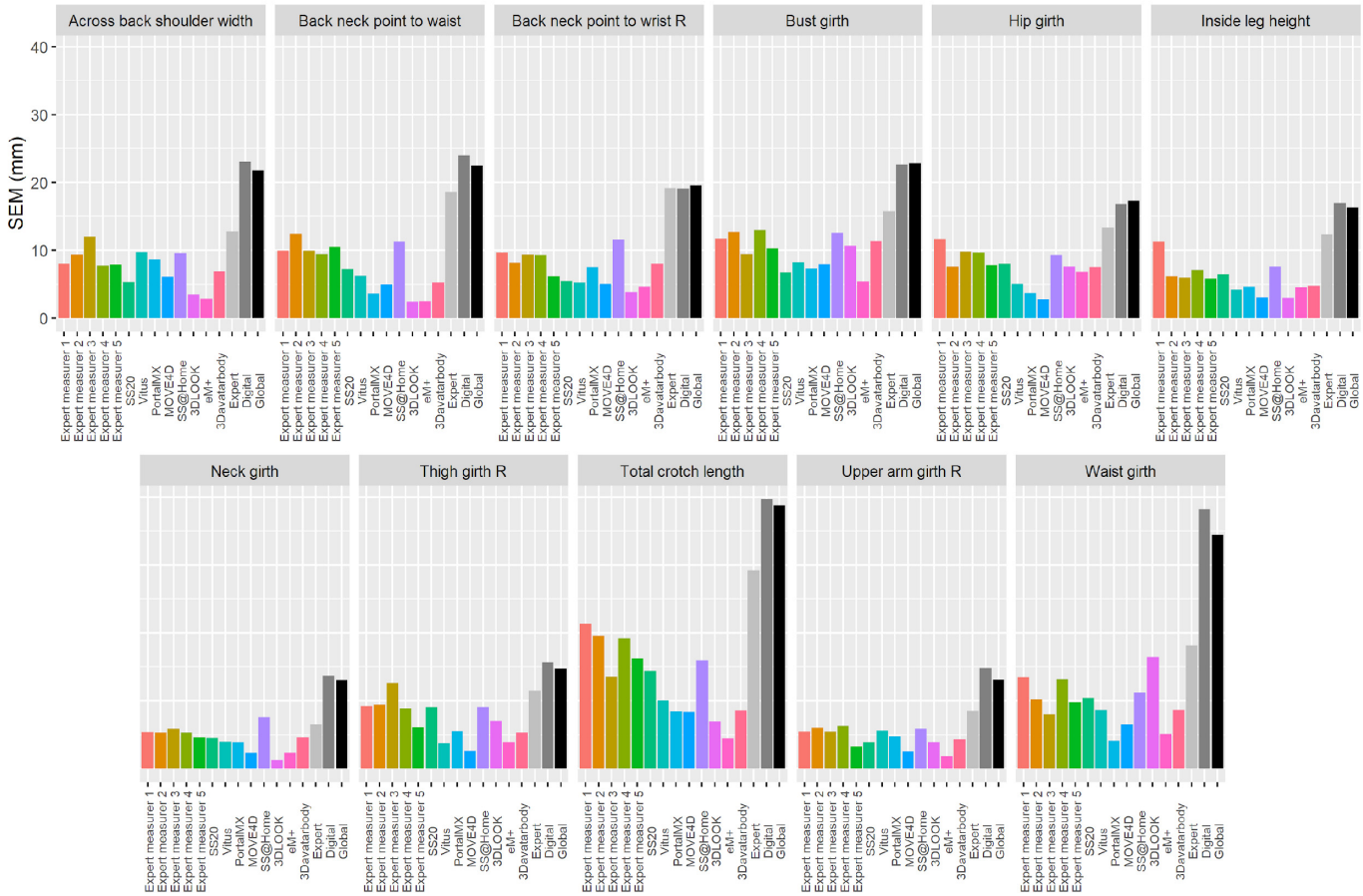


FIGURE 8 **MAD results for each anthropometric, for each measuring station, all expert measurers, all digital stations, and all stations (ANSUR tolerances illustrated by the red line)**

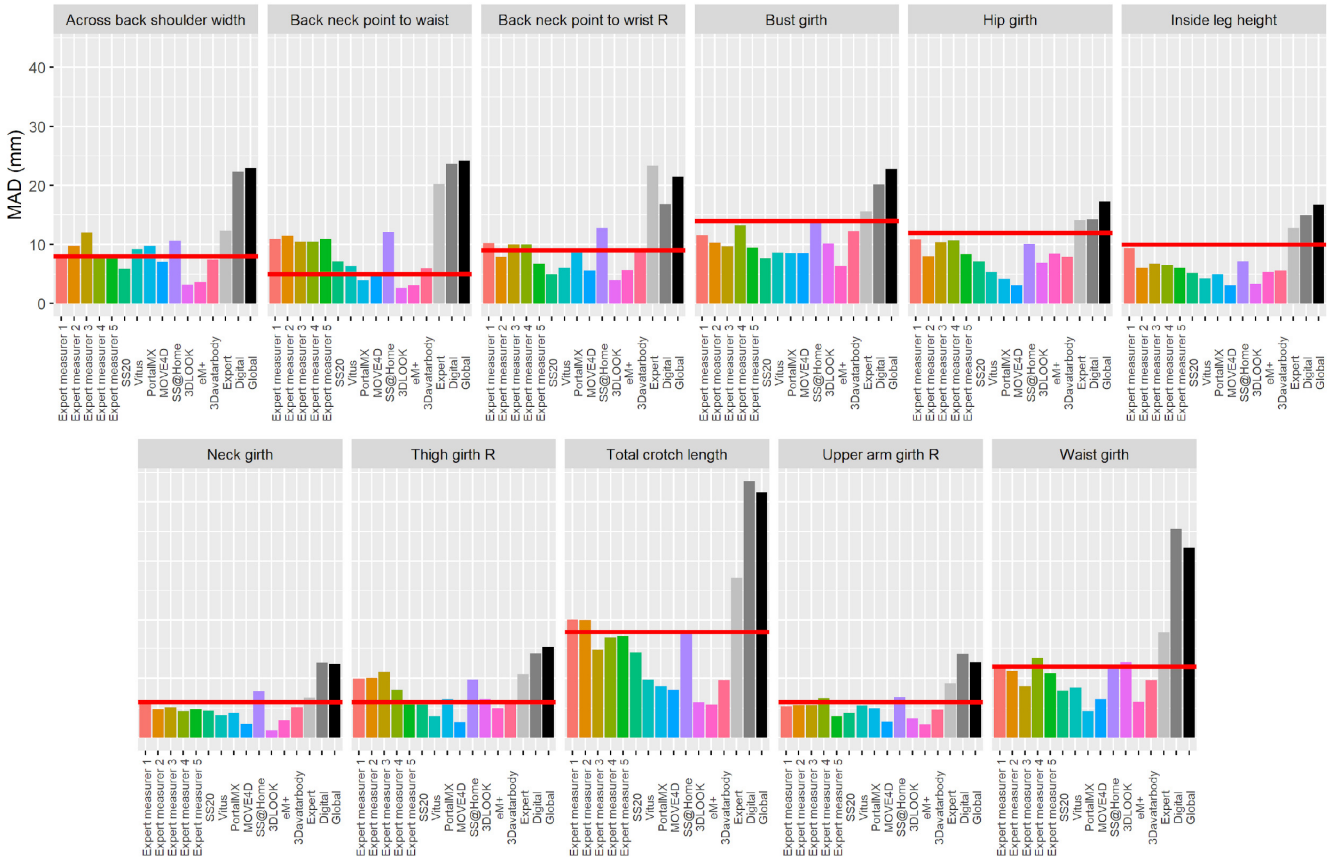


FIGURE 9 **CV (%) for each anthropometric, for each measuring station, all expert measurers, all digital stations, and all stations**

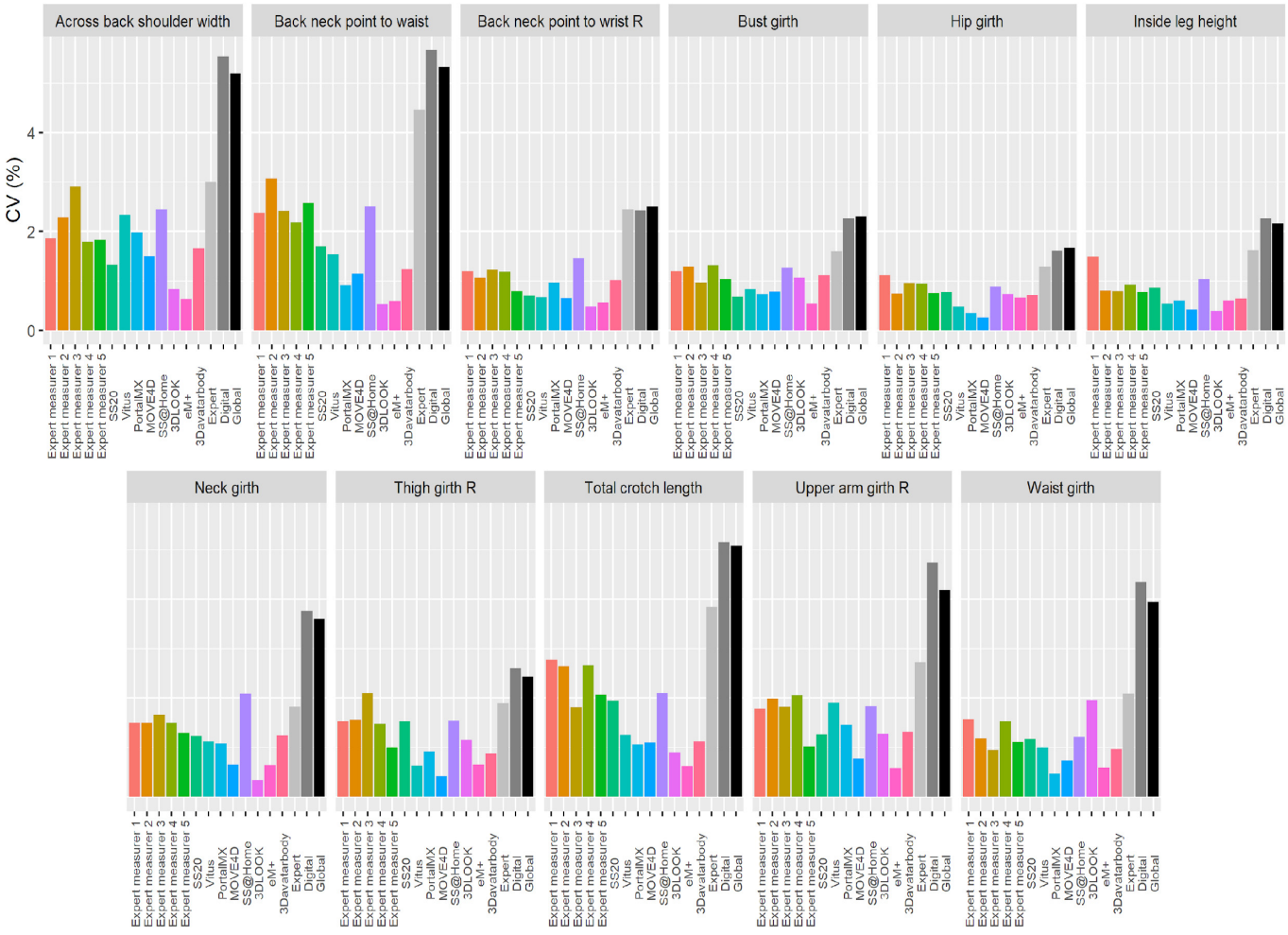
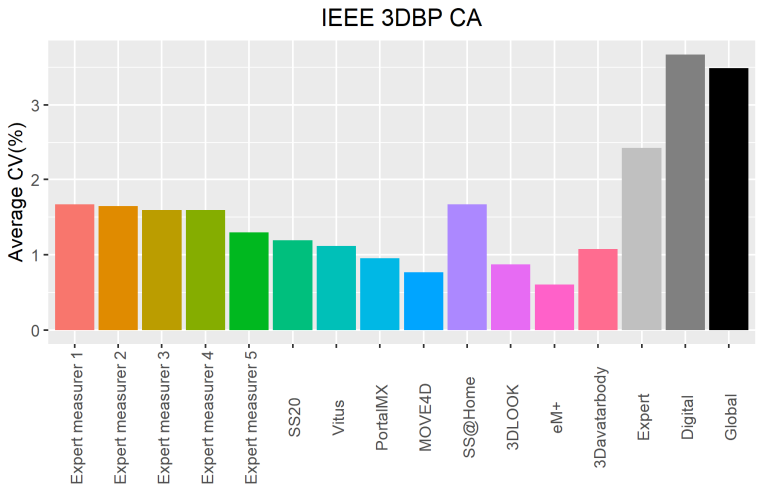


FIGURE 10 Average CV (%) for all anthropometrics, for each measuring station, all expert measurers, all digital stations, and all stations



Overall, the 13 anthropometric techniques showed high repeatability with a low error: less than 2% error (CV) and SEM values ranging from 2 mm to 20 mm, depending on the measurement. Therefore, any single expert measurer or single scanning technology is highly repeatable. All measuring stations showed repeatability MAD errors within ANSUR thresholds for most of the measurements (TABLE 7); 70% of Expert measurements were within tolerance and 84% of Digital measurements were within tolerance. However, the repeatability varied substantially depending upon the measurement and the station. Digital measuring stations demonstrated higher levels of repeatability than manual measuring stations, except for SS@Home, which was comparable to manual stations. These results align with previous investigations, showing that digital measuring methods can reduce much of the error associated with traditional data acquisition. They also reinforce the argument that manual measurements are not a suitable ground truth—or "gold standard,"—against which 3D imaging systems can be compared to determine accuracy.

It is important to note that the use of model-based approaches (like some phone apps) may impact the low level of repeatability error. Identical values of measurements between multiple scans of the same human or between multiple people for a particular measurement may indicate the use of model-derived data with reduced (or even without) actual input data to simulate input at the regions where the measurement is taken. A repeatability error of 0.00% per station is not obtainable if multiple scans of data for an individual or multiple individuals' data are being compared. Similarly, unexpectedly small repeatability errors should be regarded with caution, especially those resulting from model-based technologies that use sparse inputs like phone apps. In this dataset, for example, both the data models for eM+ and 3DLOOK produce completely symmetrical avatars, which have identical body measurements for left and right limbs. We find another example in the neck girth measurement from 3DLOOK

(FIGURE 4). It is clear that the variability of the gathered data is centered unusually in one value for males and in another value for females. It also achieves a SEM of 1.2 mm, much smaller than the SEMs for the rest of the stations, which range from 2.3 mm to 7.5 mm.

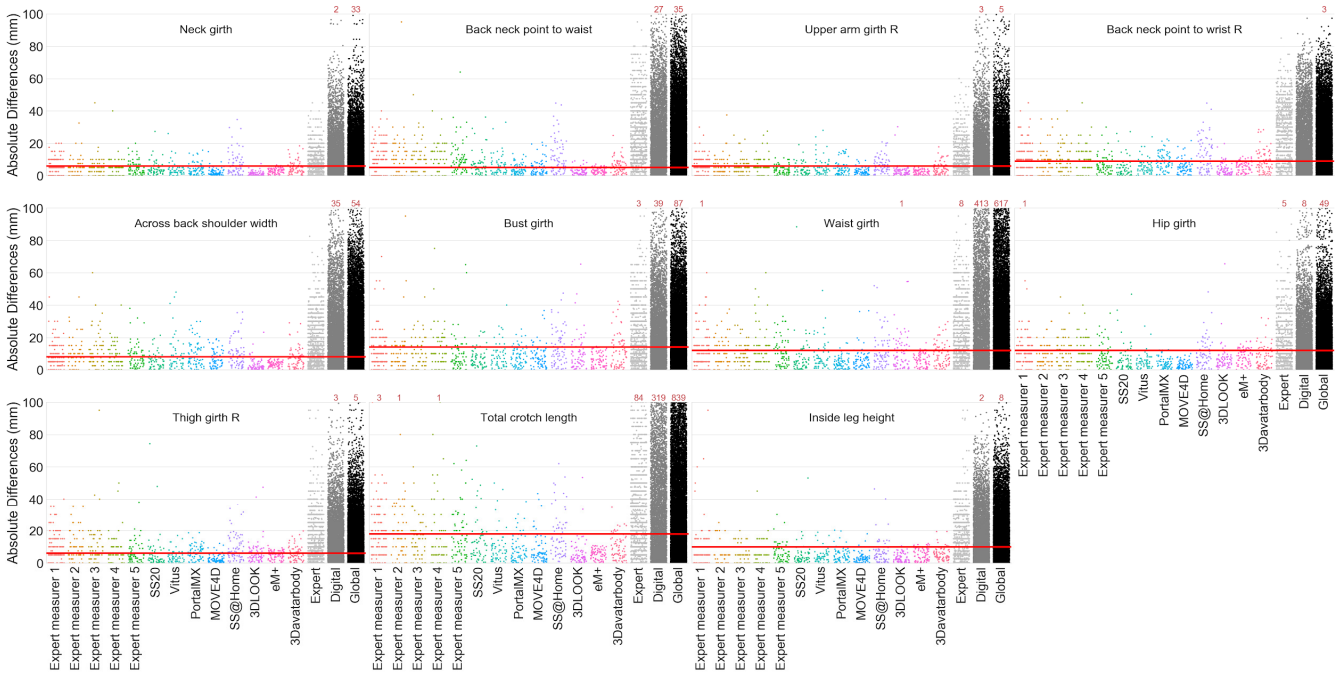
This study considered 11 body measurements common to various apparel use cases. For both digital and manual anthropometric techniques, hip girth, bust girth, and inside leg height demonstrated the highest degrees of repeatability, while back neck point to waist, total crotch length, across back shoulder width, and upper arm girth showed the lowest degrees of repeatability. This may be attributable to the ease and objectivity through which the respective anatomical references involved in each type of measurement are identified. While each measuring station was instructed to extract standardized measurements, the methods and landmarks used by each measuring station likely varied, possibly causing discrepancies in repeatability values between different stations for the same measurement. This highlights the need for clearer guidance and standardization in anthropometric definitions for landmarks.

When grouping different technologies to compute the joint repeatability errors, if the joint error is lower than each individual's then the agreement between the grouped methods is high; increasing the number of repetitions helps lower the observed error. Conversely, if the joint error increases, the agreement between methods is low. The results of this study suggest that repeatability error increases significantly when data is aggregated from any grouping of anthropometric techniques or measuring stations. This means that a team of manual expert measurers or a grouping of different scanning technologies—or a blend of the two—yields significantly worse repeatability results than any single human measurer or scanning technology. This is an important result to consider because most industry use case scenarios commonly combine multiple expert measures, expert manual and digital anthropometric techniques, or multiple digital anthropometric techniques (or measuring stations), taking for granted that, for example, measurements provided by different vendors but having the same designation are compatible. Teams of experts are also common in large anthropometric studies. Most retail contexts usually do not operate—at any medium to large scale—without a team of manual expert measurers. Although using multiple scanning technologies is a realistic option, medium- to large-scale retail contexts struggle to operate without a team of manual expert measurers. Therefore, in recognition of a tendency toward a built-in repeatability error within traditional apparel practice, the results of this study favor digital over manual anthropometrics for repeatability.

FIGURE 10 displays, per measuring station and per body measurement, a scatter-plot of all the absolute differences between the two measurements taken for each subject. For example, in the neck girth subplot, the left-most collection of dots contains dots per subject measured by Expert measurer 1. Each dot is the absolute difference between the two neck girth measurements Expert measurer 1 performed. (The red dots are the number

of times that the absolute differences were greater than 100 mm.) The y-axis upper limit was fixed at 100 mm; the red numbers above some columns count (for those measuring stations) the number of dots above 100 mm. The red horizontal line in FIGURE 11 represents the ANSUR tolerances [6] for each body measurement. More plots are available on the IEEE 3DBP website.

FIGURE 11 A scatter-plot of all the absolute differences per measuring station, per body measurement, and between two measurements taken for each subject



3.2.1. PAIRWISE COMPATIBILITY

TABLE 10 and TABLE 11 show, respectively, the bias against the grand mean for each station for Phase 1 and Phase 2. Pairwise compatibility can only be applied within phase. The bias between two specific stations is computed by subtracting the values in the table for the pair. For example, the bias between Expert measurer 1 and Expert measurer 2 for neck girth is 3.9 mm [$1.4 - (-2.5) = 3.9$]. Expert measurer 1 provided higher values for that measurement.

TABLE 10 Mean signed difference (bias), in mm, for each measuring station in Phase 1

Measuring station	Neck girth	Back neck point to waist	Upper arm girth (right)	Back neck point to wrist (right)	Across back shoulder width	Bust girth	Waist girth	Hip girth	Thigh girth (right)	Total crotch length	Inside leg height
Expert measurer 1	1.4	-9.3	3.9	5.5	7.6	-19.9	0	-0.5	2.8	6.1	0.3
Expert measurer 2	-2.5	-7.5	-2.2	-13.2	-5.7	3.3	-3.7	-6	9.5	-17.6	2.9
Expert measurer 3	-5.5	-1.6	-1.3	-18.7	-3	-5.5	-15.3	-3.3	-1.8	0	-7.7
Expert measurer 4	-1.3	18.4	1.3	8	12.5	4.8	0.6	-7.2	3.2	-36.9	7.1
SS20	10.3	9.5	8	5.1	-13.5	5.2	23.9	11	-0.7	-8.7	-3.9
Vitus	-1.8	-9.5	-8.2	12.7	2.4	9.5	-4.6	5.7	-11.6	53.4	1.2
All expert	-2.2	0.1	0.2	-4.8	2.8	-3.9	-4.9	-4.4	3.4	-12.1	0.7
All digital	4.1	-0.3	-0.3	9.1	-5.4	7.4	9.3	8.3	-6.3	23	-1.3
Grand Means	359.3	413.4	303.7	776.5	416.4	981.4	865.1	1033.2	600.5	754.9	756.4

TABLE 11 Mean signed difference (bias), in mm, for each measuring station in Phase 2

Measuring station	Neck girth	Back neck point to waist	Upper arm girth (right)	Back neck point to wrist (right)	Across back shoulder width	Bust girth	Waist girth	Hip girth	Thigh girth (right)	Total crotch length	Inside leg height
Expert measurer 1	-1.9	2.6	-9.7	22.1	9.0	-1.2	-12.0	1.3	1.4	4.8	3.0
Expert measurer 5	-6.6	-21.5	3.0	-14.4	7.6	-5.0	1.5	-12.8	5.1	13.0	12.2
PortalMX	-1.0	-31.1	5.8	-15.2	15.7	-3.3	-5.4	10.5	-1.6	27.4	19.2
MOVE4D	1.0	3.1	2.1	-21.9	-15.8	16.6	13.6	6.6	3.8	-2.5	-14.9
SS@Home	-4.7	20.0	2.2	5.0	-30.6	-3.8	47.3	7.8	-15.8	-8.8	-12.0
3DLOOK	-2.3	28.0	-12.5	-3.1	-3.9	-6.8	-32.6	-8.2	7.5	10.4	2.3
eM+	7.3	0.5	1.5	21.8	23.1	-16.3	-29.2	-11.2	-2.4	-43.5	-1.9
3Davatar body	8.2	-1.4	7.3	6.0	-5.5	19.8	17.0	6.0	1.9	-1.1	-7.8
All expert	-4.2	-9.4	-3.3	3.8	8.3	-3.1	-5.3	-5.8	3.3	8.9	7.6
All digital	1.4	3.1	1.1	-1.3	-2.8	1.0	1.8	1.9	-1.1	-3.0	-2.5
Grand means	363.9	428.9	320.0	787.6	422.9	999.9	880.6	1044.1	608.7	769.5	750.2

TABLE 12 to TABLE 19 show, for example, the PSD for several measurements (two small girths: neck girth and upper arm girth; and two large girths: waist girth and hip girth) for Phase 1 and Phase 2. PSD gives an idea of the degree of paired variability between two measuring methods for a given measurement independent of the bias value. The tables for all measurements are available at the IEEE 3DBP website.

TABLE 12 PSD (mm) for each pair of measuring stations in Phase 1 (neck girth)

	Expert measurer 1	Expert measurer 2	Expert measurer 3	Expert measurer 4	Vitus	SS20
Expert measurer 1	NA	8.0	9.1	8.9	21.6	23.2
Expert measurer 2	8.0	NA	7.8	8.2	20.3	21.9
Expert measurer 3	9.1	7.8	NA	9.1	20.7	22.2
Expert measurer 4	8.9	8.2	9.1	NA	20.0	21.1
Vitus	21.6	20.3	20.7	20.0	NA	11.7
SS20	23.2	21.9	22.2	21.1	11.7	NA

TABLE 13 PSD (mm) for each pair of measuring stations in Phase 2 (neck girth)

	Expert measurer 1	Expert measurer 5	3DLOOK	3Davatar body	MOVE4D	eM+	SS@Home	PortalMX
Expert measurer 1	NA	8.9	21.4	15.7	9.5	19.2	13.5	10.9
Expert measurer 5	8.9	NA	19.4	16.0	11.4	19.9	15.5	12.3
3DLOOK	21.4	19.4	NA	25.4	25.2	31.8	27.6	23.5
3Davatar body	15.7	16.0	25.4	NA	14.8	18.8	18.2	12.2
MOVE4D	9.5	11.4	25.2	14.8	NA	17.3	10.5	9.1
eM+	19.2	19.9	31.8	18.8	17.3	NA	20.1	16.2
SS@Home	13.5	15.5	27.6	18.2	10.5	20.1	NA	13.6
PortalMX	10.9	12.3	23.5	12.2	9.1	16.2	13.6	NA

TABLE 14 PSD (mm) for each pair of measuring stations in Phase 1 (upper arm girth, right)

	Expert measurer 1	Expert measurer 2	Expert measurer 3	Expert measurer 4	Vitus	SS20
Expert measurer 1	NA	10.7	8.8	11.8	16.3	11.2
Expert measurer 2	10.7	NA	10.1	10.9	14.9	10.4
Expert measurer 3	8.8	10.1	NA	10.4	14.6	10.7
Expert measurer 4	11.8	10.9	10.4	NA	17.1	12.4
Vitus	16.3	14.9	14.6	17.1	NA	13.6
SS20	11.2	10.4	10.7	12.4	13.6	NA

TABLE 15 PSD (mm) for each pair of measuring stations in Phase 2 (upper arm girth, right)

	Expert measurer 1	Expert measurer 5	3DLOOK	3Davatar body	MOVE4D	eM+	SS@Home	PortalMX
Expert measurer 1	NA	11.0	18.6	16.0	8.4	28.7	12.8	10.5
Expert measurer 5	11.0	NA	18.6	14.9	7.7	30.2	15.3	8.4
3DLOOK	18.6	18.6	NA	16.0	17.1	27.0	18.4	19.3
3Davatar body	16.0	14.9	16.0	NA	13.6	27.4	15.9	15.8
MOVE4D	8.4	7.7	17.1	13.6	NA	28.5	12.8	7.4
eM+	28.7	30.2	27.0	27.4	28.5	NA	26.0	31.6
SS@Home	12.8	15.3	18.4	15.9	12.8	26.0	NA	14.3
PortalMX	10.5	8.4	19.3	15.8	7.4	31.6	14.3	NA

TABLE 16 PSD in mm for each pair of measuring stations in Phase 1 (waist girth)

	Expert measurer 1	Expert measurer 2	Expert measurer 3	Expert measurer 4	Vitus	SS20
Expert measurer 1	NA	30.4	33.7	30.9	60.4	53.3
Expert measurer 2	30.4	NA	18.7	23.1	54.8	47.0
Expert measurer 3	33.7	18.7	NA	24.5	56.0	48.3
Expert measurer 4	30.9	23.1	24.5	NA	57.6	48.7
Vitus	60.4	54.8	56.0	57.6	NA	38.3
SS20	53.3	47.0	48.3	48.7	38.3	NA

TABLE 17 Pairwise standard deviation of the differences (PSD) in mm for each pair of measuring stations in Phase 2 (waist girth)

	Expert measurer 1	Expert measurer 5	3DLOOK	3Davatar body	MOVE4D	eM+	SS@Home	PortalMX
Expert measurer 1	NA	20.6	38.5	34.0	20.9	28.1	50.6	16.8
Expert measurer 5	20.6	NA	39.5	31.7	19.2	29.6	45.2	18.8
3DLOOK	38.5	39.5	NA	48.0	39.4	40.5	62.5	35.6
3Davatar body	34.0	31.7	48.0	NA	30.6	39.3	50.8	29.6
MOVE4D	20.9	19.2	39.4	30.6	NA	26.8	47.6	15.9
eM+	28.1	29.6	40.5	39.3	26.8	NA	49.6	26.0
SS@Home	50.6	45.2	62.5	50.8	47.6	49.6	NA	50.3
PortalMX	16.8	18.8	35.6	29.6	15.9	26.0	50.3	NA

TABLE 18 Pairwise standard deviation of the differences (PSD) in mm for each pair of measuring stations in Phase 1 (hip girth)

	Expert measurer 1	Expert measurer 2	Expert measurer 3	Expert measurer 4	Vitus	SS20
Expert measurer 1	NA	18.4	22.4	23.3	27.2	29.7
Expert measurer 2	18.4	NA	14.9	15.5	21.5	24.6
Expert measurer 3	22.4	14.9	NA	17.6	23.4	26.2
Expert measurer 4	23.3	15.5	17.6	NA	22.6	25.5
Vitus	27.2	21.5	23.4	22.6	NA	12.5
SS20	29.7	24.6	26.2	25.5	12.5	NA

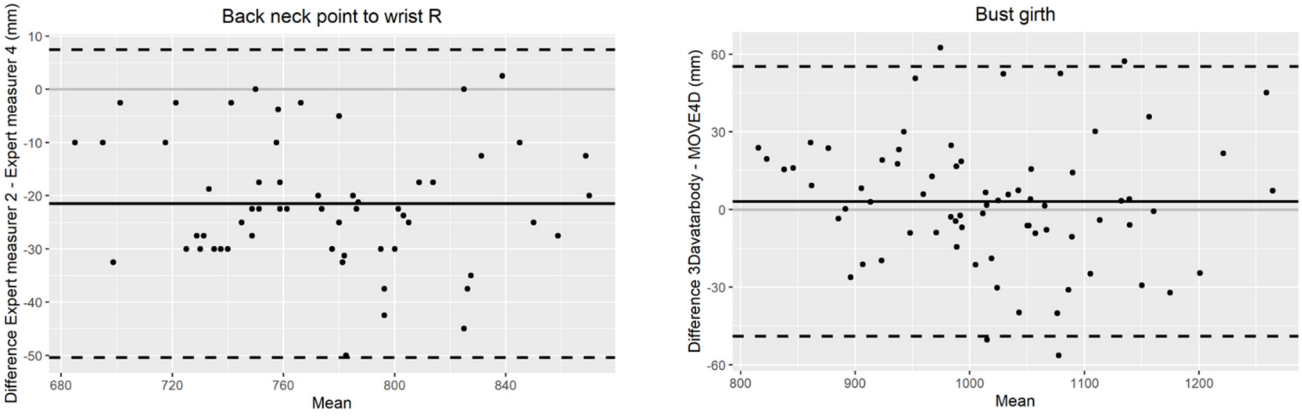
TABLE 19 PSD (mm) for each pair of measuring stations in Phase 2 (hip girth)

	Expert measurer 1	Expert measurer 5	3DLOOK	3Davatar body	MOVE4D	eM+	SS@Home	PortalMX
Expert measurer 1	NA	17.2	29.6	27.7	13.5	23.0	18.6	13.5
Expert measurer 5	17.2	NA	28.5	24.4	14.2	20.6	16.3	16.3
3DLOOK	29.6	28.5	NA	33.5	27.0	29.7	28.8	26.3
3Davatar body	27.7	24.4	33.5	NA	21.6	23.4	23.6	24.2
MOVE4D	13.5	14.2	27.0	21.6	NA	20.1	12.5	12.7
eM+	23.0	20.6	29.7	23.4	20.1	NA	20.8	21.8
SS@Home	18.6	16.3	28.8	23.6	12.5	20.8	NA	17.5
PortalMX	13.5	16.3	26.3	24.2	12.7	21.8	17.5	NA

As shown in TABLE 12 to TABLE 19, there are pairings showing low PSD values (for example, below 10 mm) and pairings with high PSD values (example, above 30 mm). The pairs of technologies with lower values could potentially be made more compatible by applying a simple shift for each measurement using the biases, providing very good reliability as a group. The pairs of technologies with high values will not show good reliability, even if biases are used to shift one technology to another. Further analyses could discover more complex compatibility functions to find n-way compatibility of different groupings of technologies.

FIGURE 12 shows two examples of Bland-Altman plots. One case shows a small bias between stations and the other a high bias between stations. The Bland-Altman plots are an illustrative way to show the agreement between two measuring methods for a pairwise comparison. The horizontal black line represents the bias, and the two dashed lines delimits 95% of the data (1.96 times the PSD). The complete set of Bland-Altman plots for all measurements are available at the IEEE 3DBP website.

FIGURE 12 Examples of the Bland-Altman plots



As previously mentioned, if two stations have both high repeatability and a low PSD, simple pairwise shifts could be used to make different stations compatible [see Equation (10)]. Such a correction would prove essential for applications combining multiple expert measures, expert measurers and digital anthropometric techniques, or multiple digital anthropometric techniques. However, applying these methods to measuring stations not considered in this study may require the gathering of a similar paired dataset.

Within this study, we applied the methods described in ISO 20685-1:2018 for assessing compatibility between manual measurements and digital measurements (thresholds are provided in TABLE 4) to each pair of stations within the same phase. TABLE 20 and TABLE 21 provide the frequency of compatible measurements according to ISO for each pair of stations within the same phase.

TABLE 20 ISO compatibility, pairwise, number of measurements compatible, Phase 1

	Expert measurer 1	Expert measurer 2	Expert measurer 3	Expert measurer 4	Vitus	SS20
Expert measurer 1	NA	0 (0%)	1 (9%)	2 (18%)	0 (0%)	1 (9%)
Expert measurer 2	0 (0%)	NA	2 (18%)	3 (27%)	0 (0%)	0 (0%)
Expert measurer 3	1 (9%)	2 (18%)	NA	1 (9%)	0 (0%)	1 (9%)
Expert measurer 4	2 (18%)	3 (27%)	1 (9%)	NA	0 (0%)	1 (9%)
Vitus	0 (0%)	0 (0%)	0 (0%)	0 (0%)	NA	1 (9%)
SS20	1 (9%)	0 (0%)	1 (9%)	1 (9%)	1 (9%)	NA

TABLE 21 ISO compatibility, pairwise, number of measurements compatible, Phase 2

	3DLOOK	Expert measurer 1	Expert measurer 5	3Davatar body	MOVE4D	eM+	SS@Home	PortalMX
3DLOOK	NA	0 (0%)	2 (18%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Expert measurer 1	0 (0%)	NA	2 (18%)	1 (9%)	3 (27%)	2 (18%)	1 (9%)	3 (27%)
Expert measurer 5	2 (18%)	2 (18%)	NA	1 (9%)	2 (18%)	1 (9%)	1 (9%)	2 (18%)
3Davatar body	0 (0%)	1 (9%)	1 (9%)	NA	3 (27%)	0 (0%)	1 (9%)	1 (9%)
MOVE4D	0 (0%)	3 (27%)	2 (18%)	3 (27%)	NA	0 (0%)	2 (18%)	2 (18%)
eM+	0 (0%)	2 (18%)	1 (9%)	0 (0%)	0 (0%)	NA	0 (0%)	1 (9%)
SS@Home	0 (0%)	1 (9%)	1 (9%)	1 (9%)	2 (18%)	0 (0%)	NA	2 (18%)
PortalMX	0 (0%)	3 (27%)	2 (18%)	1 (9%)	2 (18%)	1 (9%)	2 (18%)	NA

These results show that, generally, there is very little compatibility between the different stations according to ISO thresholds and methods. Only 46 of 473 pairs (<10%) from all the pairings between measuring systems were compatible. Only 11 of 77 pairs of expert measurers were compatible. The most compatible measurement was thigh girth at 37% of pairings, while across back shoulder width and inside leg height showed no compatibility between any pair of stations. These results put into question the suitability of the ISO methods to assess compatibility between manual methods and digital methods; even manual measurements made by experts—the "gold standard" proposed by this method—demonstrate serious difficulties in complying with the ISO standards.

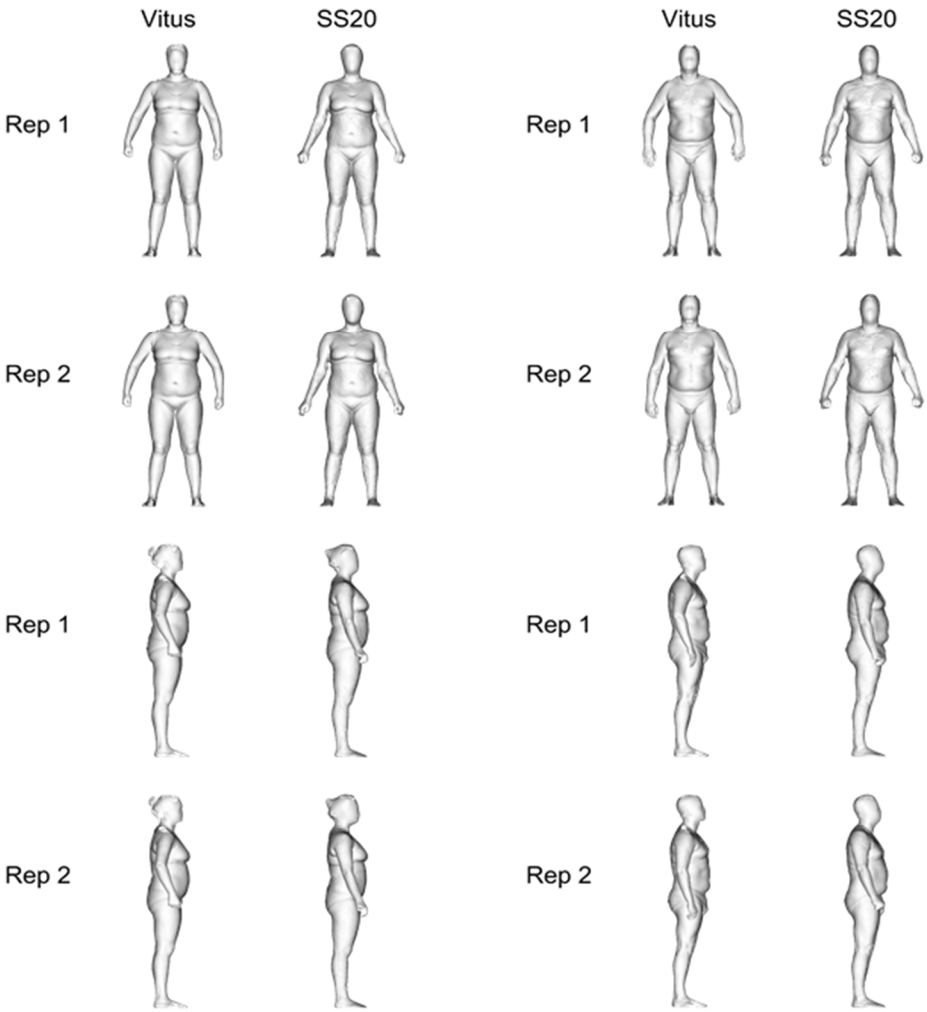
These differences are likely attributable to the variation in interpretation and landmarking used to acquire the anthropometric measurements, highlighting the need for clearer guidance and standardization in anthropometric definitions, including for landmarks.

3.3. QUALITATIVE ASSESSMENT

The framework used to create 3D images from each digital measuring station allowed visual inspection of the 3D images. Visual inspection of the 3D images created by the digital measuring stations demonstrates measuring station performance with 3D images, specifically image resolution, quality, artifacts, and accuracy. FIGURE 13, FIGURE 14, and FIGURE 15 provide examples from Phase 1 and Phase 2 with different body types. The qualitative framework outlined within this study demonstrates a method in which 3D images can be collected from comparable angles for comparison in subsequent work. Since 3D images could be viewed in real time across digital measuring stations, the benefit of consistent methods within an application or study was shown. The 3D images created by

the digital measuring stations in this study demonstrated 3D images of varying qualities, with smart phone apps producing lower fidelity images than most full body scanners. This effect seems to increase with subjects whose body shape deviates markedly from the average (obese, high muscularity, etc.). The complete set of images for all subjects is available at the IEEE 3DBP website.

FIGURE 13 3D images from the front and side aspect of one a) female participant and one b) male participant for both repetitions at the Vitus and SS20 measuring stations



a) b)

FIGURE 14 **3D images from the front and side aspect of one male participant taken from both repetitions of the MOVE4D, PortalMX, SS@HOME, 3D Avatar Body, 3DLOOK, and eM+ systems**

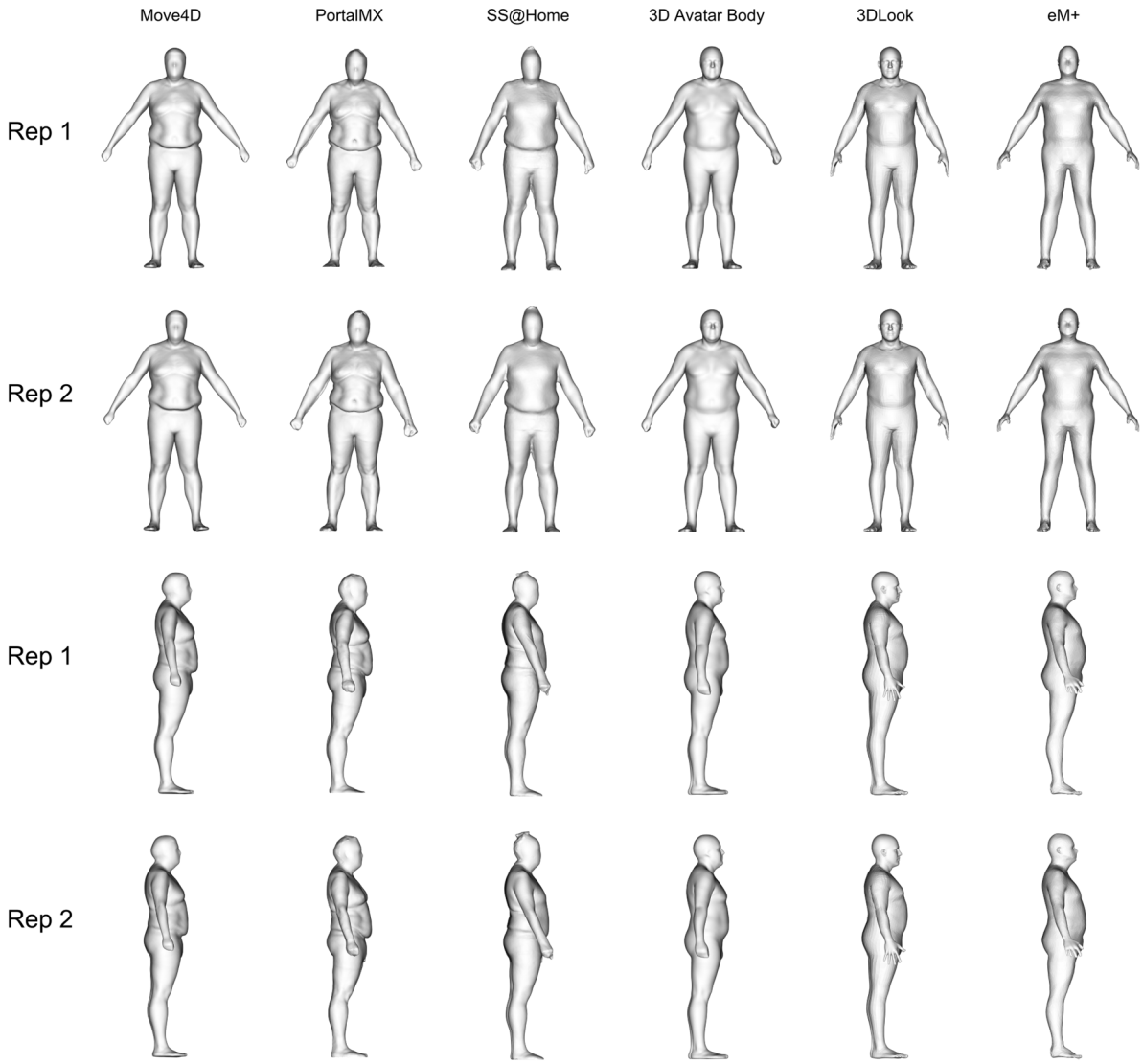
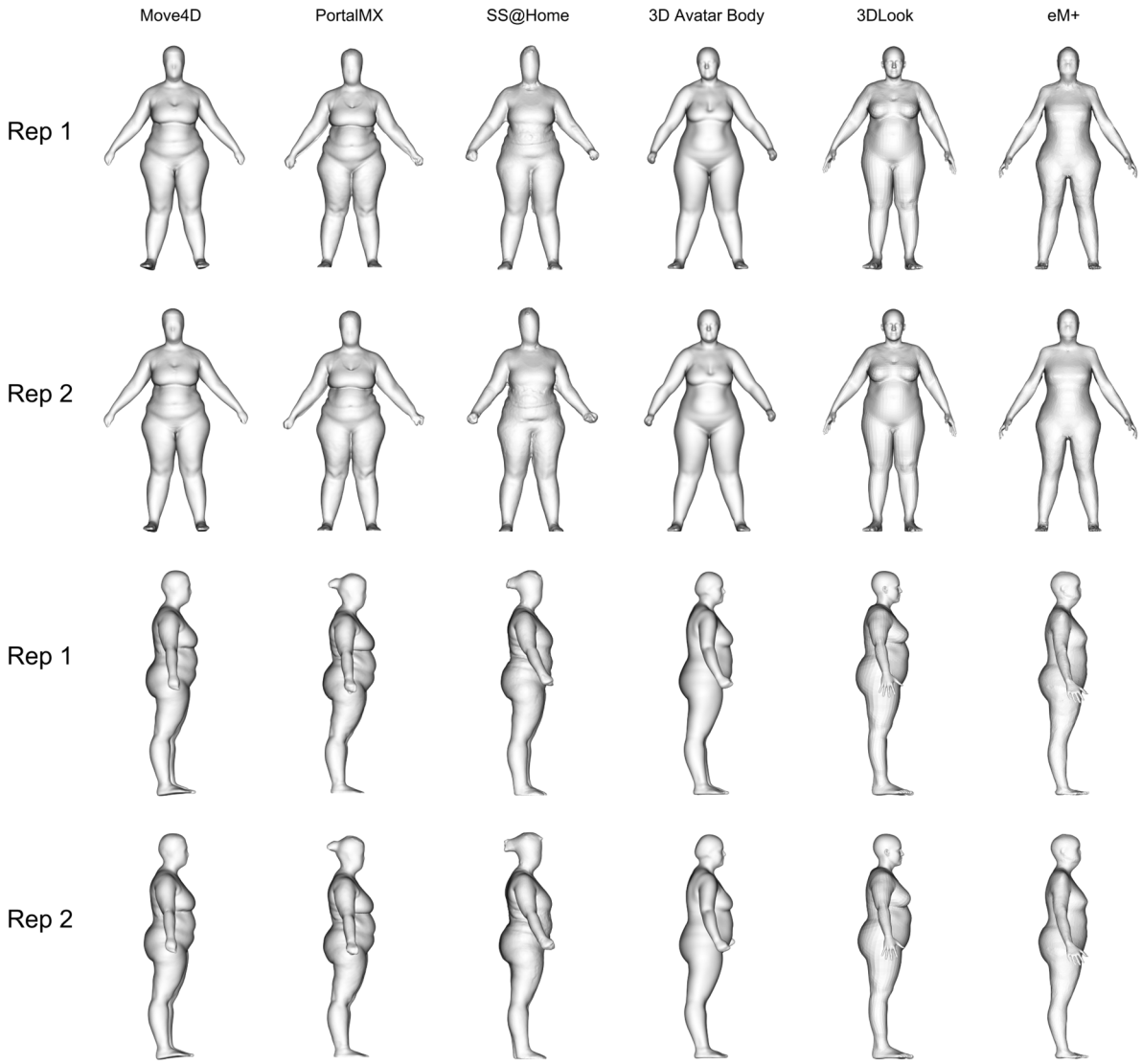


FIGURE 15 **3D images from the front and side aspect of one female participant taken from both repetitions of the MOVE4D, PortalMX, SS@HOME, 3D Avatar Body, 3DLOOK, and eM+ systems**



3.4. LIMITATIONS OF THE STUDY

This study has limitations that require consideration. First, only a limited cross section of anthropometric techniques and vendors was explored in this initial study. Second, while each measuring station was instructed to extract standardized measurements, the methods and digital landmarks used by each measuring station likely varied and were a possible cause of the discrepancies between anthropometrics. Third, this study focused on comparing acquired length and girth, which represent a very narrow subset of anthropometric measurements that can be used

in studies, such as critical shape analysis, for garment fit. In practice, length, breadth, and girth are taken along with volume and area to study morphometrics (shape analysis) to estimate the contribution of fat and muscle mass to body composition. Last, although the results of this study fall within the recommended limits of established industry standards, it is difficult to determine with confidence if the results obtained from various equipment and methods are significant enough to warrant a strong preference for any particular method or technology, given the magnitude of the reported systemic variability.

3.5. ANTHROPOMETRY AND APPAREL

The dichotomy between dimension and shape has long plagued apparel. As reconstruction algorithms evolve toward a better understanding of the distribution of measurements in 3D space, tremendous opportunities to better understand the relationships between measurements and the human form (body shape) will result. How dimensions are linked to body shape will provide useful insight for apparel fit use cases. Consequently, while the avatars resulting from this study would not be suitable for bespoke apparel use case scenarios, they have value for retail situations. Shape variations revealed in this 3D data set are very much representative of real-world garment fit situations where the “fit” of any given size must include a level of tolerance suitable for accommodating a range of morphological variations. The use case implications for such data sets are striking.

With recognition that digitization is inevitable, it is easy to discount the value of manual measuring. However, manual measuring remains critical to define heuristic practice and for understanding where and why the direct translation of body to product will not be linear.

Digitization of measurements may assist with known apparel use case scenarios plagued with garment fit concerns, such as difficulties palpating bony protrusions on heavy-set bodies. Where traditional manual landmarks do not offer the best correlation between body and garment, any additional insight could be considered a step forward. Digital measurements will also allow for a greater quantity of measurements to be taken in a short time, which can be used for the automation of pattern generation.

4. CONCLUSIONS

This study provides a unique dataset containing more than 14000 measurement points and more than 1100 3D body scans gathered by five experts, four full body scanners, and four phone apps. This dataset can be used for future analyses and studies, including segmentation by age, gender, or body mass index (BMI). All the collected

data (measurements and 3D objects) and the tables, plots, and images generated for the analyses can be accessed at the IEEE 3DBP website.

The methods proposed in this study for designing the experiment; collecting, cleaning, and managing the data; and analyzing repeatability and compatibility can constitute a paradigm for future studies, facilitating comparability of data and results. This study also provides benchmarks and reference values of repeatability and compatibility indicators for a number of state-of-the-art technologies in the sector.

The results of the study suggest that all 13 measuring stations showed high repeatability but low compatibility between them. The digital methods outperformed the manual measurers in repeatability. One potential cause for the low compatibility is the ambiguity of ISO body measurement definitions and the ISO method's low suitability for digital implementation. For certain cases, where data is available and there is high repeatability and low paired PSD, simple pairwise shifts can be applied to compensate biases between two stations, thereby rendering them compatible.

The application of ISO methods for assessing pairwise compatibility between stations proved impractical for this type of assessment, and the low compatibility found even between expert measurers highlights the need for alternative methods in real use case scenarios. The calculated reference values and benchmarks for the participating technologies could help in establishing the new criteria.

Regarding the qualitative assessment of the 3D objects, all technologies provided 3D results showing body shapes that could be understood as representative of the individual person; however, the degree of fidelity of the apps was lower than those of the actual 3D scanners. The different apps also provided different degrees of fidelity. Moreover, the technologies providing watertight, artifact-free 3D models, like the phone apps and MOVE4D, have the potential to greatly facilitate the use of 3D content in different digital applications.

Finally, having such a great number of 3D assets but no methods to quantify their quality highlights the need for consistent methods, parameters, and assessments to quantify the quality of 3D representation of human bodies.

5. FUTURE RESEARCH AND DEVELOPMENT

Future work will seek to promote the adoption of improved methods, parameters, and criteria to evaluate the repeatability and compatibility of measuring methods and to introduce superior ways to establish reference values and benchmarks when extracting other anthropometrics such as breadths, volumes, areas, morphometrics, and estimations of body composition and when cross-processing such data. Additional work will focus on exploring the effects of participant demographics (e.g., gender, age, or BMI group), creating methods for the visual inspection and grading of 3D images, implementing more cross-processing of data, and establishing more comprehensive anthropometric definitions.

The IEEE 3D Body Processing Industry Connections Quality subgroup will focus on incorporating the learnings into the work of the IEEE P3141—3D Body Processing Standards group for future body processing standards.

6. REFERENCES

[1] Stahl, M., "IEEE Industry Connections (IEEE-IC) 3D Body Processing (3DBP) Initiative—An Introduction," IEEE-SA Industry Connections White Paper, IEEE Standards Association, 2017. Accessed February 19, 2021, https://standards.ieee.org/wp-content/uploads/import/governance/iccom/3DBP-An_Introduction.pdf.

[2] Kouchi, M., "Anthropometric methods for apparel design: body measurements devices and techniques," in Anthropometry, Apparel Sizing and Design, Amsterdam: Elsevier, 2014, pp.67–94.

[3] ISO 8559-1:2017, Size designation of clothes—Part 1: Anthropometric definitions for body measurement. Accessed February 8, 2021, https://www.iso.org/standard/61686.html.

[4] ISO 7250-1:2017, Basic human body measurements for technological design—Part 1: Body measurements definitions and landmarks. Accessed February 19, 2022, https://www.iso.org/standard/65246.html.

[5] ASTM D5219-15, Standard Terminology Relating to Body Dimensions for Apparel Sizing. Accessed February 19, 2022, https://webstore.ansi.org/standards/astm/astmd521915.

[6] Gordon, C. C., T. Churchill, C. E. Clauser, B. Bradtmiller, J. T. McConville, I. Tebbetts, and R. A. Walker, "1988 Anthropometric Survey of U.S. Army Personnel: Methods and Summary Statistics, Technical Report Natick/TR-89/044," Anthropology Research Project Inc, Yellow Springs, OH, 1989, http://mreed.umtri.umich.edu/mreed/downloads/anthro/ansur/Gordon_1989.pdf.

[7] McDonald, C., A. Ballester, R. K. Rannow, M. Fedyukov, and S. Sokolowski, "Working Group Progress for IEEE P3141—Standard for 3D Body Processing, 2017–2018," in *Proceedings of 3DBODY.TECH 2018–9th International Conference and Exhibition on 3D Body Scanning and Processing Technologies*, Lugano, Switzerland, Oct. 2018.

[8] Sicotte, M., M. Ledoux, M.-V. Zunzunegui, A. A. Souleymane, and V.-K. Nguyen, "Reliability of anthropometric measures in a longitudinal cohort of patients initiating ART in West Africa," *BMC Medical Research Methodology*, vol. 10, no. 1, pp. 102–111, Oct. 2010.

[9] Chumlea, W. C., A. F. Roche, and E. Rogers, "Replicability for anthropometry in the elderly," *Human Biology,* vol. 56, no. 2, pp. 329–337, 1984.

[10] Lohman, T. G., A. F. Roche, and R. Martorell, eds., *Anthropometric Standardization Reference Manual,* vol. 177, Champaign, IL: Human Kinetics Books, 1988.

[11] Kouchi, M., M. Mochimaru, K. Tsuzuki, and T. Yokoi, "Random errors in anthropometry," Journal of Human Ergology, vol. 25, no. 2, pp. 155–166, 1996.

[12] Nada, J., Z. Putz, G. Kolev, S. Nagy, and G. Jermendy, "Intraobserver and interobserver variability of measuring waist circumference," *Medical Science Monitor,* vol. 14, no. 1, pp. CR15–CR18, Jan. 2008.

[13] Kouchi, M., and M. Mochimaru, "Evaluation of Accuracy in Traditional and 3D Anthropometry," in *Digital Human Modeling for Design and Engineering Symposium,* Pittsburgh, PA, Jun. 2008.

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

[15] Verweij, L., C. Terwee, K. Proper, C. Hulshof, and W. Van Mechelen, "Measurement error of waist circumference: gaps in knowledge," *Public Health Nutrition,* vol. 16, no. 2, pp. 281–88, February 2013.

[16] Kuehnapfel, A., P. Ahnert, M. Loeffler, A. Broda, and M. Scholz, "Reliability of 3D laser-based anthropometry and comparison with classical anthropometry," *Scientific Reports,* vol. 6, p. 26672, May 2016.

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

[18] Gordon, C. C., and B. Bradtmiller, "Interobserver error in a large scale anthropometric survey," *American Journal of Human Biology,* vol.4, no. 2, pp. 253–263, 1992.

[19] Yoon, J. C., R. G. Radwin, "The accuracy of consumer-made body measurements for women's mail-order clothing," *Human Factors.* vol. 36, no. 3, pp. 557–568, 1994.

[20] "3D Scanning Market," Allied Market Research, 2018, http://www.alliedmarketresearch.com/3D-scanning-market.

[21] Daanen, H. A. M., and G. J. van de Water, "Whole body scanners," *Displays,* vol.19, no. 3, pp. 111–120, Nov. 1998.

[22] Daanen, H. A. M., and F. B. Ter Haar, "3D whole body scanner revisited," *Displays,* vol. 34, no. 3, pp. 270–275, Oct. 2013.

[23] Arbutina, M., D. Dragan, S. Mihic, and Z. Anisic, "Review of 3D body scanning systems," *Acta Technica Corviniensis-Bulletin of Engineering,* vol. 10, no. 1, p. 17, 2017.

[24] ISO 20685-2:2015, Ergonomics—3-D scanning methodologies for internationally compatible anthropometric databases—Part 2: Evaluation protocol of surface shape and repeatability of relative landmark positions, https://www.iso.org/standard/63261.html.

[25] Kouchi, M., M. Mochimaru, B. Bradtmiller, H. Daanen, P. Li, B. Nacher, and Y. Nam, "A protocol for evaluating the accuracy of 3D body scanners," *Work*, vol. 41, S1, pp. 4010–4017, 2012.

[26] Robinson, A., M. McCarthy, S. Brown, and A. Evenden, "Improving the quality of measurements through the implementation of customised reference artefacts," in *Proceedings of the 3rd International Conference on 3D Body Scanning Technologies,* Lugano, Switzerland, Oct. 2012.

[27] Wright, W., "Man vs machine—Measuring people for the apparel industry," in *Proceedings of 3DBODY.TECH 2019 10th International Conference and Exhibition on 3D Body Scanning and Processing Technologies,* Lugano, Switzerland, Oct. 2019.

[28] Rumbo-Rodríguez, L., et al. "Comparison of Body Scanner and Manual Anthropometric Measurements of Body Shape: A Systematic Review." *International journal of environmental research and public health* 18.12 (2021): 6213.

[29] Ballester, A., A. Piérola, E. Parrilla, J. Uriel, A. V. Ruescas, C. Perez, J. V. Durá, and S. Alemany, "3D human models from 1D, 2D, & 3D inputs: reliability and compatibility of body measurements," in *Proceedings of 3DBODY.TECH 2018 9th International Conference and Exhibition on 3D Body Scanning and Processing Technologies,* Lugano, Switzerland, Oct. 2018.

[30] Dekker, L., "3D human body modelling from range data," Doctoral thesis, University of London, 2000.

[31] Robinette, K. M., and H. A. M. Daanen, "Precision of the CAESAR scan-extracted measurements," *Applied Ergonomics,* vol. 37, no. 3, pp. 259–265, May 2006.

[32] Wang, J., D. Gallagher, J. C. Thornton, W. Yu, M. Horlick, and F. X. Pi-Sunyer, "Validation of a 3-dimensional photoic scanner for the measurement of body volumes, dimensions, and percentage body fat," *American Journal of Clinical Nutrition,* vol. 83, no. 4, pp. 809–816, Apr. 2006.

[33] Lu, J.-M. and M.-J. J. Wang, "The Evaluation of Scan-Derived Anthropometric Measurements," *IEEE Transactions on Instrumentation and Measurement,* vol. 59, no. 8, pp. 2048–2054, Aug. 2010.

[34] Pepper, M.R., J.H. Freeland-Graves, W. Yu, P.R. Stanforth, J.M. Cahill, M. Mahometa, B. Xu, "Validation of a 3-dimensional laser body scanner for assessment of waist and hip circumference," J. Am. Coll. Nutr. 2010, 29, 179–188.

[35] Vonk, T. D. H. and Daanen, H. A. M., "Validity and Repeatability of the Sizestream 3D Scanner and Poikos Modeling System," in *Proceedings of the 6th International Conference on 3D Body Scanning Technologies,* Lugano, Switzerland, Oct. 2015.

[36] Wells, J. C. K., J. Stocks, R. Bonner, E. Raywood, S. Legg, S. Lee, P. Treleaven, and S. Lum, "Acceptablility, precision and accuracy of 3D photonic scanning for measurement of body shape in a multi-ethnic sample of children aged 5-11 years: the SLIC study," *PloS One*, vol. 10, no. 4, 2015.

[37] Ng, B. K., B. J. Hinton, B. Fan, A. M. Kanaya, and J. A. Shepherd, "Clinical anthropometrics and body composition from 3D whole-body surface scans," *European Journal of Clinical Nutrition,* vol. 70, no. 11, pp. 1265–1270, Nov. 2016.

[38] Bullas, Alice, Simon Choppin, Ben Heller, Sean Clarkson, and Jonathan Wheat (2014). Kinanthropometry Application of Depth Camera Based 3D Scanning Systems in Cycling: Repeatability and Agreement with Manual Methods. In: *5th International Conference on 3D Body Scanning Technologies, Lugano, Switzerland, 21-22 October 2014.* Hometrica Consulting, 290–298.

[39] Markiewicz, U., M. Witkowski, R. Sitnik, and E. Mielicka, "3D anthropometric algorithms for the estimation of measurements required for specialized garment design," *Expert Systems with Application,* vol. 85, pp. 366–385, Nov. 2017.

[40] Gleadall-Siddall, D. O., R. L. Turpin, C. C. Douglas, L. Ingle, and A. T. Garrett, "Test-retest repeatability of the NX-16: a three-dimensional (3D) body scanner in a male cohort," *Sport Sciences for Health,* pp. 1–10, 2019.

[41] Parrilla, E., A. Ballester, C. Solves-Camallonga, B. Nácher, S. A. Puigcerver, J. Uriel, A. Piérola, J. C. González, and S. Alemany, "Low-cost 3D foot scanner using a mobile app," *Footwear Science,* vol. 7, S1, pp. S26–S28, 2015.

[42] "Size Stream At Home—Mobile 3D Body Scanning," [Online]. Accessed February 18, 2022, https://www.kickstarter.com/projects/mobile-3d-scanner/size-stream-at-home-mobile-3d-body-scanning.

[43] Seo, H., Y. I. Young, and K. Wohn, "3D Body Reconstruction from Photos Based on Range Scan," in *Proceedings of Technologies for E-Learning and Digital Entertainment, First International Conference, Edutainment 2006,* Hangzhou, China, Apr. 2006.

[44] Boisvert, J., C. Shu, S. Wuhrer, and P. Xi, "Three-dimensional human shape inference from silhouettes: reconstruction and validation," *Machine Vision and Applications,* vol. 24, no. 1, pp. 145–157, Jan. 2013.

[45] Zhu, S., P. Y. Mok, and Y. L. Kwok, "An efficient human model customization method based on orthogonal-view monocular photos," *Computer-Aided Design,* vol. 45, no. 11, pp. 1314–1332, Nov. 2013.

[46] Ballester, A., E. Parrilla, J. A. Vivas, A. Pierola, J. Uriel, S. A. Puigcerver, P. Piqueras, C. Solves-Camallonga, M. Rodriguez, J. C. Gonzalez, and S. Alemany, "Low-cost data-driven 3D reconstruction and applications," in *Proceedings of the 6th International Conference on 3D Body Scanning Technologies,* Lugano, Switzerland, 2015.

[47] Ballester, A., "Data-driven three-dimensional reconstruction of human bodies using a mobile phone app," *International Journal of the Digital Human,* vol. 1, no. 4, pp. 361–388, 2016.

[48] Allen, B., B. Curless, Z. Popović, "The Space of Human Body Shapes: Reconstruction and Parameterization from Range Scans," *ACM SIGGRAPH 2003 papers,* pp. 587–594, 2003.

[49] Alemany, S., J. Uriel, A. Ballester, and E. Parrilla, "Three-dimensional body shape modeling and posturography," in *DHM and Posturography,* Academic Press, 2019, pp. 441–457.

[50] Seminati, E., D. C. Talamas, M. Young, M. Twiste, V. Dhokia, J. L. J. Bilzon, "Validity and reliablity of a novel 3D scanner for assessment of the shape and volume of amputees' residual limb models," *PloS One,* vol. 12, no. 9, 2017.

[51] Heymsfield, S. B., B. Bourgeois, B. K. Ng, M. J. Sommer, X. Li, and J. A. Shepherd, "Digital anthropometry: a critical review," *European Journal of Clinical Nutrition,* vol. 72, no. 5, pp. 680–687, 2018.

[52] Wheat, J. S., S. Clarkson, S. W. Flint, C. Simpson, and D. R. Broom, "The use of consumer depth cameras for 3D surface imaging of people with obesity: a feasibility study," *Obesity Research & Clinical Practice,* vol. 12, no. 6, pp. 528–533.

[53] Gill, S., "Human measurement and product development for high-performance apparel," in *High Performance Apparel,* Woodhead Publishing, 2018, pp. 191–208.

[54] Ashdown, S., "Full body 3-D scanner," in *Anthropometry, Apparel Sizing and Design,* Woodhead Publishing, 2020, pp. 145–168.

[55] McDonald, C., L. Oviedo, and A. Ballester, "Working Group Progress for IEEE P3141—Standard for 3D Body Processing, 2019–2020," in *Proceedings of 3DBODY.TECH 2017—8th International Conference and Exhibition on 3D Body Scanning and Processing Technologies,* Montreal, Canada, 2017.

[56] McDonald, C., Y. Wu, A. Ballester, and M. Stahl, "IEEE Industry Connections (IEEE-IC) Landmarks and Measurement Standards Comparison in 3D Body-model Processing," in *IEEE Industry Connections (IEEE-IC)*

Landmarks and Measurement Standards Comparison in 3D Body-model Processing, pp.1–34, Feb. 2018. 2018. Accessed February 2022, https://ieeexplore.ieee.org/document/8362828.

[57] ISO 13528:2015, Statistical methods for use in proficiency testing by interlaboratory comparison. Accessed February 19, 2022, https://www.iso.org/standard/56125.html.

[58] ISO 3534-1:2006, Statistics—Vocabulary and symbols—Part 1: General statistical terms and terms used in probability. Accessed February 19, 2022, https://www.iso.org/standard/40145.html.

[59] Walter, S. D., M. Eliasziw, and A. Donner, "Sample size and optimal designs for reliability studies," *Statistics in Medicine,* vol. 17, no. 1, pp. 101–110, 1998.

[60] Eliasziw, M., S. L. Young, M. G. Woodbury, and K. Fryday-Field, "Statistical Methodology for the Concurrent Assessment of Interrater and Intrarater Reliability: Using Goniometric Measurements as an Example," *Physical Therapy,* vol. 74, no. 8, pp. 777–788, August 1994.

[61] Yitzhaki, S., "Gini's mean difference: a superior measure of variability for non-normal distributions," *Metron—International Journal of Statistics,* vol. 61, no. 2, pp. 285–316, 2003.

[62] Bland, J. M., and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *The Lancet,* vol. 327, no. 8476, pp. 307–310, 1986.

[63] ISO 5725-2: 2019, Accuracy (trueness and precision of measurement methods and result—Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method. https://www.iso.org/obp/ui/#iso:std:iso:5725:-2:ed-2:v1:en.

[64] ISO 20685-1:2018, 3-D scanning methodologies for internationally compatible anthropometric databases—Part 1: Evaluation protocol for body dimensions extracted from 3-D body scans. https://www.iso.org/standard/63260.html.

[65] ISO 15535:2012 (en), General requirements for establishing anthropometric databases. https://www.iso.org/obp/ui#iso:std:iso:15535:ed-3:v1:en.

RAISING THE WORLD'S STANDARDS

3 Park Avenue, New York, NY 10016-5997 USA http://standards.ieee.org

Tel.+1732-981-0060 Fax+1732-562-1571