Protocols for High-Quality Indoor and Outdoor Scanning of Clothed People

Kristijan BARTOL, Stefan GUMHOLD
Faculty of Computer Science, Technical University of Dresden, Dresden, Germany;

https://doi.org/10.15221/23.47

Abstract

With the advancements in high-quality scanning technology, which is becoming more accurate, convenient, and available to smaller labs, the recommendations and protocols should be discussed and established. The protocols are important to make sure that the required list of steps is carried out for the desired set of goals and to minimize the overall time required to carry out the preparation, scanning, and postprocessing steps. In this paper, we propose scanning protocols for clothed people in indoor and outdoor settings. The indoor settings should be more suitable for high-quality 3D scans, which should serve as a reference to the ground-truth human body. The outdoor setting should be more suitable for providing challenging scenarios, closer to the real world. We discuss the postprocessing steps required to align indoor and outdoor datasets for the best quality of ground-truth information. Finally, we overview existing use cases and applications using scanning datasets and recommend the corresponding protocols.

Keywords: 3d scanning, 4d scanning, protocols, recommendations, indoor, outdoor, clothed people, applications

1. Introduction

Since the early stationary 3D scanners in 1990s based on laser scanning, the scanning technology has significantly advanced in terms of speed, accuracy, price, and size [1]. Studies have even shown that the scanning technology of the human body has become comparable in accuracy to the expert measurers [2, 3, 4], which are still the golden standard [5]. These remarkable technological improvements and the availability of the high-quality scanners to smaller laboratories have resulted in several, publicly available, free, and high-quality 3D and 4D datasets of human bodies in tight or minimal clothing [6, 7, 8, 9, 10, 11] and in everyday clothing [12, 13, 14, 15, 16, 17, 18]. The publicly available datasets are particularly useful for the research community, fostering advancements in anthropometry, textile research, computer vision, fitness, and medicine [1].

Nevertheless, the acquisition of high-quality human scanning datasets is still time-consuming, complicated, and potentially costly due to mandatory compliance with the data regulations, the need to hire actors which can consistently repeat certain sequences of motions and engineers which can postprocess the data, while also managing the entire process over extended periods of time [18]. To simplify scanning protocols, the standards for human body measurement have already been established [5, 19]. The standards recommend which body measurements should be obtained and how, which body poses the person should take, and how the subject should be prepared to obtain the best quality of the scans. There are several surveys, reviews, and books that overview the scanning protocols and recommend the best practices for particular applications [1, 20, 21, 22, 23]. Specific to 3D and 4D scanning of people in loose clothing, there are a few successful methods which also discuss the protocols but also the limitations of the technology and the postprocessing algorithms [12, 13, 17, 18].

In this work, we overview the scanning protocols which include the preparation stage, scanning, and postprocessing, and propose specific protocols for applications, to potentially reduce the effort and overall costs of the process. Furthermore, we distinguish between indoor and outdoor scanning environments. Indoor environments are more suitable for high-quality acquisition because the environment and the conditions can be controlled, which allows a higher level of detail and accuracy. Outdoor environments are more suitable for capturing challenging, real-world scenarios. The disadvantage is that it is generally more difficult to obtain high-quality results outdoors due to different lighting conditions and inconvenience to set up high-quality scanning equipment. In the remainder of the document, we overview previous work in terms of scanning technology and scanners, existing protocols, and datasets in Section 2. In Section 3, we overview the standard scanning protocol with the emphasis on the possibilities of combining indoor and outdoor datasets. In Section 4, we propose scanning protocols for applications such as 3D human body pose and shape estimation, garment modeling and estimation, fitness, medicine, sizing, surveying, and metaverse. We argue that each application needs more specific protocols to reduce the required resources as much as possible, in terms of time, data volume, money, and effort. In Section 5, we further discuss the importance of planning, point out current limitations of the scanning protocols, and offer potential solutions. We conclude the discussions in Section 6.

2. Previous Work

2.1. Scanning Technology and Scanners

The commercial scanners can be classified based on the technology used and mobility. In terms of technology, we distinguish passive stereo (photogrammetry), active stereo (structure light), and time-of-flight scanners. Passive stereo scanners consist of multiple, synchronized RGB cameras, positioned around the subject. The general issue with passive stereo is that they struggle with surfaces without or with repeating texture [35]. Active stereo scanners compensate for the potential lack of texture by using one or more light-pattern projectors in addition to the RGB cameras [36]. Time-of-flight scanners emit light and measure the distance between the points and the camera by measuring light travel time. They are cheaper and consume less energy, but the biggest disadvantage is their lower resolution. The most common technology used for fixed-pose, 3D scanning is structured light, while the most common technology for dynamic, 4D scanning is photogrammetry, although there are scanners which combine multiple scanning technologies [1].

2.2. Commercial Scanners

In terms of mobility, we distinguish stationary, hand-held, and mobile scanners. Stationary scanners are physically large systems consisting of multiple cameras and are intended to be fixed in a particular location for longer periods during the acquisition process. Hand-held scanners such as are used to scan the person by walking around. It requires that the person stays still for a few minutes. Mobile scanners are smaller and more convenient to set up. The mobile scanning systems usually consist of a set of RGB cameras or mobile devices such as smartphones, with or without the embedded time-of-flight cameras. The full list of commercial scanners from 2021 is available in a review paper [1].

2.3. Datasets

The list of public datasets, freely available for researchers, is provided in Table 1. The datasets are characterized by which ground-truth references they provide (pose, shape, and clothing deformations) and which protocol they use (static or dynamic, indoor-only or both indoor and outdoor). Although we highlight 15 datasets, none of these datasets contain all the characteristics, meaning they have all the ground-truth references for the body and clothes geometry, as well as clothing body and clothing dynamics, while also having corresponding outdoor scans or recordings. Therefore, we believe it is particularly relevant to discuss the corresponding protocols to help future data acquisition.

Table 1. A list of existing, public datasets.

Manuscript style component	Format				
	Pose	Shape	Clothes	Dynamic	Outdoor
HumanEva [24]	Yes	No	No	Yes	No
Total Capture [10]	Yes	No	No	Yes	No
AMASS [9]	Yes	No	No	Yes	No
Panoptic Studio [16]	Yes	No	No	Yes	No
Human3.6M [15]	Yescenter	Yes*	Yes*	Yes	No
Human4D [11]	Yes	Yes	No	Yes*	No
Dyna [8]	Yes	Yes	No	Yes	No
FAUST [6]	Yes	Yes	No	No	No
DFAUST [7]	No	No			
HUMBI [17]	Yes	Yes	No	Yes	No
3DPW [14]	Yes	Yes	Yes	No	Yes
EMDB [25]	Yes	Yes	No	Yes	Yes
ClothCap [12]	Yes	Yes	Yes	Yes	No
BUFF [13]	Yes	Yes	Yes	Yes	No
4DHumanOutfit [18]	Yes	Yes	Yes	Yes	No

3. Method

3.1. Planning

Depending on the size of the project and the application, the planning starts a few weeks or a few months before any acquisition (scanning). In research groups, it starts even a few years before, when the project proposal is submitted. Therefore, planning is a fundamental step to align with expectations as best as possible. The planning requires defining clear goals and motivation to carry out the project, specifying the list of scanning equipment which will be used or ordered, assigning the roles to the core team members, and defining a timeline with a sequence of steps to carry out the data(set) acquisition protocol. Depending on the application, it is recommended to hire actors for the acquisition (scanning) stage [18]. The actors can consistently repeat defined sequences of actions without significant oscillations. Carrying out the motion sequences consistently in and between the actors is particularly relevant when the sequences must be repeated in different clothing and when the same sequences are planned to be repeated both indoors and outdoors.

3.2. Preparation

The preparation stage is crucial to assure high quality of human scans. For 3D body scanning, the person usually takes one of the standard poses (T-pose, A-pose, I-pose, or sitting positions) [5]. Using the standard poses simplifies the postprocessing stage, body measurement extraction algorithms, and separating body from clothing [18]. For 4D scanning, the actor should be precisely instructed how to carry out the motion sequence, which includes the initial body pose, the speed and the sequence of movements (which affects the overall sequence duration), and the final pose. For certain applications, such as gait recognition [26], the actors might be instructed to move in a unique way with respect to other actors, while doing the same action.

3.3. (Indoor) Scanning

Indoor scanning of humans is a data acquisition process that is usually carried out in a controlled (laboratory) environment using high-quality, stationary or handheld scanners. The aim of indoor scanning is to obtain highly accurate data, including body pose, shape, and clothing wrinkle details. Although the environment can be controlled in terms of lightning, background, and camera locations and the technology keeps improving, obtaining high-quality 3D and 4D scans in loose clothing is still challenging.

3.4. Postprocessing

Depending on the scanner, the postprocessing stage can differ. Common high-level steps include cleaning up the point clouds, merging partial point clouds from multiple views into a single point cloud, creating meshes from point clouds, and applying the texture [22, 27]. Many high-quality scanners already include proprietary software which does the postprocessing steps automatically, so the end-user already gets clean meshes [28].

3.5. Outdoor Scanning

The equipment available outdoors is not usually suitable for accurate 3D scanning due to the varying lighting conditions and the fact that the equipment is expected not to be in a fixed position for longer periods of time. The advantage of outdoor recording is that people can engage in more complex activities and can more easily include multiple participants at the same time. Most commonly, outdoor scanning equipment consists of one or more RGB cameras, sometimes coupled with lightweight motion capture equipment such as IMU sensors mounted on person's joints [14]. Due to the lightweight equipment, outdoor scanning produces much sparser data, such as 3D joint locations. However, outdoor scanning data could be paired with higher-quality indoor scanning data, assuming the same people are used in both environments, although this hasn't been done publicly in the research community yet. The final dataset(s) should contain challenging, in-the-wild scenarios with various lighting conditions and real-world occlusions, while having reference body shape and clothing geometry, obtained in a more controlled environment with higher quality equipment.

4. Applications and Protocols

As applications require different amounts and types of data, we overview scanning protocols and provide corresponding recommendations to avoid potentially expensive acquisition of redundant or unnecessary data. By providing high-quality, public, and free scanning datasets, researchers can significantly benefit from training and evaluation of supervised deep learning models. We identify several applications for human scanning datasets and recommend protocols to assure high-quality acquisition with minimal costs.

4.1. Body Pose Estimation

Human body pose estimation [37, 38, 39, 40] is a particularly active research area and is important for several industrial applications such as fitness, sports, medicine, entertainment, etc. [1]. To obtain the corresponding 3D pose ground-truth information during scanning, previous works have either used motion capture systems such as IMU sensors [14, 25], suits [10, 11, 24], or body markers [15] or marker-less motion capture systems that rely on extracting accurate joint locations from multiple views [12, 13, 16, 17, 18]. There are many publicly available datasets, both indoor and outdoor (Section 2.3), that contain images and videos of people, as well as corresponding ground-truth 3D human pose information.

4.2. Body Shape Estimation

To obtain the corresponding 3D shape ground-truth during scanning, it is necessary to use high-quality scanners in indoor setup and is highly recommended to scan people in one of the standard poses and in minimal clothing. Additionally, it is beneficial to also take body measurements by the expert measurer to verify the quality of 3D scans by comparing the extracted and measured values. Although there are several, high-quality, indoor datasets with accurate ground-truth body shapes, there is s still only one public dataset, published in 2018, which contains 3D poses and shapes ground-truth information with the corresponding in-the-wild videos of the same subjects (see Section 2.3 and Table 1 for reference).

4.3. Clothing Modeling and Estimation

Clothing modeling and estimation is an emerging field in computer vision and graphics [41, 42, 43] and is particularly relevant for the fashion industry. Obtaining accurate clothing deformations such as wrinkles and folds for different clothing geometries and loose clothing is still challenging using the current scanning technology [18, 44]. Not surprisingly, there are not many publicly available scanning datasets of clothed people. Therefore, it is particularly relevant to discuss and recommend protocols to assure high-quality data acquisition with minimal effort for future projects.

Based on a recently published dataset [10], it has been shown that the photogrammetry-based system can be a viable choice for obtaining a diverse dataset of people in loose clothing. However, the system used [45] consists of 68 properly positioned RGB cameras in a relatively large room, dedicated only to high-quality 4D scanning. The system itself is a result of a significant effort and expertise of the corresponding team and the postprocessing algorithms are the state-of-the-art for human body reconstruction from multiple views [46]. Other commercial solutions, some of which are listed in Table 1, might also be a viable choice, although there is no direct evidence in the form of publicly available data. Finally, the latest scanning technology should work well for tight clothing. We discuss the potential possibilities for outdoor scanning in Section 5.3.

4.4. Fitness and Medicine

High-quality 3D and 4D scanning of people in clothing is relevant for modeling and manufacturing custom sportswear such as special suits and shoes [20]. Also, accurate human body shape scanning can help to semi-automate personal training and nutritional recommendations [34]. Otherwise, motion capture systems are already being developed and deployed for tracking 3D poses in sports [47] and analyze motion [7, 12, 13]. The protocols for the two latter use cases, accurate human pose and shape acquisition, have already been overviewed in Subsections 4.1 and 4.2. Regarding the scanning for modeling, manufacturing, and testing custom sportswear, it is worth noting that special conditions might be required such as large-enough scanning area to perform the motion sequence, controlled room temperature, and additional sensors for more advanced data collection [10, 18, 24].

Scanning is used in medicine mostly for orthopedic diagnosis such as scoliosis detection [32] but it is also increasingly relevant for modeling skeletal structures and muscle tissue [33].

4.5. Sizing and Surveying

Sizing and surveying projects are traditionally done by manually measuring a selected sample of the population. The set of recommended measurements is defined by the standard [5]. Surveys are important for appropriate cloth design, avoiding high return rates from online shopping, gas emission reduction from overproduction, and tracking the distribution and changes between and within the age groups, ethnicities, and countries. Although manual measurement is relatively convenient and scalable, it still requires significant expertise in measurement repetition to guarantee the best quality and comply with the standards [2, 3, 4]. The latest scanning technology offers a (semi-) automatic alternative to manual measurement. Moreover, high-quality, clothed 3D or 4D scans contain much more information than the 1-dimensional body measurements.

For sizing and surveying, we recommend protocols which will assure at least the level of quality provided by the manual measurements. This can, of course, be achieved by obtaining the measurements on top of the scans but could also be achieved directly, by careful 3D scanning in one of the standard poses, in minimal or in tight clothing [3]. The movements (4D scanning), loose clothing, and outdoor scanning are optional.

4.6. Metaverse

Metaverse is a recent technical area that involves virtual-reality (VR) environments in which users interact with the environment or with each other [29]. One of the important components, related to humans and 3D data, is accurate human pose estimation [30] as it makes a virtual experience more immersive and consistent. Based on the current focus on metaverse technologies, we believe that not only pose estimation, but also accurate shape and clothing will become a requirement for future systems. That said, scanning technology and protocols should be relevant to discuss.

Considering that VR environments should run in real-time, and it is expected that they contain more virtual objects than just human bodies and clothes, it is crucially important to make the data as simple for processing as possible. Standard models like SMPL [31] are convenient to efficiently model and control the population of human bodies in different poses. By fitting SMPL to 3D or 4D scans of people in minimal clothing, body shapes could be transferred to the virtual environment. Efficient models for clothing are still under research. We discuss the possibilities of scanning people in clothing to transfer to virtual environments in Section 5.3.

5. Discussion

5.1. The Importance of Reducing the Overall Required Resources

As reported in a recent report presenting a novel dataset of clothed people [18], the overall time required to obtain and process the dataset was more than 5000 hours over more than six months. The total volume of persisted data (the actual dataset in a compressed format) is 22TB, while the non-persisted data (the volume which was required during the acquisition) is 408TB. The dataset consists of 20 actors in 7 outfits, performing 11 actions per outfit. Proper and detailed planning, high-quality equipment, team expertise, state-of-the-art reconstruction algorithms, and good management are some of the prerequisites for a successful dataset acquisition project. Otherwise, the time frame easily extends, the total number of working hours can easily get stuck on the postprocessing step if the data was not recorded flawlessly, and it is important to be aware that the total volume of data storage is significantly larger during the acquisition compared to the final, compressed version of the dataset.

Note that the above-described dataset still consists of a relatively small number of participants, insufficient to represent a significant sample of human body population, even within groups. Also, the scanning was done only in a controlled, indoor environment. To carry out the scanning protocol further, the same actors would have to be recorded again in the outside environment. To gain as much as possible from the high quality and details of the indoor dataset, it is recommended that the actors wear the same set of clothes and repeat the same set of actions, or some of the new ones. Obviously, 3D and 4D scanning of real people has certain limitations, which we further discuss in the remainder of the section.

5.2. The Limitations of the Scanning Protocols

We identify several limitations of the scanning protocols. The list of limitations is provided below:

1. Scaling limitation. As described in Subsection 5.1, the scanning process is challenging to scale to a significant number of people, especially in cases when all the participants have to follow the same protocol, such as repeating the same sequence of motion and wear a specified set of garments. Each recording session is time-consuming, and it only becomes more challenging if the total time span for the dataset acquisition becomes longer. Not to mention that repeating the same sequences with the same set of clothing outdoors is even more difficult to organize.
2. Difficulty of perfectly replicating the motion. Even though it is recommended to hire actors for recording consistent motion sequences, it is still challenging to repeat the sequences of motion to make them consistent within and between the actors. Again, it is unreasonable to expect that the potential outdoor scanning sessions, with a possible different terrain and conditions will be perfectly repeated.
3. The impossibility of controlling clothing dynamics. The limitation of consistent 4D scanning of clothed people comes from the fact that garments are too complex to be controlled during the motion sequence. This limitation makes it more difficult to identify and track 3D points on the clothes to analyze the material behavior or clothing dynamics.
4. A protocol must remain consistent over a longer period. Given that the scanning project lasts at least a few months, it means that the protocol must not change after the acquisition has started. Otherwise, the collected that will not be consistent between the subjects which would make the final dataset less useful for applications such as body pose, shape, and clothing deformation analysis.

5.3. Potential Solutions and Alternatives

We discuss some of the original potential solutions to some of the limitation concerns raised in the previous subsection. We discuss solutions to limitations 2 and 3 (replicating the motion and controlling the clothing dynamics). In addition, we also discuss the possibilities of outdoor scanning in the context of the identified limitations and the role of synthetic data in the context of scanning, postprocessing, and applications. Note that the limitations 1 and 4 are more inherent and it is not currently clear how to non-trivially mitigate scaling issues and potentially inconsistent protocols.

To mitigate the inconsistencies in replicating motion within and between the actors, we propose to first fit SMPL or an SMPL-family [48, 49, 50] of models to the motion sequences. Now that all the sequences are expressed in a common space, the statistics of the “motion offsets” can be analyzed. Based on the "offset distribution", mean and median motions can be extracted as a reference. These reference motions, along with the actual fits, can be provided as ground-truth. Given the advancements in image editing of human poses [53] and the overall rate at which the computer vision technology advances, it might be possible to also edit the original video sequences to align with the reference motion sequences. Note that the described proposal also works well with the corresponding outdoor scanning data.

To mitigate the fact the clothing dynamics cannot be controlled during scanning, we propose to take advantage of computer graphics technology, particularly physics-based simulations (PBSs). Based on recent advancements in modeling garment geometry [43, 51, 52] and dynamics [41, 42], it might be possible to fit the geometry and material properties to the scanning data. By fitting both geometry and material, the user receives a homologous mesh with known semantics, static, and dynamic properties. Such information can be used to transfer these properties to other scanning samples of the same type, as well as transfer the clothing simulation to the outdoor scanning data. Although such data cannot be considered ground-truth, we argue it is still valuable to understand and analyze clothing behavior in different environments, within, and between different actors.

6. Conclusion

We overviewed the protocols for 3D and 4D scanning of clothed people indoor and outdoor and offer recommendations and discussions on potentially optimizing the tedious and expensive process of collecting the datasets. The emphasis is put on the planning component to prevent or mitigate overhead work, to foresee potential issues, and to point out the importance of precisely defining the scanning protocols. The protocols are difficult to change along the way, causing inconsistency between the data samples and, thus, lowering the final quality of the dataset. Even though the scanning technology is mature and the postprocessing algorithms for 3D and 4D reconstruction are constantly improving, there are still several important protocol limitations, some of which are inherent to the scanning process and will hardly ever be completely solved. For example, scalable, high-quality scanning of clothed people is challenged by the difficulty of processing a significant sample of the population and, therefore, draw meaningful general conclusions. For now, we discuss some of the viable solutions to mitigate the limitations, such as using the advancements in body and garment modeling for motion sequence alignment within and between the subjects, and physics-based simulations for understanding the clothing geometry and material properties. We hope that our manuscript will prove useful to the researchers and other practitioners in understanding the scanning protocols for their applications and avoiding some of the pitfalls.

References

[1] K. Bartol, D. Bojanić, T. Petković and T. Pribanić, "A Review of Body Measurement Using 3D Scanning," in IEEE Access, vol. 9, pp. 67281-67301, 2021, doi: 10.1109/ACCESS.2021.3076595.

[2] N. Koepke, M. Zwahlen, J. C. Wells, N. Bender, M. Henneberg, F. J. Rühli, et al., "Comparison of 3D laser-based photonic scans and manual anthropometric measurements of body size and shape in a validation study of 123 young swiss men", PeerJ, vol. 5, pp. e2980, Feb. 2017, doi: 10.7717/peerj.2980

[3] M. Kouchi and M. Mochimaru, "Errors in landmarking and the evaluation of the accuracy of traditional and 3D anthropometry", Appl. Ergonom., vol. 42, no. 3, pp. 518-527, Mar. 2011, doi: 10.1016/j.apergo.2010.09.011

[4] J. Medina-Inojosa, V. Somers, T. Ngwa, L. Hinshaw and F. López-Jiménez, "Reliability of a 3D body scanner for anthropometric measurements of central obesity", Obesity Open Access, vol. 2, no. 3, pp. 10, 2016, doi: 10.16966/2380-5528.122

[5] 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", 2018.

[6] F. Bogo, J. Romero, M. Loper and M. J. Black, "FAUST: Dataset and Evaluation for 3D Mesh Registration," 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, pp. 3794-3801, doi: 10.1109/CVPR.2014.491.

[7] F. Bogo, J. Romero, G. Pons-Moll and M. J. Black, "Dynamic FAUST: Registering Human Bodies in Motion," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 5573-5582, doi: 10.1109/CVPR.2017.591.

[8] G. Pons-Moll, J. Romero, N. Mahmood, and M. J. Black. 2015. Dyna: a model of dynamic human shape in motion. ACM Trans. Graph. 34, 4, Article 120 (August 2015), 14 pages. Doi: https://doi.org/10.1145/2766993

[9] N. Mahmood, N. Ghorbani, N. F. Troje, G. Pons-Moll and M. Black, "AMASS: Archive of Motion Capture As Surface Shapes," 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), 2019, pp. 5441-5450, doi: 10.1109/ICCV.2019.00554.

[10] Matthew Trumble, Andrew Gilbert, Charles Malleson, Adrian Hilton, John P. Collomosse: Total Capture: 3D Human Pose Estimation Fusing Video and Inertial Sensors. British Machine Vision Conference, 2017

[11] A. Chatzitofis et al., "HUMAN4D: A Human-Centric Multimodal Dataset for Motions and Immersive Media," in IEEE Access, vol. 8, pp. 176241-176262, 2020, doi: 10.1109/ACCESS.2020.3026276.

[12] G. Pons-Moll, S. Pujades, S. Hu and M. J. Black. “ClothCap: seamless 4D clothing capture and retargeting.” ACM Trans. Graph. 36 (2017): 73:1-73:15, doi: 10.1145/3072959.3073711

[13] C. Zhang, S. Pujades, M. J. Black, and G. Pons-Moll. “Detailed, Accurate, Human Shape Estimation From Clothed 3D Scan Sequences”. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 2017

[14] T. von Marcard, R. Henschel, M. J. Black, B. Rosenhahn, and G. Pons-Moll. “Recovering Accurate 3D Human Pose in The Wild Using IMUs and a Moving Camera”. European Conference on Computer Vision (ECCV), September 2018.

[15] C. Ionescu, D. Papava, V. Olaru and C. Sminchisescu, "Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 7, pp. 1325-1339, July 2014, doi: 10.1109/TPAMI.2013.248.

[16] H. Joo, T. Simon, X. Li, H. Liu, L. Tan, L. Gui, S. Banerjee, T. S. Godisart, B. Nabbe, I. Matthews, T. Kanade, S. Nobuhara, and Y. Sheikh. "Panoptic Studio: A Massively Multiview System for Social Motion Capture," 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015, pp. 3334-3342, doi: 10.1109/ICCV.2015.381.

[17] Yu, Zhixuan & Yoon, Jae Shin & Lee, In & Venkatesh, Prashanth & Park, Jaesik & Yu, Jihun & Park, Hyun. (2020). HUMBI: A Large Multiview Dataset of Human Body Expressions. 2987-2997. Doi: 10.1109/CVPR42600.2020.00306.

[18] M. Armando, L. Boissieux, E. Boyer, J. S. Franco, M. Humenberger, C. Legras, V. Leroy, M. Marsot, J. Pansiot, S. Pujades, R. Rekik, G. Rogez, A. Swamy, S. Wuhrer. “4DHumanOutfit: a multi-subject 4D dataset of human motion sequences in varying outfits exhibiting large displacements”. ArXiv preprint, June 2023, doi: arxiv-2306.07399

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

[20] N. Zakaria and D. Gupta, Anthropometry Apparel Sizing and Design, Amsterdam, The Netherlands:Elsevier, 2019

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

[22] T. Olds and F. Honey, "The use of 3D whole-body scanners in anthropometry", Proc. 9th Int. Conf. Int. Soc. Adv. Kinanthropometry, pp. 1-12, Jan. 2005

[23] R. N. Lescay, A. Becerra and A. H. González, "Anthropometry. Comparative analysis of technologies for the capture of anthropometric dimensions", Revista EIA, pp. 47-59, Dec. 2016

[24] Sigal, L., Balan, A.O. & Black, M.J. HUMANEVA: Synchronized Video and Motion Capture Dataset and Baseline Algorithm for Evaluation of Articulated Human Motion. Int J Comput Vis 87, 4–27 (2010). https://doi.org/10.1007/s11263-009-0273-6

[25] M. Kaufmann, J. Song, C. Guo, K. Shen, T. Jiang, C. Tang, J. J. Zarate, O. Hilliges. “EMDB: The Electromagnetic Database of Global 3D Human Pose and Shape in the Wild”. International Conference on Computer Vision (ICCV), 2023

[26] S. Hanisch, E. Muschter, A. Hatzipanayioti, S. C. Li, T. Strufe. “Understanding Person Identification through Gait”. ArXiv, Mar 2022, doi: 10.48550/arXiv.2203.04179

[27] T. Weyrich, M. Pauly, R. Keiser, S. Heinzle, S. Scandella, and M. Gross. “Post-processing of Scanned 3D Surface Data”. 2004, Science. Doi: 10.2312/SPBG/SPBG04/085-094.

[28] E. Parrilla, A. Ballester, F. PARRA, A. Ruescas, & J. URIEL, D. GARRIDO, and S. Alemany. “MOVE 4D: Accurate High-Speed 3D Body Models in Motion”. 2019. 10.15221/19.030.

[29] Y. Wang, Z. Su, N. Zhang, R. Xing, D. Liu, T. H. Luan, X. Shen. “A Survey on Metaverse: Fundamentals, Security, and Privacy”. IEEE Communications Surveys & Tutorials, Jan 2023, doi: 10.1109/COMST.2022.3202047

[30] T. Anvari and K. Park, "3D Human Body Pose Estimation in Virtual Reality: A survey," 2022 13th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, Republic of, 2022, pp. 624-628, doi: 10.1109/ICTC55196.2022.9952586.

[31] M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, M. J. Black. SMPL: a skinned multi-person linear model. ACM Transactions on Graphics, 2015, doi: 10.1145/2816795.2818013.

[32] Đonlić, Matea. "Three-dimensional analysis of back surface under dynamic conditions in scoliosis diagnostics." PhD thesis, University of Zagreb, Faculty of Electrical Engineering and Computing, 2019, https://urn.nsk.hr/urn:nbn:hr:168:702231

[33] Keller, Marilyn & Zuffi, Silvia & Black, Michael & Pujades, Sergi. (2022). OSSO: Obtaining Skeletal Shape from Outside. 20460-20469. 10.1109/CVPR52688.2022.01984.

[34] Casadei K, Kiel J. Anthropometric Measurement. 2022 Sep 26. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan–. PMID: 30726000.

[35] Bianco, Simone, Gianluigi Ciocca, and Davide Marelli. 2018. "Evaluating the Performance of Structure from Motion Pipelines" Journal of Imaging 4, no. 8: 98, doi: 10.3390/jimaging4080098

[36] A. H. Ahmadabadian, A. Karami and R. Yazdan, "An automatic 3D reconstruction system for texture-less objects", Robot. Auton. Syst., vol. 117, pp. 29-39, Jul. 2019, doi: 10.1016/j.robot.2019.04.001

[37] Rajasegaran, Jathushan & Pavlakos, Georgios & Kanazawa, Angjoo & Malik, Jitendra. (2022). Tracking People by Predicting 3D Appearance, Location and Pose. 2730-2739, doi: 10.1109/CVPR52688.2022.00276.

[38] Goel, Shubham, Georgios Pavlakos, Jathushan Rajasegaran, Angjoo Kanazawa and Jitendra Malik. “Humans in 4D: Reconstructing and Tracking Humans with Transformers.” ArXiv, doi: abs/2305.20091 (2023)

[39] Li, Z., Liu, J., Zhang, Z., Xu, S., Yan, Y. (2022). CLIFF: Carrying Location Information in Full Frames into Human Pose and Shape Estimation. In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M., Hassner, T. (eds) Computer Vision – ECCV 2022. ECCV 2022. Lecture Notes in Computer Science, vol 13665. Springer, Cham, doi: 10.1007/978-3-031-20065-6_34

[40] Kocabas, Muhammed, Chun-Hao Paul Huang, J. Tesch, Lea Muller, Otmar Hilliges and Michael J. Black. “SPEC: Seeing People in the Wild with an Estimated Camera.” 2021 IEEE/CVF International Conference on Computer Vision (ICCV) (2021): 11015-11025.

[41] Santesteban, Igor, Miguel A. Otaduy and Dan Casas. “SNUG: Self-Supervised Neural Dynamic Garments.” 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2022): 8130-8140.

[42] Grigorev, Artur, B. Thomaszewski, Michael J. Black and Otmar Hilliges. “HOOD: Hierarchical Graphs for Generalized Modelling of Clothing Dynamics.” 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2022): 16965-16974.

[43] Luigi, Luca & Li, Ren & Guillard, Benoît & Salzmann, Mathieu & Fua, Pascal. (2023). DrapeNet: Garment Generation and Self-Supervised Draping. 1451-1460, Doi: 10.1109/CVPR52729.2023.00146.

[44] Halimi, Oshri, Fabián Prada, Tuur Stuyck, Donglai Xiang, Timur M. Bagautdinov, He Wen, Ron Kimmel, Takaaki Shiratori, Chenglei Wu and Yaser Sheikh. “Garment Avatars: Realistic Cloth Driving using Pattern Registration.” ArXiv, doi: abs/2206.03373 (2022)

[45] Kinovis, Inria 4D modeling multi cameras platform. Online http://kinovis.inria.fr/ inria-platform, 2023. Accessed on 31st August 2023.

[46] Zins, Pierre, Yuanlu Xu, Edmond Boyer, Stefanie Wuhrer and Tony Tung. “Data-Driven 3D Reconstruction of Dressed Humans From Sparse Views.” 2021 International Conference on 3D Vision (3DV) (2021): 494-504.

[47] Badiola-Bengoa, Aritz, and Amaia Mendez-Zorrilla. 2021. "A Systematic Review of the Application of Camera-Based Human Pose Estimation in the Field of Sport and Physical Exercise" Sensors 21, no. 18: 5996, doi: 10.3390/s21185996

[48] Osman, Ahmed A. A., Timo Bolkart and Michael J. Black. “STAR: Sparse Trained Articulated Human Body Regressor.” ArXiv, doi: abs/2008.08535 (2020)

[49] Pavlakos, Georgios, Vasileios Choutas, Nima Ghorbani, Timo Bolkart, Ahmed A. A. Osman, Dimitrios Tzionas and Michael J. Black. “Expressive Body Capture: 3D Hands, Face, and Body From a Single Image.” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019): 10967-10977.

[50] Osman, Ahmed A. A., Timo Bolkart, Dimitrios Tzionas and Michael J. Black. “SUPR: A Sparse Unified Part-Based Human Representation.” European Conference on Computer Vision (2022).

[51] Ma, Qianli, Jinlong Yang, Siyu Tang and Michael J. Black. “The Power of Points for Modeling Humans in Clothing.” 2021 IEEE/CVF International Conference on Computer Vision (ICCV) (2021): 10954-10964.

[52] Corona, Enric, Albert Pumarola, G. Alenyà, Gerard Pons-Moll and Francesc Moreno-Noguer. “SMPLicit: Topology-aware Generative Model for Clothed People.” 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2021): 11870-11880.

[53] Zhang, Jinsong, Kun Li, Yu-Kun Lai and Jingyu Yang. “PISE: Person Image Synthesis and Editing with Decoupled GAN.” 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2021): 7978-7986.