BodyDigitizer: An Open Source Photogrammetry-based 3D Body Scanner

Travis Gesslein, Daniel Scherer, Jens Grubert

Department of Electrical Engineering and Computer Science
Coburg University
E-Mail: jg@jensgrubert.de

Abstract: With the rising popularity of Augmented and Virtual Reality, there is a need for representing humans as virtual avatars in various application domains ranging from remote telepresence, games to medical applications. Besides explicitly modelling 3D avatars, sensing approaches that create person-specific avatars are becoming popular. However, affordable solutions typically suffer from a low visual quality and professional solution are often too expensive to be deployed in nonprofit projects. We present an open-source project, BodyDigitizer, which aims at providing both build instructions and configuration software for a high-resolution photogrammetry-based 3D body scanner. Our system encompasses up to 96 Raspberry PI cameras, active LED lighting, a sturdy frame construction and open-source configuration software. The detailed build instruction and software are available at http://www.bodydigitizer.org.

Keywords: 3d body scanner, 3d scan, full-body scan, photogrammetry, open-source

1 Introduction

![3D scan results from the proposed BodyDigitizer setup.](image)

With the rise of consumer-oriented Augmented and Virtual Reality (AR/VR) systems and applications, there is a growing demand to represent person-specific 3D representations
of humans. Applications that can strongly benefit from person-specific human 3D representations include remote telepresence [OERF+16], games and social VR [fbs], fitness applications [nak] or medical applications [MTM+17].

There is a long history of capturing the appearance (and motion) of human bodies [FBA+94], and open-source or low-cost approaches with commodity hardware suitable for full-body scans, emerged over the last years (e.g., [TZL+12, sca, SFW+14, ZFYY14, CCNS12]). Single-sensor solutions, e.g. using commodity depth sensors like the Kinect [CCNS12] or Intel RealSense [str] result in comparatively low hardware costs but can suffer from lower spatial resolution compared to multi-view stereo setups [SJP13] and require users to stand still for a prolonged period of time. Similar, utilizing single RGB-cameras for 3D reconstruction is popular in a wide variety of capturing tasks due to their wide-availability, e.g., city-scale 3D reconstructions [AFS+11] hologram verification [HGSR13] but can suffer from similar constraints as the single-depth camera-based approaches. In contrast, multi-camera rigs allow a fast capture of human bodies but can typically require a more elaborate physical setup.

Within this paper, we focus on high-resolution reconstruction (see Figure 1) with a short acquisition time using a multi-camera rig that relies on photogrammetry as 3D measurement principle. Specifically, we make available the build instructions as well as the acquisition software open-source.

2 Related Work

Commodity and open-source 3D body scanners have become popular over the last years.

Cui et al. proposed to scan a human body with a single Kinect [CS11], but the results were of low quality as the Cui’s approach did not handle non-rigid movement or used color information. Weiss et al. [WHB11] fitted a SCAPE model [ASK+05] to the 3D and image silhouette data from a Kinect, but failed to reproduce personalized details. Tong et al. used multiple stationary depth-sensors in combination with a turntable [TZL+12]. Cui et al. simplified this setup to a single depth-sensor [CCNS12]. For this turntable-based approach, open-source build instructions have been made available [sca]. Zhang et al. [ZFYY14] relaxed scanning requirements further and presented an approach that works with a handheld scanner. However, all those approaches result in long acquisition times, which can have negative impact on the scanning result (due to human motions) or be unacceptable for a given application scenario (e.g., scanning patients with bodily disorders for medical applications).

There also have been recent advances in real-time capture [OERF+16]. However, also these state of the art approaches typically suffer from visible spatial and temporal artifacts, which might be undesirable for various applications. Further, for some applications the hard real-time constraints might not be necessary.

Offline approaches typically result in higher quality models [CCS+15] but can be cost intensive with commercial 3D scanners (e.g. Vitronic Vitus [vit]) easily reaching a cost of more than 50,000 Euro). The closest work to ours is the PI3DSCAN project [pi3, Gar14] and similar derivatives based on a Raspberry Pi infrastructure (e.g., [SK14]). The PI3DSCAN project
proposes a multi-camera setup for use with photogrammetry-based 3D reconstruction. While the PI3DSCAN project provides selected information regarding relevant hardware components, it also aims at selling products and services. Specifically, the project lacks concise assembly instructions or relevant open-source software for the management of the hardware components (e.g., scanning management software is sold for a yearly license of 750 Euro). Hence, we aim at providing both concise built instructions as well as open-source acquisition software.

3 3D Body Scanner

Figure 2: From left to right: Outside view of the scanner with person and turned on LED strips; inside view; projector camera, power supplies and switches are mounted at the frame; detailed front view of camera mount; detailed side view of camera mount.

Our 3D body scanner consists of a wooden frame, electronic components built around Raspberry PIs and an acquisition software, which are described next. The total hardware cost of the scanner amounts to 9500 Euro, which is still substantially cheaper than commercial models, and can be decreased further with a lower number of cameras.

3.1 Frame Construction and Mounts

The frame was constructed to hold a flexible number of cameras with a minimum horizontal angle between cameras of ca. 13° (except for the entrance area, facing the back of the user). The shaped frame has the dimensions (width x depth x height) of 2.90m x 2.51m x 2.10m and a total number of 24 beams to ensure that mounted cameras with standard wide angle lenses could capture body parts with sufficient overlap. The total cost of the frame (wood, metal brackets, screws) was 150 Euro.

A challenge was to create affordable camera mounts. These camera mounts are required to position each camera module interactively when building the 3D scanner, and then hold them in place tightly. With 96 required mounts, even affordable consumer-oriented solutions
like ball heads would quickly become expensive at scale (prices starting at 4 Euro per ball head). Also, 3D printing camera mounts would have been prohibitive, due to the amount of time required to print 96 holders. Instead, we opted for a simple solution consisting out of a wooden board on which the camera module was fixed with hot glue. This frame was then fixated with 3 screws (1 on the front, 2 on the back) on the underlying wooden bar, allowing to adjust the pose of the camera by screwing in or out the individual screws. The mounting procedure took on average 3 minutes and the cost was 12 Cent per mount.

For mounting Raspberry PIs to the frame, we used also used simple wooden boards. The Raspberry PIs were hot glued to the board and the board, subsequently, screwed to the wooden beams.

![Figure 3: Left: 3D model of frame. Right: Assembly steps to final frame.](image)

### 3.2 Electronic Components

As we wanted to have flexible imaging pipeline enabling both high quality reconstruction and (in the future) live streaming, we chose 96 Raspberry PIs (Model 3) (4 per beam) in combination with Raspberry PI Camera Module V1\(^1\) and a 8 GB micro-sd card. The total price of these components was 5.400 Euro. In addition, we wanted to have adjustable lighting

\(^1\)https://www.raspberrypi.org/documentation/hardware/camera/
and added 12V LED stripes with 60 LEDs and 1000-1300 lumen per meter as primary light source (see Figure 2, for close-up view).

For power supply, we used ten 300 Watt PC power supplies with a 24 pole ATX connector. We cut off the connector and used the individual wires for 5V (Raspberry PI) and 12V (LED stripes) power supply. As we wanted to minimize the number of cables used (and the price per cable), we employed CAT 5 cables both for data transfer and power supply. However, due to the small diameter of individual wires, we limited the length of the power wires to 80 cm in order to mitigate the effect of potential drop between power supply and Raspberry Pis. This also implied that the power supplies where mounted in the center of wooden beams serving power also to two beams to the left and right. With longer cables, the Raspberry Pis would occasionally turn off during runtime as the potential drop would lead to a final voltage of 4.65 V or below.

To allow for LED lighting control, we use custom MOSFET boards, which can control four stripes via commands from a single Raspberry PI (using its GPIO pins). Currently, we can control the light to be in 100%, 50% or 0% level. In future, one could extend this setting to allow for shape from shading approaches using synchronized light patterns.

Finally, four short-throw projectors (Optoma GT 760) are used to project image patterns onto the human body to allow surface reconstruction in otherwise textureless regions.

### 3.3 Data Acquisition

The acquisition software is based on node.js and follows a client-server model. The clients automatically try to connect to a known server IP at a regular interval. Using a simple command line interface the user can issue commands to all connected clients to take images, turn LEDs on / off. Once the image acquisition command has been issued, two sets of images are taken. First, an image set is created with projectors showing a black image (i.e. no projection) to be able to recreate the surface texture. Immediately after the first picture, a random dot pattern is projected onto the user and a second image is taken to allow for surface reconstruction. Then the images are automatically transferred to the PC.

For 3D reconstruction, we currently rely on a proprietary photo reconstruction software (Agisoft PhotoScan [agi]), which automatically converts the reconstructed point cloud data into textured meshes. We plan to inspect alternative approaches in the future. For further use of the avatars, we, currently, employ a manual rigging and skinning process, but plan to investigate automated approaches [FCS15].

### 4 Limitations and Future Work

While we were able to reduce the building cost to less than 10,000 Euro the required manual effort is still substantial, requiring approximately two person months to build such a system. Depending on the required model fidelity, a smaller number of cameras (the current cost for a single Raspberry PI 3 + camera + sd card is approximately 56 Euro) might be sufficient. For example, in Figure 4, right, we used 48 cameras, in comparison to the full setup with 96
cameras in Figure 4, left. Also, depending on the use case (e.g., expected clothing), the four projectors (with a total cost of 2,000 Euro) might be optional.

While we have completed an acquisition software, we still rely on commercial solutions for the actual 3D reconstruction of the models. Further, ideally, the model scanning, conversion, rigging, skinning and provisioning into the VR scene should be fully automated, in order to support swift working procedures when preparing and conducting experiments.

In future work, we want to address these issues by providing a fully automated and flexible body scanning and streaming pipeline. Specifically, we aim at the inclusion of image-based visual hull algorithms [MBR+00] employing the already installed Raspberry PI cameras or the inclusion of multiple commodity depth cameras [KHR+12] for streaming purposes. Further, we can imagine applying

5 Conclusion

We have presented an open-source project, BodyDigitizer, which aims at providing both built instructions and configuration software for a high-resolution photogrammetry-based 3D body scanner. Our system encompasses up to 96 Rasperry PI cameras, active LED lighting, a sturdy frame construction and open-source configuration software. The detailed build instruction and software are available at http://www.bodydigitizer.org.
References

[AFS+11] Sameer Agarwal, Yasutaka Furukawa, Noah Snavely, Ian Simon, Brian Curless, Steven M Seitz, and Richard Szeliski. Building rome in a day. *Communications of the ACM*, 54(10):105–112, 2011.

[agi] Agisoft photoscan. http://www.agisoft.com/. Accessed: 2017-09-11.

[ASK+05] Dragomir Anguelov, Praveen Srinivasan, Daphne Koller, Sebastian Thrun, Jim Rodgers, and James Davis. Scape: shape completion and animation of people. In *ACM Transactions on Graphics (TOG)*, volume 24, pages 408–416. ACM, 2005.

[CCNS12] Yan Cui, Will Chang, Tobias Nöll, and Didier Stricker. Kinectavatar: fully automatic body capture using a single kinect. In *Asian Conference on Computer Vision*, pages 133–147. Springer, 2012.

[CCS+15] Alvaro Collet, Ming Chuang, Pat Sweeney, Don Gillett, Dennis Evseev, David Calabrese, Hugues Hoppe, Adam Kirk, and Steve Sullivan. High-quality streamable free-viewpoint video. *ACM Transactions on Graphics (TOG)*, 34(4):69, 2015.

[CS11] Yan Cui and Didier Stricker. 3d shape scanning with a kinect. In *ACM SIGGRAPH 2011 Posters*, page 57. ACM, 2011.

[FBA+94] Henry Fuchs, Gary Bishop, Kevin Arthur, Leonard McMillan, Ruzena Bajcsy, Sang Lee, Hany Farid, and Takeo Kanade. Virtual space teleconferencing using a sea of cameras. In *Proc. First International Conference on Medical Robotics and Computer Assisted Surgery*, volume 26, 1994.

[fbs] Facebook spaces. https://www.facebook.com/spaces. Accessed: 2017-09-11.

[FCS15] Andrew Feng, Dan Casas, and Ari Shapiro. Avatar reshaping and automatic rigging using a deformable model. In *Proceedings of the 8th ACM SIGGRAPH Conference on Motion in Games*, pages 57–64. ACM, 2015.

[Gar14] Richard Garsthagen. An open source, low-cost, multi camera full-body 3d scanner. In *5th International Conference on 3D Body Scanning Technologies*, 2014.

[HGSR13] Andreas Hartl, Jens Grubert, Dieter Schmalstieg, and Gerhard Reitmayr. Mobile interactive hologram verification. In *Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on*, pages 75–82. IEEE, 2013.

[KHR+12] Bernhard Kainz, Stefan Hauswiesner, Gerhard Reitmayr, Markus Steinberger, Raphael Grasset, Lukas Gruber, Eduardo Veas, Denis Kalkofen, Hartmut Seichter, and Dieter Schmalstieg. Omnikineckt: real-time dense volumetric data
acquisition and applications. In *Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology*, pages 25–32. ACM, 2012.

[MBR+00] Wojciech Matusik, Chris Buehler, Ramesh Raskar, Steven J Gortler, and Leonard McMillan. Image-based visual hulls. In *Proceedings of the 27th Annual Conference on Computer graphics and Interactive Techniques*, pages 369–374. ACM Press/Addison-Wesley Publishing Co., 2000.

[MTM+17] S Mölbert, A Thaler, B Mohler, S Streuber, M Black, HO Karnath, S Zipfel, and K Giel. Assessing body image disturbance in patients with anorexia nervosa using biometric self-avatars in virtual reality: Attitudinal components rather than visual body size estimation are distorted. *Journal of Psychosomatic Research*, 97:162, 2017.

[nak] Naked fit. https://naked.fit/. Accessed: 2017-09-11.

[OERF+16] Sergio Orts-Escolano, Christoph Rhemann, Sean Fanello, Wayne Chang, Adarsh Kowdle, Yury Degtyarev, David Kim, Philip L. Davidson, Sameh Khamis, Ming-song Dou, Vladimir Tankovich, Charles Loop, Qin Cai, Philip A. Chou, Sarah Memmicken, Julien Valentin, Vivek Pradeep, Shenlong Wang, Sing Bing Kang, Pushmeet Kohli, Yuliya Lutchyn, Cem Keskin, and Shahram Izadi. Holoportation: Virtual 3d teleportation in real-time. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST ’16, pages 741–754, New York, NY, USA, 2016. ACM.

[pi3] Pi3d 3d body scanner. http://www.pi3dscan.com/. Accessed: 2017-09-11.

[sca] Scanoman - open source full-body 3d scanner. http://3dmaker4u.com/scanoman-project.html. Accessed: 2017-09-11.

[SFW+14] Ari Shapiro, Andrew Feng, Ruizhe Wang, Hao Li, Mark Bolas, Gerard Medioni, and Evan Suma. Rapid avatar capture and simulation using commodity depth sensors. *Computer Animation and Virtual Worlds*, 25(3-4):201–211, 2014.

[SJP13] Jan Smisek, Michal Jancosek, and Tomas Pajdla. 3d with kinect. In *Consumer depth cameras for computer vision*, pages 3–25. Springer, 2013.

[SK14] Jeremy Straub and Scott Kerlin. Development of a large, low-cost, instant 3d scanner. *Technologies*, 2(2):76–95, 2014.

[str] Structure sensor - 3d scanning, augmented reality, and more for mobile devices. http://structure.io. Accessed: 2017-09-11.

[TZL+12] Jing Tong, Jin Zhou, Ligang Liu, Zhigeng Pan, and Hao Yan. Scanning 3d full human bodies using kinects. *IEEE Transactions on Visualization and Computer Graphics*, 18(4):643–650, 2012.
3d body scanner - vitronic - the machine vision people. https://www.vitronic.com/industrial-and-logistics-automation/sectors/3d-body-scanner.html. Accessed: 2017-09-11.

[WHB11] Alexander Weiss, David Hirshberg, and Michael J Black. Home 3d body scans from noisy image and range data. In Computer Vision (ICCV), 2011 IEEE International Conference on, pages 1951–1958. IEEE, 2011.

[ZFY14] Qing Zhang, Bo Fu, Mao Ye, and Ruigang Yang. Quality dynamic human body modeling using a single low-cost depth camera. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 676–683, 2014.