

ReconstructMe – Towards a Full Autonomous Bust Generator

Christoph KOPF*, Christoph HEINDL, Martin ANKERL,
Harald BAUER, Andreas PICHLER
Profactor GmbH, Steyr (Upper Austria), Austria

<http://dx.doi.org/10.15221/14.184>

Abstract

In the most recent years, 3D data acquisition was revolutionized by consumer grade real-time capable 3D depth cameras with appropriate data quality. A wide range of different 3D applications was developed since the release of these cameras. Especially the area of 3D body scanning became cost efficiently and the number of applications started to grow. A new trend is to generate 3D self-portraits and print them on commercial available 3D printer. This paper introduces a low-cost scanning system to generate printable 3D upper torso busts at home using commodity computation hardware.

The workflow of the system is optimized to work without the need of any manual interaction with the user and any additional persons. It's designed to generate 3D printable busts automatically which allows people without 3D knowledge to create their own 3D busts. The user just has to perform a full rotation (360 degrees) to capture the upper torso data from all views. The end of the full rotation is detected automatically. When the scanning procedure is done, a post processing routine is triggered. This routine includes cleaning the scan from noise, estimating the geometry and color of unseen areas and cutting the bust to a certain height. The final result is a watertight, scaled and oriented triangle mesh. This allows the user to put the final mesh to common 3D printer drivers and immediately print its 3D bust.

In our tests, the workflow was proven to be very robust and we were able to generate hundreds of 3D busts with high recognition value of the persons. However, our tests showed different results depending on the conditions. In this paper we will introduce our experience of how the setup can be improved and thus the quality of the results can be enhanced.

In future, this principle could be extended by using more than one sensor at the same time to capture the whole body shape of a person and to generate clean 3D full body models.

Keywords: 3D, reconstruction, kinect, 3D printer, 3D bust

1. Introduction

Since the release of low-cost sensors, 3D came into our living rooms. However, industry already recognized the potential of 3D before, but for mass production, it was still too expensive. Object recognition, tracking, scanning of objects are common tasks in the 3D world, which now can enter homes.

As 3D printers are getting cheaper, the number of 3D online printing services starts to grow and the low-cost sensors fill the market. This also means, that 3D is getting more popular and it's up to technicians and visionaries to bring these 3D devices together and enable powerful applications. However, creativity is an important factor to motivate people to use new technology. Thus, intuitive and creative applications are required to make people familiar with 3D. Many creative applications are about scanning full body shapes and use them for tasks people are familiar with. Common usage examples can be found in apparel, medicine, sports industry or simply to make a physical 3D model of themselves.

In our case, we have developed an application that is able to generate 3D busts of persons entirely on their own. Only common computer hardware and a 3D sensor are required. Due to the intuitive process design, home users without any knowledge of 3D technologies are now able to generate their own 3D busts. No additional persons and no post processing is required. The next chapters will introduce our system.

2. Related Work

The release of 3D low-cost sensors made capturing objects in 3D affordable and thus, many people started to get interested in this technology. Since then, the number of 3D applications has been growing, especially when it comes to applications connected to human body scanning.

* christoph.kopf@profactor.at; www.profactor.com

Human body scanning became a famous use case in apparel, medicine, sports, creativity and other fields. However, there are many different solutions in creating 3D body shapes from real persons. Many applications use low-cost sensors to measure 3D body shapes with either one sensor [1, 2, 3, 4, 5, 6] or multiple sensors [7, 8]. Using multiple sensors simultaneously leads to more data from different views at less time, but system complexity and calibration efforts grow. Scanning with a single sensor on the other hand has less complexity and lower calibration effort, but often leads to holes in the final model. These holes can be fixed with reconstruction algorithms. The most popular reconstruction algorithm is Poisson Reconstruction [9], as used in some approaches [6, 1]. Besides closing holes, the surface becomes smoother.

The scan setup of the approaches differs too. In many scan setups, the user has to turn in front of the sensor on his own [6, 4, 2, 1], others use turntables to rotate the user [8]. For all scan setups, it is necessary for the scanned person to remain still for more than a few seconds, which is especially hard when they have to change their relative positions to the sensor during scanning. Thus, many approaches use non-rigid point cloud registration [6, 4, 2, 8, 1] and loop closure techniques to deal with the non-rigid nature of human bodies. Other approaches use rigid alignment and fuse the generated point clouds into a single volume representation [3, 7]. The volume approach of [3] has the advantage that it filters noise from sensor data, since the single volume representation averages the captured data. Thus, small movements get filtered too.

The goal of the approach described here is to develop an intuitive and autonomous bust generation algorithm. Thus, we only need to scan the upper part of the body and the head. The sensor should be placed in a position where it can capture most data. From literature, this will lead to holes in the final model and a reconstruction algorithm will be required. Otherwise the output won't be printable without 3D knowledge and manual post processing. Since we provide the software to users without 3D knowledge, the process has to be intuitive. The proposed solution uses one sensor that has to be placed at about eye level to capture as much surface data as possible. However, holes can't be avoided and will exist in the raw scan result. These holes will be closed by applying a reconstruction algorithm [6, 1, 9]. Finally, the reconstructed model has to be cut to a bust and colored to produce printable outputs.

3. Method

A composition of 3D based algorithms leads to a printable bust model of any person. Since the system will be used by a wide range of people, we cannot assume that these people has knowledge of 3D technologies or algorithms. Thus, the main goal of the system is to tailor the process to the proposed use case and reduce the user interaction to a minimum. This should enhance the user experience and avoid faults.

This chapter will introduce the main ideas of the system.

3.1. Use Case

Since the complexity of the algorithms should be hidden from the user, the process has to be automatic. Therefore, a robust pipeline is required.

The overall goal is to create a printable 3D bust of a single user without the need of any additional operators. The previous developments allowed to use portable low-cost hardware to create scans of people [3]. However, the process was designed to scan people by moving the sensor around and the final result was the raw scan. The scan didn't finished automatically after a full rotation and the scans were not processed at all. Manual post processing was required to make the models printable. The proposed scanning solution can perform starts and stops automatically. Afterwards, a post processing pipeline automatically creates printable busts.

3.1.1. Scan Setup

The system is based on previous developments [3] and requires common computer hardware, as shown in Fig. 1:

- Processing Unit (Computer with high-end GPU)
- Sensor
- User

The user will be scanned in front of the sensor should be at a distance of about 1.5 meter. The sensor has to be placed horizontally at about eye-level. It is connected to the processing unit which requires a consumer product GPU for real-time processing. Since the sensor will capture data from every object inside its field of view, a scan area with enough space has to be provided that includes the user but not the surroundings. Otherwise, these objects could influence the sensor tracking and the tracking algorithm might fail. When a suitable location is found, the user has to perform a 360 degree rotation in front of the sensor to be scanned from all views.

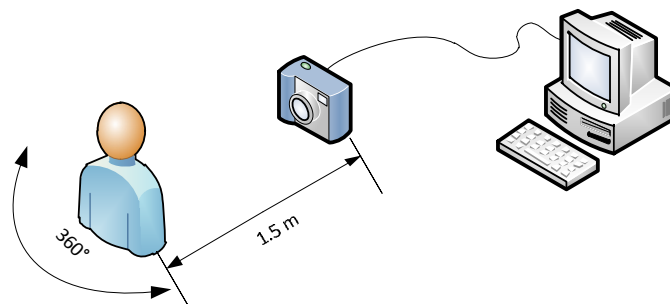


Fig. 1 System Components

3.1.2. Process Overview

Most parts of the overall process are automated. There is only a single user interaction required: Click a button to start the capturing process. This autonomous process is split into two main parts. These are **scanning** on the one hand and **post processing** on the other hand. See Fig. 2 for an overview of the process.

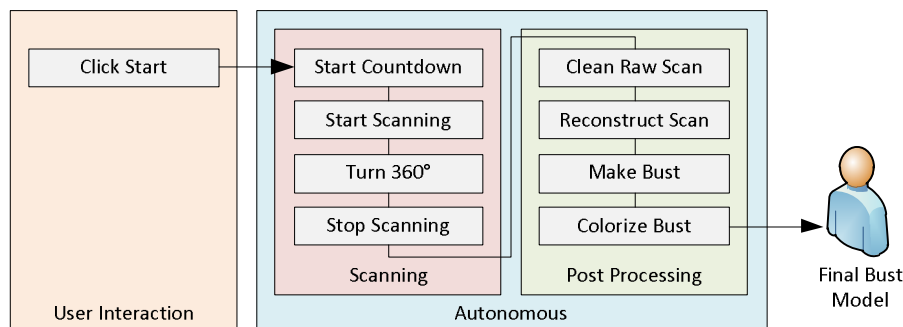


Fig. 2 Process Overview

When the scanning process has been triggered, a countdown is started. This allows the user to enter the scan position. When the countdown finishes, the scan is started and the user has to perform a full 360 degree rotation, either to his left or right. It is important that the user tries to keep his upper body static, since the approach requires rigid movements. The system is able to detect a full rotation automatically, see section 3.2.1 for more details. When a full rotation is detected, scanning will be stopped and a post processing pipeline is started. The goal of the post processing is to automate the bust generation from the raw scan. Find more details about post processing in section 3.3. Finally, the output of the autonomous process is a printable 3D bust. The bust is watertight and scaled to a common size usable for 3D printing. Since the model fulfills the requirements for printing, it can be loaded by common printing drivers.

3.2. Scanning

As Fig. 2 shows, scanning contains of basically 4 steps. This section introduces the automatic rotations detection that is required for this process. For more details about scanning itself see [3].

3.2.1. Detection of Rotation Angle

The user has to perform a full rotation in front of the sensor. It is not reasonable to let the user manually stop the scan process, since he would have to leave the scan position – while scanning. This would lead to artefacts in the scan. Thus, this solution automatically detects a full rotation of the user. To calculate the current rotation angle, we assume the origin of the world frame in the center of the scan volume. Since the user's position is in the center of the scan volume, we can further assume that the sensor path will be a circle around the center axis of the scan volume. Fig. 3 shows this setup.

The direction of the axis will be defined as union normal \vec{n}_r . Also, a plane p_r through the world frame origin with normal \vec{n}_r is defined. Consider two camera positions in the world frame without orientation (only translation) t_{c1} and t_{c2} , these points can be projected on p_r , resulting in t_{p1} and t_{p2} . Assuming the origin as center, the angle $\alpha_{t_{p1}t_{p2}}$ between t_{p1} and t_{p2} can be calculate. A full rotation is detected when $\left| \sum \alpha_{t_{pi}t_{pi+1}} \right| \geq 360$.

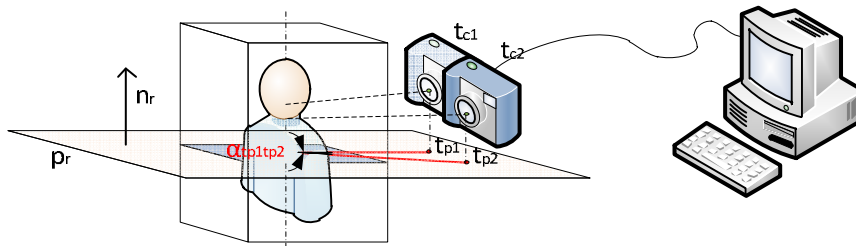


Fig. 3 Determining rotation angle

3.3. Post Processing

When the scanning process is finished, a raw model is generated. Printing a raw model is not possible, since there are artefacts, different isolated pieces have to be removed and the model is not watertight. While scanning, the user was only captured from one height, which might lead to holes on areas typically not visible like the top of the head, the chin and of course the bottom. The goal of the post processing is to prepare the scan in such a way, that it is printable. The following subsections introduce the required steps to get a printable bust model and Table 1 (see next page) provides an overview of the intermediate results.

3.3.1. Cleaning and Reconstruction

After scanning, a raw model is generated. This model consists of isolated segments due to sensor noise and the surface extraction algorithm [3]. First, these outliers get removed and a clean model is provided for reconstruction. In our implementation, we used the Poisson Surface Reconstruction algorithm described in [9].

This algorithm describes an oriented point cloud through a function and is thus able to estimate the surface at holes. At the same time, the model becomes smoothed at different levels, depending on the density of the point cloud, noise of the point cloud, and the used parameter set. Especially two parameters influence the quality a lot.

First, there is the **depth** parameter for the Octree, which is used for the marching cubes algorithm to produce polygon data. Usually this parameter has a value of the interval [6; 10]. A small value causes a coarse grained mesh with smoothed details, whereas a high value preserves details due to the higher depth of the Octree for the meshing.

The second parameter is called **samples per node** and describes how many points for an Octree leaf should be taken for smoothing. This parameter usually has values of the interval [1; 5] for noise free models and [6; 10] otherwise. This parameter determines the smoothness at a certain depth.

Since the generated model of our scanning method is very detailed and noise free, a medium depth of 8 and samples per node of 1 has been experimentally chosen. This leads to a satisfying result within a reasonable time frame.

Results of the cleaning and the reconstruction algorithm can be found at row 2 and 3 of Table 1.

3.3.2. Cut Model

Reconstructing the model leads to a smoothed mesh with closed holes. However, a cut is required to put the bust on flat surfaces. Therefore, the Constructive Solid Geometry (CSG) module of ReconstructMe SDK 2.1.845 is used [10]. CSG provides Boolean operation such as intersection, union, or difference on geometries such as spheres, boxes or planes. These operations can be used on implicit surfaces too. ReconstructMe SDK 2.1.845 is able to load closed manifold triangle meshes into a voxel grid and perform CSG operations on it.

While scanning, two volumes were created: A geometry volume containing the scan as implicit surface and a color volume containing averaged color information. Both volumes are in the same world frame and are stored on the GPU memory, since the scanning makes heavy use of GPU programming. The cleaning and reconstruction step described above is performed on the CPU, so was downloaded to the CPU's RAM. Now, the goal is to cut off the bottom of the scan to generate the reconstructed and watertight model. To do this, the geometry volume on the GPU first gets reset to initial values.

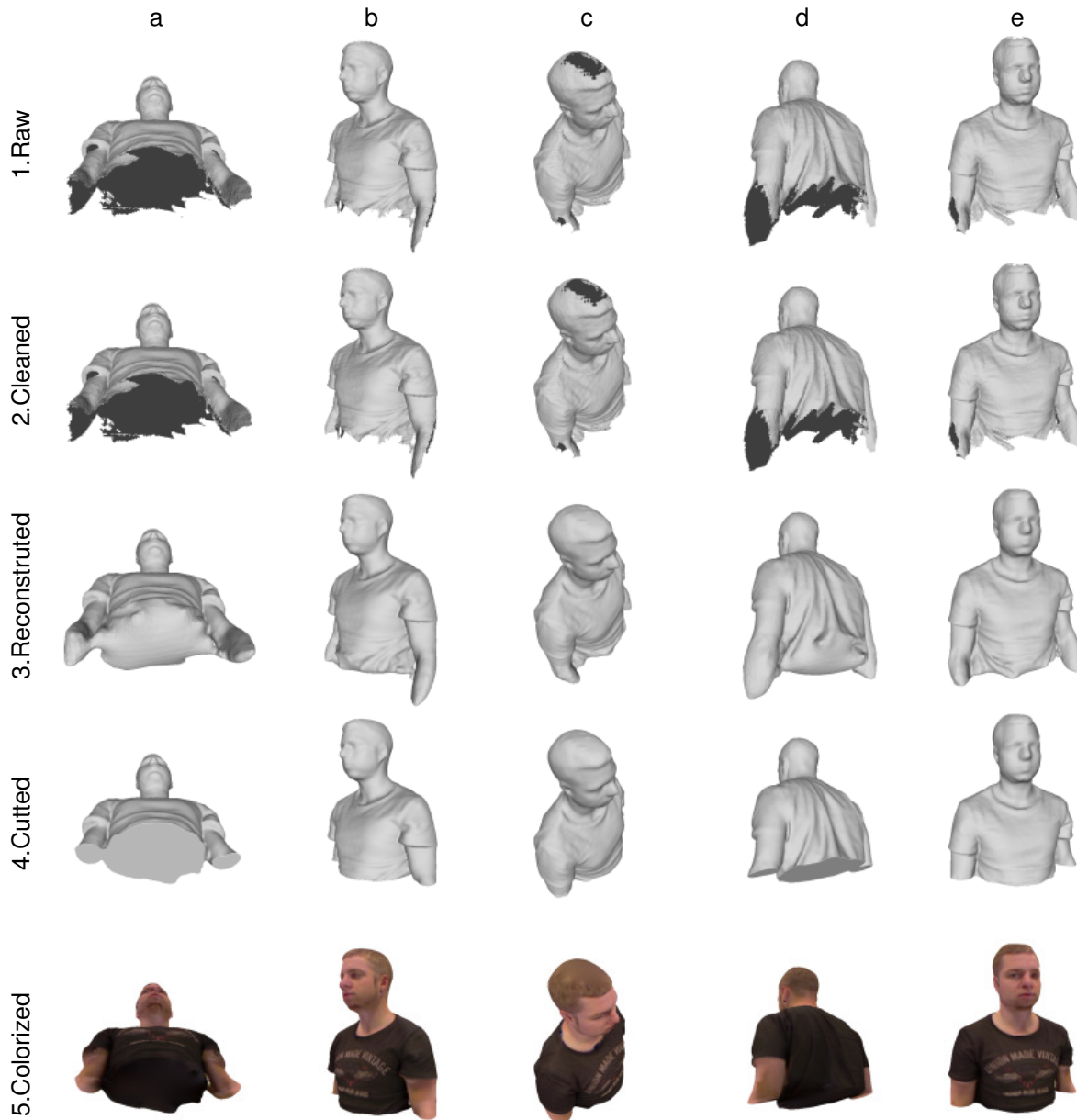


Table 1 Post processing of a model from different views at different steps

Afterwards, the reconstructed surface is loaded into the geometry volume and CSG operations can be performed on the GPU. Now it's an easy task to cut the model at a certain height. A subtraction operation with a plane is performed, such that the side of the negative normal vector of the plane is deleted. After this step, the model can be polygonized again, and downloaded to the RAM memory for further post processing.

For results of the CSG operation see row 4 of Table 1.

3.3.3. Inpaint Model

At this point, a colorless bust model is available. While scanning, the color was captured in a separate volume. Instead of geometry information, the color volume contains RGB information at world space. Thus, for each point of the final bust mesh, the color volume can be queried for an RGB value. However, since the reconstruction of the surface closed holes, there is no color information for these regions available and the missing information has to be estimated. To do this, we have developed a basic color inpaint algorithm to estimate the colors for areas without color information. The proposed algorithm first identifies groups of connected points without color information \mathbb{G} . Each group $g_i \in \mathbb{G}$ has a certain number of border points with color information. Now, each point in the group g_i takes

color information of the border points into account and generates a new color depending on the distance to the border point. This produces smoothed colorization on uncaptured areas. Since these areas are either not seen (bottom platform) or very homogenous (top of head), this simple approach produces satisfactory results. For results, please have a look at the 5th row of Table 1.

4. Results

Due to the presented method in Chapter 3, generating a 3D bust became very simple. Due to extensive testing we have identified more factors that can influence the scan setup in a positive way. The scan method based on [3] uses rigid transformations only. Thus, a user is advised to avoid big movements while scanning. When doing a colorized scan, it is also important that the user's pupils do not move, otherwise the eyes could be blurred in the final bust. If the facial expression changes during scanning, this could lead to bad results as well.

Another important property to the setup is to have enough space. The user has to ensure that no other object is without the scan volume except the user itself. Otherwise, sensor tracking might fail or the result might contain artefacts.

When performing a colorized scan, it is important to have constant lighting conditions around the scan setup. The best case of lightning is diffuse light that is typically used in changing rooms or at professional photographers. This kind of light will generate a homogenous lighting without much shadow which is required to have the captured colors appear natural. In practice, we found that closed curtains and indoor light with different light sources – like in most offices – generate good results.

Another optimization for bust generation concerns camera position and user position. The camera should be at about eye level and look horizontally, see Fig. 1. The user has to be in a standing or sitting position in front of the camera. This is required to generate an upright model.

Fig. 3 shows a scan and a printed model of a ReconstructMe member.



Fig. 4 Left, 3D bust model. Right, printed 3D bust.

The right picture of Fig. 3 shows the 3D print which was printed by an online service. Since the number of 3D printing services is grow, the models can nowadays be printed with several different materials and in color.

5. Conclusion

Thanks to our method, generating a 3D bust has become a very simple task. A tailored and automated post processing routing enables intuitive and robust generation of 3D busts. Our tests showed best practices which can be applied to the setup to improve the results.

We hope to bring 3D technologies to a wider range of users as it is a powerful technology with high potential. Luckily, 3D sensors have become so cheap that they are affordable for any interested individual.

In future, our system can be extended to capture full body shapes instead of scanning only busts. The colorization pipeline can be improved by using texture mapping strategies instead of vertex colorization.

Using texture mapping will enable us to capture high quality images instead of low resolution images provided by the current sensors.

Finally, to deal with movements of the human body, non-rigid alignment and loop closure strategies could be used to improve the results as well.

6. References

- [1] Y. Cui and D. Stricker, "3D Body Scanning With One Kinect," in *2nd International Conference on 3D Body Scanning Technologies*, Lugano, Switzerland, 2011.
- [2] Y. Cui, W. Chang, T. Nöll and D. Stricker, "KinectAvatar: Fully Automatic Body Capture Using a Single Kinect," in *Proceedings of the 11th International Conference on Computer Vision - Volume 2*, Daejeon, Korea, 2013.
- [3] C. Kopf, C. Heindl, M. Rooker, H. Bauer and A. Pichler, "A Portable, Low-Cost 3D Body Scanning System," in *4th International Conference on 3D Body Scanning Technologies*, Long Beach CA, 2013.
- [4] A. Weiss, D. A. Hirshberg and M. J. Black, "Home 3D body scans from noisy image and range data," in *Int. Conf. on Computer Vision (ICCV)*, Barcelona, Spain, 2011.
- [5] C. Yan, S. Schuon, S. Thrun, D. Stricker and C. Theobalt, "Algorithms for 3D Shape Scanning with a Depth Camera," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 5, pp. 1039-1050, 2013.
- [6] H. Li, E. Vouga, A. Gudym, L. Luo, J. T. Barron and G. Gusev, "3D Self-portraits," *ACM Trans. Graph.*, vol. 32, no. 11, pp. 187:1--187:9, 2013.
- [7] B. Kainz, S. Hauswiesner, G. Reitmayr, M. Steinberger, R. Grasset, L. Gruber, E. Veas, D. Kalkofen, H. Seichter and D. Schmalstieg, "OmniKinect: Real-time Dense Volumetric Data Acquisition and Applications," in *18th ACM Symposium on Virtual Reality Software and Technology*, Toronto, Canada, 2012.
- [8] J. Tong, J. Zhou, L. Liu, Z. Pan and H. Yan, "Scanning 3D Full Human Bodies Using Kinects," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 4, pp. 643--650, 2012.
- [9] M. Kazhdan, M. Bolitho and H. Hoppe, "Poisson Surface Reconstruction," in *Fourth Eurographics Symposium on Geometry Processing*, Cagliari, Sardinia, Italy, 2006.
- [10] "ReconstructMe SDK Online Documentation," Profactor GmbH, [Online]. Available: <http://reconstructme.net/wp-content/uploads/ReconstructMe/doc/index.html>. [Accessed 30 July 2014].