Design of an Automated Extraoral Photogrammetry 3D Scanner

Musa Alyaman[1], Alaa Abd-Raheem[2], and Farah AlDeiri[3]
[1]Department of Mechatronics Engineering, the University of Jordan, Jordan
[2]Nanolab, German Jordanian University, Jordan,
[3]Levant Engineering Company, Jordan

Abstract: *Dental anatomy is a field of anatomy dedicated to the study of tooth structure, it uses 3D physical model for teeth and jaw. Creating a computer aided design model from an existing teeth or jaw is called 3D scanning. This can be accomplished using different techniques like: LASER, camera, and many advanced optical techniques. In this work, Photogrammetry will be used, it is based on camera scanning technique. The input is a set of photographs taken by a camera from a predefined positions and orientations, and the output is a 3D model of a real-world teeth or jaw. Structure from Motion algorithm is used to 3D reconstruction from 2D images, it produces the point cloud and eventually the complete textured mesh. Our system is closed, where it controls scanning conditions in terms of lighting and angles of captured images.*

Keywords: *3D scanning, photogrammetry, Structure from Motion, 3D reconstruction, dental anatomy, teeth impression.*

Received September 29, 2018; accepted January 23, 2019

1. Introduction

Dental impression is an imprint of hard teeth and soft tissues in the mouth, in which to form re-productive teeth or cast. It is a container which is designed to roughly fit over the dental trays is used to make an impression. The materials that are used in this process are designed to be liquid or semi-solid when mixed and placed in the tray, after few minutes this material will be solidified; leaving an imprint of the structures in the mouth [5].

Nowadays, there is a trend to increase the use of scanners in dental operation. Digital teeth impressions can be produced using a scanner and by creating point cloud that can be reconstructed using specific software, it will appear as a 3D- digital impression [15]. Digitalized impressions process are significantly advantageous compared to traditional counter-parts[15, 17] such as; easy to store the patient's impressions, efficiency, productivity, and accuracy [17].

The introduction of Computer-Aided Design/Computer-Aided Manufacturing (CAD/CAM) concepts into dental applications was first presented by Dr. François Duret in his thesis in 1973. Whereas the emergence of the first digital intraoral scanner for restorative dentistry was in the 1980s by Dr. Werner Mormann and Eng. Marco Bradenstini [7].

There are many CAD/CAM scanning systems that are constructed based on the traditional impression. This is what is called extraoral scanning. It involves making a plaster cast from the silicone impression before scanning it [9]. Extraoral scanning is used in dental laboratories to digitally save scans of traditionally acquired teeth impressions. Yamany *et al*., [19] suggested a method to extract the three-dimensional teeth model of the patient by designing a system that consists of an intraoral Charge Coupled Device (CCD) camera with white light that is held on a five-link 3D digitizer arm. The 3D model reconstruction process contains three phases: Data Acquisition, Data Processing, and surface reconstruction [1].

3D scanners are devices that create a digital map of the surface of an object and collect data on its three-dimensional shape [6]. Reverse engineering, quality control and dimension measurements are the main concepts behind 3D scanning. 3D scanning is used in industrial applications to check the quality of the final product [20]. There are several other applications as well as intelligent robot control and vehicle obstacle detection [2]. 3D scanning is defined as the process of generating a computer aided design model from an existing physical part [14].

3D scanning using photogrammetry (Imagery based 3D scanning) has emerged as a proven and affordable alternative to LASER scanning. What makes camera scanning unique is the fact that it does not need additional editing regarding holes in the scanned model unlike other methods [13]. Hence if there is an object with spaces, camera and image processing algorithms can be used for accurate recognition.

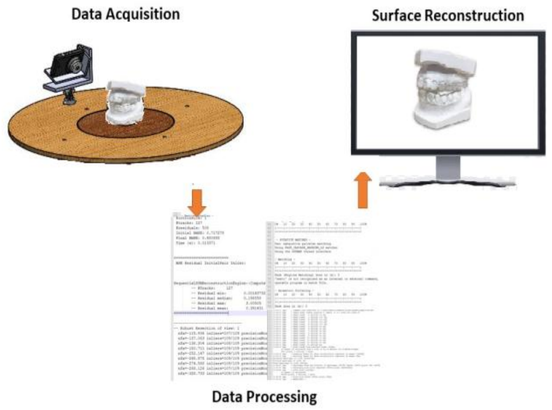


Figure 1. 3D reconstruction process steps.

The 3D model reconstruction process contains three steps; Data Acquisition, Data Processing, and Surface Reconstruction [10] as shown in Figure 1. The data acquisition step provides information of the physical objects. The input to this step depends on the scanning method used in this system the input is camera images. In the data processing step the initial point cloud, extracted from images, is constructed. In the final step the generated point cloud, from previous step, is further processed to reconstruct the surface of the scanned object. This is known as surface reconstruction.

2. Methods

The automated 3D scanner system is based on both hardware and software; the hardware is used in data acquisition step while the software is used in data processing and surface reconstruction as will be discussed further on.

2.1. Data Acquisition

The main hardware components of the 3D scanner are shown in Figure 2. Where a camera is used to capture total of 24 images taken on two elevations. This is achieved using a rotating disc, where the camera and a linear actuator are placed on. The disk is moved 30 degrees at a time around the object, which is placed on a stationary fixed disc in the center. After completing a full rotation (360 degrees), the electrical linear actuator will rise its full stroke length. This process will repeat itself on the second elevation.

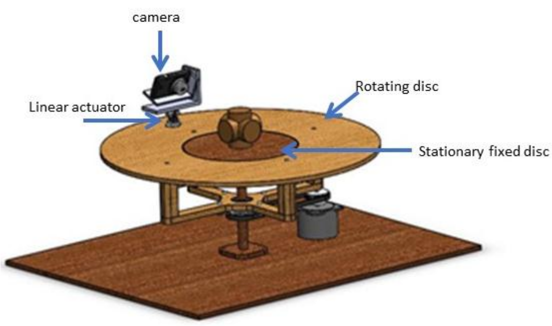


Figure 2. Hardware design.

During the initial scanning tests, many errors took place due to inconvenient light condition as well as the algorithm seeking to extract unwanted features from the background. These errors cause the reconstruction process to be more time consuming. A closed environment for scanning was made as a solution to eliminate errors in the reconstruction process. The time consumed is mainly concerned with the software part. It is defined as the time period between loading images in the algorithm and the final stage where a 3D object is displayed. Furthermore, a closed environment is useful to carefully control the lighting conditions on the object, which can significantly affect the results because shadows can create false meshes from the 3D point cloud.

The idea behind any automated system is to eliminate human intervention in the process. Therefore, an automated system should have controllers to perform the function. The controller used was Arduino Uno which controls all hardware components according to their mentioned sequence of operation, also it is used to interface with the PC. Figure 3 shows the block diagram for hardware components. Firstly, a personal computer is used for running the algorithm on photos for further images processing and it interfaces between the Graphical User Interface (GUI) and Arduino. Secondly, stepper motor driver is for operating the stepper motor which is connected with the rotating disc. Thirdly, relay driver of the linear actuator which allows the camera connected to the linear actuator to have two different elevations. Finally, servo motor is connected to mechanical arm that presses the capture button of the camera every 30 degrees.

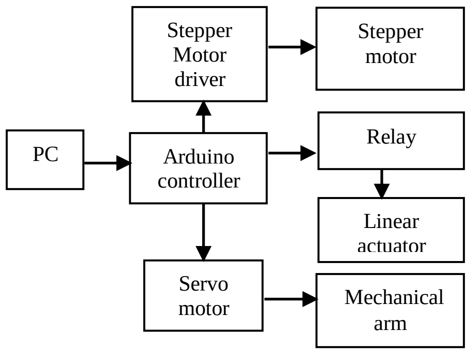


Figure 3. Hardware connections.

Once the data acquisition is done, the user is alerted on the screen to press Wi-Fi button transferring the images to the processor (PC). Then photos will be automatically copied to the required folder location. 3D reconstruction button initiates the second step which is data processing.

2.2. Data Processing

Photogrammetry is the art and science of using overlapping photographs to reconstruct three dimensional scenes or objects. The main algorithm used in photogrammetry is the Structure from Motion (SfM) where camera positions and feature points are integrated in acquiring the final 3D model [18]. The Structure from Motion algorithm is broken down into two parts: Structure which represents the scene; it works by finding the corresponding 3D points from given 2D point matches in two or more images. The second part is Motion, which represents the camera matrices for the views [4]. The views represent the translation and rotation camera matrices.

3D reconstruction in this system works by capturing sequential overlapped images. These images will be read to detect their significant features. These significant features may be edges, blobs or corners. One of the algorithms that are used in extracting features in computer vision is the Scale Invariant Feature Transform (SIFT) [8]. With every two consecutive images taken using a calibrated camera; the features extracted are then matched. Iterations are made to each extracted set of point in order to remove outliers (incorrect matches) using the Random Sample Consensus (RANSAC) algorithm [11]. Figure 4 illustrates the concept of how 2D points are matched and located on the 3D space, where P1 and P2 are sequential images for the same object. p is the estimated reference of the two images.

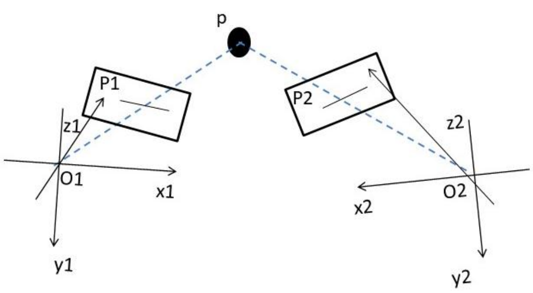


Figure 4. Acquiring 3D point from matched 2D points.

Triangulation is the method of estimating a point's location in 3D space when it is seen from multiple views. To understand how the x, y, z coordinates of the new point are acquired, there are several elements and equations. These equations consist of the camera intrinsic and extrinsic matrices, and the translation and rotation of image views.

In RANSAC algorithm it should determine set of parameters to achieve the concept of it, using line fitting method that should specify a threshold, number of sampled point and outlier ratio to determine the number of trials to repeat the RANSAC algorithm on the same sampled points to eliminate the error, p is the probability usually it is 95%[4].

$$T = \frac{\log(1-p)}{\log(1-(1-e)^s)} \tag{1}$$

In this system the size of each image taken is 5184 x 3888 pixels. Assuming all 24 calibrated images are accepted to the algorithm, a total of 483.83 MPixels (20.16 MPixels per images), it represents the estimated set of point cloud if all pixels were extracted and matched.

Feature detection and extraction is a very important process as it involves reducing the amount of resources required to describe a large set of data. After acquiring local point cloud patches from the 24 images surrounding the object, these groups are all combined together forming the final global point cloud of the object. This is the first step which is known as sparse dense point cloud. It is noticed that in this step that the points representing the 3D model are unevenly distributed and suffer from large gaps among them. The next step will be densification of these points to acquire more points that help in estimating the shape of the object.

In this system, 3D reconstruction was done using Open Multi-View Geometry (OpenMVG) [12], along with another library the OpenMVS (Multi-View Stereovision). As the output of OpenMVG is a dense point cloud; OpenMVS continues to recover the full surface of the scene so that the final output is a textured mesh.

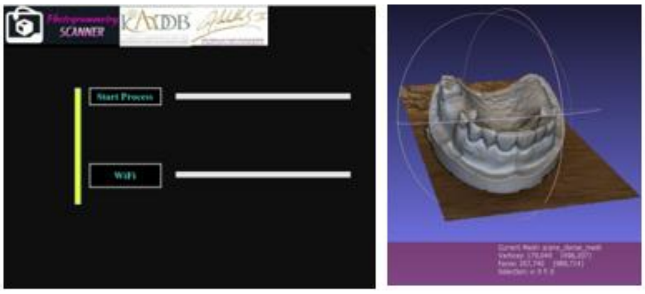
2.3. Surface Reconstruction

The next step is to fit a surface to the generated point cloud. This step is known in the computer vision field as surface fitting. As the point cloud are a set of discrete points in the 3D space, surface fitting algorithms are divided into interpolation and approximation methods. The difference lies in the fact that the surfaces reconstructed in interpolation methods pass through the data points while the surfaces reconstructed in the approximation methods are close to the data points[16]. This collection of points represent vertices, edges and faces that form the three-dimensional object hence forming what is called a mesh.

Finally, all points resampled in 3D space are displayed in color with the help of the point pixel values extracted in step B (Data processing). This process is known as mesh texturing and it makes the 3D model more realistic as it matches its true colors [3].

3. Results

The 3D scanner software consists of single user interface which is user friendly and is less sophisticated as shown in Figure 5-a. It has 2 main buttons the start process and the WIFI. The start process button used to start the data acquisition step to take the 24 images. Wi-Fi button used to transfer the images to the required location and to start the remaining data processing and surface reconstruction steps. After the process is finished, the 3D model is displayed on a separate screen as shown in Figure 5-b.



a) Graphical User Interface b) 3D model.

Figure 5. System software.

The quality of the results depends on the nature of the object, whether the object contains many features or it is shiny object. All resulting models in the system are reconstructed from a fixed number of images (24 images) which are taken from the camera, which through experiments were found to be the least number of images to gain satisfactory reconstruction result. The results are divided into hardware and software; the data acquisition results cover the final hardware of the system while the data processing and surface reconstruction cover the software results.

3.1. Data Acquisiton

The assembly of the system includes; rotating disc holding the camera and the linear actuator, stationary central disc with the object placed on it, and the stepper motor rotating the system (rotating disc). Figure 6-a shows the fabricated model showing the hardware components. To make sure that the algorithm will recognise the object from its background and ensure good distribution of light on the objects in such a way that no shadows affect the quality of images. Therefore, the final system was covered as shown in Figure 6-b.

a) Fabricated hardware model. b) Covered final model.

Figure 6. Final system design.

The outcome of this stage of the system yields sequential 24 images that cover the entire object with a very good overlap ratio between the images. The overlap ratio of these multiple images must as large as possible to assist the algorithm in order to reduce the consuming time required to match the images. Figure 7 shows the total 24 images of a sample object.

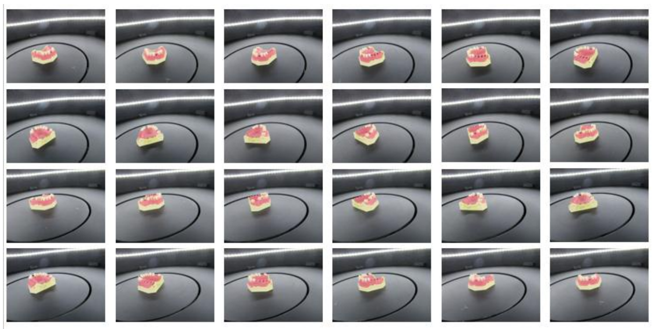


Figure 7. Data acquisition images results.

3.2. Data Processing

The 3D reconstruction accuracy is evaluated depending on the residual error and the number of 3D points. This is all dependent on the key point and matching steps. After each object is processed in the algorithm, the number of calibrated images is counted based on the fact if there are sufficient features or not. Also, to determine the quality of reconstructed 3D model the number of basic sparse point and dense point cloud is counted.Table1 below summarizes the results of three different objects, noting that all the images were taken using the system's camera.

Table 1. Data processing results.

Object	Number of Calibrated Images	Sparse Point Cloud	Dense Point Cloud
	23/24	1598	354331
	24/24	65832	119814
	8/24	563	3809
	10/24	213	1548

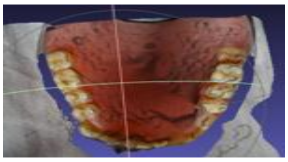
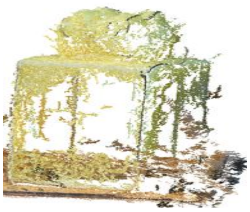
The teeth surface is not modeled as a smooth surface. It has a lot of details that are important such as depressions and elevations. Extracting these features depends on the teeth impression materials that is made of. In reconstruction process the number of accepted images does not always equal the number of images captured due to some factors. These factors are based on the object features such as color, shininess and the distinction of its features. In general, objects to be scanned should not contain transparent and reflective surfaces as these objects result in high rejection ratio of input images and false meshes because the features are wrongly interpreted [11].

3.3. Surface Reconstruction

Since the main objective of the system is converting 2D images into a 3D model. The table below shows different teeth impression models with different features (e.g., size, color) used in education purposes. It shows the final 3D model of an object compared to its image. The first model was used for the purpose of testing how the algorithm works and recognize whether there is a tooth or not. The second object represents half of a jaw, the result was rich in features.

The third object was a shiny object where the mesh was not smooth and filled. The last object is a soap tooth model which is used by first year dental students. The problem with this object was the lack of features and no contrast in color.

Table 2. Results of 3D models.

Image of Object	3D Reconstruction Model
	
	
	
	

In this part, the time consumed in the 3D scanning process will be discussed in more detail. A comparison between the four mentioned objects is held in terms of time consumption and result analysis stages which contain the three steps; data acquisition, data processing and surface reconstruction, as shown in Table 3.

Table 3. Time consumption comparison in data processing step.

Image of Object	Time consumption (hours)
	2:45
	1:49
	1:20
	00:45

As a qualification part to our scanning models, 3D printing of one model was done to visualize the measurement of the resultant 3D model. Figure 8 shows the 3d printed model.

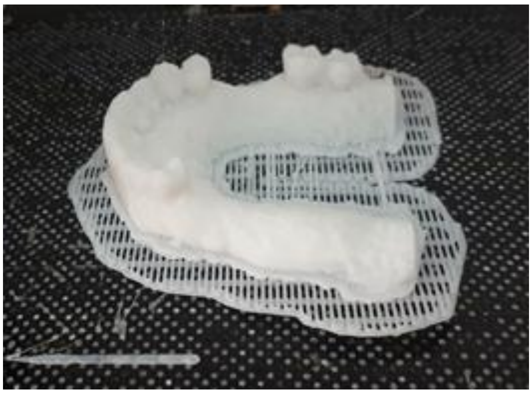


Figure 8. The resultant 3D printed model.

In general, each step of reconstruction process consumed different amount of time depends on number of iterations as in the two first steps; spares point cloud and the dense point cloud most the time sense point cloud consumed more time. As of the surface reconstruction, the mesh refinement step consumed more time (usually exceed at least one hour) to fill the holes and smooth the surface.

4. Conclusions

This paper discusses the design of a fully automated imagery-based 3D scanning system where images are captured at specific angles and elevations surrounding the model from all different views. This was done to ensure good percentage of overlap between images for the feature extraction and matching steps. Moreover, features from the background scene were excluded by designing a closed system and to also ensure proper distribution of light. The system experienced some constraints in some areas regarding the quality of the scanned object such as shiny, transparent and featureless object. It is suggested to be further improved ahead so that feature light is projected on such objects in order to add features that can be extracted later as points to improve the quality of the final mesh.

Acknowledgement

The authors would like to express thanks to King Abdullah Design Development Bureau (KADDB) for their support and supervision and also King Abdullah Fund for Development Research and Innovation Support Program, for financial and logistics support.

References

[1] Ahmed M., Yamany S., Hemayed E., Ahmed S., Roberts S., and Farag A., "3D Reconstruction of the Human Jaw from a Sequence of Images," *in Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Juan, pp. 646-653, 1997.

[2] Chen F., Brown G., and Song M., "Overview of 3-D Shape Measurement Using Optical Methods," *Optical Engineering*, vol. 39, no. 1, pp. 10-22, 2000.

[3] Chen Z.,"Robust Reconstruction of Textured Surfaces From 3D Point Clouds," Thises, National University of Singapore, 2013.

[4] Förstner W. and Wrobel F., *Photogrammetric Computer Vision Statistics, Geometry, Orientation and Reconstruction*, Springer, 2016.

[5] Glenner R., http://www.glidewelldental.com/chairside-magazine/volume9-1/history-dental-impressions.aspx, Last Visited 2018.

[6] Ireland A., McNamara C., Clover M., House K., Wenger N., Barbour M., Alemzadeh K., Zhang L., and Sandy J., "3D Surface Imaging in Dentistry-What We are Looking at," *British Dental Journal*, vol. 205, no. 7, pp. 387-392, 2008.

[7] Logozzo S., Franceschini G., Kilpela A., Caponi M., Governi L., and Blois L., "A Comparative Analysis of Intraoral 3D Digital Scanners for Restorative Dentistry," *The Internet Journal of Medical Technology*, vol. 5, no. 1, pp. 1-18, 2011.

[8] Lowe D., "Object Recognition from Local Scale-Invariant Features," *in Proceedings of the 7th IEEE International Conference on Computer Vision*, Kerkyra, pp. 1150-1157, 1999.

[9] Luthardt R., Weber A., Rudolph H., Schone C., Quaas S., and Walter M., "Design and Production of Dental Prosthetic Restorations: Basic Research on Dental CAD/CAM Technology," *Intenational Journal of Computerized Dentistry*, vol. 5, no. 2-3, pp. 165-176, 2002.

[10] Moulon P., Monasse P., and Marlet M., "Global Fusion of Relative Motions for Robust, Accurate and Scalable Structure from Motion," *in Proceedings of IEEE International Conference on Computer Vision*, Sydny, pp. 3248-3255, 2013.

[11] Moulon P., Monasse P., and Marlet R., "Adaptive Structure from Motion with a Contrario Model Estimation," *in Proceedings of Asian Conference on Computer Vision*, Daejeon, pp. 257-270, 2012.

[12] Moulon P., Monasse P., Perrot R., and Marlet R., "OpenMVG: Open Multiple View Geometry," *in Proceedings of Reproducible Research in Pattern Recognition*, Cancun, pp. 60-74, 2017.

[13] Pawar S., Mithaiwala H., Gupta A., and Jain S., "Review Paper on Design of 3D Scanner," *in Proceedings of International Conference on Innovative Mechanisms for Industry Applications*, Bangalore, pp. 650-652, 2017.

[14] Peterson E., Klein M., and Stewart R., "Whitepaper on Structure from Motion (SfM) Photogrammetry: Constructing Three Dimensional Models from Photography," Technical Report, Research and Development Office Science and Technology Program Final Report ST-2015-3835-1, 2015.

[15] Rondon N., http://www.yourdentistryguide.com/digital-impressions, Last Visited 2018.

[16] Söderkvist I., "Introductory Overview of Surface Reconstruction Methods," Technical Report, 1999.

[17] USA S., http://www.sundentallabs.com/digital-impressions-versus-traditional-impressions/, Last Visited 2018.

[18] Westoby M., Brasington J., Glasser N., Hambery M., and Reynolds J., "Structure-from-Motion photogrammetry: A Low-Cost, Effective Tool for Geoscience Applications," *Geomorphology*, vol. 179, pp. 300-314, 2012.

[19] Yamany S., Farag A., Tasman D., and Farman A.,"A 3-D Reconstruction System for the Human Jaw Using a Sequence of Optical Images," *IEEE Transactions on Medical Imaging*, vol. 19, no. 5, pp. 538-547, 2000.

[20] Yao A., "Applications of 3D Scanning and Reverse Engineering Techniques for Quality Control of Quick Response Products," *The International Journal of Advanced Manufacturing Technology*, vol. 26, no.11-12, pp. 1284-1288, 2005.

Musa AlYaman received the B.Eng. degree in computer engineering from University of Jordan, in 2004, the M.Eng. degree in Industrial Automation from Yarmouk University, in 2006, and the PhD in Image processing and Artificial Intelligence at Newcastle University, in 2015. He is currently an assistant professor and head of the department of Mechatronics Engineering at the University of Jordan. He has published several scientific papers in high impact factor and archival technical journals and conferences. His research interests include Embedded Systems Applications in Biomedical field, Industrial Automation, Intelligence Transportation Systems**.**

Alaa Abd-Raheem received the B.Eng. degree in mechatronics engineering from University of Jordan, in 2018. She currently works as a research assistant at the German Jordanian University. Shehas published one scientific paper in high impact factor journal in elevator traffic design. Her research interests include Biomedical field, Industrial Automation, Deep Learning, Image Processing and Drives.

Farah AlDeiri received the B.Eng. degree in mechatronics engineering from University of Jordan, in 2018. She currently works in the industrial automation field as well as in STEM education. Her research interests include Image Processing, Industrial Automation, Drives and Networking.