Role of Background Subtraction in Creating Human Body Point Clouds from Photographs

Dinu DRAGAN*[1], Srdan MIHIC[2], Zoran ANISIC[1], Ivan LUKOVIC[1]
[1] Faculty of Technical Sciences, University of Novi Sad, Serbia;
[2] DOOB Innovation Studio DOO, Novi Sad, Serbia

Abstract

In this paper we research the influence of background subtraction on photogrammetry pipeline when creating 3D print ready human body data. Background subtraction is a technique in image processing where image background is removed from the image and only foreground is left for further processing. The goal of the paper is to assess whether background subtraction could influence positively or negatively the photogrammetric processing of photographs. The research is aimed at the freely available software that natively does not support background subtraction, but also does not forbid the use of background subtraction. We aim to find out whether the software could benefit from adding background subtraction algorithms into their processing pipelines.

Keywords: Image Background Subtraction, Photogrammetry Pipeline, Human Figurine Production, Photogrammetry Software, Human Body Point Cloud Quality Evaluation

1. Introduction

From our experience, processing of the entire photogrammetry pipeline—which results in 3D print ready human body model—can be time consuming and could last for several hours. Anything that could speed up the production of 3D human body model from a set of co-related photographs is needed and welcomed. The photogrammetric processing of the co-related photographs results in a cloud of 3D points. Many points in a 3D point cloud are not needed and should be removed [1]. Most of these points do not belong to the scanned human body; they are extracted from the background of the photographs. Points belonging to the background are usually removed manually [1]. Some tutorials on photogrammetry, such as [2], suggest removing the background from the photographs. Image processing technique where image background is removed from the image and only foreground is left for further processing is called background subtraction.

In this paper we explore background subtraction influence on the photogrammetry pipeline in creating 3D print ready human body model. Background subtraction is employed in different ways in software that performs photogrammetric processing of photos and creates 3D human body data. In some software, background subtraction is not included directly in the photogrammetry pipeline, but it is also not forbidden, such as Insight3D [3] and VisualSFM [4]. In others background subtraction is an integral part of the photogrammetry pipeline, such as Agisoft PhotoScan [5]. There is also photogrammetry software that suggests keeping the background subtraction entirely as image processing is based on the image background as much as on the image foreground, such as Autodesk Memento [6].

The goal of the paper is to investigate whether background subtraction could improve the photogrammetry pipelines of the freely available photogrammetry software that natively does not include background subtraction and that does not explicitly prohibit the use of background subtraction. Two photogrammetry software environments were chosen: Insight3D and VisualSFM. Both are free software. During the evaluation of the background subtraction influence over the photogrammetry pipeline, we had issues in processing human body photographs using Insight3d. Thus, the evaluation also helped in finding out whether these programs are suitable for processing human body co-related photographs.

Two methods are used to assess how background subtraction could influence the photogrammetry pipeline. The methods are chosen based on our experience in human figurine production. The first method is based on execution time—time needed to create point cloud with or without background. As stated before, time is essential for the human figurine production, because production could last for couple of hours, and anything that speeds up the process is always welcomed. The second method is empirical quality evaluation of created point clouds with or without background subtraction.

* dinud@uns.ac.rs, d.dragan@doobinnovation.com; +381-21-485-2421

For 3D printing of human figurine, based on previous experience, we devised a minimal set of characteristics that a point cloud should possess: a point cloud should be dense enough, there should be enough points in each part of the body, and points in head region should be of sufficient quality and density.

We have chosen ten typical scans from our daily production [7]. This set included both subjects well dressed for the photogrammetry processing—they were dressed in cloths with good texture, and subjects that were not dressed suitable for the photogrammetry processing—they had white cloths that did not have good texture. We did not test different algorithms for automatic image background subtraction, but only the influence of background subtraction on the photogrammetry pipeline. Therefore, all the photographs used in test where manually preprocessed—their background subtracted—before they were introduced into the photogrammetry pipeline. Choosing the adequate algorithm for automatic image background subtraction for the photogrammetry pipeline is different topic altogether which we plan to address in the future.

The organization of the paper is as follows. The background on the photogrammetry pipeline, the image background subtraction, and the potential role of background subtraction in the photogrammetry pipeline is described in Section 2. In section 3 we describe human body scanning environment, photogrammetry software that we used for testing, and the methods used to evaluate background subtraction influence on the photogrammetry pipeline. Testing results are described and discussed in Section 4. Section 5 concludes the paper.

2. Background

The field of human body scanning/digitalization is a mature one [8] and it is constantly improving. There are complete systems for the production of 3D human body digital models [8]. The systems that were previously dominated by large enterprises and academic research are becoming more and more available and they are gaining in popularity [9] due to lowering cost of the equipment and appropriate software. Different technologies are employed for 3D human body scanning [9] that could be broadly categorized in two distinct groups:

- Laser based scanning technologies which require purposely built scanning environment (laser emitters and collectors), they require more time to collect scanning data, but they demonstrate high accuracy and lesser need for the post-processing of the scanned data.
- "White light" scanning technologies which are less accurate and more process-demanding, but scanned data is collected faster (in some scanning facilities data are acquired in less than a second) and scanning facility can be build fast using low-cost equipment.

The "white light" scanning technologies are based on photogrammetry [10]. Usually the photogrammetric pipeline showed in Fig. 1 is applied to create a digitalized human body model [1].

Fig. 1. The photogrammetry pipeline in human model production.

A scanning facility is built using a dozen cameras or several cameras with depth sensors, such as MS Kinect [11]. Systems based on a single camera moved around the scanning object also exist [12]. No matter how the "white light" scanning system is built, the initial scanning process usually results in a set of co-related photographs containing the scanning object taken from different angles.

The multi-camera photograph sets are then processed using software for point cloud generation [13] most of which are based on structure from motion (SFM) approach. SFM is used to extract 3D information from 2D images. This usually includes feature extraction and matching based on SIFT descriptors, external photographs parameters bundle adjustment, orientation and calibration of various photographs [14]. The entire process ends with photograph matching which results in a dense cloud of 3D points.

The generated point cloud is further processed to eliminate 3D points that are not of interest such as background points, outliners, and noise (which are result of previous steps in the photogrammetry pipeline). Also, duplicates are eliminated, different point clouds are combined, etc [1]. After the point cloud is processed, the scanned human body mesh is created by generating polygon surfaces base on the previously generated point cloud. The process of mesh generating is error prone and requires further editing, mainly filling holes and removing spikes, which are followed by smoothing and data

reduction and structuring [1]. At the end of the photogrammetry pipeline texture is mapped onto the human body mesh to add color to the mesh and to increase its realism.

The entire production of human body models is usually not covered using a single software product [2], and most often several software products are used through the entire production. Easily imaginable scenario [2] is using three software products: for photographs acquisition, for mesh creation, and for mesh editing.

As stated in the introduction, points that are generated from the photographs background are removed manually from the point cloud. One way to speed up the photogrammetry pipeline is to eliminate the removal of these points from the pipeline by not including them in the beginning—meaning that the background should be subtracted from the starting photograph. However, subtraction of the background could influence the quality of the resulting model, because the data used to build the dense point cloud is now reduced. Including background subtraction into the photogrammetry pipeline will be useful only if the quality of the resulting model is not reduced.

Also, it is necessary to take into the consideration the time needed to remove the background from all the photographs in the scan. If the time needed to complete the production—with and without background subtraction—is similar, than it is not justified to use background subtraction. Of course, all of this should be concluded in the context of the software tested.

3. Scanning environment, photogrammetry software, and evaluation methods

Human body scans have been acquired using DOOB-LICATOR [15], a room size scanning booth containing 54 DSLRs cameras arranged in nine columns of six cameras each [15], Fig. 2. DOOB GROUP AG is a German based global provider of 3D technology products and services [16]. DOOB™ scanning system and 3D print figurine stores are already present across Europe, USA, Japan, and Australia [7]. Applications of DOOB™ 3D technology include the creation of digital avatars and the mass customization of consumer products [7].

Fig. 2. DOOB-LICATOR – 3D scanning booth used for the test photograph sets acquisition
(with permission of DOOB GROUP AG [7]).

For the testing purposes, we have chosen 10 typical scans from our daily production. We included scans containing:
- adults,
- children,
- mannequins.

Due to customer contract restrictions, we cannot show the scanned photographs of customers. An example of the photographs is shown in Fig. 3 which contains one of the photographs from the mannequin scan. We have chosen both the scans suitable for the photogrammetry pipeline—that have good texture and a lot of details, and the scans that are less suitable for photogrammetry pipeline as they have uniformly colored cloths with less details. As it can be observed from the Fig. 3, the background in photographs does not have many details and as such it does not contribute too much to the quality of the resulting camera alignment and the resulting point clouds.

We have chosen Insight3d and VisualSFM software to test the influence of background subtraction on the photogrammetry pipeline. Both are freely available. However, Insight3d is mostly recommended for 3D reconstruction of scenes containing buildings, while VisualSFM is more a generic 3D reconstruction software. Both are based on SFM. VisualSFM exploits multicore parallelism for feature detection, feature matching, and bundle adjustment [4].

Based on our experience in 3D human figurine production, we chose two methods to assess background subtraction influence over the photogrammetry pipeline:

1. Execution time – we measured the time that is needed to align the cameras and the time needed to generate dense point cloud.
2. Point cloud empirical quality evaluation – we evaluated the density of the entire point cloud, whether there had been enough points in each part of the body, and we evaluated the point density in the head region and the quality of the points in the head region—the visibility of the customer's face. Based on our experience in working with scans for 3D human figurine production, we devised a numerical scale that describes the dense point cloud feature which will enable figurine printing of the best quality—where 0 is not usable, 1 is bad, 2 is useful, 3 is very good. Quality of the point cloud was evaluated by our expert staff and grades are reached in consensus.

Fig. 3. An example of the photographs from the testing scan – mannequin scan
(with permission of DOOB GROUP AG [7]).

Tests were conducted on two computer configurations:

- Low-end computer configuration: Intel T6400 2GHz with 2 cores, 4GB of RAM and GeForce 9600M GT GPU.
- High-end computer configuration: i7 3.6GHz with 8 cores, 32 GB of RAM and GeForce GTX 750 Ti GPU.

In addition to background subtraction influence on the photogrammetry pipeline, we wanted to test how computer configuration influences the results. We expected that processing will be faster on the high-end computer, but that there would be no changes in perceived quality of the point clouds.

4. Results and discussion

We tested the Insight3d first. However we got very bad results on both low-end and on high-end computer. None of the point clouds generated have been useful for human figurine production. The point clouds did not contain enough points to generate meaningful and useful mesh. The most useful point cloud was generated for scan no. 3 which had the best texture, Fig. 4. As it can be observed from Fig. 4 the point cloud is not dense enough and human form cannot be observed in it.

We got those results from Insight3d for two reasons:

- Insight3d is mostly used for 3D reconstruction of scenes containing buildings and not for scenes containing humans.
- To achieve the best results Insight3d requires that scans have far better coverage of the scene than we achieve with 54 cameras. It would probably generate point clouds of better quality if we had bigger number of overlapped photographs of the subjects.

Based on the test scans employed we concluded that Insight3d is not suitable for the photogrammetry pipeline employed in production of the 3D printable human model, because it did not resulted in usable 3D human models.
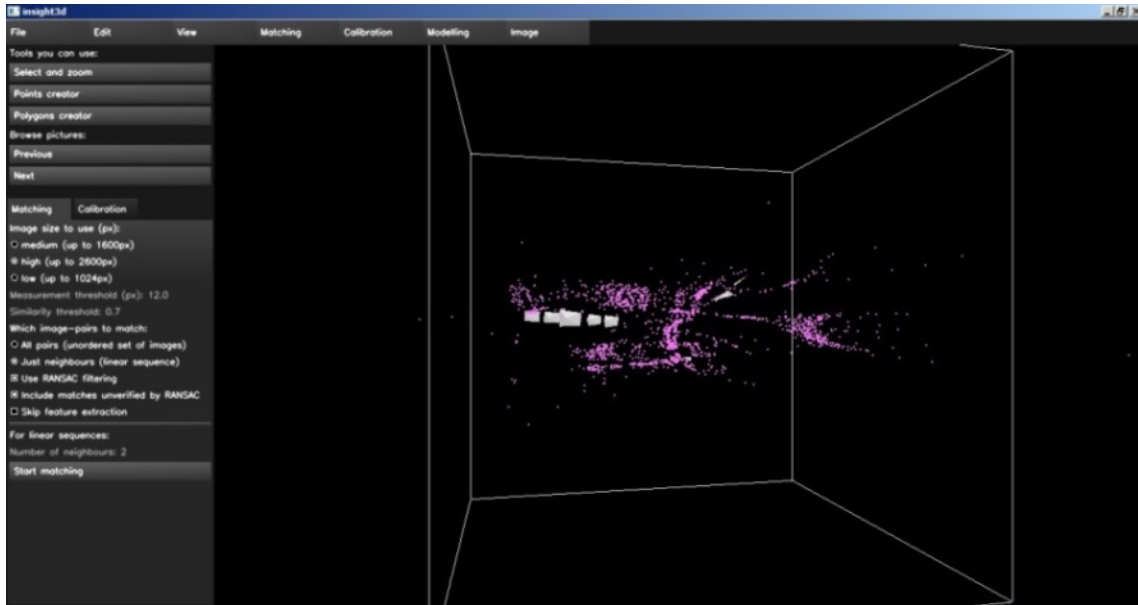
Fig. 4. An example of one of the most useful point clouds generated from the test scan using Insight3d.

VisualSFM achieved far better results in regarding to the photogrammetry pipeline employed in human figurine production. For the scans with subjects dressed in good textured cloth, all of the cameras were aligned right and density of the point clouds was at least useful. An example of a point cloud generated using VisualSFM—that is good for the 3D printing of the human figurine—is shown in Fig. 5.
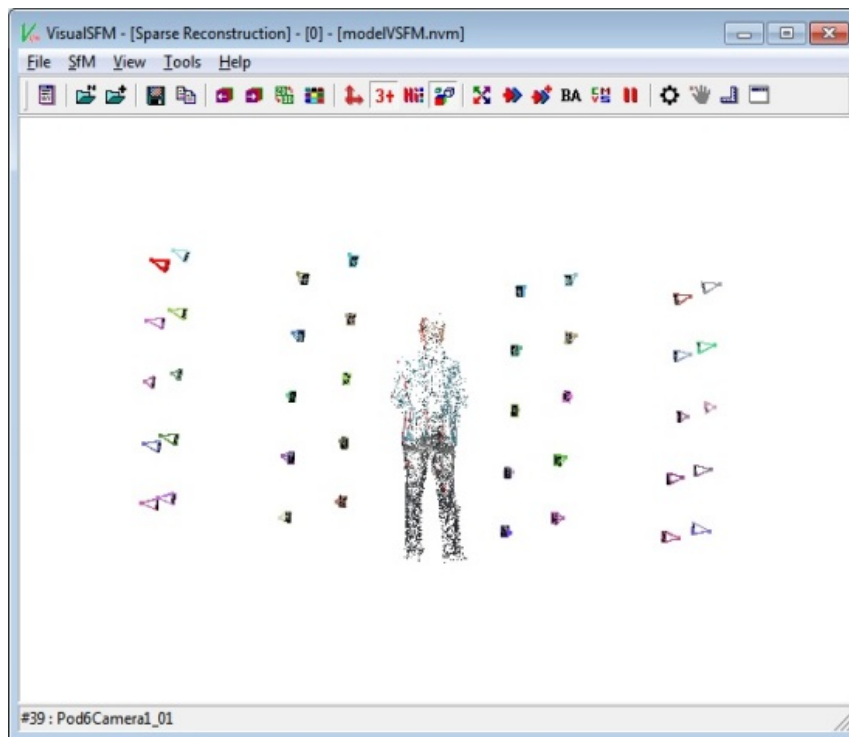

Fig. 5. An example of one of the better point clouds generated from the test scan using VisualSFM and background subtraction.

The VisualSFM test results on the low-end and the high-end computers are listed in Table 1 and 2.

Table 1. The VisualSFM results on low-end computer configuration.

Scan		Time (s)			Point cloud empirical quality evaluation (0 – 3)				
		Camera alignment	Point cloud generation	Overall	Density	Point coverage	Head region density	Head region quality	Average
1.	**BS**	**253**	**47**	**300**	**3**	**3**	**2**	**2**	**2.5**
	original	**254**	**72**	**326**	**3**	**3**	**3**	**3**	**3.0**
2.	BS	331	63	394	2	1	1	1	1.3
	original	245	90	335	1	1	1	1	1.0
3.	**BS**	**208**	**59**	**267**	**3**	**3**	**3**	**3**	**3.0**
	original	**319**	**53**	**372**	**3**	**3**	**3**	**3**	**3.0**
4.	BS	261	49	310	1	2	2	2	1.8
	original	160	93	253	1	2	2	2	1.8
5.	BS	47	12	59	2	2	1	1	1.5
	original	38	19	57	2	2	1	1	1.5
6.	BS	128	53	181	2	2	2	2	2.0
	original	208	81	289	2	2	2	2	2.0
7.	BS	46	13	59	2	2	1	1	1.5
	original	172	68	240	2	2	1	1	1.5
8.	**BS**	**102**	**47**	**149**	**2**	**2**	**3**	**3**	**2.5**
	original	**132**	**48**	**180**	**2**	**2**	**2**	**2**	**2.0**
9.	BS	47	62	109	2	2	2	2	2.0
	original	62	65	127	2	2	2	2	2.0
10.	BS	61	67	128	2	2	2	2	2.0
	original	63	76	139	2	2	2	2	2.0
Av.	BS	148.4	47.2	195.6	2.1	2.1	1.9	1.9	2.0
	original	165.3	66.5	231.8	2.0	2.1	1.9	1.9	2.0

Table 2. The VisualSFM results on high-end computer configuration.

Scan		Time (s)			Point cloud empirical quality evaluation (0 – 3)				
		Camera alignment	Point cloud generation	Overall	Density	Point coverage	Head region density	Head region quality	Average
1.	**BS**	**14**	**11**	**25**	**3**	**3**	**2**	**2**	**2.5**
	original	**17**	**18**	**35**	**3**	**3**	**3**	**3**	**3.0**
2.	BS	15	40	55	2	1	1	1	1.3
	original	22	28	50	1	1	1	1	1.0
3.	**BS**	**20**	**11**	**31**	**3**	**3**	**3**	**3**	**3.0**
	original	**19**	**11**	**30**	**3**	**3**	**3**	**3**	**3.0**
4.	BS	13	18	31	1	2	2	2	1.8
	original	25	19	44	1	2	2	2	1.8
5.	BS	6	6	12	2	2	1	1	1.5
	original	9	6	15	2	2	1	1	1.5
6.	BS	19	14	33	2	2	2	2	2.0
	original	19	18	37	2	2	2	2	2.0
7.	BS	6	4	10	2	2	1	1	1.5
	original	1	12	13	2	2	1	1	1.5
8.	**BS**	**14**	**12**	**26**	**2**	**2**	**3**	**3**	**2.5**
	original	**21**	**14**	**35**	**2**	**2**	**2**	**2**	**2.0**
9.	BS	7	14	21	2	2	2	2	2.0
	original	11	14	25	2	2	2	2	2.0
10.	BS	9	19	28	2	2	2	2	2.0
	original	10	15	25	2	2	2	2	2.0
Av.	BS	12.3	14.9	27.2	2.1	2.1	1.9	1.9	2.0
	original	15.4	15.5	30.9	2	2.1	1.9	1.9	2.0

As it can be observed from Tables 1 and 2, in average VisualSFM achieved the same quality of the generated point clouds regardless of presence of background subtraction in the photogrammetry pipeline. There are some fluctuations in the results—sometimes the results are better with background subtraction, sometimes not, and sometimes the results were the same. Scans with good texture achieved good results regardless of background subtraction presence. For difficult scans, background subtraction helped camera alignment process putting focus on the subject.

Use of background subtraction accelerated the photogrammetry pipeline by 14%, Fig. 6a and 6b, and also there had been far less points to process in the Pont cloud processing phase of the photogrammetry pipeline when background subtraction was used.
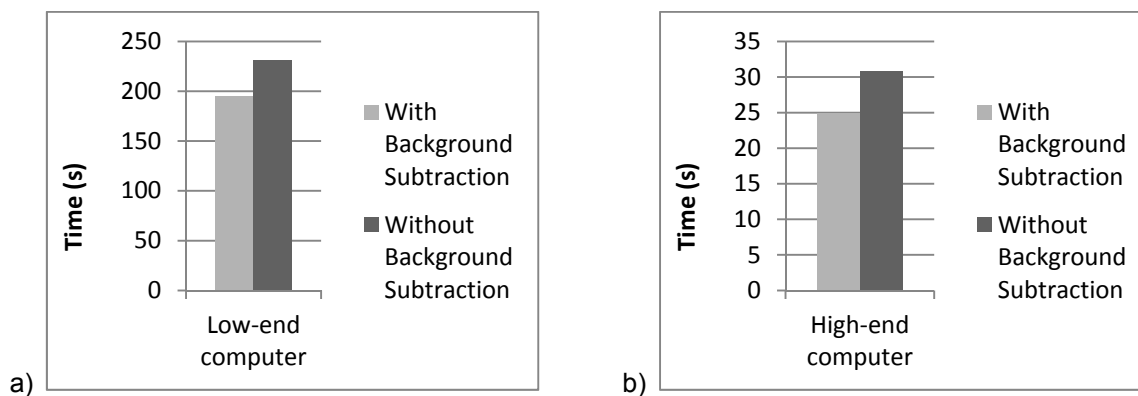

Fig. 6. Average time of the VisualSFM photogrammetry pipeline processing with/without background subtraction; a) Low-end computer, b) High-end computer.

We can conclude that—with VisualSFM—there is a place for background subtraction in the photogrammetry pipeline of the human model production. Whether there is a background subtraction algorithm that can be efficient and still support the time improvement is completely different question which we plan to address in the future.

5. Conclusion

Our testing of background subtraction role in the photogrammetry pipeline using Insight3d and VisualSFM on human scans led to several conclusions. Insight3d was not suitable for the photogrammetry pipeline and the current scanning system we employ in human figurine production.

VisualSFM is suitable for the scanning system as all of the scans tested proved usable for human figurine production. However, we did not test the scans in latter stages of production such as surface generation and editing and texture mapping. From our experience we could say that the resulting point clouds could be used for the latter stages of production and human figurine 3D printing.

Results presented in the paper demonstrate that background subtraction could be used as a part of the photogrammetry pipeline, at least for VisualSFM. If the background of the entire scan could be subtracted in reasonable time—couple of seconds, this would speed up the processing of the entire pipeline and lead to a faster human figurine production. We did not observe that background subtraction had negative impact on the quality of the generated point clouds. Even more, in some cases—3 scans—background subtraction increased the perceived overall quality of the point clouds.

It is not an easy task to find the background subtraction algorithm which is suitable for the photogrammetry pipeline both in quality and speed. In the future we plan to test different algorithms for automatic background subtraction for the photogrammetry pipeline in human figurine production. We also plan to increase the number of scans we use in test, and to test additional photogrammetry software, both free and licensed.

Acknowledgment

6. References

[1] N. D'Apuzzo, "Photogrammetry and Machine Vision - 3 Point cloud processing, surface generation, texturing", in *Photogrammetry and Machine Vision Class Slides*, http://www.igp.ethz.ch/photogrammetry/education/lehrveranstaltungen/photogrammetryandmachi nevision/courematerial/PMV03.pdf accessed 2015.

[2] B. Blizard, "The Art of Photogrammetry: Introduction to Software and Hardware", *TESTED.com*, http://www.tested.com/art/makers/460057-tested-dark-art-photogrammetry/ accessed 2015.

[3] L. Mach, "Insight3d", *Opensource image based 3D modeling software*, http://insight3d.sourceforge.net/ accessed 2015.

[4] C. Wu, "VisualSFM: A Visual Structure from Motion System", http://ccwu.me/vsfm/, accessed 2015.

[5] Agisoft PhotoScan, http://www.agisoft.com, accessed 2015.

[6] Autodesk Memento Beta, https://memento.autodesk.com/about, accessed 2015.

[7] DOOB™, http://www.doobgroup.com/en/discover/discover/, accessed 2015.

[8] N. D'Apuzzo, "Recent Advances In 3d Full Body Scanning With Applications To Fashion And Apparel", In Gruen, A., Kahmen, H. (Eds.), Optical 3-D Measurement Techniques IX, Vienna, Austria, 2009.

[9] J. Straub and K. Scott, "Development of a large, low-cost, instant 3d scanner", Technologies Vol.2, No.2, 2014, pp.76-95.

[10] E. Gattet, J. Devogelaere, R. Raffin, L. Bergerot, M. Daniel, Ph Jockey, and L. De Luca, "A versatile and low-cost 3D acquisition and processing pipeline for collecting mass of archaeological findings on the field", ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences ,1Vol., 2015, pp.299-305.

[11] E.E. Hitomi, J.V. da Silva, and G.C. Ruppert, "3D Scanning Using RGBD Imaging Devices: A Survey", In Proceedings of the Computational Vision and Medical Image Processing IV: VIPIMAGE 2013, Funchal, Madeira Island, Portugal, 14–16 October 2013; pp. 197–202.

[12] C. Munkelt, B. Kleiner, T. Torhallsson, P. Kühmstedt, and G. Notni, "Handheld 3D Scanning with Automatic Multi-View Registration Based on Optical and Inertial Pose Estimation", In Fringe 2013; Springer: Berlin/Heidelberg, Germany, 2014; pp. 809–814.

[13] G. Guidi, S. Gonizzi, and L. L. Micoli, "Image pre-processing for optimizing automated photogrammetry performances", ISPRS Annals of The Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol.2, No.5, 2014, pp.145-152.

[14] F. Remondino, M.G. Spera, E. Nocerino, F. Menna, F. Nex, and S. Gonizzi-Barsanti, "Dense image matching: Comparisons and analyses", Digital Heritage International Congress (Digital Heritage), vol. 1, 2013, pp.47-54, doi: 10.1109/DigitalHeritage.2013.674371.

[15] M. Dunn, "A company can provide customers with a 3D figurine made to look just like them", *NEWS.COM.AU* http://www.news.com.au/technology/innovation/a-company-can-provide-customers-with-a-3d-figu rine-made-to-look-just-like-them/story-fnjwucti-1227224660191 accessed 2015.

[16] B. Upbin, "My 3-D Printed Self Is Creepily Accurate And Cheaper Than Therapy", Forbes/Tech, http://www.forbes.com/sites/bruceupbin/2015/08/20/my-3-d-printed-self-is-creepily-accurate-and-cheaper-than-therapy/, accessed 2015.