

## THE PERFORMANCE EVALUATION OF MULTI-IMAGE 3D RECONSTRUCTION SOFTWARE WITH DIFFERENT SENSORS

V. Mousavi<sup>a</sup>, M. Khosravi<sup>a</sup>, M. ahmadi<sup>a</sup>, N. Noori<sup>a</sup>, A. Hosseini naveh<sup>a</sup>, M. Varshosaz<sup>a</sup>

<sup>a</sup> K.N.Toosi University of Tecknology, Faculty of Surveying Eengeeniring ,Tehran,Iran  
vmoosavy@mail.kntu.ac.ir, khosravi.msd@gmail.com, ahmadii.eng@gmail.com  
negar\_noori123@yahoo.com, haghshenas.shiva@yahoo.com, ali\_hosseini\_naveh@yahoo.com  
varshosazm@kntu.ac.ir

**KEY WORDS:** sensor, software, Image-based 3D reconstruction, Data comparison, point cloud, three-dimensional model

### ABSTRACT:

Today, multi-image 3D reconstruction is an active research field and generating three dimensional model of the objects is one the most discussed issues in Photogrammetry and Computer Vision that can be accomplished using range-based or image-based methods. Very accurate and dense point clouds generated by range-based methods such as structured light systems and laser scanners has introduced them as reliable tools in the industry. Image-based 3D digitization methodologies offer the option of reconstructing an object by a set of unordered images that depict it from different viewpoints. As their hardware requirements are narrowed down to a digital camera and a computer system, they compose an attractive 3D digitization approach, consequently, although range-based methods are generally very accurate, image-based methods are low-cost and can be easily used by non-professional users. One of the factors affecting the accuracy of the obtained model in image-based methods is the software and algorithm used to generate three dimensional model. These algorithms are provided in the form of commercial software, open source and web-based services. Another important factor in the accuracy of the obtained model is the type of sensor used. Due to availability of mobile sensors to the public, popularity of professional sensors and the advent of stereo sensors, a comparison of these three sensors plays an effective role in evaluating and finding the optimized method to generate three-dimensional models. Lots of research has been accomplished to identify a suitable software and algorithm to achieve an accurate and complete model, however little attention is paid to the type of sensors used and its effects on the quality of the final model. The purpose of this paper is deliberation and the introduction of an appropriate combination of a sensor and software to provide a complete model with the highest accuracy. To do this, different software, used in previous studies, were compared and the most popular ones in each category were selected (Arc 3D, Visual SfM, Sure, Agisoft). Also four small objects with distinct geometric properties and especial complexities were chosen and their accurate models as reliable true data was created using ATOS Compact Scan 2M 3D scanner. Images were taken using Fujifilm Real 3D stereo camera, Apple iPhone 5 and Nikon D3200 professional camera and three dimensional models of the objects were obtained using each of the software. Finally, a comprehensive comparison between the detailed reviews of the results on the data set showed that the best combination of software and sensors for generating three-dimensional models is directly related to the object shape as well as the expected accuracy of the final model. Generally better quantitative and qualitative results were obtained by using the Nikon D3200 professional camera, while Fujifilm Real 3D stereo camera and Apple iPhone 5 were the second and third respectively in this comparison. On the other hand, three software of Visual SfM, Sure and Agisoft had a hard competition to achieve the most accurate and complete model of the objects and the best software was different according to the geometric properties of the object.

### 1. INTRODUCTION

Nowadays, due to advances in technology, generating of 2D and 3D products with significant geometric accuracy and detail, has become possible. A detailed three-dimensional model for representation and documentation of different objects are needed in different branches of engineering.(Hess, Robson, & Hosseininaveh Ahmadabadian, 2014)

Three-dimensional modelling of scenes and objects in different scales is performed using both image-based methods (passive) and range-based methods (active). Range-based or active methods (Vosselman & Maas, 2010) like structured light systems and laser scanner are common systems for point cloud generation. Image-based or passive methods (Remondino & El-Hakim, 2006) which have made significant progress with the development of computer (Furukawa, Curless, Seitz, & Szeliski, 2010; Pollefeys et al., 2008) are currently regarded as automatic methods for image orientation (Remondino, Del Pizzo, Kersten,

& Troisi, 2012) and three-dimensional reconstruction in different scales (Haala, 2013; Lafarge & Mallet, 2012).

A variety of algorithms can be used for three-dimensional modelling of different objects or scenes. One of the most common algorithms in this field is SFM (Structure from Motion) algorithm (Agarwal et al., 2011). For this purpose, an archive of images of the object or scene is needed. In SFM method, a large number of images is oriented without any knowledge of the internal parameters of the camera (Barazzetti, Scaioni, & Remondino, 2010) . Orientation of the images is performed automatically through a series of specific points with identical features. One of the most commonly used operators in this operation is SIFT (Scale Invariant Feature Transform) (Fahmi, 2011) , which extracts identical features between images with different positions, scales and lighting.

Different software, based on this algorithm or similar algorithms, are available as open source (VisualSFM, Bundler,

Apero, Insight3D, etc.), commercial (Agi Soft, Photo Modeller) and Web service (Microsoft Photosynth Autodesk123D Catch Beta, My 3D Scanner, Hyper3D, Arc3D, etc.) for users (Remondino et al., 2012). Web-based three-dimensional software services create three-dimensional models quickly and without any special knowledge. (Cavas-Martínez et al., 2014)

In previous studies, effective factors such as network design (Ahmadabadian, Robson, Boehm, & Shortis, 2014) are examined for reducing image distortion of the object. But the type of sensor used has not been evaluated so far. Due to the availability of mobile sensors for everyone and propagation of professional sensors and the advent of stereo sensors, comparing these three types of sensors is an effective step in the evaluation and finding optimal methods for creation detailed three-dimensional models. The question here is whether the type of sensors used in the model can affect the accuracy of the model?

The purpose of this paper is to identify the best software and sensors for generating a complete and accurate three-dimensional model from small objects. For this purpose, a detailed overview of recent research in this field is presented in section 2. According to the mentioned three types of sensors and appropriate identified software, precise geometric evaluation is performed based on evaluation standards of point clouds in section 4.

## 2. REVIEW OF PREVIOUS WORK

Active sensors have been used in various fields since 2000. Although these sensors have made significant progress in recent years, a lot of problems occur when using them; for instance, they are heavy and non-portable and they cost a lot. (Remondino, Spera, Nocerino, Menna, & Nex, 2014)

Today, the integration of computer vision and photogrammetry methods has led to creation of image-based modelling processes. (Remondino & El-Hakim, 2006) In this method, multiple images of a scene turn into point cloud by correspondence algorithms and then a uniform three-dimensional model of this point cloud is generated. Perhaps it can be said that correspondence is one of the key steps in image-based modelling. Finally, using this method we can extract semantic information from images. In recent years, due to the development of computer and related technology and the emergence of some algorithms like SFM, image-based methods is more economical. SFM algorithm includes simultaneous determination of internal and external parameters of the camera and reconstructing the three-dimensional structure of the scenes. In image-based methods first the correspondence between two images is examined by extracting key features. to do this, the area-based or pixel-based algorithms are used. SIFT (Scale Invariant Feature Transform) and SURF (Speeded Up Robust Feature) (Bay, Tuytelaars, & Van Gool, 2006) are the most common algorithms. Then, using SFM algorithm or its integration with other algorithms like DMVR (Dense Multi View Reconstruction), orientation of cameras and internal and external parameters of the camera are calculated and a sparse point cloud is created. Many systems use Bundle Adjustment (Engels, Stewénius, & Nistér, 2006) in order to improve the accuracy of camera direction and reduce image distortion. In some software, additionally, an algorithm is added to create a dense point cloud.

Examining the image-based methods and also comparing these methods with range-based methods has been in the spotlight of researchers in recent years.

These days, with the advances in technology and the advent of a variety of algorithms, different methods and software for 3-D modeling is available: free Open source software, commercial software which is for professional modeling, and web-based software that make 3-D modeling very simple. Each of these applications use various algorithms in different stages of modeling process and as a result, network design, the number of images, system and even the time needed for processing a model is different. In this article we tried to use all three kinds of software and also different algorithms in the modeling process. Table 1 shows the summary of the software used in recent researches.

Software	Image Matching	Spars Reconst ruction	Dense Reconst ruction	Author
Visual SFM	SIFT	SFM	CMVS/ PMVS	(Wu, 2011)
Agisoft	SIFT-like	SFM	SGM	2014
SURE	-	-	SGM	(Rothermel, Wenzel, Fritsch, & Haala, 2012)
Arc3D	SURF	SFM	MVS	K.U.Leuven (2005)
Photosynth	SIFT	SFM	-	Microsoft (2008)

Table 1 - Used Software in recent studies

Preparing a three-dimensional model of an object depends on photogrammetric methods, geometric accuracy of the obtained model, the quality of the object texture, resolution of the images and network design (Skarlatos & Kiparissi, 2012). So the quality and resolution of the image used, is an important factor in three-dimensional model generation. A good camera certainly brings better results in comparison with a cheaper one.

The greater the number of pixels in an image, the more the number of created points in final model, and thus, the time needed for model generation increases. However, due to the heavy cost of professional cameras, the question is that which resolution is sufficient to provide an appropriate model. If the camera lens is better in its image quality, the amount of light received by the CCD becomes more and the lens distortions becomes less. Both these factors improve accuracy of the modeling. The amount of noise positioned to image by sensor or its efficiency in terms of its stability in internal parameters of the camera is an important issue which has not been studied in recent years yet. Due to the availability of mobile sensors for everyone and popularity of professional sensors and emergence of stereo sensors. Comparing these three is an effective step in evaluating and finding an optimal procedure to produce detailed three-dimensional models. Therefore, in this study, three different sensors are used in the process of point cloud generation and three-dimensional modelling.

## 3. MATERIALS AND METHODS OF THE STUDY

In order to evaluate the performance of different software as well as various sensors in terms of optimal model generation, four different objects have been used in the evaluation process.

Three of the objects have regular shape and the one of them is a statue with a more complexity in geometric shape.



Figure 1 –pictures of studied objects

### 3.1 Data collection

Data collection consists of two main steps. The first step involves the preparation of three-dimensional model of objects by active scanner sensor ATOS Compact Scan 2M which is related to Gom Inspect company. It conducts three-dimensional model by ToF<sup>1</sup> method.

	<b>Atos compact scan 2M</b>
	The number of camera pixels: 2 x 2 000 000 The dimensions of the sensor: 340 mm x 130 mm x 230 mm

The second step is image acquisition. In this stage, the network design is performed according to different algorithms, so that it is compatible with all software and algorithms. In this stage, 69 images are taken in a ring with the convergence angle of about 5 degrees.

### 3.2 Image acquisition stage

If we divide image sensors to three categories, these categories include professional sensors, emerging new stereo sensors and current mobile sensors.

In this study, one sensor is chosen from these three categories and each of them are evaluated. Among professional cameras, Nikon D3200 and among stereo cameras, Fujifilm FinePix Real3D W3 and among mobile cameras, iPhone 5 was chosen.



Figure 2 - Sensors that used in this article

Photogrammetry algorithms in processing step may face two problems: one of them is original image quality (noise, low radiometric quality of the images, shadows, etc.), second is the surface of the object (the shiny surface of an object or lack of texture). (Remondino et al., 2014)

In relation to the quality of the original image or the amount of noise, three kinds of sensors are investigated in this study and the amount of noise and the efficiency of each sensor for 3-D model generation are examined. But 3-D model generation of a bright object <sup>2</sup> is still one of the problems in photogrammetry that needs more research. To solve this problem, in this study, a solution of titanium dioxide powder and alcohol is used. The

reason for using this solution is the fine particles of the powder, therefore it doesn't harm the accuracy of the model.

In the process of point cloud generation of an object, at first matching process between original images is needed. Matching refers to determining the key and corresponding points between images. So according to this definition, key points must be found in images. But in objects with regular geometric shape or without texture on its surface, this process fails or if done, the result is a totally sparse point cloud. To create an artificial texture on these kinds of objects, optical methods are usually used. This means that a texture is fitted on the surface by a projector. The minimum number of the projectors is two, but in this study, three projectors were used to prevent the shadow.

Also, due to this fact that objects used in this study are metal and each metal has a thermal expansion coefficient, it is necessary to keep the ambient temperature fixed for image acquisition and also similar temperature while generating 3-D model from active sensor. In this study this temperature was 22 °C.

### 3.3 The process of three-dimensional modelling in software

In the process of creating three-dimensional model of mentioned objects, at first the surface of the object is covered with titanium dioxide and then by an appropriate network design, images are taken. For projecting texture to the surface of each objects, three video projectors are used which are located at 120 degrees of each other and a specific pattern is projected on the object. According to the dimensions of the object and the distance of projectors, the distance between the camera and the object is calculated as well as the number of image stations according to this distance. Distance from the camera to the desired object varies around 70 to 100 cm and 68 image stations was considered. This process was conducted for each sensor and each object and at the end, 48 data sets were obtained. Each of software was run 12 times for 4 objects and 3 sensors and the results were compared with true data. The results of the comparison are examined in the next section.

## 4. RESULTS

After the mentioned stages for each object and each sensor, raw data is obtained. These raw data are run in various software and the point cloud for each object is achieved in three different sensors. Then, these different sensors should be compared in terms of accuracy and quality of the obtained model.

object	Evaluation criteria between two point clouds
Stone statue	Fitting two plates to each cloud point and comparing the distance between these two in two different cloud points
Sphere	fitting a sphere to each cloud point and comparing obtained spheres
cylinder	Fitting a cylinder to a section of cloud point and comparing two cylinders
hexagon	Fitting two plates to opposite sides of each other for each cloud point and comparing the distance between these two in two cloud points

<sup>1</sup> Time of flight

<sup>2</sup> shiny

#### 4.1 The comparison of sensors

According to the selected criteria for each object, a special shape is fitted to objects and the comparison has been made between them. For each point cloud, standard deviation of all points of the fitted shape is considered as quality criteria for the obtained point cloud. The result of the comparison is illustrated in following tables.

Accuracy	Agi Soft	VSFM	SURE	Arc3D
Stone statue	0.41	0.48	0.39	No model
Sphere	0.17	0.09	0.10	No model
Cylinder	0.03	0.15	0.16	No model
Hexagon	0.08	0.08	0.17	No model

Table 2- Accuracy for Iphone5 sensor

Accuracy	Agi Soft	VSFM	SURE	Arc3D
Stone statue	0.40	0.48	0.1	0.37
Sphere	0.17	0.09	0.07	0.25
Cylinder	0.03	0.14	0.09	0.38
Hexagon	0.03	0.07	0.06	0.07

Table 3 - Accuracy for Nikon D3200 sensor

Accuracy	Agi Soft	VSFM	SURE	Arc3D
Stone statue	0.39	0.59	0.40	0.90
Sphere	0.17	0.13	0.01	No model
Cylinder	0.06	0.14	0.33	0.43
Hexagon	0.13	0.49	0.18	No model

Table 4 - Accuracy for Fuji film FinePix Real 3D W3 sensor

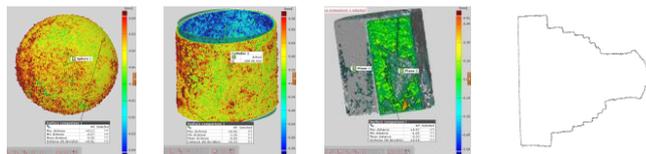


Figure 3 – Models that obtained from Stereo sensor and SURE algorithm

#### 5. DISCUSSION AND CONCLUSION

Today, due to advances in computer science, three-dimensional modeling of objects and scenes at different scales has become very popular. Research has been done in this field. Three-dimensional modeling algorithms vary and they are examined and compared by researchers in various fields in recent years.

As noted above, the geometric accuracy of an obtained model depends on several general parameters. Examining these parameters alongside algorithms can help the accuracy of the obtained model. One of the parameters involved in the model is resolution or image quality which has not been investigated in the recent years. Therefore, in addition to comparing several algorithms for three-dimensional modeling, we also examine the effect of different sensors in the obtained model. In general, three types of sensors are available in the market, which are all used in this study.

According to the conducted survey we can clearly see that professional sensors lead to better accuracy and density in the final model and then stereo sensors and mobile sensors have lower accuracy and produce models with more noise. But the obtained accuracies in different sensors have slight differences relative to each other. In general, according to the table of accuracy in different sensors and algorithms in similar condition, we can see that how much the sensors used in three-dimensional modeling depend on the used sensor or algorithm. According to the mentioned conditions, there is a limitation for modeling objects with the image-based method. Some of these limitations are the brightness of the object, darkness of the object or lack of texture in the surface. These limitations create some bugs in modeling and sometimes make it fail.

In the future we can eliminate these limitations by providing methods to model dark or bright objects or solve the problem of objects without texture.

#### REFERENCES

- Agarwal, S., Furukawa, Y., Snavely, N., Simon, I., Curless, B., Seitz, S. M., & Szeliski, R. (2011). Building rome in a day. *Communications of the ACM*, 54(10), 105-112.
- Ahmadabadian, A. H., Robson, S., Boehm, J., & Shortis, M. (2014). Stereo-imaging network design for precise and dense 3d reconstruction. *The Photogrammetric Record*, 29(147), 317-336.
- Barazzetti, L., Scaioni, M., & Remondino, F. (2010). Orientation and 3D modelling from markerless terrestrial images: combining accuracy with automation. *The Photogrammetric Record*, 25(132), 356-381.
- Bay, H., Tuytelaars, T., & Van Gool, L. (2006). Surf: Speeded up robust features *Computer vision–ECCV 2006* (pp. 404-417): Springer.
- Cavas-Martínez, F., Pérez-Sánchez, C., Adrián-Sáez, J., Cañavate, F., Nieto, J., & Fernández-Pacheco, D. (2014). USE OF DIGITAL IMAGES FOR THREE-DIMENSIONAL RECONSTRUCTION OF MECHANICAL COMPONENTS.
- Engels, C., Stewénius, H., & Nistér, D. (2006). Bundle adjustment rules. *Photogrammetric computer vision*, 2.
- Fahmi, H. M. (2011). Scale Invariant feature transform. *noppa.lut.fi*.
- Furukawa, Y., Curless, B., Seitz, S. M., & Szeliski, R. (2010). *Towards internet-scale multi-view stereo*. Paper presented at the Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on.
- Haala, N. (2013). The landscape of dense image matching algorithms.

Hess, M., Robson, S., & Hosseininaveh Ahmadabadian, A. (2014). A contest of sensors in close range 3D imaging: performance evaluation with a new metric test object. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1, 277-284.

Lafarge, F., & Mallet, C. (2012). Creating large-scale city models from 3D-point clouds: a robust approach with hybrid representation. *International Journal of Computer Vision*, 99(1), 69-85.

Pollefeys, M., Nistér, D., Frahm, J.-M., Akbarzadeh, A., Mordohai, P., Clipp, B., . . . Merrell, P. (2008). Detailed real-time urban 3d reconstruction from video. *International Journal of Computer Vision*, 78(2-3), 143-167.

Remondino, F., Del Pizzo, S., Kersten, T. P., & Troisi, S. (2012). Low-cost and open-source solutions for automated image orientation—A critical overview *Progress in cultural heritage preservation* (pp. 40-54): Springer Berlin Heidelberg.

Remondino, F., & El-Hakim, S. (2006). Image-based 3D modelling: A review. *The Photogrammetric Record*, 21(115), 269-291.

Remondino, F., Spera, M. G., Nocerino, E., Menna, F., & Nex, F. (2014). State of the art in high density image matching. *The Photogrammetric Record*, 29(146), 144-166.

Rothermel, M., Wenzel, K., Fritsch, D., & Haala, N. (2012). *Sure: Photogrammetric surface reconstruction from imagery*. Paper presented at the Proceedings LC3D Workshop, Berlin.

Skarlatos, D., & Kiparissi, S. (2012). Comparison of laser scanning, photogrammetry and SfM-MVS pipeline applied in structures and artificial surfaces. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3, 299-304.

Vosselman, G. V., & Maas, H.-G. (2010). *Airborne and terrestrial laser scanning*: Whittles.

Wu, C. (2011). VisualSFM: A visual structure from motion system. URL: <http://homes.cs.washington.edu/~ccwu/vsfm>, 9.