Preservation of cultural heritage: A comparison study of 3D modeling between laser scanning, depth image and photogrammetry methods

Mohamad Haziq Bin Ahmad Yusri,
Centre of Postgraduate Studies, Universiti Teknologi MARA, Selangor, Malaysia

M. A. Johan
School of Mechanical Engineering, College of Engineering, Universiti Teknologi MARA, Terengganu, Malaysia

*M. H. M. Ramli**
Sports Engineering and Artificial Intelligence, College of Engineering, Universiti Teknologi MARA, Selangor, Malaysia
Email: haniframli@uitm.edu.my

ABSTRACT

This paper presents a multi-technique approach for cultural heritage recording and documentation. The focal point of the study is to generate three-dimensional (3D) model representations of a physical object and use Augmented Reality (AR) to document the design. Three main methods are used to generate 3D object, which are laser, depth image, and photogrammetry. A comparison analysis is designed to evaluate each method accordingly. For evaluation purposes, a case study is designed where a scale model of a 25-cm-long famous historic Portuguese Indian Armada, the "Flor de la Mar" is used as a sample for generating 3D model records. The comparison analysis shows that the method of photogrammetry is superior in term of details, precision, and visualization, while laser scanning and depth image methods are capable of previewing the data into point cloud but with less accuracy. The results show that the photogrammetry method achieves 97% of accuracy in terms of dimensions and shapes.

ISSN 1823- 5514, eISSN 2550-164X

Received for review: 2020-xx-xx
Accepted for publication: 2020-xx-xx
Published: 2020-xx-xx

Keywords: *Augmented Reality (AR); Photogrammetry; 3D model; laser; depth camera; maritime.*

Introduction

Terengganu located the eastern Peninsular Malaysia is renowned for its ship-craftsmen. The famous pinas sailboat, is one of two types of junk rigged schooners, built in Terengganu. This kind of vessel was built of Chengal wood by the Malays since the 19th century [1] and roamed the South China Sea and adjacent oceans as one of the two types of traditional sailing vessels the late Malay maritime culture. Pinas was so well-known during Malay maritime culture that British postage include depiction of a Malay pinas in its stamp, in 1955 [2].

These ship-craftsmen build sailboats from their mind and heart without referring to any plan or schematics. This traditional boat crafting requires time, high wood skills, and techniques and most of the young generation does not find it interesting. Furthermore, Chengal woods are scarcely available and expensive due to aggressive deforestation. In this regard, it is important to preserve the designs as they are among the country's heritage. The preservation work was mostly done through many methods; among them are documentation in books or manuscripts (picture and text), video (in the case of cultural information: song, dance, or even documentary), and replication of the artifact [3], [4]. It can be seen that a number of methods are used [4], [5] to record and preserve historic monuments including ships. More often than not the time-of-flight ranging with laser scanning and image-based 3D reconstruction (photogrammetry) are used to reconstruct the historic monument or object by collecting and processing the obtained 3D data.

This study focuses on three different methods of generating 3D models. First is laser scanning method where a light detection and ranging (LiDAR) sensor is used to scan the 3D object. The second one is depth image method where it utilizes the time-of-flight principle with an imaging sensor for object scanning via a Microsoft Kinect depth camera. Both of these methods generate point cloud in 3D digital space to a construct 3D model. The last method is photogrammetry. This method uses 2D images to generate the 3D models through Structure-from-Motion (SFM) imaging technique that utilize feature-detection algorithms to automatically match and align overlapping multi-image datasets. This method requires a high rendering computation for processing the 3D models. The most common software used for this method includes Agisoft Photoscan, Reality Capture, AliceVision, VisualSFM, Regard3D and Agisoft Metashape [6]. Reality Capture and Agisoft Metashape software standout better than the rest in terms of output quality, fast meshing compilation, and processing speed [7].

In this study, Agisoft Metashape and Reality Capture are chosen due to their merits. For evaluation purposes, a case study is designed where a scale model of a 25-cm-long famous historic Portuguese Indian Armada, the “Flor de la Mar” is used as a sample for generating 3D model records.

3D Object Reconstruction methods

Laser Scanning

Laser scanning uses time-of-flight (ToF) concept in obtaining distance data of an object. It uses a laser beam to measure how pulses of light bounce off and return to the starting point. This determines the distance of the object.

In the experiment, a TF-mini-LiDAR sensor is utilized to scan the 3D object. TF-mini-LiDAR operating range is within 12 meters and is suitable for indoor application. The setup of the proposed scanning system is shown in Figure 1. The scanner consists of a stepper motor and a motor controller, a servo motor, a LiDAR sensor, and a micro-controller. All the parts are assembled and encapsulated viz. Additive Manufacturing (AM) technology. The stepper and servo motors are used for back-and-forth movement and tilting mechanism as illustrated in Figure 2.

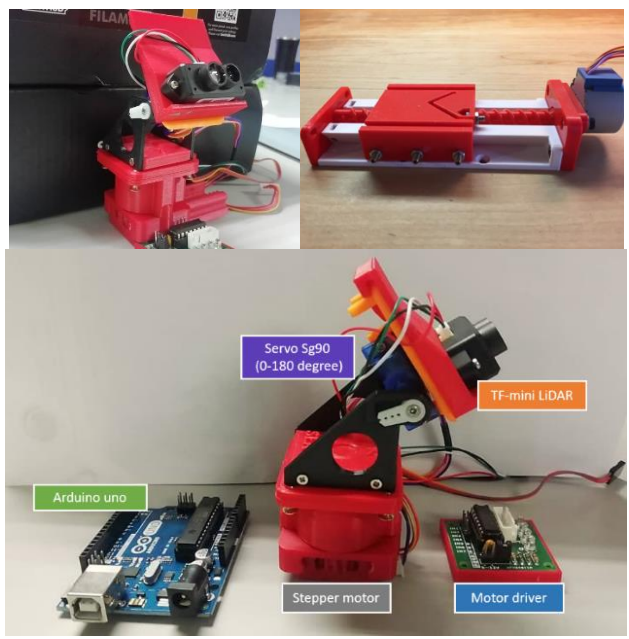

Figure 1: The proposed laser-based scanner setup.

The scanner setup is quite similar in a study done by R. Athavale et al. [8]. However, in this study a track is used to guide the scanner linear movement on Y-axis, while in [8], R. Athavale et al. mounted the LiDAR on a mobile robot for scanning tasks and led to wide field of view (FoV) and it is not suitable for this project, as the focus is scanning a single object.

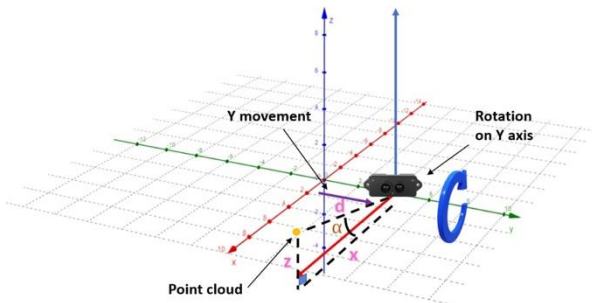

Figure 2: 3D space point cloud plot by the laser scanner

Eqn. (1) is used to generate point clouds as the LiDAR sensor slides on Y-axis. The distance, d is obtained from halve of the distance travel.

$$x = \cos\alpha \,.\, d \quad (1a)$$
$$y = linear\ incremental\ by\ 5mm \quad (1b)$$
$$z = \sin\alpha \,.\, d \quad (1c)$$

Two main software are used in the laser scanning method, one is for control the scanner, i.e., Arduino IDE, the other is for data processing which is known as Processing software. Processing software is generally used for architectural 3D modeling [9] to construct 3D shape and control shape transformations from the generated point clouds. However, the software only can render a simple 3D object, i.e., maximum point clouds is 1000 points. The overall process flow is shown in Figure 3. Scanned data is filtered before it is imported to processing software. Some of the data is overshoot and out of the chart due to sensor's defect and noises [10]. Lastly, data is processed and rendered to produce point clouds. Once a 3D object construction is completed it is converted to OBJ file format for further object manipulations.

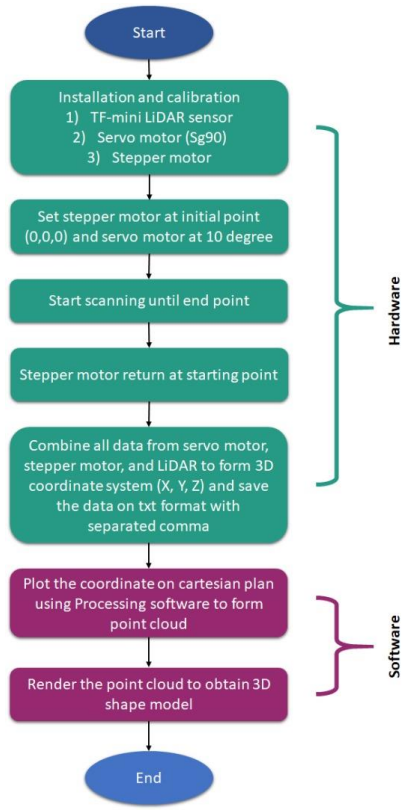

Figure 3: Laser scan process flow-chart

Depth Camera Scanning

A depth camera involves active 3D imaging sensor technology, characteristically capable of high measurement rates, producing several tens of images with depth information per second [11]. Depth cameras also utilize ToF principle or structured light to scan a 3D object. Following the introduction of Microsoft Kinect, a large amount of research is stimulated because it combines the technology of RGB-D cameras, infrared projectors, detectors that mapped depth through ToF at a low price [12]. Kinect was quickly adopted by researchers as a mean to further understand 3D data range in generating the virtual world. Some of the researchers use it to differentiate between living and non-living things by object recognition method [13].

In this study, both Kinect V1 and V2 are used for 3D scanning experiments to evaluate their usability and quality of 3D scanning. Kinect V1 has camera of is 320 x 240, while Kinect v2 has improved to 512 x 424 at 30 frames per second (fps) and maximum depth distance is 4.5 meters.

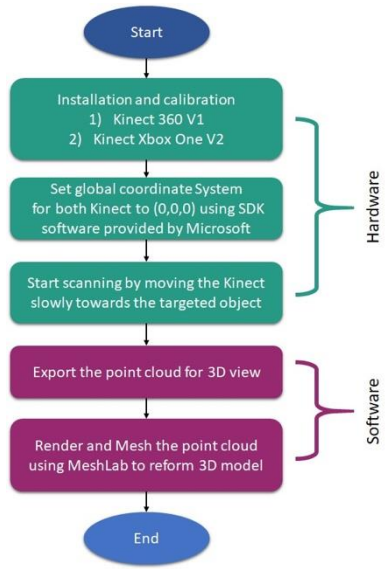

Figure 4: 3D scanning flowchart for depth camera

The overall process flow is shown in Figure 4. The process starts with the installation and calibration of Kinect V1 and V2 via the provided Software Development Kit (SDK). Next is to initialize the global coordinate system and followed by scanning process. In the first scanning process, the Kinect will automatically set its coordinate system at the origin in 3 dimensions while capturing the depth image. Once the scanning process is completed, 3D point cloud data is generated and converted into OBJ format for further processing. Finally, the OBJ data is processed using meshing software such as MeshLab to smoothen and harden the surface area of the 3D scanned object.

Photogrammetry

Digital photogrammetry is a technique to obtain 3D model by overlapping the captured two-dimensional (2D) images in measuring the size, shape and 3D geometric position of an object. The principle of operation is based on the assigned point located on the images. The points connection needs to be at least two images that could create the 3D object coordinates. The 3D model quality improvement could be made by increasing the intersection numbers of assigned points. The above-mentioned process is proven to be effective as can be seen in [14] and [15].

It is pointed out that the camera position significantly affects the construction of the 3D model [16] where its position is dependent on the size and shape of the object. In addition, different type of camera selection yield

different 3D model because each camera has dedicated pixels and focal length and sometime adjustable.

In this experiment, smart phone Honor 9 Lite with 2 MP embedded with depth sensor and 26mm focal length are used to take pictures and generate 3D model via SFM based software. The overall process flow is shown in Figure 5. The process is quite simple because most of the processing and rendering efforts are done by the SFM software from calculating the point localization for each image, images orientation, depth images and blur pictures filtering and finally create a precise 3D model. In order to generate a smooth surface 3D model, the lighting must be properly controlled during the task of taking pictures [17].

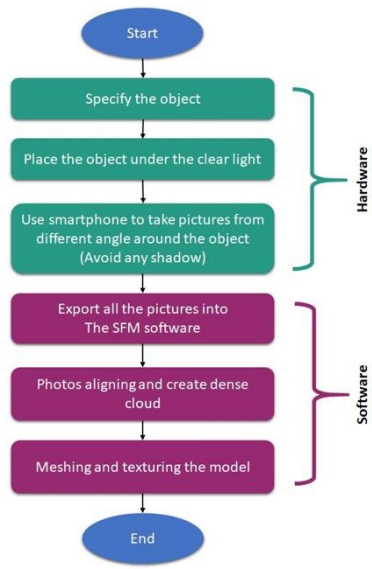

Figure 5: 3D modelling process using photogrammetry method

During photography task the camera is set to a constant focal length i.e., without zooming as illustrated in Figure 6 and Figure 7. First image is captured from perpendicular direction of the object and denoted as k^{th} position. The following position is $k+1^{th}$ until $k+n^{th}$, with n is the total number of images required to generate the 3D model. The camera position and orientation can be calculated by eqns. (2)-(3). Further illustration of coordinate and orientation changes from position k^{th} to $k+1^{th}$ is depicted in Figure 8.

Linear position in (X, Y, Z) coordinate

$$X_{k+1} = X_k + (R_{lx} - P_{lx}) \quad (2a)$$

$$Y_{k+1} = Y_k + (R_{ly} - P_{ly}) \quad (2b)$$

$$Z_{k+1} = Z_k + (R_{lz} - P_{lz}) \quad (2c)$$

Angular position in $(\theta_x, \theta_y, \theta_z)$

$$\theta_{(k+1),x-axis} = \theta_{k,x-axis} + \theta_x \quad (3a)$$
$$\theta_{(k+1),y-axis} = \theta_{k,y-axis} + \theta_y \quad (3b)$$
$$\theta_{(k+1),z-axis} = \theta_{k,z-axis} + \theta_z \quad (3c)$$

Where;

R_{lx}, P_{lx} is the base length (m) at k^{th} and $k+1^{th}$ position respectively, R_{ly}, P_{ly} is the Mainmast height (m) at k^{th} and $k+1^{th}$ position respectively, R_{lz}, P_{lz} is Base depth (m) at k^{th} and $k+1^{th}$ position respectively, θ_x is angle difference between R_{lx} and P_{lx}, θ_y is angle difference between R_{ly} and P_{ly}, and θ_z is angle difference between R_{lz} and P_{lz}.

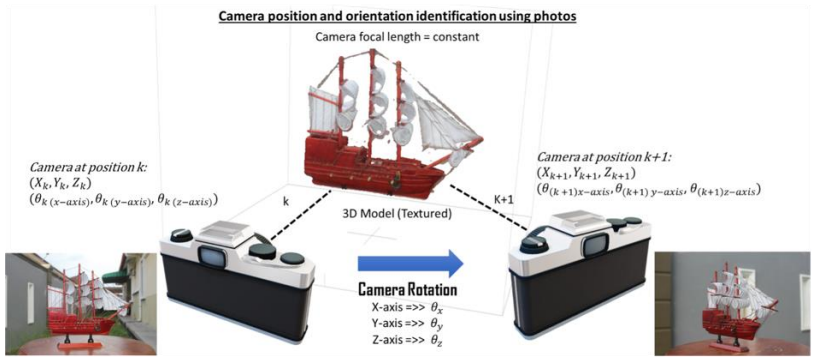

Figure 6: Illustration of photography scene at different angle and orientation

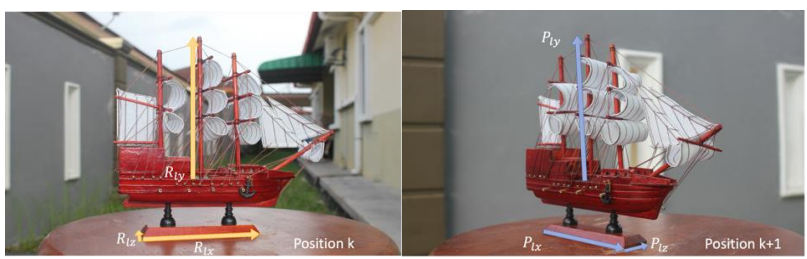

Figure 7: Camera view at (Left) k^{th} position, (Right) $k+1^{th}$ position

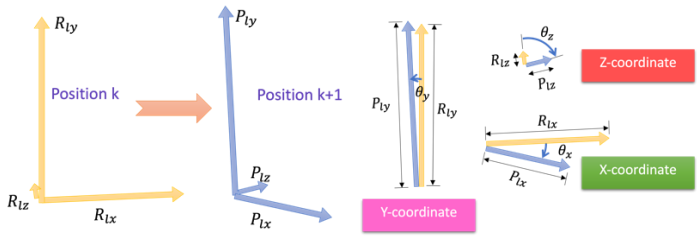

Figure 8: (Left) Illustration of the camera position at k^{th} to $k+1^{th}$ in 3D space (Right) Orientation angles obtained from the position changes plotted on 3D space

Results and Discussions

This section provides 3D scan result of each method. Evaluation of each method is properly carried out and also discussed in this section. In order to establish a unbiassed comparison analysis, each method use the same object in in the first experiment.

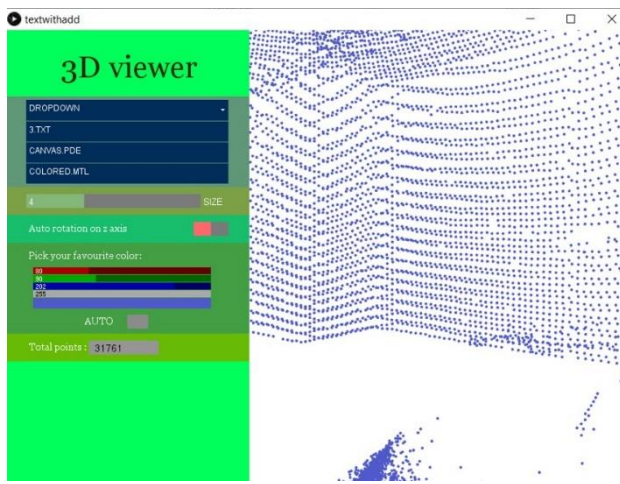

Figure 9: A view of 3D point cloud output of a room via Processing software

Laser scanning method

Figure 9 shows 3D point cloud output of a room via laser scanning approach and processed using Processing software. The total generated point cloud are 31761 points. The point cloud is represented as a dot and can be manipulated in order to further smoothen the output of 3D point cloud. However, the software only can render a simple 3D object, i.e., maximum point clouds is

1000 points and during the conversion to OBJ file, the computer freeze because of the data is too large for the software to handle.

Depth Image method

In depth image method, both Kinect V1 and V2 depth cameras are used in the evaluation experiment. The results of both cameras are shown in Figure 10. Similar to the first approach, the same room is used in the experiment.

Figure 10: 3D scan of a room using (Left) Kinect V1, (Right) Kinect V2

It can be seen in Figure 10 that the generated 3D model by both cameras is much better compared to the first method. However, latest version of camera, i.e., V2 is stands out compared to V1. Nevertheless, the quality, accuracy, and precision are very poor and could not be used for 3D model reconstruction of a physical object.

Photogrammetry method

Figure 11 shows the generated 3D model of a room using photogrammetry method. A total of 136 photos were taken and used to generate this model. A careful observation on the output, photogrammetry method surpasses both laser and depth camera methods in terms of reconstruction quality, accuracy, and precision.

Figure 11: Generated 3D model of a room using photogrammetry method

To evaluate further the capability of photogrammetry method, a scale model of a 25-cm-long famous historic Portuguese Indian Armada, the “Flor de la Mar” is used as a sample for generating 3D model record. A total of 148

photos were taken at various angles to ensure all sides of the model are captured. In addition, Alignment and sequence of photography are carefully planned and sorted out to optimize the output of 3D model. In this case, a circular photography method was applied in anti-clockwise direction. Figure 12 shows the output of 3D reconstruction model using Reality Capture Software.

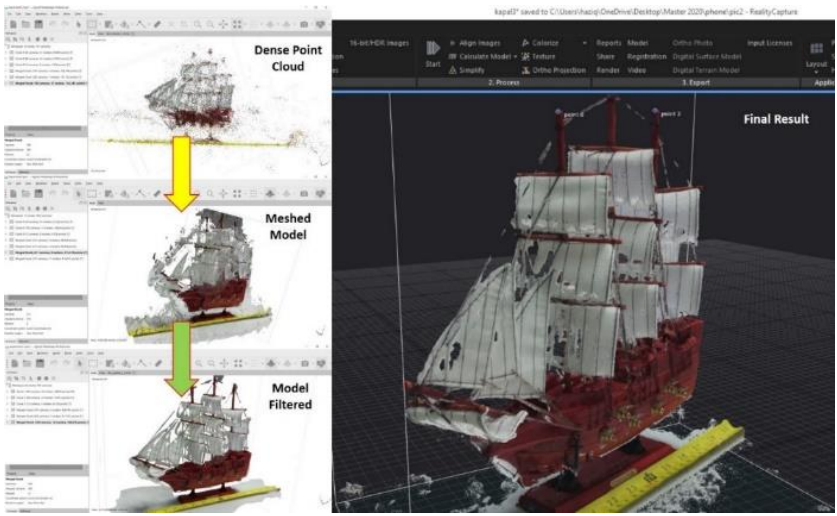

Figure 12: 3D ship model result generated by Reality Capture software

To obtain the final result as shown in Figure 12, the captured photos undergo a number of processes. First, a method called scale-invariant feature transform (SIFT) is applied to match the correct position of imported images as depicted in Figure 13. In simple terms, SIFT is a feature detection algorithm in computer vision to detect and describe local features in images and then match the adjacent images to construct a correct 3D model. With this method, the SFM software can generate dense point cloud by creating gradient lines on the picture and labelled the directions based on feature point localization from image's pixels. It has been used in a gamut of applications including object recognition, robotic mapping and navigation, image stitching, 3D modeling, gesture recognition, video tracking, and so forth and proven to be effective [18]. The point cloud is then undergone meshing process followed by the texture projection to construct a proper 3D model. These processes take some time to complete depending on the available computing power and resources. Lastly, the finished 3D model will be transferred to a Computer Aided Design (CAD) software e.g., CATIA software for model finetuning, record keeping, and blueprint process.

Figure 13: The captured images are aligned according to SIFT algorithm

Analysis of Model Confidence Level

Figure 14 depicts a comparison between physical model, the photogrammetric point cloud and mesh model coloured according to absolute deviation in percentage, one on the starboard (right) side and the other is on port (left) side. The blue shaded colour on the model represents the ideal value, which is about 80% to 100% confidence of measurements between the real object and the 3D model, followed by cyan colour with 10% of confidence, green colour (5% of confidence), and red shaded colour with the estimated measurement of only 1% to become true model. The greater the percentage values, the higher the accuracy to become true model. Most of the red shaded colour are on the right side of the model and at the bottom of the sailcloth. This may due to the excessive light and shadow during the photography session [9] because the images were taken in the afternoon at the open space area, with cloudy weather. Therefore, there are some surface areas are blocked and covered with shadows that lead to error in estimation. In addition, the thin white sailcloth may become transparent during the photo taking process as the light pass through the sailcloth. The transparent object will reduce the percentage of accuracy drastically during the alignment of images and lead to estimation of wrong values. Therefore, to reduce the percentage errors, points markers were used to help the SFM software to identify the positions, orientations and depth of the model in the images as illustrated in Figure 15. In this case, the alignment of images is focused on three main points 0, 1, and 2 which are ultimate reference for the alignment using SIFT. Consequently, the confidence level increases as more reference points are introduced. However, it is a trade-off between computation time and confidence level.

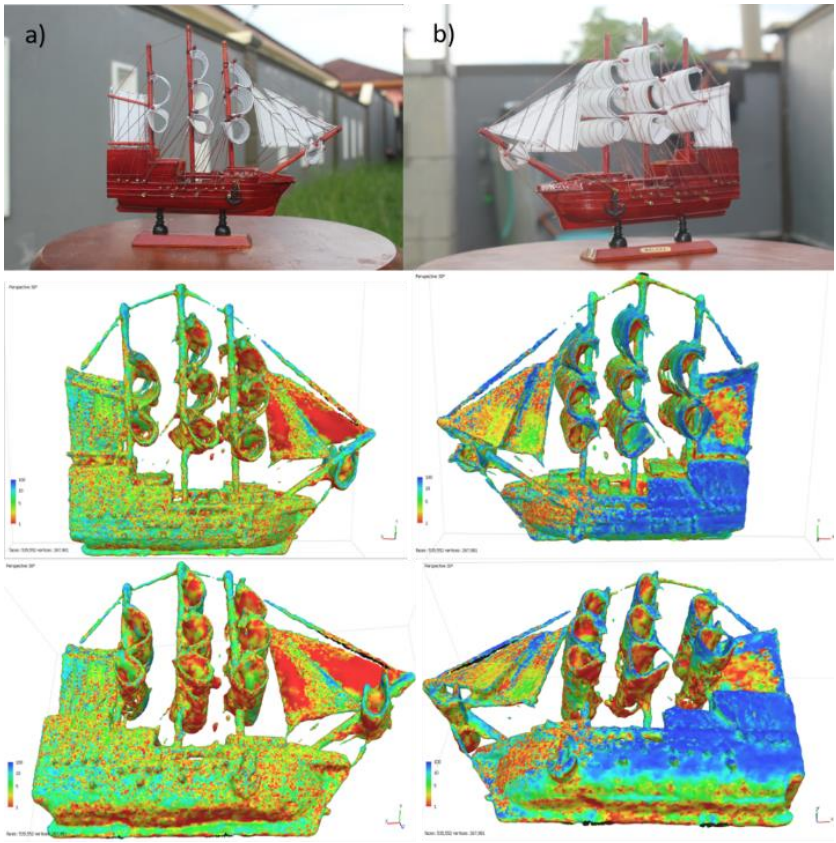

Figure 14: A comparison of the photogrammetric point cloud and mesh model coloured according to absolute deviation in percentage, a) starboard side b) port side.

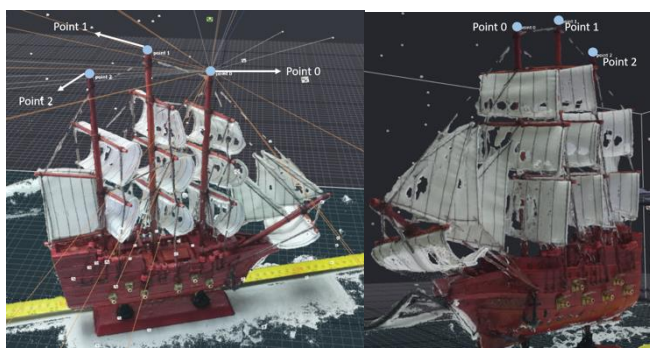

Figure 15: The alignment of images is focused on marked points using SIFT algorithm to increase confidence level

Model Precision Analysis

Once the alignment of the pictures has been finalized, a polygonal model (structured data) is generated followed by texture in order to produce the best digital representation of the physical model. The produced model further fine-tuning to improve the state of the reconstructed 3D model. Therefore, the 3D model dimensions need to be evaluated and compared with the actual dimension of the physical model in order to justify that this method is effective to be used as one of the preservations and record documentation method of objects.

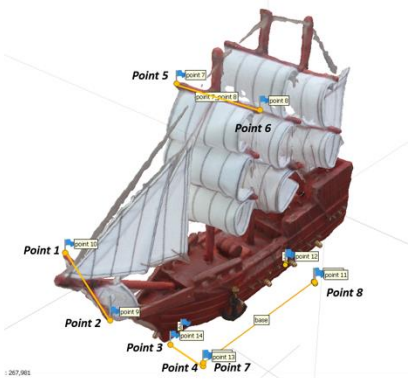

Figure 16: Marked points on the reconstructed 3D model to facilitate precision analysis

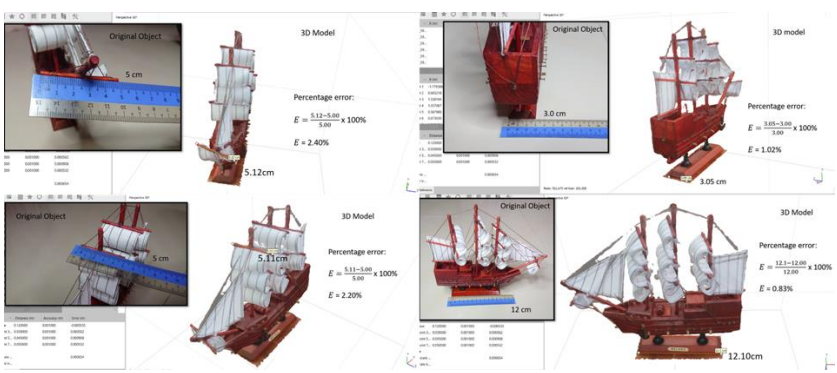

Figure 17: Dimension accuracy of the selected sections of the reconstructed 3D model

To simplify the analysis of model precision, a number of points is marked on the reconstructed 3D model as illustrated in Figure 16. A total of 4 sections are identified as the potential candidates for precision analysis which

are Sailcloth holder length (pt. 1 to pt. 2), Base Width (pt. 3 to pt. 4), Bowsprit length (pt. 5 to pt. 6), and Base length (pt. 7 to pt. 8)

Table 1: Percentage error between physical and 3D reconstructed models

Part	Actual Dimension	Model Dimension	% Error
Sailcloth holder length	5.00cm	5.11cm	2.20
Bowsprit length	5.00cm	5.12cm	2.40
Base length	12.00cm	12.10cm	0.83
Base Width	3.00cm	3.05cm	1.02

Table 1 shows the percentage error on selected sections between physical and 3D reconstructed models. The lowest percentage error is measured at Base length with 0.83%. While the maximum percentage error is found at Bowsprit length with 2.40%. In average, the percentage error is about 1.6125%. The tabulated percentage error values indicates that the photogrammetry method combined with SFM software capable of reconstructing high quality and precision 3D model.

SFM Software Performance: Reality Capture vs Agisoft Metashape

For the model of "Flor de la Mar", 148 images were captured and brought into the two different processing platforms which are Reality Capture and Agisoft Metashape software. In terms of processing speed, Reality Capture surpasses Metashape. Metashape takes an average of 154 minutes or almost 3 hours to complete all the processes to reconstruct the 3D model. While in Reality Capture, processing time is much lower with average of 74 minutes (roughly one hour) with the same settings and configurations applied in Metashape. In terms of output quality, Reality Capture produces better output details and less open holes than Metashape as can be seen in Figure 18.

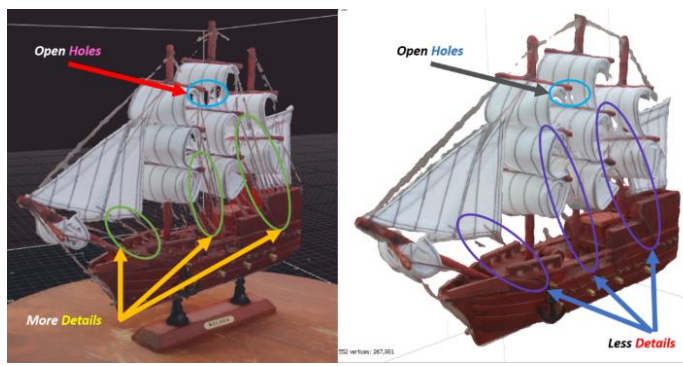

Figure 18: The comparison output result between Reality Capture (left) and Metashape (right).

Conclusion

This paper addresses the capacity of laser scanning, depth camera, and photogrammetry methods of towards digital reconstruction of 3D models. The findings indicate that laser scanning is the worst method for reconstruction of 3D model. Experiment done using TF-mini-LiDAR shows that it can only generate point cloud and unable to produce 3D model. While in depth camera method, the generated 3D model is very rough with limited surface details and distorted segments. Kinect sensor is weak against light and have limited scan area. Photogrammetry method yields a good result with high accuracy of dimensions and good quality of 3D texture. This achievement will uphold the process of preserving national heritages in AR environment and also provide a more practical way to produce shape design blueprints.

In terms of SFM software, i.e., photogrammetry processing software Reality Capture software is better compared to Agisoft Metashape in two main characteristics, one is processing speed and the other is output quality. Each software has the ability to scale the model into original size, which is 1:1 scalar with the precision of 97% successful rate for the ship model.

Acknowledgment

This work is supported in part by Ministry of Higher Education grant (Fundamental Research Grant Scheme).

References

[1] PISOL MAIDIN, “Tukang Timbal Membina Perahu : Tradisi dan Inovasi,” *Sari - Int. J. Malay World Civilis.*, vol. 21, pp. 39–56, 2003.

[2] “Koleksi Setem, Filateli & Sampul Surat Hari Pertama.pdf.” .

[3] Y. Alshawabkeh, M. El-Khalili, E. Almasri, F. Bala’awi, and A. Al-Massarweh, “Heritage documentation using laser scanner and photogrammetry. The case study of Qasr Al-Abidit, Jordan,” *Digit. Appl. Archaeol. Cult. Herit.*, vol. 16, no. December 2019, p. e00133, 2020, doi: 10.1016/j.daach.2019.e00133.

[4] P. Sapirstein, “Accurate measurement with photogrammetry at large sites,” *J. Archaeol. Sci.*, vol. 66, pp. 137–145, 2016, doi: 10.1016/j.jas.2016.01.002.

[5] Q. Dong, Q. Zhang, and L. Zhu, “3D scanning, modeling, and printing of Chinese classical garden rockeries: Zhanyuan’s South Rockery,” *Herit. Sci.*, vol. 8, no. 1, pp. 1–15, 2020, doi: 10.1186/s40494-020-

00405-z.

[6] K. Kingsland, “Comparative analysis of digital photogrammetry software for cultural heritage,” *Digit. Appl. Archaeol. Cult. Herit.*, vol. 18, no. April, p. e00157, 2020, doi: 10.1016/j.daach.2020.e00157.

[7] K. Kingsland, “Comparative analysis of digital photogrammetry software for cultural heritage,” *Digit. Appl. Archaeol. Cult. Herit.*, vol. 18, no. July, p. e00157, 2020, doi: 10.1016/j.daach.2020.e00157.

[8] R. Athavale, D. S. Harish Ram, and B. B. Nair, “Low cost solution for 3D mapping of environment using 1D LIDAR for autonomous navigation,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 561, no. 1, 2019, doi: 10.1088/1757-899X/561/1/012104.

[9] R. A. Galantucci and F. Fatiguso, “Advanced damage detection techniques in historical buildings using digital photogrammetry and 3D surface anlysis,” *J. Cult. Herit.*, vol. 36, pp. 51–62, 2019, doi: 10.1016/j.culher.2018.09.014.

[10] G. Ariante, U. Papa, S. Ponte, and G. Del Core, “UAS for positioning and field mapping using LIDAR and IMU sensors data: Kalman filtering and integration,” *2019 IEEE Int. Work. Metrol. AeroSpace, Metroaerosp. 2019 - Proc.*, pp. 522–527, 2019, doi: 10.1109/MetroAeroSpace.2019.8869696.

[11] J. P. Virtanen, K. N. Antin, M. Kurkela, and H. Hyyppä, “The feasibility of using a low-cost depth camera for 3D scanning in mass customization,” *Open Eng.*, vol. 9, no. 1, pp. 450–458, 2019, doi: 10.1515/eng-2019-0056.

[12] H. L. L. Wellens, H. Hoskens, P. Claes, A. M. Kuijpers-Jagtman, and A. Ortega-Castrillón, “Three-dimensional facial capture using a custom-built photogrammetry setup: Design, performance, and cost,” *Am. J. Orthod. Dentofac. Orthop.*, vol. 158, no. 2, pp. 286–299, 2020, doi: 10.1016/j.ajodo.2020.01.016.

[13] Y. Zhang, T. Jia, Y. Chen, and Z. Tan, “A 3D Point Cloud Reconstruction Method,” *9th IEEE Int. Conf. Cyber Technol. Autom. Control Intell. Syst. CYBER 2019*, pp. 1310–1315, 2019, doi: 10.1109/CYBER46603.2019.9066630.

[14] C. Georgia, T. Luhmann, M. Chizhova, and D. Gorkovchuk, “Fusion of UAV and Terrestrial Photogrammetry with Laser Scanning for 3D Reconstruction of Historic.”

[15] I. Reljić and I. Dunder, “Application of photogrammetry in 3D scanning of physical objects,” *TEM J.*, vol. 8, no. 1, pp. 94–101, 2019, doi: 10.18421/TEM81-13.

[16] R. Pacheco-Ruiz, J. Adams, and F. Pedrotti, “4D modelling of low visibility Underwater Archaeological excavations using multi-source photogrammetry in the Bulgarian Black Sea,” *J. Archaeol. Sci.*, vol. 100, no. May, pp. 120–129, 2018, doi: 10.1016/j.jas.2018.10.005.

[17] H. El-Din Fawzy, “3D laser scanning and close-range

photogrammetry for buildings documentation: A hybrid technique towards a better accuracy," *Alexandria Eng. J.*, vol. 58, no. 4, pp. 1191–1204, 2019, doi: 10.1016/j.aej.2019.10.003.

[18] T. Lindeberg, "Scale Invariant Feature Transform," *Scholarpedia*, vol. 7, no. 5, p. 10491, 2012, doi: 10.4249/scholarpedia.10491.