

## EXPLORATORY STUDY OF 3D POINT CLOUD TRIANGULATION FOR SMART CITY MODELLING AND VISUALIZATION

Shakirah Amirah Mohd Ariff<sup>1</sup>, Suhaibah Azri<sup>1</sup>, Uznir Ujang<sup>1</sup>, Amalin A'ishah Mohd Nasir<sup>1</sup>, Nursyahira Ahmad Fuad<sup>2</sup> and Hairi Karim<sup>2</sup>

<sup>1</sup> 3D GIS Research Group, Geoinformation, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia – samirah33@live.utm.my, suhaibah@utm.my, mduznir@utm.my, lynnalin.95@gmail.com

<sup>2</sup> I Net Spatial Sdn Bhd, 41A & 41B, Jalan Flora 1/9, Taman Pulau Flora, 81300 Skudai, Johor, Malaysia. – nursyahiraaf@inetspatial.com, wmhairi@inetspatial.com

### Commission VI, WG VI/4

**KEY WORDS:** 3D Modelling, 3D Visualization, Airborne Laser Scanning, Mobile Mapping System

#### ABSTRACT:

The current trends of 3D scanning technologies allow us to acquire accurate 3D data of large-scale environment efficiently. The 3D data of large-scale environments is essential when generating 3D model is for the visualization of smart cities. For the seamless visualization of 3D model, large data size will be used during the 3D data acquisition. However, the processing time for large data size is time consuming and requires suitable hardware specification. In this study, different hardware capability in processing large data of 3D point cloud for mesh generation is investigated. Light Detection and Ranging (LiDAR) Airborne and Mobile Mapping System (MMS) are used as data input and processed using Bentley ContextCapture software. The study is conducted in Malaysia, specifically in Wilayah Persekutuan Kuala Lumpur and Selangor with the size of 49km<sup>2</sup>. Several analyses have been performed to analyse the software and hardware specification based on the 3D mesh model generated. From the finding, we have suggested the most suitable hardware specification for 3D mesh model generation.

### 1. INTRODUCTION

3D smart city is an integration between smart city and IT technology for 3D city (Leng, Xiong et al. 2010). To allow 3D model in achieving the concept of smart city, a high-quality 3D urban surface models are required. Development of 3D urban surface modelling requires a large-scale environment 3D data to execute it. In the present, there are various ways of acquiring large scale of environment 3D data, such as Terrestrial Laser Scanning (TLS), Airborne Laser Scanning (ALS), Mobile Mapping System (MMS) and Unmanned Aerial Vehicles (UAV). Nevertheless, it is relatively challenging to acquire a complete large-scale environment 3D spatial data using a single type of sensor (Cheng et al., 2018) because of the limitations like only from a single perspective view (Cheng et al., 2013). Thus, the integration from different types of dataset or sensor is necessary. TLS and MMS are capable in acquiring side view data. However, TLS is a ground-based scanning method that is suitable for dense area. On the other hand, MMS is a vehicle mount scanning method, making it suitable to acquire data in urban road (Liang et al., 2020). While ALS and UAV are auxiliary data obtained from high position that are capable to acquire top view data for a large-scale area.

TLS, ALS, MMS and UAV are the list of data collection techniques in acquiring 3D data. Due to each respective differences, it resulted into difficult integration between each techniques. A powerful technique is required to integrate different types of dataset. The 3D data acquired from various positions or systems requires a registration process as the data processing and a procedure to integrate different dataset into a complete 3D model. Nonetheless, there are room for improvement regarding the robustness on variations of the scanned objects and the environments, and the computational efficiency in visualization of 3D model. This research proposes

an integration method between ALS and MMS in generating a high-quality 3D urban surface model, the hardware and software specification in generating 3D mesh model.

#### 1.1 Airborne Laser Scanner

ALS or LIDAR is an advanced technique that provides a good set of three-dimensional data with X, Y and Z axes to generate a DEM, including other information to assess and monitor landslides such as colour, intensity, geologic and geomorphic using DEM's derivatives. In a few minutes, millions of data measurements can be provided using this technique which is commonly denoted as '3D Point Clouds'. ALS technique can make data rapidly captured with high data density, 3D object modelling as well as an user-friendly procedure. In particular, ALS captured high - density 3D points can provide an opportunity to identify the detailed and distinctive characteristics in partial areas (Pirasteh & Li, 2017). ALS has become one of the most popular method chosen in recent years because it provides a rapid 3D data collection over a massive area. Moreover, the captured 3D data contains terrain models, forestry, 3D buildings and so on. ALS requires three main parts for acquiring 3D data which are laser scanning system, global positioning system (GPS) and inertial measuring unit (IMU).

#### 1.2 Mobile Mapping System

Properties for simple building do not require high accuracy sensor as it simply measurable using naked eye whilst complex model captures using various sensors such as TLS or GPS (Uden and Zipf 2013). However, most of the modern smart phones today are provided with sensors, making them a mobile multi-sensor system. MMSs have been developed and used in several fields such as urban planning, 3D city modelling, virtual heritage conservation, augmented reality, transportation and forestry

(Yang, 2019). MMSs are initially used to extract detailed 3D data at a high resolution and accuracy for numerical city modelling while providing spatial data in the most effective way and possible for better understanding of urban environments. Also, because of MMSs are mounted on vehicle, it enable to provide high-quality road-related data to improve 3D object modelling. Thus, MMSs are used to capture spatial information to assist in mapping or navigating in urban areas.

The operational MMSs were first developed in the 1990s by the Centre for Mapping at the Ohio State University to automate and improve the efficiency of data collection for digital mapping. This group used a vehicle-equipped with a GPS, charge-coupled device cameras, colour video cameras and several dead reckoning sensors. In the 2000s, to meet the increasing demand for high-quality 3D urban data to delineate road details and man-made structures, MMSs were used to measure highway assets, indivisible or abnormal load route planning and 3D city modelling (Karimi & Grejner-Brzezinska, 2004). Moreover, such systems provide information on building facades or power lines. When these developments were taking place, commercial use of MMSs (the StreetMappersystem) increased. In 2007, Google Maps Street View was generated using a vehicle-based survey and began to provide street-level images to supplement the Google’s maps. Street-level images enable people to improve their spatial perception or awareness in urban areas. The effects of the virtual reality interactive screen images available in widespread areas from Google Street View have affected virtual tourism and geo-gaming.

## 2. 3D MESH MODEL

A 3D mesh is one of the ways to represent 3D model and made up by a geometric data structure that allows the representation of surface subdivisions by a set of polygons. Further analysis by using this 3D model can be achieved as demonstrated by other researchers (Azri et al., 2018; Salleh et al., 2018; Yusoff et al., 2011). 3D mesh model is mainly used to discretize a continuous or implicit surface. The generation of 3D mesh model are divided into several phases: data capture, data processing and data validation. Details on each phases can be found in the next section.

### 2.1 Methodology to Generate Mesh

Three main phases are followed to generate the 3D mesh model which are data capture, data processing and data validation. Figure 1 shows the flowchart for the study.

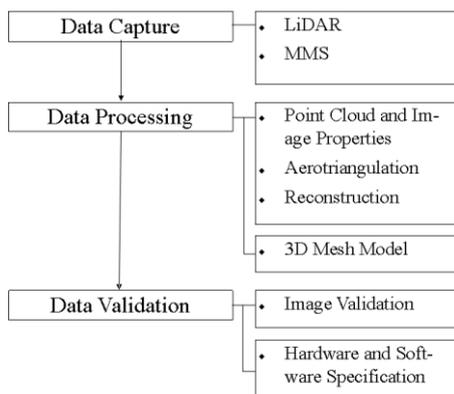


Figure 1: Flowchart of the methodologies

### 2.1.1 Data Capture

Two types of technique were used that are LiDAR point cloud data and MMS 360 image. The following procedure must be adhered during the data acquisition, data overlapping, camera model, projected pixels size, focal length, exposure, lighting, image retouching, photogroup and masks.

### 2.1.2 Data Processing

Before any processing can be done, the data acquired has to be verified whether the necessary requirement for ContextCapture software is fulfilled. The properties of point cloud and image are necessary for the accuracy of the 3D model. Important properties for point cloud are the trajectory files that linked the trajectories and point cloud through time stamp. MMS image required spatial reference systems, X (easting), Y (northing), Z (height or altitude) and time. The properties are necessary for the images, position and rotation or pose, pose metadata, component and mask. The image position and rotation delineate the estimation for the 3D reconstruction. Accurate estimation of the image is important for the 3D model accuracy. Pose is entered using initialization from pose metadata or entered manually. The pose metadata is the metadata of the position and rotation which identify the known pose of the photograph. This metadata can be imported from GPS tags, third-party data, or during block import. Component is an important as only image in main component of the block can be used for 3D reconstruction.

Aerotriangulation process is the process of knowing the accurate photogroup properties of each input photogroup and the pose of each input photograph. It will compute position and rotation of every images and then all of the images are computed in the main component for the reconstruction process. Every image position and rotation will be calculated from the metadata to be used in the reconstruction process. As the image is already in one component, the software automatically grouped them in the main component. 337 images for the Taman Perindustrian Suajana Indah with no control points and tie points are used for this study. Control points are entered manually, or some are imported from columned file to support accurate geo-referencing and avoid long-range metric distortion image. It can only be used in the aerotriangulation process if it consists of three or more control points, with each of the control points has two or more image measurements. In the aerotriangulation process in the camera calibration, grid distortion (Figure 2a), photo position uncertainty (Figure 2b), scene coverage (Figure 2c), tie point uncertainties (Figure 2d), number of image observing the tie points (Figure 2e), reprojection error (Figure 2f) and point resolution (Figure 2g), survey can be identified.

The result of the camera calibration result can be seen from the distortion grid in Figure 2a. The grey lines represent the zero-distortion grid, while blue lines represent the real camera values. The image is distorted to the upper right. The distortion is caused by the relative camera position. A photo position uncertainty is the software ambiguities position of the image acquired, that occurs because of the uncertain control and tie point. The table 1 below shows the minimum and maximum values of the image position and the Figure 2b shows the top view (XY plane), side view (ZY plane) and front view (XZ plane) of computed photo positions (black dots). Blue ellipses indicate the position uncertainty, scaled for readability.

	X [units]	Y [units]	Z [units]
Minimum	0.00024	0.00028	0.00034
Mean	0.00092	0.00143	0.00124
Maximum	0.00941	0.04629	0.01038

Table 1: Photo position uncertainties

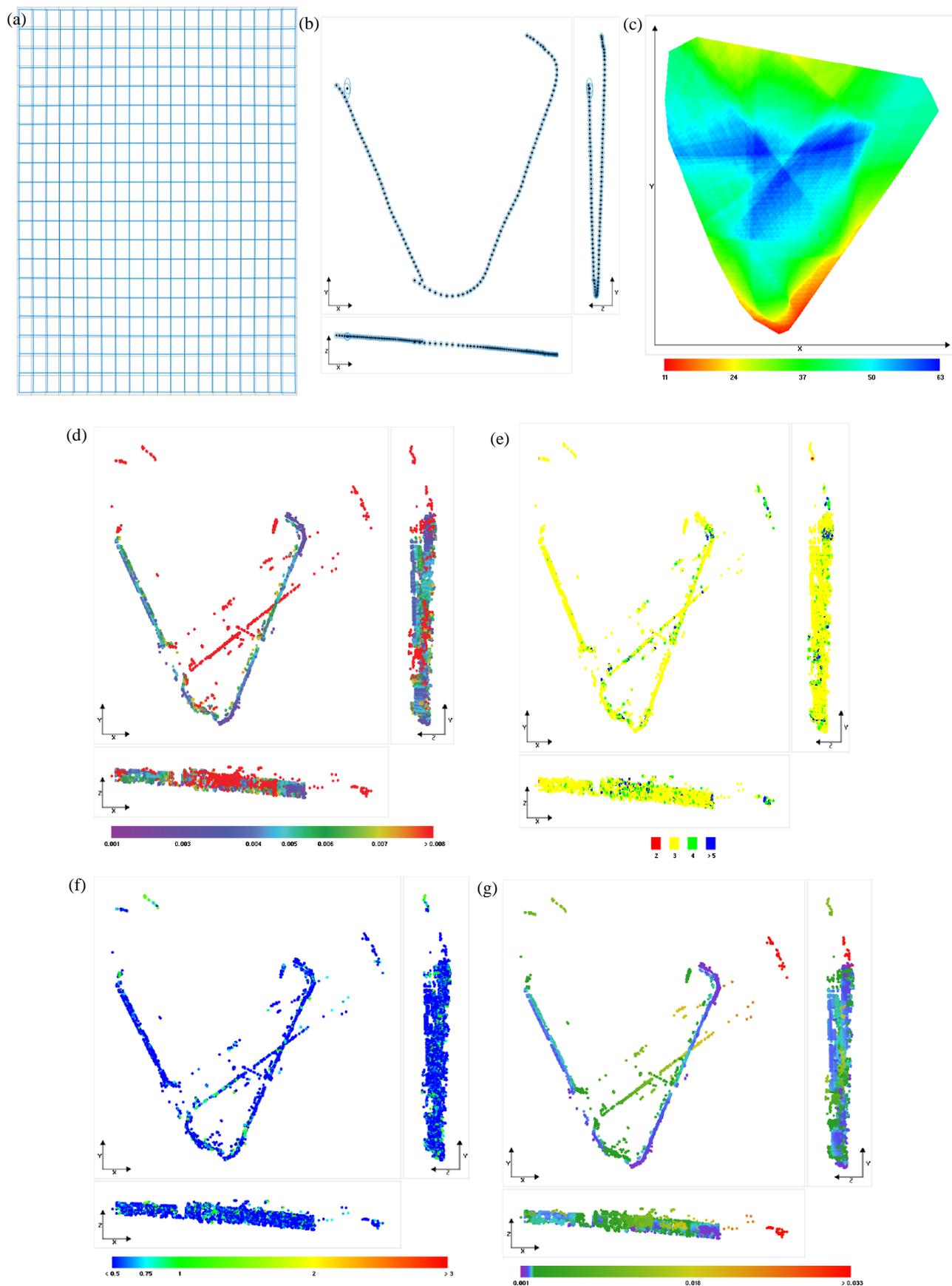


Figure 2a-g shows the aero-triangulation process report; (a) Distortion grid, (b) Photo position uncertainties, (c) Scene coverage, (d) Point uncertainties, (e) Number of photo observing tie points, (f) Reprojection error, and (g) Tie point resolution.

Scene coverage is the coverage that an image can see. Figure 2c shows the top view (XY plane) display of the scene, with colours indicating the number of photos that theoretically see each areas. There is a minimum of 11 images that cover the front view of the location indicating low coverage image. Tie points is the 2D correspondences resembling to the same physical point with an unknown coordinate that is defined manually. During the aerotriangulation process, software able to automatically generate a large quantity of automatic tie points. However, the tie point that automatically generated have lower photo matching than manually inserted.

No. of Points	Median No. of Photos per Point	Median No. of Points per Photo	Median Reprojection Error [pixels]	RMS of Reprojection Error [pixels]	RMS of Dist. To Rays [meters]
24869	3	630	0.2	0.43	0.00246

Table 2: Generated tie points

Tie point uncertainties derived from the position vagueness. Figure 2d shows the top view (XY plane), side view (ZY plane) and front view (XZ plane) displays all of tie points, with colours representing ambiguity in the individual point position. The values are in meters, with a minimum uncertainty of 0.00109 meters and a maximum of 1.1168 meters. The median position uncertainty equals 0.00434 meters. Number of observations per tie point in Figure 2e shows the top view (XY plane), side view (ZY plane) and front view (XZ plane) displays all of tie points with colours representing the number of photos that have been used to define each point. The minimum number of photos per tie point is 2 and the maximum is 9. The average number of photos observing a tie point is 3. Reprojection errors per tie point in Figure 2f shows the top view (XY plane), side view (ZY plane) and front view (XZ plane) displays all of tie points, with colours representing the reprojection error in pixels. The minimum reprojection error is 0.00 pixel and the maximum is 1.84 pixels. The average reprojection error is 0.33 pixels. Resolution in Figure 2g shows the top view (XY plane), side view (ZY plane) and front view (XZ plane) displays all of tie points, with colours representing resolution in the individual point position. The values are in meters, with a minimum resolution of 0.00124 meters and a maximum of 0.03283 meters. The median resolution equals to 0.00211 pixel. Survey is the position constraints based on the position, orientation or scale priors based on user tie points. They are used to perform a rigid registration of the block during aerotriangulation. The four types of constraints are origin constraints, scale constraints, axis constraints and plane constraints.

The reconstruction process is to manage a 3D reconstruction framework. The reconstruction is defined by the following properties; firstly, spatial framework that defines the spatial reference system, region of interest and tiling. Secondly, the reconstruction constraints that allows the use of existing 3D data to control the reconstruction and to avoid reconstruction errors. Next, the reference 3D model that acts as the reconstruction sandbox, that stores a 3D model in native format which is progressively completed as 3D model productions progress. Lastly, the processing settings that sets the geometric precision level and other reconstruction settings. They are five ways to represent 3D model using the ContextCapture software; firstly, 3D mesh that is generated into a 3D model, 3D point cloud that produces a coloured point cloud, orthophoto or DSM, that manufactures to interoperable raster layers. The 3D mesh, 3D point cloud and orthophoto are optimized for visualization and analysis. Next, retouching the 3D mesh which generated a 3D

model can be edited via third-party software and then imported it back into ContextCapture software for subsequent productions. Lastly, the only reference 3D model that produces a 3D model can only be used via ContextCapture Master software due to quality control and as a cache for subsequent productions.

### 2.1.3 Data Validation

The data are validated using the image that were obtained from the Google Maps. The 3D mesh model result and Google Maps images are being compared as to comprehend the differences and the factors that affect the results. Then, based on the 3D mesh model, the hardware and software specification are also suggested.

## 2.2 Results and Validation for 3D Mesh Model

### 2.2.1 3D Mesh for Putrajaya

This 3D mesh model for Putrajaya is using a point cloud dataset with the area size of 64 km<sup>2</sup> with data size of 2.77GB. The orthomosaic image has the area size 358 km<sup>2</sup> with data size of 1.24 GB.



Figure 3: 3D Mesh Model for Putrajaya

The Figure 3 above shows the 3D Mesh result with the user specification of 64-bit operating system (OS), Windows 10 version, Intel Core i7 CPU, Intel HD Graphics 4000 GPU, 4 GB RAM memory, 578 GB hard disk free space to produce the 3D Mesh model. It takes 23 hours to completely produce the 3D Mesh model, the result is a success as the 3D Mesh of whole area is successfully shown and the blank area is for the water body. During the reconstruction process, adaptive tiling is used to adaptively subdivide reconstruction into boxes to meet targeted memory usage. As, the computer use has a 4 GB RAM so it was appropriate to choose this type of tiling. This is a suitable tiling method to reconstruct a 3D model with a highly non-uniform resolution such as when reconstructing Putrajaya area from the aerial images and ground images of a few of landmarks. In such a case, it is not possible to find a regular grid size adequate for all areas. However, the minimum memory required by the software to process the data is 5.9 GB. Furthermore, as the data given are unicolor orthomosaic image, the result will show in monochrome colour. To produce the best visualization of 3D Mesh, colour orthophoto is needed with suitable memory to avoid slowing down the process.

### 2.2.2 3D Mesh for The National Heart Institute of Malaysia

The 3D mesh model for The National Heart Institute of Malaysia, only using one type of dataset that is a point cloud data with the area size of 0.2456 km<sup>2</sup> and data size of 0.031 GB.



Figure 4: 3D Mesh Model for The National Heart Institute of Malaysia.



Figure 5: Image for The National Heart Institute of Malaysia, (Source: Google Maps)

The Figure 4 above shows 3D Mesh result for The National Heart Institute of Malaysia, Wilayah Persekutuan Kuala Lumpur based on user specification of 64-bit operating system (OS), Windows 10 version, Intel Core i7 2.5 GHz CPU, NVIDIA GEFORCE GTX 850M GPU, 16 GB RAM memory, 268 GB hard disk free space and Figure 5 shows the comparison image from Google maps. To produce 3D Mesh in The National Heart Institute of Malaysia area, it takes 37 minutes to complete. As the data size is low, no tiling is required to process the data since the expected memory usage to produce the model is 1.2 GB and allows extra precision of the processing mode. Unfortunately, as only one type of data is given, the best visualization cannot be produced. The 3D Mesh result shows several blank spots in the upper area of the building that need to be covered using other data such as orthophoto. To produce the best visualization of 3D Mesh, it is best to combine several types of data from different sources such as from orthophoto, aerial image and point cloud. To support the statement, Figure 5 is captured by Google Maps to compare with Figure 4. The blank spot is mostly at the top-view and side-view of the building.

### 2.2.3 3D Mesh for Taman Perindustrian Saujana Indah

The 3D mesh model for Taman Perindustrian Saujana Indah is using MMS image as the dataset with the ground coverage of 10370.51km<sup>2</sup> and data size of 1.24GB.



Figure 6: 3D Mesh Model for Taman Perindustrian Saujana Indah.



Figure 7: Image for Taman Perindustrian Saujana Indah, (Source: Google Maps)

The Figure 6 above shows 3D Mesh result for Taman Perindustrian Saujana Indah, Shah Alam, Selangor based on the same user specification in Figure 4 and Figure 7 that show the comparison of images from Google Maps. To produce 3D Mesh Taman Perindustrian Saujana Indah, it takes 37 minutes to complete. Regular planar grid is used as the tiling and extra precision of the processing mode. However, as the only one type of dataset is used, the 3D visualization shown is terrible. When the 3D Mesh result is compared to the Figure 5, it obviously shows several blank spots in the upper area of the building and the side view of the building. The result also shows the difference between the dimension between the 3D mesh model and real image.

## 3. HARDWARE AND SOFTWARE SPECIFICATION

The hardware and software specification are done based on the operating system, processor, storage, memory, graphic card of the hardware and the data size, memory, graphic card, tiling and precision of the software while producing 3D mesh model based on the experimental data.

### 3.1 Data size

All of the data are processed using the same computer with the specification of operating system Window 10 x64, Processor Intel i7, memory of 16GB and Graphic Card NVIDIA GEFORCE GTX850M. The comparison is made from the different ground area and the processing time. Extra precision processing mode is used for all of the ground area size, however, as for ground area with small size, no tiling is required. However, adaptive tiling are used for moderate and large ground areas.

It takes 15 minutes for 3D Mesh modelling process to finish for a small ground area (0.1346 km<sup>2</sup>). Then, for moderate (0.2456 km<sup>2</sup>) and large (4.551 km<sup>2</sup>) ground area are using adaptive tiling and takes 2 hours and 25 minutes and 6 hours 12 minutes to finish processing respectively. Based on the information, the Figure 8 below shows the comparison between processing time and the size of data.

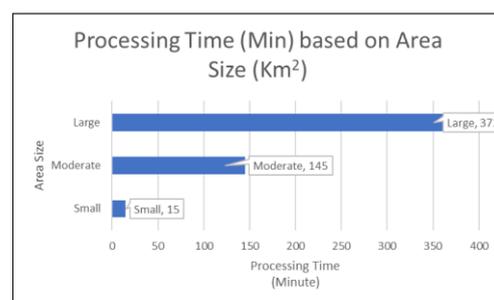


Figure 8: Comparison between processing time based on area size.

The Figure 8 shows that, the bigger the area size, the higher the time taken to process the data. The area size is classified as small, moderate and large for easier comparison between the area size and the processing time. The small area with the size of 0.1346 km<sup>2</sup> requires 15 minutes for the 3D Mesh model to complete while the moderate area with the size of 0.2456 km<sup>2</sup> takes 145 minutes for 3D Mesh model to complete and lastly, the large area with the size of 4.551 km<sup>2</sup> takes 372 minutes for 3D Mesh model to complete.

Size for moderate area is doubled from the size of small area with the difference of 0.111 km<sup>2</sup> while the time taken for moderate area size to finish processing is 130 minutes longer than the small area. This is related to the types of tiling as small area size does not use any tiling while moderate size uses adaptive tiling. The tiling processing mode contributes to the increase of processing time. Furthermore, by comparing between large area size and moderate area size, the different between two area size is only 4 km<sup>2</sup>. However, the processing time is differed by 225 minutes. Even though the different between the area size is large, the processing time is lesser than when comparing it to small area size. This is shown that the area size and the type of tiling are affected by the processing time.

### 3.2 Memory

The processing time between two memories are being compared and the time taken for data processing to finish is noted. The data is using extra precision processing mode and as The National Heart Institute of Malaysia has small area size, no tiling is needed. Processing time for 4GB memory and 16 GB memory to finish processing 3D Mesh model are 15 minutes and 19 minutes, respectively. A graph is made to show the processing time taken based on the memory.

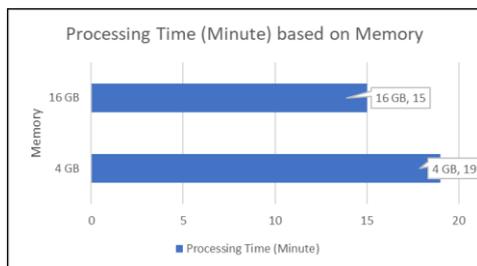


Figure 9: Comparison between processing time based on memory.

As shown on the Figure 9 above, the higher the memory, the lesser the time taken to process the data. The two types of computers that were used for data processing and 3D modelling have two memory size which are 4 GB and 16 GB. When generating 3D mesh models, the 16 GB memory spends lesser time in aerotriangulation process compared to 4 GB memory size. Aerotriangulation is a process needed by ContextCapture software to perform the 3D reconstruction from photographs. ContextCapture software required accurate photo group properties for each input photo group and the pose of each input photograph. The bigger the area size, more time is needed for aerotriangulation, thus increases the processing time to create 3D Mesh model.

### 3.3 Graphic card

Two different computers with the same memory of 16GB and a good quality of graphic cards are being compared. Each computer is processing the same dataset using an extra precision processing

mode and adaptive tiling. The result shows that the processing time for NVIDIA GEFORCE GTX 1070 is 2-hour 25 minutes, while NVIDIA GEFORCE GTX 850M is 3-hour 17 minutes. The result is shown in the figure below.

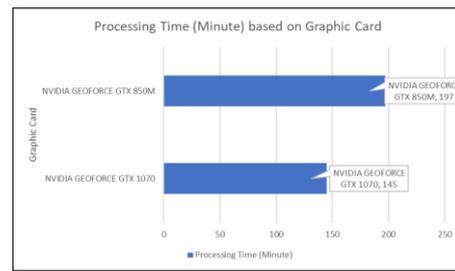


Figure 10: Comparison between processing time based on graphic card.

Figure 10 shows that NVIDIA GEFORCE GTX 1070 process 3D mesh model faster than NVIDIA GEFORCE GTX 850M. When comparing between the performance of both graphic cards, NVIDIA GEFORCE GTX 1070 spends lesser time in reconstruction and production process when producing 3D Mesh Model. Thus, a graphic card with a higher graphic and computing performance is essential for smoother and faster visualization of 3D model.

### 3.4 Tiling

The data is processed using the same computer with the specification of operating system Window 10 x64, Processor Intel i7, memory of 16GB and Graphic Card NVIDIA GEFORCE GTX850M. Same dataset is used with extra precision processing mode. The large tiles area of 51 meters with adaptive tiling takes 4 hours and 46 minutes to process 102 tiles, moderate tiles size of 25 meters takes 6 hours to process 388 tiles whereas small tile size of 6.6 meters needs 31 hours and 25 minutes to process. Both moderate and small tiles size use the regular volumetric tiling.

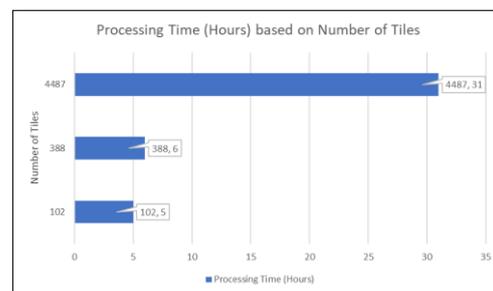


Figure 11: Comparison between processing time based on number of tiles.

Figure 11 shows that the higher the number of tiles, the longer the processing time is. Two types of tiling used are adaptive tiling and regular volumetric tiling. For adaptive tiling, it adaptively subdivides reconstruction into boxes to meet a target memory usage. Thus, the software will be automatically calculate based on the size of memory, what is the maximum number of tiles. However, minimum memory required by ContextCapture for smooth processing is 5.9 GB. Regular volumetric tiling will divide reconstruction into cubic tiles. Reconstruction process will be made based on the tiles area size inputted. The bigger the area tile size, the lower the number of tiles thus the processing time

will decrease. Hence, the smaller the tile size, the lower the area with blank spot is.

### 3.5 Processing mode

Same computer is used to process the same dataset with the different types of precision, which are medium, high, extra and ultra. No tiling is required for medium precision, thus it only takes 31 minutes and 50 seconds to complete the 69% process of utilizing CPU usage. High precision is using regular planar grid tiling to produce 4 tiles with 200-meters grid size each and 70% CPU utilizing usage. It takes 34 minutes and 43 seconds processing time. Extra and ultra-precision is also using regular planar grid tiling with 1 tile with the grid size of 400 and 16 tiles with the grid size of 100 meters each. Extra precision takes 35 minutes 57 seconds and ultra-precision takes 3 hours and 30 minutes.

Precision Mode	Type of Tiling	Number of Tiles	Grid Size (meter)	Processing Time
Medium	No Tiling	0	-	31Min 50Sec
High	Regular Planar Grid	4	200	34Min 43Sec
Extra		1	400	35Min 57Sec
Ultra		16	100	3H 30Min

Table 3: Comparison on processing time and the precision mode.

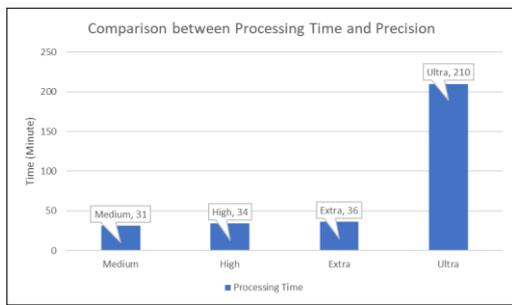


Figure 11: Comparison between processing time based on number of tiles.

Figure 11 shows that the higher the precision, the longer the time taken to process the data. The precision is classifying as medium, high, extra and ultra-based on the reconstruction process in ContextCapture. As shown above, medium precision takes 31-minute, high precision takes 34-minute, extra precision takes 36 minute and ultra-precision take 210-minute processing time. The medium, high, and extra precision have similar processing time which in between 31 minutes to 36 minutes, which each take about 1 to 2-minute differences compare to ultra-precision that take 190 minutes longer. Based on the table above, it shows that the ultra-precision has the highest number of tile (16 tiles) and the lowest grid size (100 meters).

Medium precision has the fastest time because it takes less memory to process. It was suited for orthophoto and DSM productions as it has no different for the location accuracy and use 2-pixel tolerance for the input image. For the high precision, it is suitable for small input data size as it uses 1-pixel tolerance for the input image while extra precision is suitable for large file size with 0.5-pixels tolerance in input image. The high and extra precision modes are similar in terms of computation time and memory consumption. Thus, the high and extra precision only

have a 2-minute time difference. Lastly, the ultra-precision take the longest time as it uses higher memory and computation time.

### 3.6 Hardware suggestion specification

From the studies, the following hardware specification are suggested to ensure uninterrupted data processing. The Windows 64-bit operating system has an user-friendly environment and highly compatible to perform with different software. As for the processor, the latest Intel Core i9 is known for its 3D modelling. However, the price of Intel Core i7 is more economic whereas it can support complex processing and modelling. As for the memory, Windows 64 Bit Intel Core i7 requires a minimum of 32 GB to work well, Required storage is vary on the size of data, however for optimized stage, the storage remained must be doubled than the data size. 64 GB RAM allows smooth rendering for processing work whereas from the studies, it shows that the higher the RAM, the lower the processing time in data processing. Lastly, the graphic card Nvidia Quadro P2000 is much cheaper than Nvidia GeForce Series with the same functional and is suitable for various multi-purpose computer workstations and mid-range rendering, CAD work and design.

## 4. DISCUSSION

Development of 3D modelling and visualization can be associated to the development of smart cities (see Azri et al., 2016; Azri et al., 2014). Decision making for the development and maintenance of the smart cities can be affected without accurate and complete 3D model. The use of spatial data in making complex decision have been proved to be resourceful (Mohd et al., 2016). Thus, a way for the complete 3D model to be developed is important. Azri et al. (2015) demonstrates that this 3D model which classified as vector spatial data can be quantified for further analysis. However, while developing the 3D model, several technical issues and challenges have occurred, affecting imperfect 3D model. The ContextCapture has described the issues while acquiring the 3D data in acquisition report as shown in the picture below.



Figure 3: Warning issues by the software in acquisition report.

From the acquisition report, it can be understood that the issues that occur in data processing is connected to the flawed technique in data acquisition. For example, images in the folder are in portrait orientation as the image rotation or the auto rotation camera are not deactivated during data acquisition. These can affect the 3D model precision and performances during data processing. Another issue is when data failure in rigid registration

from photo positioning because of incoherent GPS tags as the ground control points are not accurately pointed. Effecting the reconstruction process as some photo cannot be used because of unknown coordinate points. Thus, the final 3D mesh model has blank or empty spot.

To counter this issue, this point must be highlighted while acquiring 3D data, firstly, image must be overlapped during the acquisition. Each part of the subject or object should be photographed from at least three distinct but not radically different viewpoints. The overlapping between consecutive photographs should typically exceed two thirds. Different viewpoints of the same part of the subject should be less than 15 degrees apart. For simple subjects, it can be achieved by taking approximately 30-50 evenly spaced photographs all around the subject. For aerial photography, a longitudinal overlap of 80% and lateral overlap of 50% or more are recommended. To achieve best results, acquire both vertical and oblique photographs in order to simultaneously recover building facades, narrow streets and courtyards. In addition, to ensure best performance and precise 3D model, the geo-referencing, and ground control point must be accurate and the photo orientation should be in landscape orientation. Secondly, to ensure perfect coverage of 3D model, the use of several data acquisition techniques is required. Point cloud data can be used together with aerial image from UAV and MMS. When dataset from several sources is combined, it can cover each other for smoother 3D model as when only one type of dataset is being used, the result of 3D mesh model is containing missing or empty spot.

## 5. CONCLUSION

In the changing urban environment, modern technologies are necessary for the economics, environment and social relations. The smart cities are supporting the economics decision, sustainable environments and navigating social relations (Vinod Kumar, 2020). To achieve that, various initiatives have been done by the government and public sector to implement public 3D visualization (Zakhary et al., 2020; Tuncer et al., 2019; Hosseinali et al., 2019). The main purpose of the 3D model is to allow smoother visualization of 3D model for public services, transportation and security surveillance.

Furthermore, the complement of suitable hardware and software are important for the 3D model outcome to be perfect. Most of the software is written about the minimum requirement to install the software but does not mention the effect on the configuration of the hardware when using the application or software. When developing a 3D model, other than the data acquired/collection// technique and data processing procedure, the hardware and software are also important. Based on this research, when using higher capability of computer hardware, the results are produced in shorter time and lesser technical issues. However, the results on the visualization of 3D mesh model is almost the same or not different from each other.

## ACKNOWLEDGMENTS

This research was partially funded by UTM Research University Grant, Vot R.J130000.7652.4C332 and Vot Q.J130000.3552.06G41.

## REFERENCES

Azri, S., Anton, F., Ujang, U., Mioc, D., Rahman, A.A., 2015. Crisp Clustering Algorithm for 3D Geospatial Vector Data

Quantization, Lecture Notes in Geoinformation and Cartography. Springer Verlag, pp. 71-85.

Azri, S., Ujang, U., Abdul Rahman, A., 2018. Dendrogram Clustering for 3D Data Analytics in Smart City. International Archives of Photogrammetry, Remote Sensing and Spatial Information Science. XLII-4/W9, 247-253.

Azri, S., Ujang, U., Anton, F., Mioc, D., Rahman, A.A., 2014. Spatial Access Method for Urban Geospatial Database Management: An Efficient Approach of 3D Vector Data Clustering Technique, 9th International Conference on Digital Information Management (ICDIM). IEEE, Bangkok, Thailand.

Azri, S., Ujang, U., Castro, F.A., Abdul Rahman, A., Mioc, D., 2016. Classified and clustered data constellation: An efficient approach of 3D urban data management. ISPRS Journal of Photogrammetry and Remote Sensing 113, 30-42.

Bentley Development Team, 2019. Bentley ContextCapture Master Software. <https://docs.bentley.com/LiveContent/web/> (11 January 2019).

Cheng, Liang & Tong, Lihua & Li, Manchun & Liu, Yongxue. 2013. Semi-Automatic Registration of Airborne and Terrestrial Laser Scanning Data Using Building Corner Matching with Boundaries as Reliability Check. Remote Sensing. 5. 6260-6283. 10.3390/rs5126260.

Cheng, Liang & Chen, Song & Liu, Xiaoqiang & Xu, Hao & Wu, Yang & Li, Manchun & Chen, Yanming. 2018. Registration of Laser Scanning Point Clouds: A Review. Sensors. 18. 1641. 10.3390/s18051641.

Hosseinali, F., Khosravi Kazazi, A., 2019. Developing a VGI method for 3D city modeling based on CityGML and Open Data Kit. Earth Observation and Geomatics Engineering 3, 54-63.  
Lafioune, N. and M. St-Jacques (2020). "Towards the creation of a searchable 3D smart city model." Innovation & Management Review.

Liang, F., Yang, B., Dong, Z., Huang, R., Zang, Y., & Pan, Y. 2020. A novel skyline context descriptor for rapid localization of terrestrial laser scans to airborne laser scanning point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, 165, 120-132.

Leng, B., et al. (2010). "A 3D shape retrieval framework for 3D smart cities." *Frontiers of Computer Science in China* 4(3): 394-404.

Mohd, Z.H., Ujang, U., 2016. Integrating Multiple Criteria Evaluation and GIS In Ecotourism: A Review. International Archives of Photogrammetry, Remote Sensing and Spatial Information Science. XLII-4/W1, 351-354.

Pirasteh, Saied & Li, Jonathan. 2016s. Landslides investigations from geoinformatics perspective: quality, challenges, and recommendations. Geomatics, Natural Hazards and Risk. 8. 1-18. 10.1080/19475705.2016.1238850.

Salleh, S., Ujang, U., 2018. Topological information extraction from buildings in CityGML. IOP Conference Series: Earth and Environmental Science 169, 012088.

Tse, Rebecca & Gold, Christopher & Kidner, Dave. 2008. 3D city modelling from LIDAR data. 10.1007/978-3-540-72135-2\_10.

Tuncer, O., Ergun, G., Mustafa, O., Osman, S., Nilhan Ciftci, S., Ayhan, T., Erdal, Y., Alpaslan, T., 2019. Oblique Aerial Image Acquisition, 3D City Modeling, 3D City Guide Project for Konya Metropolitan Municipality, in: Information Resources Management, A. (Ed.), Architecture and Design: Breakthroughs in Research and Practice. IGI Global, Hershey, PA, USA, pp. 1371-1385.

Uden, M. and A. Zipf (2013). Open Building Models: Towards a Platform for Crowdsourcing Virtual 3D Cities. *Progress and New Trends in 3D Geoinformation Sciences*: 299-314.

Vinod Kumar, T.M., 2020. Smart Environment for Smart Cities, in: Vinod Kumar, T.M. (Ed.), Smart Environment for Smart Cities. Springer Singapore, Singapore, pp. 1-53.

Vosselman, George & Maas, Hans-Gerd. 2014. Airborne and Terrestrial Laser Scanning.

Yang, Byungyun. 2019. Developing a Mobile Mapping System for 3D GIS and Smart City Planning. *Sustainability*. 11. 3713. 10.3390/su11133713.

Yoshimura, R., Date, H., Kanai, S., Honma, R., Oda, K., & Ikeda, T, 2016. Automatic registration of MLS point clouds and SfM meshes of urban area. *Geo-spatial Information Science*, 19, 171 - 181.

Yusoff, I.M., Ujang, U., Rahman, A.A., Katimon, A., Ismail, W.R., 2011. Influence of georeference for saturated excess overland flow modelling using 3D volumetric soft geo-objects. *Computers & Geosciences* 37, 598-609.

Zakhary, S., Rosser, J., Siebers, P.-O., Mao, Y., Robinson, D., 2020. Using unsupervised learning to partition 3D city scenes for distributed building energy microsimulation. *Environment and Planning B: Urban Analytics and City Science*, 2399808320914313.