Design and Prototyping of a 3D Scanner

Maida Talat
Department of Electrical and Computer Engineering COMSATS University, Islamabad Park Road, Islamabad, 45550, Pakistan.
maidatalat@gmail.com

Maryam Zulfiqar
Department of Electrical and Computer Engineering COMSATS University, Islamabad Park Road, Islamabad, 45550, Pakistan.
maryamzulfiqar55@gmail.com

Dr. Omar Ahmad
Department of Electrical and Computer Engineering COMSATS University, Islamabad Park Road, Islamabad, 45550, Pakistan.
omer_ahmed@comsats.edu.pk

Asma Ramay
Department of Electrical and Computer Engineering COMSATS University, Islamabad Park Road, Islamabad, 45550, Pakistan.
asma.ramay@comsats.edu.pk

Abstract— **3D scanning is the process of converting a real-world object or an entire structure, into a three-dimensional virtual model. The primary goal of a 3D scanner is to create point clouds based on geometric samples on the subject's surface. 3D scanners are extremely expensive and they are not made in Pakistan. Importing them here costs hefty delivery charges which varies according to the size and weight of the scanner, as well as a prolong waiting time of around a month. The scope of this project was to design and develop a low cost 3D scanning device capable of generating a 3D point cloud model of the object. The design included a turntable on which the object to be scanned was positioned, as well as an arm holding the camera that was used to scan the object. Scanning techniques like photogrammetry using smartphone and DSLR and 3D scanning using Kinect were used to maintain the cost and performance of scanner. In photogrammetry software Meshroom and VisualSFM were used to create a 3D point cloud of the object from 2D images. In 3D Scanning Kinect with Skanect software was used to generate a 3D model of objects of different nature under different lighting conditions. The results of photogrammetry indicate that minimum focal length required for the creation of point cloud is 35mm below this focal length Meshroom showed no result while VisualSFM gave poor results. However with photographs having a focal length of 39mm good results were obtained in both softwares. The results of the 3D Scanning showed that Kinect works well for scanning smooth surfaces, but requires sufficient lighting condition also shinny and transparent surface objects can not be modeled directly without utilizing any scanning spray or powder, or in low illumination situations. Kinect can scan objects as small as 2x2 inches in size and as large as a room. The resulted model can be exported to any other software for further use and also for 3D printing.**

Keywords— 3D Scanning, Photogrammetry, Meshroom, VisualSFM, Skannect software, Kinect.

I. INTRODUCTION

A 3D scanner is a device that creates digital 3D models of real-world objects or environments based on information collected about their shape and possibly appearance (i.e. color).The main objective of a 3D scanner is to generate point cloud based on the geometric samples on the surface of the subject. The shape of the subject is determined through these points by a process called reconstruction. If color information is gathered at each point, the colors on the subject's surface can also be evaluated.

3D scanners are very similar to cameras. Like cameras they have a cone-shaped field of view and can only capture information about unobscured surfaces. Each point in the 3D scanner image provides details about the distance to the subject's surface as well as identifying the three-dimensional position of each point in the image.

3D scanning is currently one of the world's most rapidly evolving technologies. They reduce time in the design phase, speed the prototype process, quickly and fully regulate quality, their ability to reproduce parts, and the capacity to compare designs with manufactured goods is very efficient. Various work has been done in the field of 3D scanning.

Jeremy Straub described the effectiveness of the scanner for both non-living and human scanning. Given the significance of lighting to visible light scanning systems, the scanner's output under a variety of lighting configurations was evaluated, defining its sensitivity to lighting design (Jeremy Straub, 2015)[1]. Suavarna Raude designed and created a 3D scanner that used an infrared distance sensor to scan the object by plotting 90*90, or 8100 points [1][2]. A mesh was created using an algorithm based on the idea of a point cloud which was rewritten in the format of a '.STL file'. Guilia Rogati designed and validated the new Microsoft Kinect sensor-based 3D foot scanner to provide a 3D scan of plantar feet that showed optimal intra-operational repeatability for many clinical applications, was perfect to produce 3D foot plantar surface scans and aided to develop individual sets and shoes for pricey laser scanners, as a legitimate low-cost solution. (Giulia Rogati, 2019)[3].

II. BACKGROUND AND MOTIVATION

3D scanners are extremely expensive and they are not made in Pakistan. Importing them here costs hefty delivery charges, which varies according to the size and weight of the sanner, as well as prolong waiting time of around a month. Following table shows some 3D Scanners models on amazon and their prices.

Table 1 3D Scanner Models and Prices (Amazon Best Sellers)[4]

3D Scanners Models	Prices
SIMAX3D Array Block Education Desktop 3D Scanner	$1,199.00
Newest EinScan SE Desktop 3D Scanner-Dual-Mode Fixed and Auto Scan	$1,399.00
2021EinScan SP Desktop 3D Scanner	$2,499.00

The reasons behind the expensive costs of 3D Scanners are:

- Firstly the whole money we pay for 3d scanners is not available to manufacturers because they don't directly sell them to customers. Instead, they build up a worldwide distribution network.
- Secondly, hidden costs, which include the production and marketing of the gadget, are included in the scanning price.
- Lastly, the most famous and successful professional 3D scanners do not sell the millions they sell the dozen every month and produce them in low number.

The inspiration of doing this project was the high cost of 3D Scanners and their unavailability in Pakistan. During this project we have designed and developed a low cost 3D scanning device capable of generating a 3D point cloud model of the object. The design included a turntable on which the object to be scanned was positioned, as well as an arm holding the camera that was used to scan the object.

In order to maintain the cost and performance of 3D Scanner following technologies were tested:

i. The photogrammetry is to take several photographs in various angles of an object. This can include a conventional camera, no requirement for infrared sensor fantastic cameras to achieve good results.

ii. A Kinect 3D scanning used to determine the depth fields of a single object using the same technology that could have a medium range 3D scanner, camera, and infrared camera.

To guarantee the final goals were attained, following objectives were kept in mind.

- A 3D scanning device connected to a PC.
- Software for interfacing the scanner to the PC.
- Total cost of the system to be kept below $150.

III. TECHNIQUES

During our project we have designed and developed an active 3D scanner that uses Kinect for 3D scanning which employs structured light method along with triangulation to scan the object and uses the concept of stereo vision to create the 3D model of the object.

A. Active Scanners

Active scanners transmit some sort of radiation or light and detect its reflection or radiation passing through object to test an object or atmosphere.

B. Structured Light 3D Scanners

In structured light scanners a light pattern is projected on the surface of the object and a camera slightly offset from pattern projector is used to look for the deformation of the pattern caused by the object and thus position of each point in its field of view is calculated and distance to an object is determined.

C. Stereo vision

Stereo vision is the process of extracting 3D information from digital images. Two cameras separated horizontally from one another are used in traditional stereo vision to obtain two different views of a scene. Thus by comparing information about a scene from two vantage points 3D information can be extracted by examining the relative positions of objects in the corresponding image planes.

The figure 1 below in left is an epi-polar geometry which have two cameras positioned at O and O' respectively. Point O is the left camera projection center and O' is the right camera projection center these two cameras are separated from one another by a known distance. Points X, X1 and. X2 are the object points in real world. The plane that is formed by joining the points of camera projection centers and the 3D point X is called as epi-polar plane. Epipolar lines are formed by joining the points where this epi-polar plane intersects the image planes.

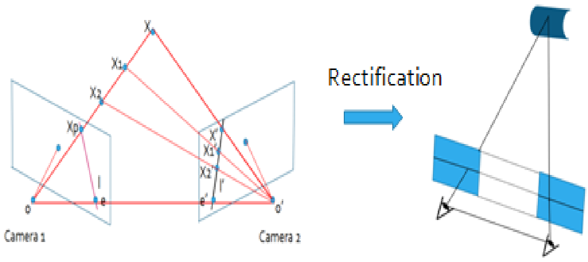

Figure 1 Stereo Vision Algorithm

The figure 1 shows two epi-polar lines l and l' and also two epi-poles e and e' in the left and right image planes respectively. An epi-pole is an intersection point of image plane with the baseline. Corresponding to the three different object points there is single projection point xp in the left image plane showing that the depth information is lost here since the camera could not differentiate 3 different points in the real world along the same line and at what distance they are from each another. This problem is resolved in stereo vision systems by the use of second camera. As the figure shows the right image plane has three different projection points x', x1' and x2' corresponding to the three object points x, x1 and x2 respectively. So, from this information depth information can be inferred from these 2D images by using triangulation process. This is the fundamental concept of stereo vision.

This configuration is called the general camera configuration in which the epi-polar lines are not parallel, so in order to find correspondence between the points in images and to find depth of the object, 2D search is required i.e. in horizontal and vertical directions. This problem of correspondence search can be simplified by image rectification. Rectification corresponds to the parallel camera configuration. It makes all the epi-polar lines parallel, so the image planes of cameras are parallel to each other and to the baseline.

The right figure 1 shows the canonical stereo configuration obtained from that rectification. In this case the 2D searching problem is converted to 1D search so correspondence between points in images can be find out by seraching along the particular epi-polar lines fall which falls along the horizontal scan line of the images. The corresponding relationship between the projection points of the same point in different images can be found after obtaining the projection of the object point on both image planes. This is known as stereo matching.

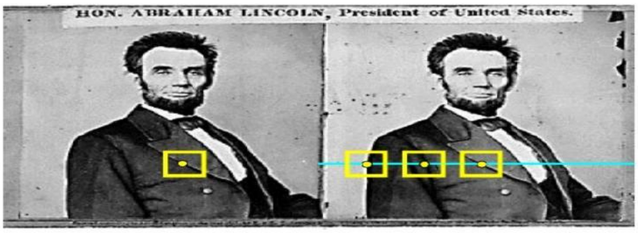

Figure 2 Stereo Matching Algorithm

In the right image of figure 2 locate the epi-polar lines for each corresponding pixel in the left image. Then choose the best match by inspecting all of the pixels along the epi-polar line. In the end compute disparity by triangulating the matches.

D. Triangulation

Depth can be computed from disparity (the distance between a pixel and its horizontal match in the other image) using triangulation, which is the process of calculating the 3D scene point from its projected points in the two image planes.

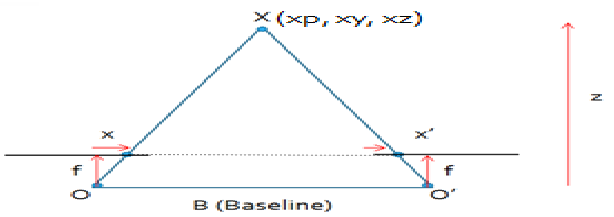

Figure 3 Depth from Disparity

The figure 3 shows the stereo geometry in the rectified configuration where cameras are positioned at same height and have identical focal length to view or capture the scene. Points O and O' are the projection centers of left and right camera respectively, which are separated by a known distance of baseline B. Point x and x' there are the projection points of object point X on left and right image planes and are also called image points.

A view at cameras situation from the bird-eye perspective can give a connection between x, x' and z in order to find depth of object point X. For each camera equate the ratios of horizontal (x) and vertical (z) distances at which cameras see the projection of object point to find the depth equation. The object point x has three co-ordinate points xp, yp and xz

.

$$\frac{z}{f} = \frac{xp}{x} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 1$$

$$\frac{z}{f} = \frac{xp-B}{x'} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 2$$

Here equation 1 is written from the perspective of left camera while equation 2 is written from the perspective of right camera which is at a distance B from the left camera. In these equations xp is the horizontal co-ordinate point of object point x in the real world. Now extracting the values of xp from equation 2 and substituting in equation 1 will give the next equation.

$$\frac{x-x'}{O-O'} = \frac{f}{z}$$

This equation x- x' is the disparity i.e. horizontal distance between similar image points in left and right image planes. O-O' is the distance between camera projection centers, f is the focal length which is defined as the distance between the camera projection center and the image plane and z is the distance between camera projection center and the object point x in 3D. z is also known as depth.

By rearranging the above equation disparity can be find out which is inversely proportional to depth. Hence depth of object point x can be determined.

$$\text{Disparity} = x\text{-}x' = \frac{Bf}{z}$$

$$\text{Depth} = z = \frac{Bf}{x-x'}$$

E. Kinect Working

In our project 3D scanning is being done by kinext which is a depth and motion sensing device produced by Microsoft which employs the use of an RGB camera and a depth sensor to scan the object. Normally in stereo vision two cameras are used to facilitate the triangulation process however kinext uses structure light method to enable the use of a depth sensor comprising of an IR projector and an IR sensor to capture depth of the objects.

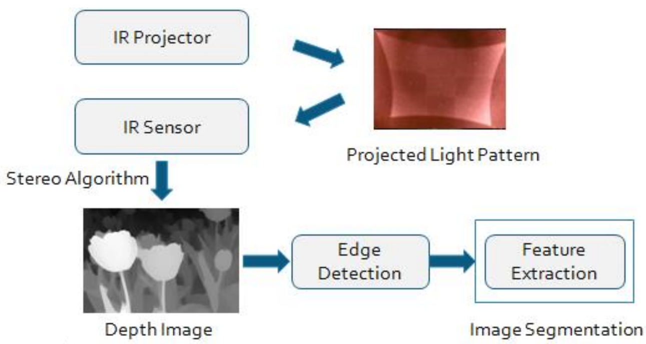

Figure 4 Kinect Working

The IR projector projects a wide near-infrared pattern facing the Kinect which is recieved by Ir sensor. A rather low profile of the items in front of it distorts the light pattern and a stereo-algorithm produces the profound image (Lun, 2018)[5]. A stereo triangulation technique extracts a deeper image. For Kinect's pics of a pixel depth image, Kinect employs a certain edge detection to distinguish close things from a backdrop, incorporating an input of the conventional visible light camera. The division of pictures is one of the most important strategies utilized for the sensing and tracking of object movements. For further analysis, the image is divided into small segments or groups of subsets of pixels. Image segmentation is done which involves feature extraction which divides an image into segments or a set of pixels to allow for a separation of the required functionality from the background or surrounding items.

We have also done work on passive scanners based on the photometric systems using the concept of photogrammetry to create 3D point cloud of the objects.

F. *Passive Scanners*

Passive Scanners do not emit any sort of radiation but rather relies on the detection of ambient radiation as reflected. Most such scanners detect visible light since the ambient radiation is readily accessible. Other radiation types could also be used, such as infrared. This approach can be very inexpensive, since these kinds of scanners do not require specific hardware but simple digital cameras in most cases.

G. *Photometric Systems*

They generally utilize a single camera, but capture many images under varied lighting circumstances. These methods attempt to reverse the picture formation model in order to determine the surface orientation of each pixel.

IV. HARDWARE

The hardware block diagram of 3D scanner provides the gist of its working. Processor is the most important component here which controls the hardware. A processor is a piece of hardware that decodes and processes an action whenever an instruction is received and delivers an output which in this case is driving the motor controller assembly.

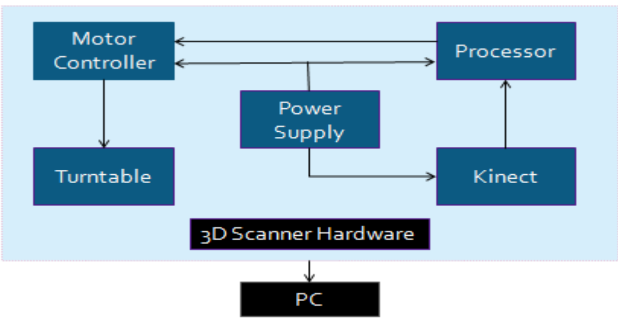

Figure 5 Block Diagram of 3D Scanner Hardware

The motor controller is an assembly of Arduino connected to the driver which drives the motor which in turn rotates the turntable. As the turntable rotates to change the position of the object the Kinect will scan the object and the processor will process it through software instructions for further creation of point cloud or 3D model in PC. Power supply is given to each and every component to work at their full potential.

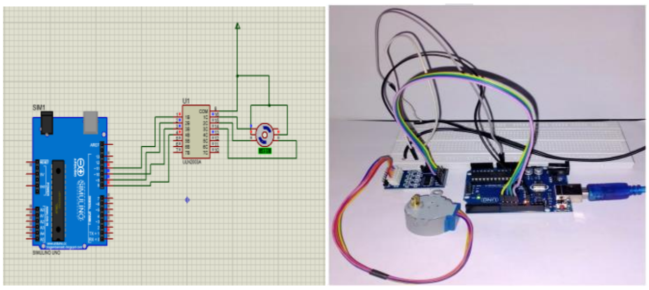

Figure 6 Proteus Stimulation and Hardware Implementation of Turntable Controller

Proteus is an electrical circuit design, sketching, and stimulation tool. The turntable controller module has been tested on Proteus first before its implementation on hardware because it takes less time than practical circuit construction and also possibility of error is lower in software stimulation.

In our design of turntable controller, four output pins (8, 9, ~10, ~11) of Arduino are connected to the four input pins (4B, 2B, 3B, 1B) of ULN2003 driver board and the motor cable from the 28BYJ-48 stepper motor is hooked on the motor connection area of driver board.

The Arduino board is energized by the PC which in turn energizes the driver. But as the motor needs an external power source capable of providing sufficient current for achieving the desired load torque. An extra 5 Volt source is provided to it in common with the driver. The code was uploaded on Arduino, the driver gets signal from Arduino and converts it to the motor driving steps and will drive the motor accordingly which will in turn control the turntable.

Figure 7 shows the hardware design of 3D Scanner designed on CAD software Inventor. An object placed on the turntable of 3D scanner will be scanned by Kinect. The depth images of the object from different angles will be captured by Kinect and processed in conjunction with the software to create 3D model of the object. This 3D model can further be send to the 3D printer for 3D printing.

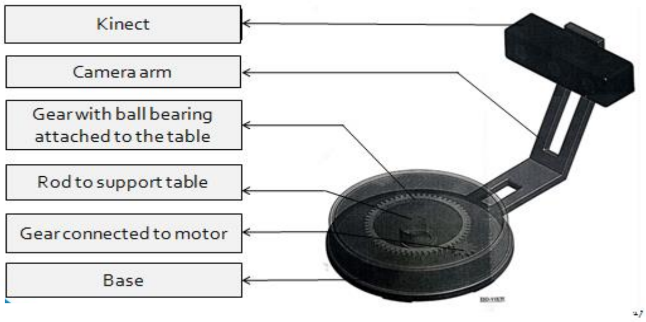

Figure 7 Initial Hardware Design of 3D Scanner

The turnatable has a base plate with a rod fixed at the center to hold the rotating plate. In order to allow rotation of rotating plate a large gear with ball bearing inside is attached to the rod. The ball bearing is free from outside where it is attached to the gear and is fixed from inside where the rod mounted to the table is attached. The rod is to hold the table and the ball bearing is to reduce rotational friction and to support the load of the table. The large gear is fixed with the table and also it is attached to the small gear the small gear is mounted over the stepper motor. When the stepper motor rotates, the small gear will rotate which in turn rotates the large gear. As the large gear is fixed with the table so due to the rotation of large gear table will rotate.

The initial hardware design had a camera arm placed at a distance of 16" from the turntable and at a height of 3" from the base. Kinect was not scanning well with these dimensions of camera arm only few object points were scanned that too erroneously while completely missing the others. So, new hardware camera arm was designed and build with a distance of 25" from turntable and at a height of 1.5" from the base while keeping the already build turntable of the previous design. The Kinect was places on camera holder of width 3” and a length of 4” at a distance of 22” from the turntable. These dimensions were obtained by placing the Kinect at various distances and heights and the dimensions at which it scanned well were recorded.

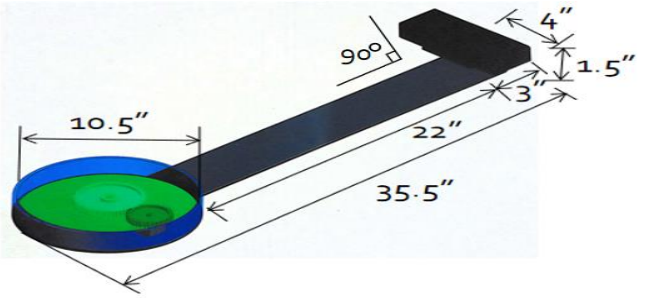

Figure 8 Modified Hardware Design of 3D Scanner

Figure 9 shows the original hardware of 3D Scanner build during the course of the project which is fully functional mainly comprised of turntable controller assembly and a camera arm holding Kinect.

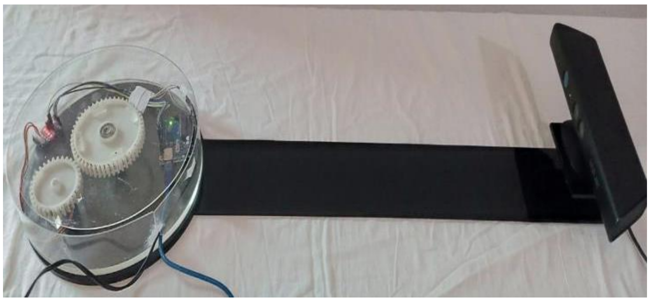

Figure 9 3D Scanner Hardware

V. SOFTWARE

We have used two different techniques to create 3D point cloud and model of an object. The first is photogrammetry, while the second is 3D scanning. Photogrammetry is the process of collecting accurate information about physical objects and the environment by capturing, measuring and analyzing photographic images. It involves two main steps, capturing photographs and then constructing 3D point cloud from those photographs. We have used two different softwares for that purpose Meshroom software and VisualSFM software. For both softwares two cameras of different focal length were used for photography, smartphone camera of OPPO A5 2020 with focal length 2.75mm and DSLR camera of focal length 39mm.

The second technique 3D Scanning uses 3D Scanner which in some ways uses video of the thing and used to measure the depth of the field using infra-red sensor. It will have to be linked concurrently to the computer software that transforms the real time video into a 3D model. We have used Kinect with Skanect software for 3D scanning. Skanect software transforms the depth scan captured by Kinect into 3D model. We have used two different methods for 3D model generation. First keeping Kinect still and rotating object using turntable and second, keeping object still and moving Kinect around it. We have also tested scanning for shinny and transparent objects and in low lighting situations.

H. Photogrammetry using Meshroom software

Meshroom is an open-source, free, aerial, close-range photogrammetry software. It uses following functions to create 3D point cloud from images.

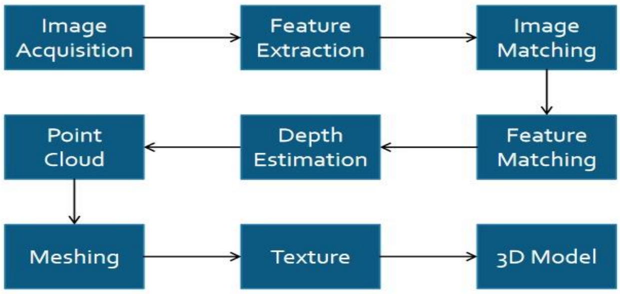

Figure 10 Processing Flow Chart of Meshroom Software

We simply drag and drop already taken and saved images from folder to the “Images” area of Meshroom software. Software will then extract features from images, the objective of feature extraction is to extract group of pixels or features that are invariant to changing camera view points during photography. After feature extraction, image matching will be done by matching corresponding features between images captured at different angles of the same object or scene. Then depth will be estimated by computing depth value of each pixel of the scene to create an accurate and detailed digital surface model of the object or scene under consideration. All depth estimated points will be combined to create a 3D point cloud. After that, meshing will be done by creating polygons that link in a sequence of lines and points to approximate the geometry of a digital model. Texturing is the process of creating and adding textures (set of color pixels that make up the image) to a three-dimensional model. Texturing gives the ultimate 3D model.

I. Photogrammetry using VisualSFM software

VisualSFM is a 3D reconstruction graphical user interface based on structure from motion (SFM). It uses following functions to create a 3D point cloud from 2D images.

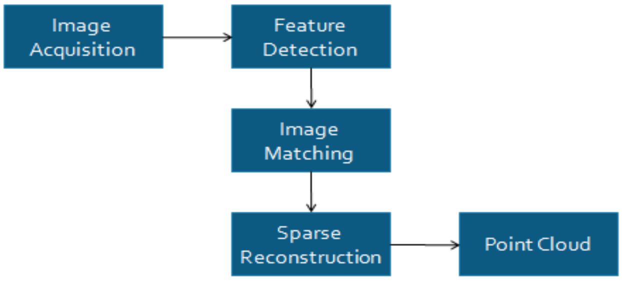

Figure 11 Processing Flow Chart of VisualSFM Software

We need to add already taken photographic images to the workspace of VisualSFM software and wait for the images to load. The Log Window monitors the process. Images should show in the workplace once this procedure is completed. Software will then recognize features on the images using the SIFT technique in order to match the images. After image matching, during sparse reconstruction bundle adjustment of the set will be performed. The images are displayed in their real orientation in the workspace, and a point cloud for each image is generated and shown.

J. 3D Scanning using Kinect with Skanect software

The object to be scanned is placed on the turntable which is rotating at very low speed. Kinect is focused on the turntable and connected to the computer to use with the Skanect software to scan the object. Skanect software uses 3D cameras, such as the structure sensor and structure core to create 3D model of 3D objects which cut hardware costs to a fraction of the prior 3D scanning solutions and also allow for online sharing of models.

We have also done scan by letting object still and moving Kinect around it for the sake of comparison. The distance between Kinect and the center of the turntable is approximately 0.548 meter. Objects of different size and nature have been be scanned in different lighting conditions.

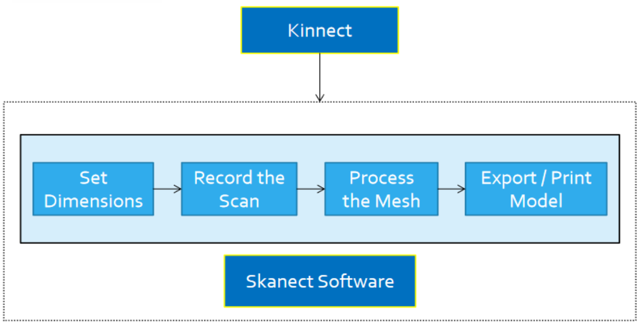

Figure 12 Block Diagram of 3D Scanning using Kinect with Skanect Software

To begin power on the Kinect and connect it to any USB port on the computer, then check in the Skanect software; if the software detects the Kinect, it indicates it is ready for scanning. We are now working with the Skanect software. To begin, we must start a new project and adjust the dimensions to match the size of the object to be scanned. The next step is to record the scan. Kinect will scan the object and determine its depth and colors. Skanect software will stitch together depth images captured by Kinect to generate a 3D mesh. The mesh obtained after recording has many irregularities that must be removed by post processing. It takes a lot of steps such as filling holes, remove parts, move, crop, colorize etc. After post processing, model will be ready to export or print.

VI. RESULTS

We have used two different techniques to create 3D point cloud and model of an object. The first is photogrammetry, while the second is 3D scanning. Results of which are as follow:

K. Photogrammetry

For Photogrammetry two different software's has been used Meshroom software and VisualSFM software, for both software's two different cameras were used for photography Smartphone camera of OPPO A5 2020 and DSLR camera.

1) USING SMARTPHONE CAMERA OF FOCAL LENGTH 2.75mm

Figure 13 depicts computation error caused by Meshroom software with images taken from smartphone camera because the images lacked necessary information required to create point cloud due to shorter focal length of camera.

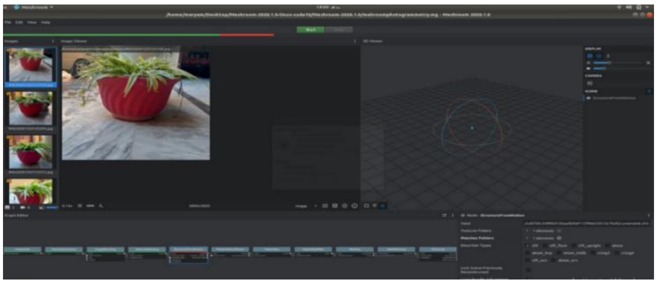

Figure 13 Computation Error

Figure 14 depicts object and its incorrect or sparse point cloud generated by VisualSFM software.

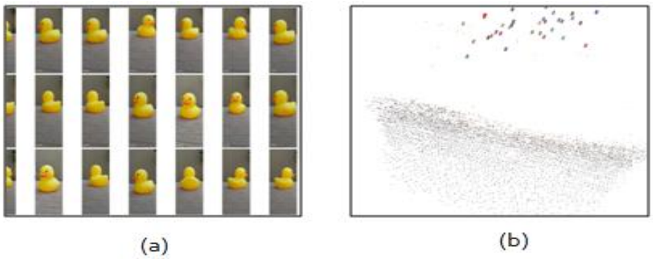

Figure 14 Shows (a) Object, (b) Sparse Point Cloud

2) USING DSLR CAMERA OF FOCAL LENGTH 39mm

Figure 15 depicts object and the resulted point cloud formed by Meshroom software.

Figure 15 Shows (a) Object, (b) Point Cloud

Figure 16 depicts object and the resulted sparse and dense point cloud generated by VisualSFM software.

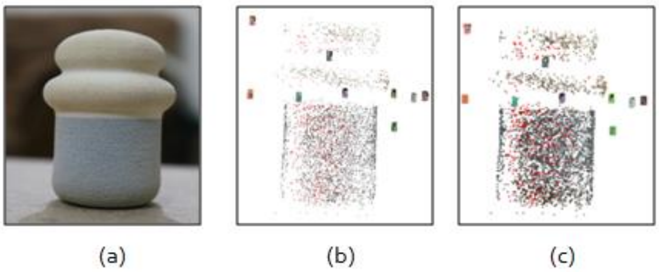

Figure 16 Shows (a) Object, (b) Sparse Point Cloud, (c) Dense Point Cloud

The results of photogrammetry indicate that minimum focal length required for the creation of point cloud is 35mm below this focal length Meshroom showed no result while VisualSFM gave poor results. However with photographs having a focal length of 39mm good results were obtained in both softwares. This means that Photogrammetry with Meshroom/VisualSFM software's is simple to conduct with professional cameras or high-quality smartphone cameras.

L. 3D Scanning using Kinect with Skanect software

For 3D scanning with Kinect using Skanect software. various types of objects were scanned in various lighting conditions by either rotating the object and leaving the Kinect steady, moving the Kinect and leaving the object stationary.

Figure 17 depicts 3D model generated keeping Kinect still and rotating object using turntable.

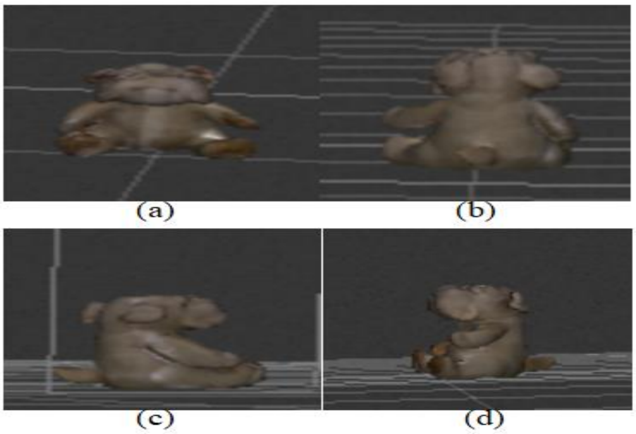

Figure 17 3D Model of Toy Dog (a) Front, (b) Back, (c) Right and (d) Left Views

Figure 18 depicts 3D model generated keeping object still and moving Kinect around it.

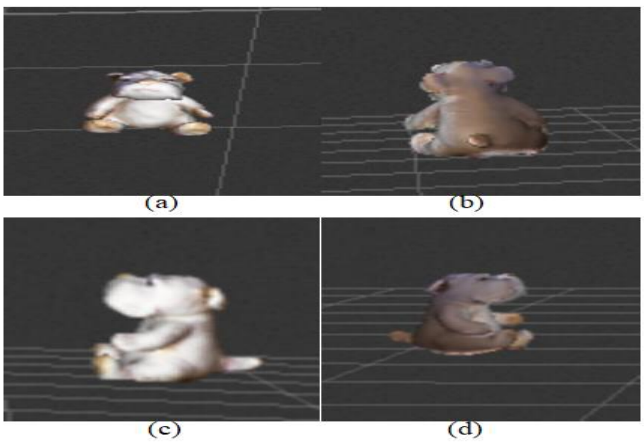

Figure 18 3D Model of Toy Dog (a) Front, (b) Back, (c) Left and (d) Right Views

Figure 19 depicts results of scanning transparent, shinny objects and scanning in low lighting situation.

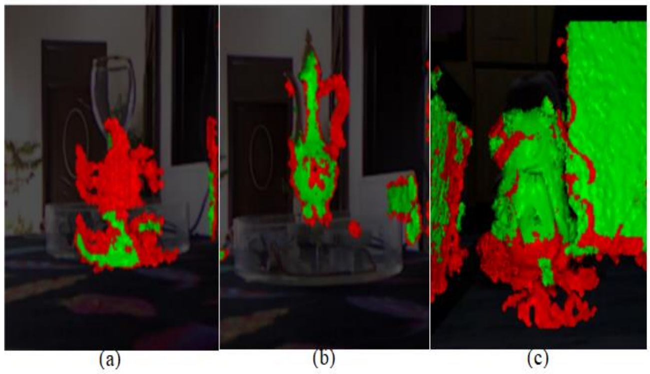

Figure 19 Scanning (a) a Transparent Glass, (b) a Shinny Kettle, (c) Toy Dog in Dark

In case of 3D scanning results showed that kinect works well for scanning smooth surfaces, but it cannot model shinny or transparent surface objects directly without the use of a scanning spray or powder and it also does not work well in low lighting conditions.

VII. Conclusion

To generate the 3D model of the object, various approaches such as photogrammetry using Meshroom and VisualSFM software's and 3D scanning with kinect using Skanect software were employed. Various types of objects were scanned in various lighting conditions by either rotating the object and leaving the Kinect steady, moving the Kinect and leaving the object stationary, but scanning can be done easily by leaving the Kinect stationary and rotating the object with a turntable. Kinect can scan things as small as 2x2 inches in size and as large as a room. The model created can be exported to any other software for further use, including 3D printing. The total cost of the 3D Scanner hardware turned out to be $140 which is very cost effective compared to readymade 3D Scanners.

Our current project is the initial prototype of a product that we intend to release as a product; many enhancements can be made over time and in response to customer feedback. Following are the improvements that could be made to this project in future:

- Using Pro version of photogrammetry/scanning software's will allow you to export or print 3D model with full geometry.
- We built the turntable to hold objects weighing roughly 1KG however it may be adapted to hold heavier objects.

References

[1] Jeremy Straub, B. K. (2015). Characterization of a Large, Low-Cost 3D Scanner. *Technologies*, 19-36.

[2] Suvarna Rode, A. K. (2019). Design and Development of 3D Scanner. *6th International Conference On Production and Industrial Engineering (CPIE-2019)*. Jalandhar-144011, India: Suvarna Rode.

[3] Giulia Rogati, A. L. (2019). Validation of a novel Kinect-based device for 3D scanning of the foot plantar surface in weight-bearing. *Journal of Foot and Ankle Research*.

[4] *Amazon Best Sellers*. (n.d.). Retrieved from Amazon: https://www.amazon.com/Best-Sellers-Industrial-Scientific-3D-Scanners/zgbs/industrial/8481412011.

[5] Lun, R. Z. (2018). *Human Activity Tracking and Recognition Using Kinect Sensor*.