Anbar Journal of Engineering Science (AJES) **Vol. 14 , No. 1 (2023) , pp. 31 ~ 39**
P-ISSN: 1997-9428 ; E-ISSN: 2705-7440 DOI: http://doi.org/10.37649/aengs.2023.137118.1027

Unviersty of Anbar

Anbar Journal of Engineering Science©

journal homepage: https://ajes.uoanbar.edu.iq/

3D Reconstruction by Structure from Motion Approach for Heritage Documentation

Hameed Sarhan Ismael[a], Salam Qaddoori Al-Zubaidi[b], , Khalida Kadhim Mansor[c], Atiya Al-Zuheri[d]

[a] *Department of Civil Engineering / University of Technology / Iraq*
[b]*Industrial Engineering Department / University of Pisa /Italy*
[c,d] *Department of Production Engineering and Metallurgy / University of Technology / Iraq*

PAPER INFO

Paper history:
Received 26/11/2022
Revised 20/01/2023
Accepted 28/02/2023

Keywords:
3D Reconstruction; Simulation; Heritage Documentation; Structure From Motion; Image Processing.

ABSTRACT

A fundamental score of this paper is to explain in detail how to create a 3D-provided modeled scene by data obtained at minimal cost to the client or users by manufacturing a smart, automated system for heritage documentation (SAS-HD). The steps can be classified by manufacturing, parts connection and simulation, selection of work sites, and obtaining data. The most important acquiesced data are digital images, which are fundamentally used by MATLAB's Structure from Motion (SfM) approach. The obtained images were subjected to sequenced tips by getting 3D sparse points of each object. This article has considered two objects in an indoor case study: the first feature is Ishtar Gate, and the second one is the winged ball inside the Iraqi museum in the Baghdad capital. The results are promising; hence Structure from Motion SfM method has been utilized to document heritage by manipulating 3D models on MATLAB interphase, which is approved for its efficiency and quick, super advanced processing steps. Heritage must be documented to provide a chance for restoration when needed to protect it from extinction; at times it is considered a tourist interface that reflects the culture of countries. So the proposed system (SAS) will be used for heritage documentation. The proposed approach allows training a wider audience of those interested in Iraqi and even global heritage by digitizing historic places with high-level accuracy, low cost, and flexible strategy.

1. Introduction

In Computer Vision (CV) and visual perception, structure from motion SFM points to the generation of three-dimensional scenes of targets by testing local motion signals. Structure from motion generates the 3D model from the motion cue. Moreover, it shows the position and orientation of the cameras used to capture the scene; this the SfM algorithm has computed this information in 3D reconstruction. (Three-dimensional digitalization)

* Corresponding author: Salam Qaddoori Al-Zubaidi; salam.al-zubaidi@phd.unipi.it ; +393276366264

[1]. In this context, active approaches consist of conserving the 3D structure (three-dimensional point cloud) by using its energy source as 3D laser scanners or structured light, which are highly cost techniques and difficult to adapt [7].

Additionally, passive methods allow the recovery of the 3D structure only from images taken from different viewpoints by one or more cameras. The considered approach in this paper is recovering passive method since a single camera is utilized for digital image acquisition at sequenced viewpoints. The work aims to build a smart system to document heritages in Baghdad capital of Iraq. Three locations have been elected to scan. The manufacturing of the system is an invention to convoy laser scanning concepts in imaging while the procedure of obtaining data to be effective with images is as same as for Kinect, that there are a group of laser sensors or Rangefinders that measure the distance to specific points on an object in addition to RGB-color images. So these data will participate together to obtain 3d scaled images of scanned object [2].

Computer vision-based approaches are popular when applied to combine various sensors in one system. Computer Vision (CV) can be used for a range-sensor multi-sensor system [3]. Recently, combining a laser sensor or Laser Range Finder (LRF) with a camera using the computer vision method modeling is a popular technology [4].

2. 3D Reconstruction Techniques

There are various human vision methods to acquire depth information using different sensory cues, such as the stereo capturing method, 3D laser scanner, 3D from shading, shape, Kinect technology, and our studying by utilizing structure from motion (SfM). The image of any observed object is projected from different views on both retinas, and their displacement can be used to triangulate the object's position. Many people think this is the only source of depth information, but this is false. The human vision system can be obtained normally. The human visual system constantly adapts to different light levels when viewing scenes. Provides a visual adaptation model that supports the generation of 3D model.

2.1 Structure from Motion (SfM)

Structure from Motion (SfM) is a fundamental approach that has participated in the heritage documentation revolution. SFM represents a method for fully automatic and robust estimation of two-view geometry, auto-calibration, and 3D metric reconstruction from point correspondences in images taken by cameras with a wide circular field of view. Many programs were invented to implement 3D reconstruction by (SFM) approach [9]. More precisely, SFM systems take as input a set of images and generate 3D point clouds or sparse points of the scene, the position from which each photo has been taken, and the optical parameters of the camera. Our approach depends on a single camera with four Laser Range Finder (LRF) sensors. These digital parts move in front of an object around its rotation axis while settling on a tripod, as in Figure 1.

[a] [b]

Figure 1. (a) SAS-HD testing, (b) leveling bubbles for balancing with the horizon

In more detail, SFM will clarify in the following sections to generate 3D sparse points.

2.1.1 Features Detection

In computer vision and image processing, the concept of features is generally used for denoting a single piece of information related to solving the computational tasks related to a certain application. There are various algorithms for feature detection, but the most popular are based on Scale-invariant feature transform (SIFT) [8].

(SIFT) is a feature detection algorithm utilized to detect and describe local features in images. David Lowe discovered it in 1999. SIFT key points of objects are first extracted from a set of reference images and stored in a database for objects recognized in an image by individually comparing each feature from the new image to this database and finding candidate-matching features based on the Euclidean distance of their feature vectors. Object matches that pass all these tests can be identified as correct with high confidence [6]

2.1.2 Matching features process

Matching features is an algorithm that leads to recognizing a feature in the rest of the available images. This is a necessary process due to the necessity of computing the triangulation in a later step. One way to avoid the search for all pixels in every image is to use Epipolar lines (Torres et al., 2012). Hereon a Method has been followed to produce 3D range reconstruction, generating 3D ranges by analyzing the surrounding snapshots from different views. Good matching can establish density depth clouds, rendering it more accurate and easier to smooth and merge the surface of the 3D range [10].

3. Materials and Methods

The work scheme aims to manufacture a Smart Automated System for Heritage Documentation or by symbols SAS-HD. The building step is boiled down to assembling electrical and mechanical subsystems and the main digital parts (Digital Camera DC and Laser Rangefinders LRF sensors); secondly, it's important to automate and simulate these parts together to control them properly. Moving parts (DC and four LRF sensors) were settled on a tripod; that stepper motor (SM) is responsible for their motion by 15-degree angle horizontally to capture sequenced digital images. Testing of SAS-HD has been achieved in Iraqi museums by selecting many sculptures. Hereon, Ishtar gate and winged ball sculptures have been considered by this article.

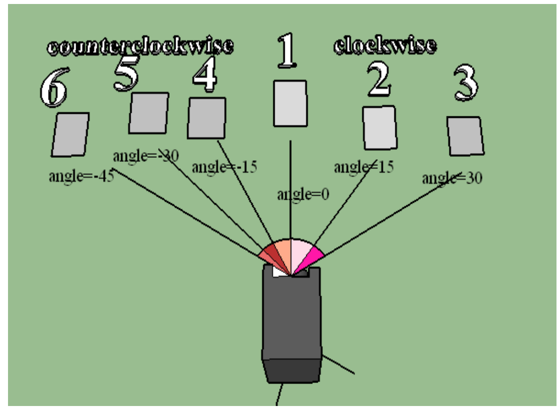

Figure 2. Imaging Technique.

Figure 2 shows that the work technique has been implemented on an object in two cases; a clockwise direction and a counterclockwise one. The stepping angle is equal to a 15-degree angle. The acquired images are shown in figure (3), which is the Ishtar Gate object in Iraqi Museum (indoor case).

Figure 3. SAS-HD testing – digital images obtained on Ishtar gate feature in the Iraqi museum (indoor area study)

The manufacturing of the system is an invention to convoy laser scanning concepts in imaging while the procedure of obtaining data to be effective with images is as same as for Kinect, that there are a group of laser sensors or Rangefinders that measure the distance to specific points on an object in addition to RGB-color images. So these data will participate together to obtain 3d scaled images of scanned objects. Computer vision-based approaches are popular when applied to combine various sensors in one system. Computer Vision (CV) can be used for arrange-sensor multi-sensor systems. Recently, combining a laser sensor or Laser Range Finder (LRF) with a camera using the computer vision method modeling is a popular technology.

3.1 Design and Implementation Process of the proposed Approach

The proposed approach structure is depicted in figure 4. The forthcoming subsections include details about this structure's components and how the included components can integrate into each other in the process. As we stated before, the whole approach aims digitization of heritage places cartography by a non- contact and harmless scanning approach.

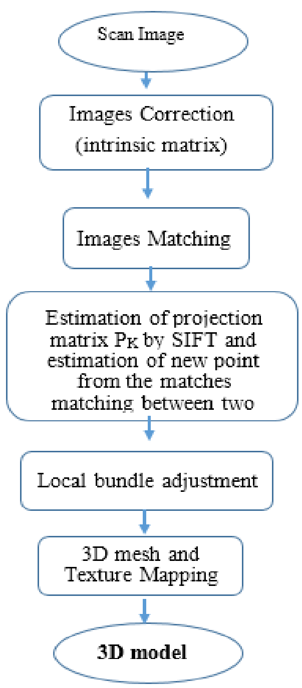

Figure 4. The proposed approach structure

3.1.1 Interest Points Detection and Matching

The suggested approach is based on detecting and matching interest points between consecutive pairs of images (Ik-1, Ik) that shows in eq (1). So, the Harris detector was used for the extraction of interest points. Since false matches (outliers) may accurse, a matching accuracy detector such as normalized cross-correlation NCC detector exists. SURF feature detector algorithm has been utilized to solve this problem and to estimate the fundamental matrix FK-1K. This estimation is dependent on the Epipolar constraint defined by the equation:

$$m^{T}_{iK} F_{K-1K} m_{ik-1} = 0 \ ….. \quad (1)$$

With (m_{ik-1}, m_{ik}) pair of matched points between images pair {I_{k-1}, I_k}.

3.1.2 Initialization

The initialization of the reconstruction process has been achieved by utilizing two images (I1, I2). It's used to recover a set of 3D sparse points from the result of the detected interest points matching (matches) and by the camera intrinsic parameters.

3.1.3 Intrinsic Parameters

Camera parameters were computed by camera configuration (camera calibrator) in MATLAB. This configuration leads to the computation of the intrinsic matrix parameters of the camera, presented by the matrix K.

K= fx 0 cx 0 fy cy 0 0 1 ……… (2)

3.1..4 Extrinsic parameters

The fundamental matrix F12 between the images I_1 and I_2 is recovered during the matching stage by the SURF algorithm, then the calculation of the essential matrix E_{12} by the formula (3)

$$E_{12} \sim KT_2 F_{12} K_1 \ ……….. \quad (3)$$

The extrinsic parameters, which are defining rotation (R) and translation (T) between two views (R_{12}, T_{12}) can be calculated from the essential matrix by equation (4):

$$E_{12} \sim [T_{12}]x \ R_{12} \ ……….. \quad (4)$$

3) Triangulation: this step allows to recover the 3D coordinates from points matched between images {I1, I2} and the projection matrices P1 and P2 defined by formula (5):

$$P1 \sim K1\,[I\ 0] \text{ and } P2 \sim K2\,[R12\ T12] .... \quad (5)$$

4. Demonstration of the Proposed Approach

The demonstration of the proposed approach is presented in two case studies related to heritage sites: Ishtar Gate and Human-Headed Winged Bull (Lamassu).

4.1 Case 1: Ishtar Gate –Heritage Site- Indoor Case

SFM is done by using two images from different views. Firstly entering images captured by various view angles equal to 15o degrees. Figure (5) shows entered images.

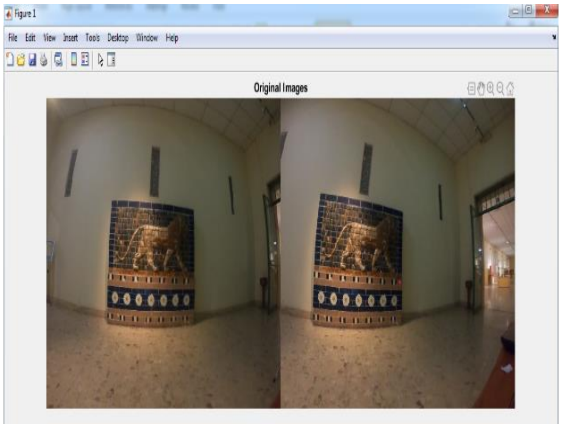

Figure 5. 1st and 2nd original Images (rotation angle=15°)

Remove distortion by intrinsic camera matrix to show the result in figure (6) below:

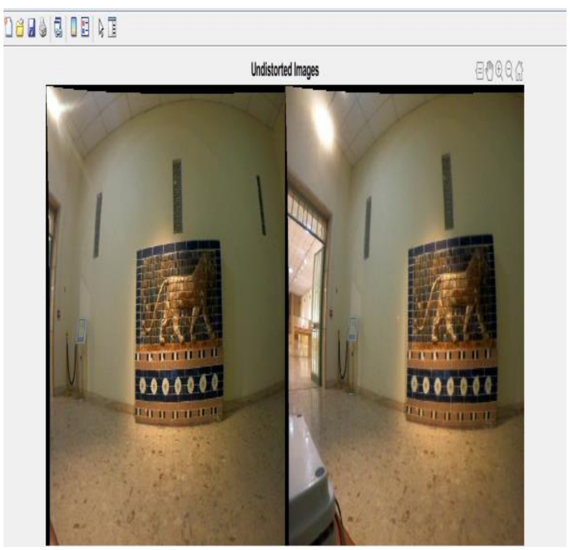

Figure 6. Undistorted Images by Intrinsic Parameters due to camera Calibration

By correcting images (figure 7) , one image is used to detect points by SIFTdetectorr:

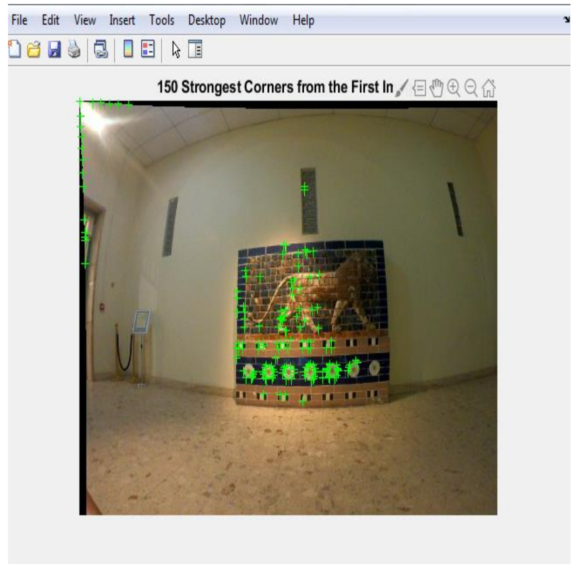

Figure 7. Auto-Selection of Edge Corner Points in 1st image

So in order to match them to their correspondings in the second image. By matching the implementation between features at each image, the reconstruction of the 3D scene will be applied as in the figure 9. The resulting 3D points (figure 9) is about 3671 point by their three coordinates (X, Y, Z).

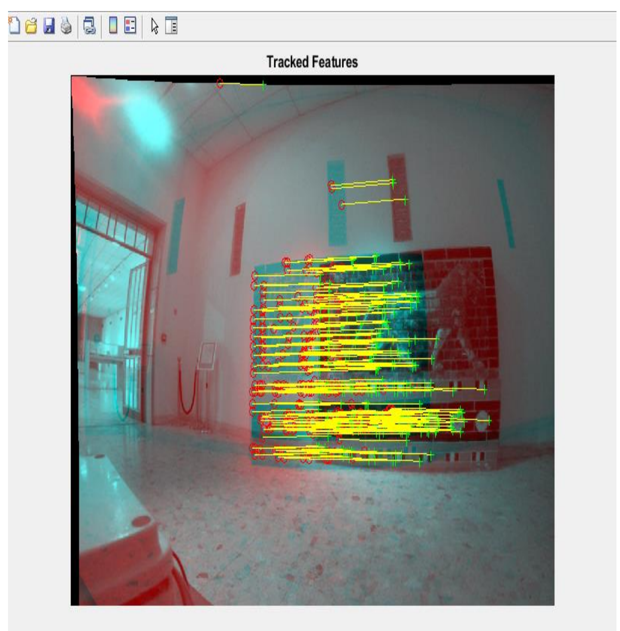

Figure 8. Tracking of Matching Points in First Image by their Correspondings in the Second Image.

Figure 9. 3D Reconstruction of Scene- Indoor Heritage Sites

4.2 Case 2: The Human-Headed Winged Bull (Lamassu)- Indoor Casethe – the second feature
The second feature is also inside the Iraqi-Museum corridor. As implementation of reconstruction or modeling approach; (SfM), has been implemented by two images also of different views. The camera's rotation angle concerning each sequence of two views equals 15o degrees. Figure (10) shows the utilized images.

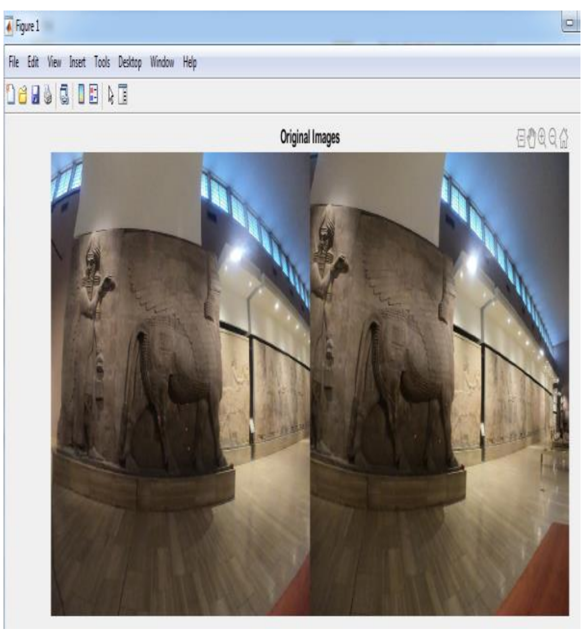

Figure 10. The original first and second images Matlab R2018b)

Also, eliminating distortion from two intrinsic camera parameters, shown in figure (11):

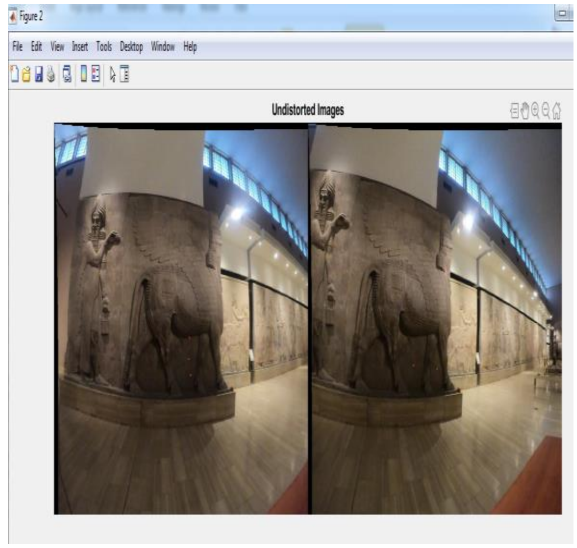

Figure 11. Undistorted Images by Intrinsic Camera Parameters.

Then point detection is done to create a matching between two images. By these points, a detector facilitates the authorship of correspondent features in a second added image (figure 13).

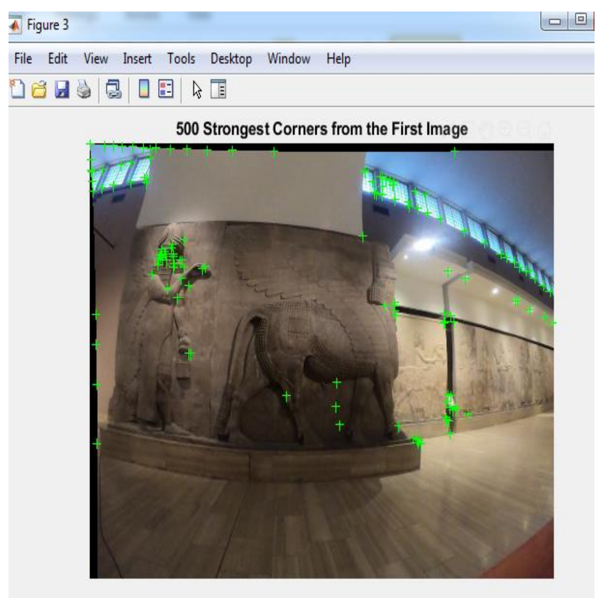

Figure 12. Auto-selection of Points to Match Images

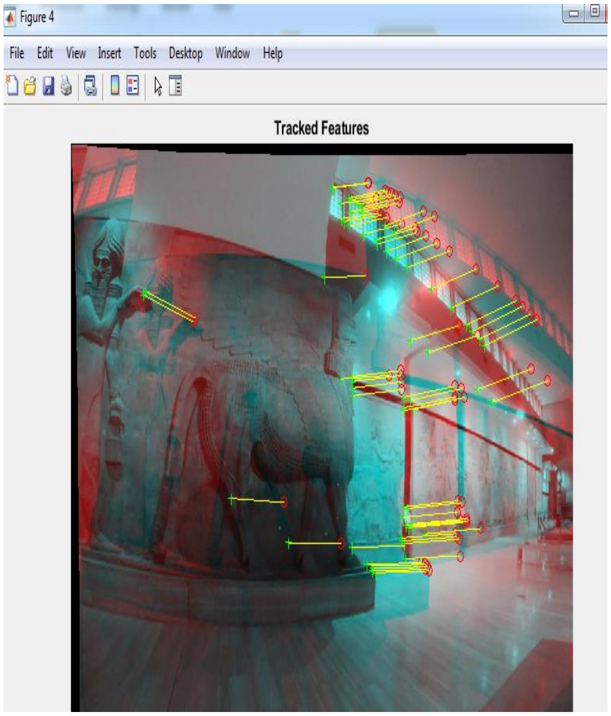

Figure 13. Tracking Features (points) –Matching Images

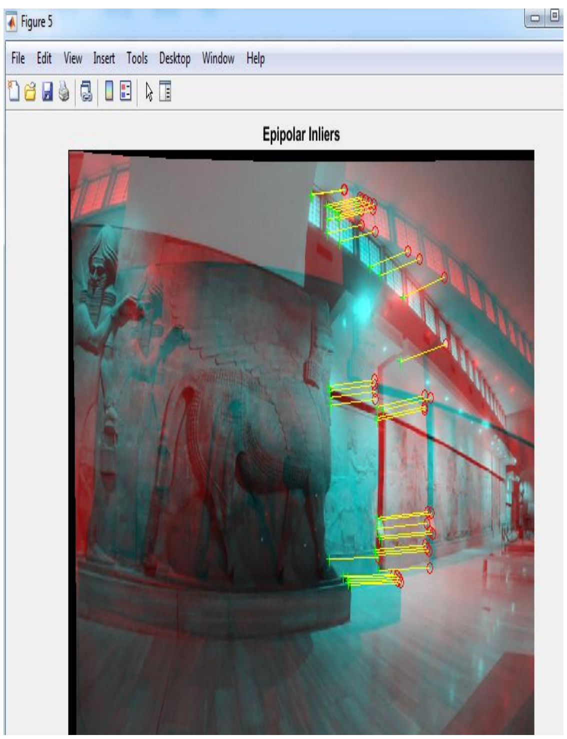

Figure 14. Matching Implementation Technique

The modeled object or imaged sculpture will be represented as a sparse points or points cloud as in figure (15)

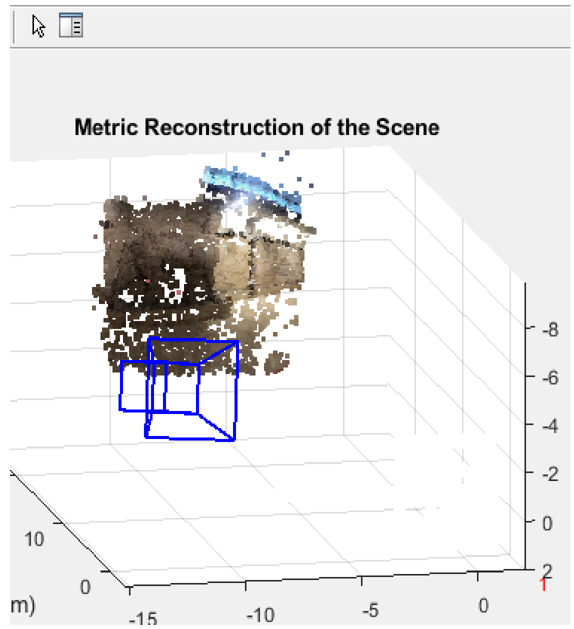

Figure 15. 3D Reconstructed Scene- Heritage Documentation

5. Main key points

The fragility of the materials used in the Mesopotamian civilization and the poor weather conditions were and still are an obstacle to contact with the components of this ancient civilization. Therefore, it was useful to adopt a digital, non-contact approach to avoid damage that could be caused to these unique assets belonging to our heritage. From the application of the proposed approach in the previous section, it can be said that the proposed system is a digital photogrammetric approach, a reliable, and sustainable system for digitizing historical cartography. It is evident in its suitability when a standard planar scanner cannot be used, or other sensors require direct contact with the fragile map support. The proposed system in this paper offers the major advantage of being a rapid, contactless, and harmless method when documenting heritage components. Among the most key points noted when implementing this approach:

- Computer vision is an active tool applicator that helps feature detection on image processing steps, influencing the output model. The concept of feature modeling is generally used to denote a single of information relevant to solving the

computational tasks related to a certain application.
- The structure From the Motion SfM method has been utilized to document heritage by manipulating 3D models on MATLAB interphase. It is approved for its efficiency and quick, super-advanced processing steps. SFM in this research is done by using two images from different views. The first step includes entering the images captured by various angles equal to 15°.
-More used images to produce 3D images lead to more accurate information using the SfM approach on MATLAB interphase approach.
- The usability of the output 3D image is active to secure sculptures, heritage, or any phenomena that can benefit public society by remotely sensed applications with considered accuracy and efficiency.
- in this research, a 3D laser scan device was designed and tested at a 360-degree angle using a camera that rotates at a certain speed to build a three-dimensional shape, and the result was geometrically acceptable.

6. Conclusions

It has become possible to confirm that the packages currently provided by technology allow the building of new systems for the traditional three-dimensional documentation and communication of architectural and cultural heritage. The system proposed in this paper can be successfully used for the 3D reconstruction of archaeological heritage in 3D measurement projects with high uncertainty. As evidence of the above, automation is an effective part that contributed to the communication between the mechanical and electrical subsystems to properly build the ingenious system with motion guidance until the data acquisition stage is passed. These were validated based on the modified system motion error in the data acquisition procedures (images) and how these defects were detected in the imaged regions. Since many of these procedures use a rangefinder laser, Errors may occur due to spatial resolution errors. For this, the treated skin area is taken into an image by a digital camera and entered into MATLAB. What needs to be found in the area of the overlapping points (used in the procedure). An algorithm can determine this. Correcting these errors while working on the image processing steps until obtaining a 3D image is necessary. Sharing the data provided by 3D laser scanning, there will certainly be an increase in collaborative projects worldwide.

References

[1] S. El-Hakim, J.A. Beraldin M. Picard and G. Godin “Detailed 3D reconstruction of large-scale heritage sites with integrated techniques” Computer Graphics and Applications, IEEE, Vol. 24, pp. 21-29, 2004.

[2] G.L. Foresti, “A Line Segment Based Approach for 3D Motion Estimation and Tracking of Multiple Objects”,Vol. 12, pp. 881-900. 1998.

[3] M. Farenzena, A. Fusiello, and R. Gherardi, “Structure-and-motion pipeline on a hierarchical cluster tree”, IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, pp. 1489-1496, 2009.

[4] J.Krivic and F. Solina “Part-level object recognition using superquadrics, Computer Vision and Image Understanding”, Vol. 95, pp. 105-126, 2004.

[5] B. Bhanu, “Representation and Shape Matching of 3-D Objects, IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-6”, pp.340-351, 1984.

[6] S. Budzan, and J. Kasprzyk, “Fusion of 3D laser scanner and depth images for obstacle recognition in mobile applications”, Optics and Lasers in Engineering, Vol.77, pp.230–240, 2016.https://doi.org/10.1016/j.optlaseng.2015.09.003

[7] S. El Hazzat, A. Saaidi, and K. Satori, “Structure from motion for 3D object reconstruction based on local and global bundle adjustment”, Proceedings of 2015 IEEE World Conference on Complex Systems, WCCS 2015. https://doi.org/10.1109/ICoCS.2015.7483225

[8] J. C. Torres, G. Arroyo, C. Romo, J. C. Torres, G. Arroyo, C. Romo, and J. de Haro, “3D

Digitization using Structure from Motion", 2012.

[9] M. J. Westoby, J. Brasington, N. F. Glasser, M. J. Hambrey, and J. M. Reynolds, "Structure-from-Motion" photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology, vol. 179,pp. 300–314, 2012. https://doi.org/10.1016/j.geomorph.2012.08.021

[10] B. Zhao, "AIFD Based 2D Image Registration to Multi-View Stereo Mapped 3D Models", Neural Processing Letters, Vol.48(3), pp.1261–1279. 2018.https://doi.org/10.1007/s11063-018-9816-6

[11] C. Solomon, and T."Breckon, Fundamentals of digital image processing : apractical approach with examples in Matlab", WILEY-BLACKWELL, 2013

[12] Wijesinghe, J. Mannapperuma, C. Kumarasinghe, D."Liyanage, ImageProcessing in Bio-medical Area (Literature)", 2016.

[13] P.A. Galindo, "Image matching for 3D reconstruction using complementary optical and geometric information" Université de Lorraine",2015.

[14] F. Bruno, S. Bruno, G., De Sensi, M.L. Luchi, S. Mancuso, and M. Muzzupappa, "From 3D reconstruction to virtual reality: A complete methodology for digital archaeological exhibition", Journal of Cultural Heritage, Vol. 11, pp.42-49, 2010

[15] R. S. Priya, S. L. Aarthy, C. Gunavathi, P. Venkatesh, S. Koppu, and X. Z. Gao "3D reconstruction of a scene from multiple 2D images". International Journal of Civil Engineering and Technology, Vol.8(12),pp. 324-331., 2010.

[16] S.C. Kuzminsky, and M.S. Gardiner, "Three-dimensional laser scanning: potential uses for museum conservation and scientific research" Journal of Archaeological Science, Vol. 39, pp. 2744-2751, 2012.

[17] Abdelhafiz, "3D Heritage Documentation JES", Journal of Engineering Sciences, 2012..