JEODEZİ VE JEOİNFORMASYON DERGİSİ
JOURNAL OF GEODESY AND GEOINFORMATION

ISSN : 2147-1339
e-ISSN: 2667-8519

Araştırma Makalesi / *Research Article*
Yıl / *Year:* 2023 **Cilt /** *Volume:* 10 **Sayı /** *Issue*: 2
Sayfa / *Page*: 112-123
Dergi No / *Journal No*: 118
Doi: 10.9733/JGG.2023R0008.E

Research on 3D reconstruction of small size objects using structure from motion photogrammetry via smartphone images

Duygu Arican[1*], Ferat Furkan Goksu[1], Nursu Tunalioglu[1], Taylan Ocalan[1]

[1]*Yildiz Technical University, Campus of Davutpaşa, Faculty of Civil Engineering, Department of Geomatic Engineering, Esenler, Istanbul, Türkiye.*

Abstract: *According to the recent developments on image acquisition via relatively low-cost devices such as smartphones or tablets, researches on Structure-from-Motion (SfM) photogrammetry-based 3 dimensional (3D) model reconstruction become more popular and wide frequent practice especially for study areas in geology/geomorphology, historical heritage, forestry, etc. Thus, this paper demonstrates accuracy assessment of SfM photogrammetry-based 3D model reconstruction of small size objects by altering number of the source imagery captured by smartphone and represents change detection of the generated datasets as (i) cloud-to-cloud and (ii) mesh-to-cloud comparisons. The number of images was decreased as 25% of each datasets belonging three different detail types of small size objects as the Safranbolu miniature house (SMH), a trinket made of a combination of historical buildings in Rome (HBR), and a wooden object (WO). A total of 12 datasets were generated and 9 cloud-to-cloud, and 21 mesh-to-cloud comparisons were performed. Since the obtained results show that the quality of 3D models of objects varies according to their shapes and sizes, change detection analyses show that the detail level of the objects are highly correlated with the resultant model accuracy.*

Keywords: Structure-from-Motion (SfM) photogrammetry, Smartphone imagery, 3D Model reconstruction, Change detection

Küçük boyutlu nesnelerin 3 boyutlu modellemesi için akıllı telefon görüntüleri kullanılarak hareketten nesne oluşturma fotogrametrisine dayalı bir araştırma

Öz: *Son yıllarda, akıllı telefon ya da tabletler gibi nispeten düşük maliyetli cihazlarla görüntü elde etme alanında yaşanan gelişmeler sonucunda, Hareketten Nesne Oluşturma (HNO) fotogrametrisine dayalı 3 boyutlu (3B) modelleme popüler olmuş ve özellikle jeoloji/jeomorfoloji, kültürel miras, orman vb. alanlarda yaygın bir uygulama haline gelmiştir. Bu çalışmanın amacı, akıllı telefon kamerası ile elde edilen görüntülerden değişen görüntü sayılarına dayalı 3B model oluşturma potansiyelinin araştırılmasıdır. Bu amaçla, küçük boyutlu nesnelere ait akıllı telefon kameraları kullanılarak elde edilen farklı sayıdaki görüntüler kullanılarak HNO fotogrametrisine dayalı 3B veri setleri oluşturulmuş ve bu veri setleri için (i) nokta bulutları arasında, (ii) üçgen model ile nokta bulutu arasında, olmak üzere sapma analizleri gerçekleştirilmiştir. Safranbolu minyatürü (SMH), Roma biblosu (HBR) ve ahşap nesne (WO) olmak üzere farklı detay düzeylerinde seçilen bu üç farklı boyuttaki nesne için elde edilen görüntüler her veri setinde %25 oranında azaltılmıştır. Bu şekilde, toplam 12 farklı veri seti oluşturulmuş ve bu veri setleri için nokta bulutları arasında 9, üçgen model ile nokta bulutu arasında ise 21 farklı karşılaştırma gerçekleştirilmiştir. Buna göre, elde edilen sonuçlar dikkate alındığında 3B modellerin doğruluğunun, modellemesi yapılan nesnelerin şekilsel ve boyutsal farklılığına göre değiştiği görülmekle birlikte, sapma analizine dayalı elde edilen sonuçlar nesnenin detay düzeyi ile sonuç model doğruluğu arasında yüksek ilişki olduğunu göstermiştir.*

Anahtar Sözcükler: Hareketten nesne oluşturma (HNO), Akıllı telefon görüntüsü, 3B Model oluşturma, Değişim analizi

* Sorumlu Yazar/Corresponding Author: Tel: +90 212 383 5298

Geliş Tarihi/Received: 09.02.2023
Kabul Tarihi/Accepted: 06.04.2023

0000-0002-4618-4357, duyguaricann@gmail.com (Arican D.) *
0000-0002-8173-4836, feratgoksu@gmail.com (Goksu F.F)
0000-0001-9345-5220, ntunali@yildiz.edu.tr (Tunalioglu N)
0000-0003-0861-013X, tocalan@yildiz.edu.tr (Ocalan T.)

1. Introduction

Computer-based 3 dimensional (3D) modeling has been a long-term purpose especially in accurate and realistic model generation (Tanskanen et al., 2013). Recent technological developments on image acquisition via relatively low-cost equipment and recent development of photogrammetric technique (Prosdocimi et al., 2017) for 3D photogrammetric modeling bring acceleration to this essential task. With the development of image processing methods, generating accurate digital surface models at relatively low-cost and in less time compared to other methods has become a current research topic in this field (Carbonneau, Lane, & Bergeron, 2003; Chandler, Ashmore, Paola, Gooch, & Varkaris, 2002; Lane et al., 2010). Moreover, use of metric and non-metric cameras on digital photogrammetry-based modeling extents the implementation areas to different studies (Mali & Kuiry, 2018). 3D reconstruction of scenes and objects has widely been used in many different applications such as reverse engineering, medicine, security, crime investigation (Chen et al., 2021), geomorphological studies, 3D topography reconstruction (Ding, Zheng, Zhou, Xiong, & Gong, 2018), historical heritage inventory facilities (De Reu et al., 2014). Recently, the feasibility of image acquisition based on smart devices such as smartphones and tablets have been introduced broader fields of applications (Jasińska, Pyka, Pastucha, & Midtiby, 2023). Many studies such as in digital landform generation purposes (Micheletti, Chandler, & Lane, 2015), volumetric earthwork computations (Jeong, H. Ahn, Shin, Y. Ahn, & Choi, 2019; Wróżyński, Pyszny, Sojka, Przybyła, & Murat-Błażejewska, 2017), artefacts of cultural heritage and documentation (Barszcz, Montusiewicz, Paśnikowska-Łukaszuk, & Sałamacha, 2021) have been conducted by the researchers for 3D model reconstruction following SfM photogrammetry, where the images collected via smartphones. Guidi, Micoli, Gonizzi, Brennan and Frischer (2015) aimed to analyze the model quality of small-medium size objects based on SfM and Image Matching (IM) with different shooting configurations. The results with greater overlap considerably decrease in measurement uncertainty. Clini, Frapiccini, Mengoni, Nespeca and Ruggeri (2016) conducted a study for 3D digital documentation of small sized (a few cm) archeological artifacts following image-based SfM methodology and tested focus stacking, and concluded that effective results regarding high definition 3D models may be provided. Collins, Woolley, Gehlken and Ch'ng (2019) investigated 3D small artifact generation using a turntable and an automated procedure based on SfM photogrammetry captured by ultra-low cost smartphone and digital camera. The automated procedure achieved high precision and visually qualified textures. Jasińska et al. (2023) focused on deformation reduction of the 3D models generated via smartphone photogrammetry and conducted a two-stage study as first they controlled the performance of stability of the internal orientation parameters by self-calibration and then they modelled small objects by selected images. According to the results, even limitations smartphone photogrammetry has introduced potential usage areas.

As it can vary according to the aim of the study, the accuracy investigation of the resultant 3D model is an ongoing issue in terms of data source collected via mobile low-cost devices. Since modeling with relatively low-cost devices, decrease in time can be searched via the number of the images to be collected for modeling the scene or the object. Although some factors can affect the quality of the resultant in terms of complexity, lighting, and materials, there are no guidelines for deciding the minimum number of images needed to reconstruct the scene. Although, each feature must be visible in a minimum of three images (Westoby, Brasington, Glasser, Hambrey, & Reynolds, 2012). Thus, this study aims to investigate the feasibility of using low-cost device collected images for 3D model reconstruction of small size objects with altering the number of images to assess the accuracy of SfM photogrammetry-based 3D model reconstruction. As high-level capacity of SfM photogrammetry in large scale modeling facilities is a well-known topic, according to the small scale details of smaller objects in shape needs to be investigated. Three small size objects were selected regarding level of detail and texture and datasets were generated by altering number of the source imagery captured by smartphone. The change detection of the generated datasets was performed following two strategies as (i) cloud-to-cloud and (ii) mesh-to-cloud comparisons. The number of

images was decreased as 25% of each datasets belonging three different detail types of small size objects. A total of 12 datasets were generated and 9 cloud-to-cloud, and 21 mesh-to-cloud comparisons were performed. The results of the study were represented in the following sections.

2. Materials and Methods

Traditional photogrammetry needs 3D camera positions and coordinates of the ground control points (GCP) to build scene triangulation and reconstruction. Contrary, the SfM methodology automatically and simultaneously uses multiple overlapped images to estimate positions of the camera and scene geometry (Yang, Chao, Huang, Lu, & Chen, 2013). This estimation is carried out by the matching features in the set of overlapped images. These features are followed image by image and make it possible initial detections of camera positions and object coordinates (Westoby et al., 2012). This is the main issue of the SfM, dealing with the determination of the 3D location of matching features in multiple images taken from different angles. Identification of the features from each image that is used for image matching is the first step of the solution. The model geometry and camera position information are solved simultaneously and automatically in SfM technique, consequently camera calibration is not required. The most popular solution to this problem is the Scale Invariant Feature Transform (SIFT) method (Lowe, 1999). Features in each image are identified scale and rotation invariant in different illumination conditions using this method. Matching points or "key points" are automatically identified in all scales and locations. Then, a feature descriptor is created. This is calculated as converting local image gradients into a representation. As should be noted here, this is largely insensitive to variations in illumination and orientation. These descriptors are unique enough to allow features to be matched in large datasets (Westoby et al., 2012).

Usually, the number of key points is related to the image quality, texture, sharpness, and resolution of the dataset. Therefore, these qualities determine the quality of the resultant point cloud. Complex images produce more key points. Correspondingly, the closer distance between the object and the camera stations increases the spatial resolution of the image, hence it increases the spatial density and resolution of the point cloud. Since complexity, lighting, and materials could affect the image texture, there is no guidance on the minimum number of images required for scene reconstruction. Although, each feature must be visible in a minimum of three images (Westoby et al., 2012). After key point identification and descriptor assignment, the position of the camera station is estimated using bundle adjustment and a low-density or sparse point cloud is produced. The key points from multiple images are matched using the approximate nearest neighbor method. Minimum of two key points and three images are required for the reconstruction of the point cloud (Westoby et al., 2012). Multi-View Stereo (MVS) algorithms are based on the correlations between measurements from a set of images at once to acquire 3D surface information. Many of the MVS methods aim to reconstruct all images simultaneously. Thus, this approach is not practical while working with more extensive datasets. Rather selecting a subset from the dataset is more important and clustering them into convenient samples (Furukawa, Curless, Seitz, & Szeliski, 2010). Clustering View for Multi-View Stereo (CMVS) approach clusters and selects the optimal viewpoint by a four-step iterative approach. These steps are (1) merging the sparse point clouds generated from SfM, (2) removing low-quality images, (3) clustering images into smaller components, and finally (4) adding images to weak clusters (Mahami, Nasirzadeh, Hosseininaveh Ahmadabadian, & Nahavandi, 2019).

Patch-based Multi-View Stereo (PMVS) represents the surface with a local tangent plane approximation of a surface. This algorithm is easy to implement and effective. The first step of this approach is creating a small 3D patch in 3D space, for each matched feature point. The center of this patch is the triangulation point from all the matched feature points from the reference and visible images. Afterward, the 3D patch is oriented around the center, which also has to be parallel to the current reference image and one of its edges is parallel to the x-axis of the reference image. Then the next step, each point

inside of the patch is projected back to the reference image and all visible images. After that, a small two-dimensional patch is obtained from each image (Yao, AliAkbarpour, Seetharaman, & Palaniappan, 2018). PMVS uses undistorted images, the orientation parameters of these images, sparse point clouds, and projection matrices to determine a dense and accurate set of rectangular patches. PMVS consists of matching, expansion, and filtering steps (Mahami et al., 2019). Then estimations are refined iteratively using non-linear least-squares minimization. Initially, the 3D point cloud is produced in a relative coordinate system. It must be aligned to a real-world coordinate system. The GCPs provide the scale and orientation. Hence, in transforming a relative coordinate system to the real-world coordinate system, GCPs are required (Westoby et al., 2012).

In this study, VisualSFM was used for generating point clouds from images. Following, CloudCompare was used to georeference point clouds, then subsequently to remove the outliers and noises. The 3D model reconstruction process was utilized in Meshlab. Cloud-to-cloud and cloud-to-mesh distances were calculated in CloudCompare. Mean distances and standard deviations were calculated in Matlab.

3. Case Study

In this study, the effects of low-cost sensor image and its amounts on point cloud generation were investigated. Three different datasets were acquired from iPhone 6S Plus (pixel resolution: 1080x1920; pixel size: 1.5 µm; f: 2.2 mm) for evaluation. The datasets were generated by taking images of objects with different shapes. The modeling procedure was implemented to small size objects (circa Safranbolu miniature house 12x25x25; trinket made of a combination of historical buildings in Rome 6x7x6; and wooden object 7x7x10 cm) represented in Figure 1. The objects that were selected for the evaluations depending on their detail level are the Safranbolu miniature house (SMH) as complex, a trinket made of a combination of historical buildings in Rome (HBR) as sculptural, and a wooden object (WO) as primitive. The workflow was followed in this study shown in Figure 2.

Figure 1: *The objects that were selected to generate the 3D point clouds: a) Safranbolu miniature house (SMH), b) Historical buildings in Rome Trinket (HBR), and c) Wooden object (WO)*

GCPs were marked in the experimental area and measured before acquiring the data. The images were acquired as overlapped from 8 angles at 44 camera stations. Thus, 352 images were obtained from approximately the same camera stations for each object. After the image acquisition, the datasets were reduced by 25%, 50%, and 75%, respectively according to the resampling method, and thus three more datasets, a total of four datasets, were generated for each object. Reduction procedure is implement due to investigation of the number of images to be required for 3D modeling association with the accuracy level of model. The number of images in all datasets is shown in Table 1.

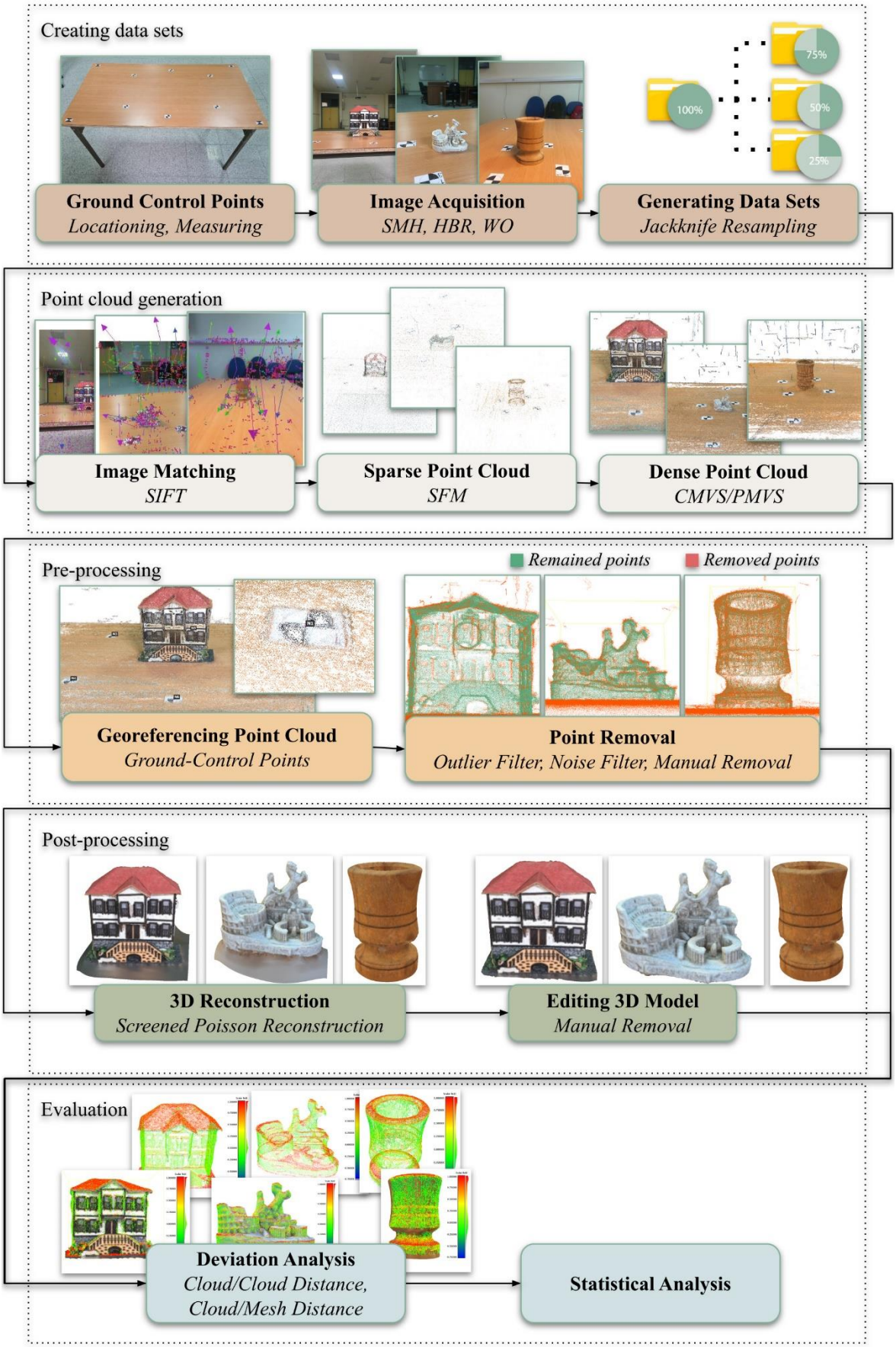

Figure 2: *Study workflow*

Table 1: *Number of images in each dataset*

Definition	Dataset	Number of Images
Safranbolu Miniature House	SMH 100	352
	SMH 75	264
	SMH 50	176
	SMH 25	88
Historical Buildings in Rome Trinket	HBR 100	352
	HBR 75	264
	HBR 50	176
	HBR 25	88
Wooden Object	WO 100	352
	WO 75	264
	WO 50	176
	WO 25	88

SfM/CMVS/PMVS approach was followed to produce dense point clouds for all datasets. In this method, SIFT algorithm was utilized for image matching. Following image matching, camera stations are estimated, and low-density or sparse point cloud occurs. The number of points in the sparse point cloud is less than in the dense point cloud, hence it is difficult to recognize the points, which belong to the object in this step. The estimated camera stations of reference datasets for each object are shown in Figure 3.

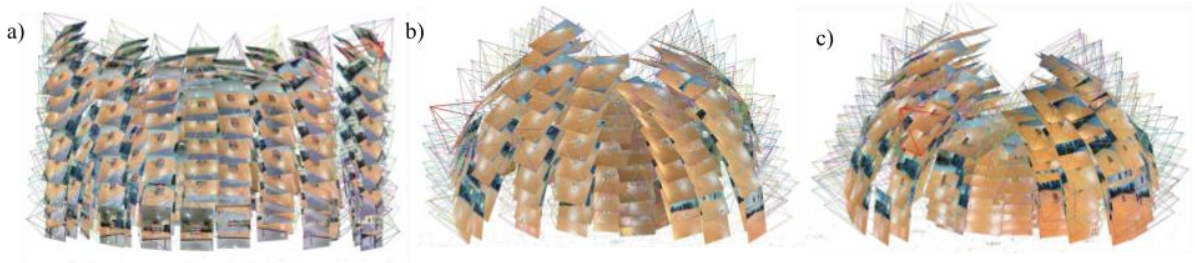

Figure 3: *Estimated camera stations: a) SMH 100 b) HBR 100 c) WO 100*

After that, dense point clouds are created by CMVS/PMVS approach. Each point cloud initially has a relative coordinate system and scale. For each point cloud, the same 6 GCPs in the experiment area were used to transform a relative coordinate system into a real-world coordinate system based on 3D similarity transformation.

The point clouds created from images have also points that do not belong to the objects. These points were removed by utilizing Statistical Outlier Removal (SOR) filter, noise filter, and manual removal in CloudCompare software, respectively. SOR Filter estimates the average distance of each point to its neighbors considering the k-nearest neighbors algorithm. Then the points farther than the average distance are removed. 50 nearest neighbors were used in this study to perform the SOR filter. The noise filter of CloudCompare is also similar to the SOR filter but it uses the underlying surface instead of the average distance estimated by neighbors. The underlying surface is estimated by the radius or the number of neighbors. In this study, 50 neighbors were used to calculate the surface. If the point is too far away from the locally fitted plane, it was removed. After the utilization of point cloud removal filters, it is also necessary to clean point clouds manually since they have other objects in the experimental area. Therefore, all the point clouds were manually cleaned. Table 2 shows the number of point clouds subsequent to each step, where the first two columns represent the number of sparse and dense point clouds generated and the last three columns demonstrate the filtering stages applied with remaining numbers of points in each step. Since the point clouds generated involve points that do not belong to the objects and the objects subjected to this study are small size, the differences between the numbers of points have been occurred.

Table 2: *The number of points in point clouds obtained after each step*

Dataset	Number of points in point clouds				
	Sparse	Dense	SOR Filtered	Noise Filtered	Manually Cleaned
SMH 100	74 552	6 350 941	5 629 279	4 133 005	384 925
SMH 75	63 406	6 082 690	5 396 043	4 015 097	338 743
SMH 50	45 950	3 318 940	2 915 703	2 107 655	199 920
SMH 25	22 912	1 508 194	1 309 501	938 760	94 454
HBR 100	40 518	6 270 809	5 931 189	4 335 087	201 545
HBR 75	33 212	4 063 899	3 818 454	2 777 767	135 317
HBR 50	24 506	3 168 460	2 981 414	2 157 836	102 585
HBR 25	10 719	1 643 116	1 506 311	1 074 670	57 076
WO 100	26 554	5 205 536	4 954 481	3 639 560	292 625
WO 75	20 667	4 254 933	4 022 772	2 935 886	252 776
WO 50	14 978	2 673 994	2 490 451	1 812 470	152 976
WO 25	20 518	4 016 088	3 855 715	2 797 078	236 328

Following the pre-processing steps, 3D Reconstruction was carried out by Screened Poisson Surface Reconstruction method (Kazhdan & Hoppe, 2013). There are triangles extending objects outward in some areas on the 3D models. These triangles were manually removed by cropping. The number of faces of 3D meshes is shown in Table 3 as the second column in number of faces section represents the total number and the third column shows the edited faces that can be deleted or changed in shape.

Table 3: *The number of faces in 3D meshes*

Dataset	Number of Faces	
	Screened Poisson Reconstruction	Edited Mesh
SMH 100	316 206	305 348
SMH 75	313 173	300 981
SMH 50	311 228	298 943
SMH 25	292 952	282 844
HBR 100	309 177	286 718
HBR 75	313 655	294 400
HBR 50	309 244	292 441
HBR 25	271 781	256 658
WO 100	433 146	432 467
WO 75	445 172	444 246
WO 50	399 448	396 623
WO 25	450 594	449 974

4. Results and Discussion

In this study, the effects of the number of images on the point clouds and mesh generation by using relatively low-cost devices, here a smartphone camera, were investigated. Twelve datasets were utilized as three of them were acquired and nine of the datasets were derived from those three for three differently shaped objects. Point clouds are generated by implementing SfM method. Lastly, meshes are generated by Screened Poisson Reconstruction method.

Generated clouds are compared by acquired datasets as reference. Cloud-to-cloud distances were calculated based on the least square plane and colored by scalar field, which is shown in Figure 4. Further, mean distances and standard deviations are calculated by the cloud-to-cloud distance shown in Table 4. In all comparisons, the reference model is generated by using 100% imagery in all objects. As SMH is a detailed object, the maximum standard deviation is computed between SMH 100 and SMH 50 as 0.978 although the minimum standard deviation is computed as 0.601 mm between reference and SMH 75

as expected. For second object, namely HBR, also involved complex details, the minimum standard deviation is computed as 0.270 mm between reference and HBR 25 and maximum standard deviation is as 0.348 mm between reference and HBR 75. This can be caused by due to the complexity of the object that a proper cloud dataset could not be formed hence the object HBR is the most complex one among the objects selected for the test. This can be supported by the results obtained for the third object, namely WO because the object WO is the plainest one among them. Thus, the standard deviation and mean distance values are computed very close to each other results when compared between reference and WOs 75, 50, and 25. This also shows the number of decreasing images may not affect the resultant model if small size and plain featured object is selected for 3D modeling.

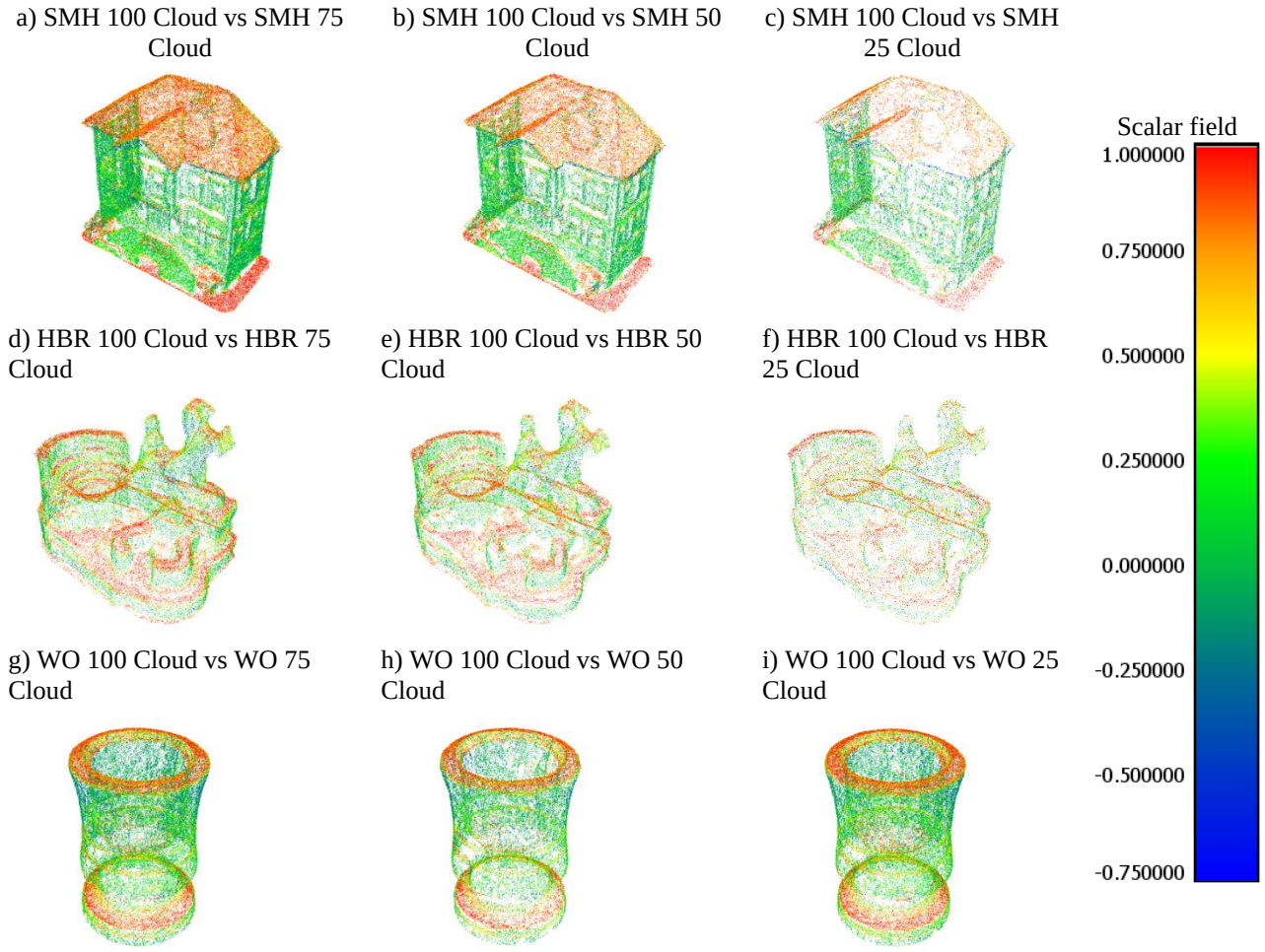

Figure 4: *Cloud-to-cloud distances (for visual representation scalar field: blue (-0.75) to red (1.00))*

Table 4: *Mean distances and standard deviations obtained from cloud-to-cloud comparison*

Reference Point Cloud	Compared Point Cloud	Mean Distance (mm)	Std. Deviation (mm)
SMH 100	SMH 75	0.471	0.601
	SMH 50	0.639	0.978
	SMH 25	0.710	0.913
HBR 100	HBR 75	0.278	0.348
	HBR 50	0.219	0.282
	HBR 25	0.211	0.270
WO 100	WO 75	0.521	0.656
	WO 50	0.512	0.652
	WO 25	0.509	0.637

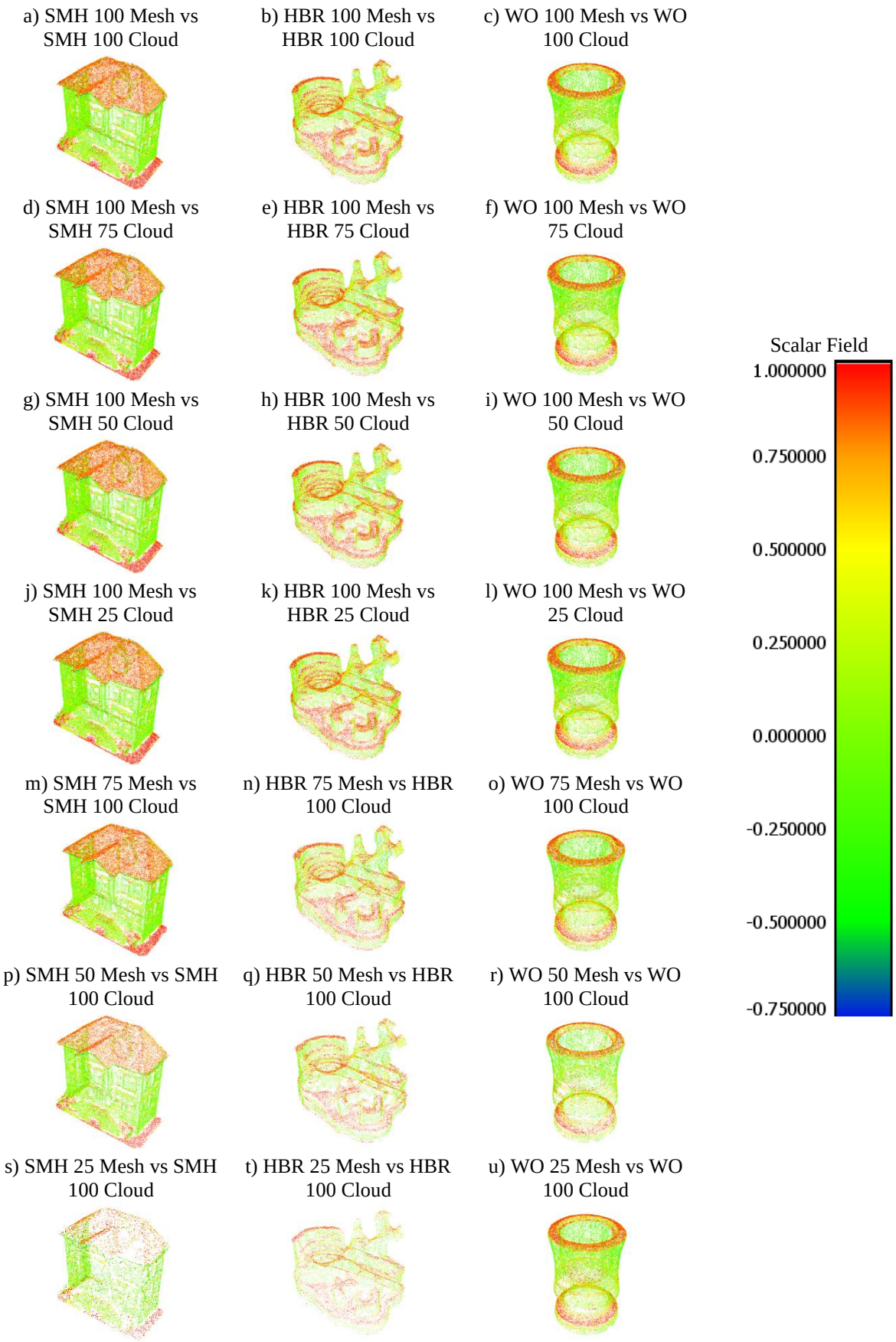

Figure 5: *Compared point clouds by mesh-to-cloud distances (for visual representation scalar field: blue (-0.75) to red (1.00))*

Mesh-to-cloud distances were also calculated by the meshes set as reference. Compared datasets were colored by scalar field. Colored point clouds were shown in Figure 5. Moreover, mean distances and standard deviations are calculated by the mesh-to-cloud distance shown in Table 5. As a total of 12 datasets were generated, 21 mesh-to-cloud comparisons were performed. Here, mesh models of the objects were used for comparison issues. The reference mesh model has been changed as 100, 75, 50, and 25% usage of images captured by smartphone. For the object SMH, the minimum standard deviation is computed between SMH 100 mesh model and SMH 100 point cloud as 0.660 mm and the mean distance of the deviation is computed as 0. Contrary to this the maximum differences in standard deviation for this object are as 1.782 and 1.659 mm between SMH 50 mesh model and SMH 100 point cloud and SMH 25 mesh model and SMH 100 point cloud, respectively. That can be concluded as an approximately three times worse models have been obtained when decreasing the number of images up to 25%. Among the comparisons results represented in Table 5, minimum standard deviation and mean distance values are listed for the object HBR. However, as mentioned before this object is the most complex one of them. When investigating the mesh models of this object, it can be seen that the details could not be defined and formed properly in the mesh models. Thus, compared parts of the object on mesh models has matched. Even so, the minimum and maximum standard deviations of this comparison are computed as 0.310 and 0.599 mm between HBR 100 mesh model and HBR 100 and HBR 100 mesh model and HBR 75 point clouds, respectively. Finally, for the object WO, the minimum standard deviation value is as 0.196 mm between WO 100 mesh model and WO 100 point cloud that value is also the most accurate one among these 21 comparisons.

Table 5: *Mean distances and standard deviations obtained from mesh-to-cloud comparison*

Reference Mesh	Compared Point Cloud	Mean Distance (mm)	Std. Deviation (mm)
SMH 100	SMH 100	0.000	0.660
	SMH 75	-0.015	1.002
	SMH 50	0.084	1.544
	SMH 25	0.127	1.520
SMH 75	SMH 100	0.037	1.078
SMH 50		-0.002	1.782
SMH 25		-0.202	1.659
HBR 100	HBR 100	0.025	0.310
	HBR 75	0.314	0.599
	HBR 50	0.248	0.491
	HBR 25	0.222	0.490
HBR 75	HBR 100	-0.226	0.585
HBR 50		-0.203	0.505
HBR 25		-0.200	0.514
WO 100	WO 100	0.005	0.196
	WO 75	0.059	0.854
	WO 50	0.077	0.887
	WO 25	-0.080	0.828
WO 75	WO 100	-0.089	0.810
WO 50		-0.073	1.019
WO 25		0.068	0.809

5. Conclusion

Today's technological developments on image acquisition and recent development of photogrammetric technique accelerate the computer-based accurate and realistic 3D modeling. Recently, due to the widespread increasing usage capacity of the smart devices such as mobile phones, tablets etc., and their sensors with different functions i.e. cameras, navigation chipset, result in new scientific research areas raised. In terms of 3D photogrammetric modeling, generating accurate digital surface models at relatively low-cost and in less time compared to other methods has become a significant tool in many different applications such as reverse engineering, medicine, security, crime investigation, geomorphological studies, historical

heritage inventory facilities, 3D topography reconstruction, search and rescue facilities etc. As the detail level of objects has increased while decreasing the shape sizes, the utility of SfM photogrammetry in terms of providing effective modeling solutions using smart devices may reduce the observation cost with meeting the required accuracies that may be changed by the aim of study. This study aims to investigate the feasibility of modeling small size objects with SfM based photogrammetry using images captured by a low-cost device. Three small size objects were selected regarding level of detail and texture (plain to complex) and datasets were generated by altering number of the source imagery captured by smartphone. The change detection of the generated datasets was performed following two strategies as (i) cloud-to-cloud and (ii) mesh-to-cloud comparisons. The number of images was decreased as 25% of each datasets belonging three different detail types of small size objects. A total of 12 datasets were generated and 9 cloud-to-cloud, and 21 mesh-to-cloud comparisons were performed. When comparing based on cloud-to-cloud methodology, it can be concluded that the standard deviation and mean distance values are computed very close to each other results when considering plain and medium detailed object forms. Moreover, decreasing the number of images may not change the resultant model if small size and plain-medium featured object is selected for 3D modeling. Additional, comparing mesh models by point clouds results in similar findings with cloud-to-cloud comparisons. Lastly, it can be concluded that the obtained results show that the quality of 3D models of objects varies according to their shapes and sizes, change detection analyses show that the detail level of the objects are highly correlated with the resultant model accuracy.

Acknowledgements

Duygu Arican is a Ph.D. scholarship holder from the Council of Higher Education (YÖK) in the field of "GIS and Informatics Applications", which is one of the 100 national priority areas determined by YÖK within the scope of the YÖK 100/2000 Doctorate Program.

Author Contribution

Duygu Arican: Conception, Design, Data collection and processing, Analysis, Writing, Literature review. **Ferat Furkan Goksu:** Data collection and processing, Literature review. **Nursu Tunalioglu:** Conception, Design, Analysis, Writing, Literature review, Supervision. **Taylan Ocalan:** Design, Analysis, Writing, Supervision.

Declaration of Competing Interests

The authors declare that they have no known relevant competing financial or non-financial interests that could have appeared to influence the work reported in this paper.

References

Barszcz, M., Montusiewicz, J., Paśnikowska-Łukaszuk, M., & Sałamacha, A. (2021). Comparative analysis of digital models of objects of cultural heritage obtained by the “3D SLS” and “SfM” methods. *Applied Sciences*, *11*(12), 5321.

Carbonneau, P. E., Lane, S. N., & Bergeron, N. E. (2003). Cost-effective non-metric close-range digital photogrammetry and its application to a study of coarse gravel river beds. *International journal of remote sensing*, *24*(14), 2837-2854.

Chandler, J., Ashmore, P., Paola, C., Gooch, M., & Varkaris, F. (2002). Monitoring river-channel change using terrestrial oblique digital imagery and automated digital photogrammetry. *Annals of the Association of American Geographers*, *92*(4), 631-644.

Chen, H., Xu, F., Liu, W., Sun, D., Liu, P. X., Menhas, M. I., & Ahmad, B. (2021). 3D reconstruction of unstructured objects using

information from multiple sensors. *IEEE Sensors Journal*, *21*(23), 26951-26963.

Clini, P., Frapiccini, N., Mengoni, M., Nespeca, R., & Ruggeri, L. (2016). SFM Technique and Focus Stacking for Digital Documentation of Archaeological Artifacts. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *41B5*, 229–236.

Collins, T., Woolley, S. I., Gehlken, E., & Ch'ng, E. (2019). Automated low-cost photogrammetric acquisition of 3D models from small form-factor artefacts. *Electronics*, *8*(12), 1441.

De Reu, J., De Smedt, P., Herremans, D., Van Meirvenne, M., Laloo, P., & De Clercq, W. (2014). On introducing an image-based 3D reconstruction method in archaeological excavation practice. *Journal of Archaeological Science*, *41*, 251-262.

Ding, Y., Zheng, X., Zhou, Y., Xiong, H., & Gong, J. (2018). Low-cost and efficient indoor 3D reconstruction through annotated hierarchical structure-from-motion. *Remote Sensing*, *11*(1), 58.

Furukawa, Y., Curless, B., Seitz, S. M., & Szeliski, R. (2010). Towards internet-scale multi-view stereo. *2010 IEEE computer society conference on computer vision and pattern recognition*, 1434-1441.

Guidi, G., Micoli, L. L., Gonizzi, S., Brennan, M., & Frischer, B. (2015). Image-based 3D capture of cultural heritage artifacts an experimental study about 3D data quality. *2015 Digital Heritage*, 321-324.

Jasińska, A., Pyka, K., Pastucha, E., & Midtiby, H. S. (2023). A Simple Way to Reduce 3D Model Deformation in Smartphone Photogrammetry. *Sensors*, *23*(2), 728.

Jeong, H., Ahn, H., Shin, D., Ahn, Y., & Choi, C. (2019). A Comparative Assessment of the Photogrammetric Accuracy of Mapping Using UAVs with Smart Devices. *Photogramm. Eng. Remote Sens*, *85*, 889-897.

Kazhdan, M., & Hoppe, H. (2013). Screened poisson surface reconstruction. ACM Transactions on Graphics (ToG), 32(3), 1-13.

Lane, S. N., Widdison, P. E., Thomas, R. E., Ashworth, P. J., Best, J. L., Lunt, I. A., Sambrook, S., & Simpson, C. J. (2010). Quantification of braided river channel change using archival digital image analysis. *Earth Surface Processes and Landforms*, *35*(8), 971-985.

Lowe, D. G. (1999). Object recognition from local scale-invariant features. *Proceedings of the seventh IEEE international conference on computer vision*, 1150-1157.

Mahami, H., Nasirzadeh, F., Hosseininaveh Ahmadabadian, A., & Nahavandi, S. (2019). Automated progress controlling and monitoring using daily site images and building information modelling. *Buildings*, *9*(3), 70.

Mali, V. K., & Kuiry, S. N. (2018). Assessing the accuracy of high-resolution topographic data generated using freely available packages based on SfM-MVS approach. *Measurement*, *124*, 338-350.

Micheletti, N., Chandler, J. H., & Lane, S. N. (2015). Investigating the geomorphological potential of freely available and accessible structure-from-motion photogrammetry using a smartphone. *Earth Surface Processes and Landforms*, *40*(4), 473-486.

Prosdocimi, M., Burguet, M., Di Prima, S., Sofia, G., Terol, E., Comino, J. R., Cerdà, A., & Tarolli, P. (2017). Rainfall simulation and Structure-from-Motion photogrammetry for the analysis of soil water erosion in Mediterranean vineyards. *Science of the Total Environment*, *574*, 204-215.

Tanskanen, P., Kolev, K., Meier, L., Camposeco, F., Saurer, O., & Pollefeys, M. (2013). Live metric 3D reconstruction on mobile phones. *Proceedings of the IEEE International Conference on Computer Vision*, 65-72.

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). 'Structure-from-Motion'photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, *179*, 300-314.

Wróżyński, R., Pyszny, K., Sojka, M., Przybyła, C., & Murat-Błażejewska, S. (2017). Ground volume assessment using'Structure from Motion'photogrammetry with a smartphone and a compact camera. *Open geosciences*, *9*(1), 281-294.

Yang, M. D., Chao, C. F., Huang, K. S., Lu, L. Y., & Chen, Y. P. (2013). Image-based 3D scene reconstruction and exploration in augmented reality. *Automation in Construction*, *33*, 48-60.

Yao, S., AliAkbarpour, H., Seetharaman, G., & Palaniappan, K. (2018). 3D patch-based multi-view stereo for high-resolution imagery. *Geospatial Informatics, Motion Imagery, and Network Analytics VIII 10645*, 146-153.