ENHANCING CLOSE-UP IMAGE BASED 3D DIGITISATION WITH FOCUS STACKING

G. Kontogianni[1*] , R. Chliverou[1], A. Koutsoudis[2], G. Pavlidis[2], A. Georgopoulos[1]

[1] Laboratory of Photogrammetry, School of Rural and Surveying Engineering, National Technical University of Athens –
15780 Zografou Athens, Greece
gkondog@central.ntua.gr - regina@survey.ntua.gr - drag@central.ntua.gr
[2] Athena Research Centre, University Campus at Kimmeria, PO BOX 159, GR-67100, Xanthi, Greece
akoutsou, gpavlid@ceti.gr

KEY WORDS: Focus Stacking, 3D reconstruction, small artifacts

ABSTRACT:

The 3D digitisation of small artefacts is a very complicated procedure because of their complex morphological feature structures, concavities, rich decorations, high frequency of colour changes in texture, increased accuracy requirements etc. Image-based methods present a low cost, fast and effective alternative because laser scanning does not meet the accuracy requirements in general. A shallow Depth of Field (DoF) affects the image-based 3D reconstruction and especially the point matching procedure. This is visible not only in the total number of corresponding points but also in the resolution of the produced 3D model. The extension of the DoF is a very important task that should be incorporated in the data collection to attain a better quality of the image set and a better 3D model. An extension of the DoF can be achieved with many methods and especially with the use of the focus stacking technique. In this paper, the focus stacking technique was tested in a real-world experiment to digitise a museum artefact in 3D. The experiment conditions include the use of a full frame camera equipped with a normal lens (50mm), with the camera being placed close to the object. The artefact has already been digitised with a structured light system and that model served as the reference model in which 3D models were compared and the results were presented.

1. INTRODUCTION

The depth of field (DoF) defines a range of distances, in which a scene appears to be acceptably sharp and is affected mainly by the camera's focal length, the lens aperture and the focusing distance. The sharpness in images due to the DoF does not change abruptly, but occurs as a gradual transition, which means that everything in front or behind the focus plane appears to gradually lose sharpness. The effect is not always detectable by the naked eye or indeed by the resolution of the camera especially when the pixel pitch is large. The Circle of Confusion (CoC) defines the maximum diameter of a circle formed by a point imaged blurred, beyond which the point is perceived as non-sharp. When the CoC is visible to the human eye, the region is outside of the depth of field and it is depicted blurry (Cambridge in Colour, 2017).

Applications in close range photogrammetry are adversely affected by a shallow DoF, which is in most cases prominent. Specifically, in the case of small artefacts, it is very difficult to capture all the details of the objects and the alignment process with a shallow DoF is expected to be based on a limited amount of salient points from regions in focus. This may affect the point matching procedure, the number of conjugate points determined and the resolution of the 3D model. In addition, it may lead to a very sparse set of points with areas on the surface of an object with no point correspondences.

In this paper, we investigate exactly this eventual effect of 3D reconstruction by applying automatic image based modelling using SfM and MVS techniques to images of "best focus" and images created with focus stacking techniques.

The rest of the paper is organised as follows: section 2 describes the focus stacking technique and section 3 presents related work of the focus stacking in 3D reconstruction; section 4 presents the experimental setup and section 5 describes how these data were integrated; finally, the evaluation of the results and some concluding remarks are presented.

2. FOCUS STACKING

Focus stacking is a technique in computational photography, according to which a number of images of typically shallow DoF and taken with different focusing distances are being used to create a new image with a longer DoF (theoretically targeting an infinite DoF). This technique is mainly used in close-up, macro, and landscape photography because the smallest appropriate aperture f-value (narrow aperture) is often inadequate to capture the whole range of distances required (Cambridge in Colour, 2017). Focus stacking consists of three basic steps:

(1) in the first step, several images are taken using different focusing distances
(2) in the second step these images are aligned since the changes in the focusing distance may cause misalignments
(3) the final step is the creation of the final focus stacking image with the use of the appropriate method and algorithm

---

* Corresponding author

Images in different focusing distances can be taken with several methods. As a first approach the camera is fixed on a tripod and the focusing distance can be changed manually by rotating the focusing ring. The rotation of the focusing ring can also be automated with the use of software like Helicon Remote, ControlmyNikon, Canon EOS Utility. In both cases, it is necessary to focus the nearest and the farthest point to estimate the boundaries of the focusing distance. Secondly, the focusing distance can be changed when the camera is being moved on a micrometer rail and a fixed-focus lens is being used (Clini et al. 2016). In addition, the focusing distance can be adjusted by utilising the tilt movement of a tilt-shift lens.

Many algorithms and methods have been developed to create a focus stacked image. Commercial software (Zerene Stacker, Helicon Focus etc.) rely on three basic approaches (Dpreview, 2017):

- The *pyramid* approach: the approach utilizes a pyramidal decomposition method from image processing, dividing image signals into high and low frequencies. This method gives good results in complex cases (intersecting objects, deep stacks) though increases contrast and glare.
- The *depth map* approach: the method selects the source image containing the sharpest pixel and uses this information to reconstruct the depth map. In this method images should be taken in a consecutive order from front to back or vice versa and it can perfectly render texture on smooth surfaces.
- The *weight pixel* approach: the *weight pixel* approach computes a weight for each pixel based on its contrast in the set of images and then forms the weighted average of all pixel from all the images. This approach works better in small stacks and preserves contrast and colour.

3. RELATED WORK

The focus stacking technique has been mainly used in 3D digitisation of small objects, such as archaeological artefacts, insects, etc. Clini et al. (2016) used focus stacking for the creation of a 3D model of the Venus of Frasassi statue (8,7x2.6 cm). They utilised a full frame camera equipped with a macro lens and they compared the result with the 3D model created with the Konica Minolta Range 7 triangulation laser scanner. Brecko et al. (2014) present a focus stacking system, which is part of commercial photographic equipment. They tested this system and compared the results with several different focus stacking software packages, including CombineZP, Auto-Montage, Helicon Focus and Zerene Stacker. They also created 3D models of the insects using Agisoft Photoscan and images created with the focus stacking technique. Gallo et al. (2012; 2014) proposed a new methodology for the reconstruction of small artefacts. The reconstruction was based on the multi-view-stereo algorithm (MVS) by using open source SfM-MVS software. Sequence of images were captured with the use of a macro lens and focus stacking was used for extending the DoF. The models that were created were compared with models from a NextEngine triangulation laser scanner. Focus stacking and photogrammetry were also used to record the 3D information about the millimetric and submillimetric details in 3D of prehistoric petroglyphs and paintings (Plisson & Zotkina 2015). Kontogianni et al. (2017) used focus stacking to investigate the effect of the technique on SfM-MVS 3D reconstruction. The results were compared with the ground truth model from a structured light 3D scanner and the first results were presented.

4. EXPERIMENTAL SETUP

4.1 Case Study Object

The object that was used in the experiment is a replica of a Native American bird effigy bowl, which was discovered in the upper Nodena site in Arkansas, USA. The original artefact is exhibited in the Hampson Archaeological Museum (Hampson Archaeological Museum, 2016) and is considered a strong representative of a distinct shape category found in the museum. The replica was created by a sculptor within the framework of the PRESIOUS EU project (PRESIOUS project, 2017), based on subjective and objective criteria such as visual and metric information that was extracted by studying the original artefact's 3D replica in MeshLab software (MeshLab, 2017).

4.2 Structured light scanner

The evaluation tests were based on the 3D model that has been created with a structured light scanner. Specifically, the replica was digitised in 3D with a Smartscan3D-HE produced by Breuckmann (Breuckmann, 2017) (Figure 1).
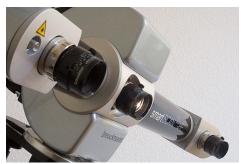

Figure 1: SmartSCAN 3D-HE structured light system
(Breuckmann, 2017)

For the efficient digitisation of the replica, a computer-controlled turntable was used to semi-automate the acquisition process and thus reduce the data collection time. To produce a complete 3D model, several partial scans were performed, which is a common procedure when working with structured light scanning. Eighteen partial scans were performed in total. The post-processing (cleaning, alignment and merging) of the partial scans was performed in the Optocat software suite that was provided by Breuckmann along with their digitisation systems. The average distance between two consecutive points on the 3D model was $\sim 200 \mu m$ and the accuracy of each measurement was $\sim 40 \mu m$. Figure 2 depicts some of the partial scans along with the complete 3D model produced. Colours in Figure 2 indicate data from different partial scans.

Figure 2: Partial scans and the complete 3D model that served as the reference model (ground truth).

4.3 Focus stacking

The experimental setup included a complete photographic studio setting. A Canon EOS 1Ds Mark III camera with a CMOS full frame sensor of 36x24mm was used. A Canon fixed focal length 50mm EF L USM lens was used providing a f/1.2 maximum aperture. This specific focal length of a normal lens, which resembles human vision, was chosen according to the size of the object, to create a tight frame without distortion (imposed by wide angle lenses) or decreased DoF (typically imposed by telephoto lenses). Shooting with a normal lens does not require setting up the camera at a long distance from the object to fill the frame, which limits the need for large space available for the shooting; space is a very important factor when a digitization project takes place in a museum. The camera was fixed on a tripod and was placed at a distance of 62cm from the object, which changed to 70cm and a higher position to be able to capture views from above. The sensor's resolution is 21 MP with pixel pitch 6.4 microns. Shots were taken using raw format.

Special lighting equipment consisting of two Bowens Duo Lite studio flashlights and their control panel (Bowens lights, 2017) with 40x50cm soft boxes were utilized to light the scene. The number of the flashlights as well as the size of the soft box was set to a minimum, as discussed earlier, so that space limitations in museums could be emulated. The use of soft boxes, while shooting museum artefacts, is a prerequisite, as any other lighting without a protective cabinet is forbidden to compensate for the invasive nature of high power lighting on the artefacts. In addition, this type of illumination ensures soft diffused lighting and the elimination of unwanted shadows and glazes on the objects' surface.

During the flash metering, the flash synchronization shutter speed of 1/60 was selected and the flashlights were adjusted to their maximum power, in order to achieve the narrowest available aperture of f/16. This combination of shutter speed and aperture value secured the largest depth of field. Sensor sensitivity was also set at ISO 100 (camera nominal value) to ensure the lowest digital noise level attainable and to provide clear detailed images. The flashlights were placed to the right and left of the object, at an angle of 45 degrees to achieve equable illumination of the object. The camera was stationary, mounted on a tripod, whereas the object was placed on a turntable and was rotated using 10-degrees steps for a whole 360-degree rotation (36 photos per rotation).

Although the use of a normal lens and its narrowest aperture contribute to a wide depth of field, it was apparent that the closest and furthest points of the object in each captured image were not sharp as those in mid-distances (the perfect focused distance). Focus stacking was called upon as a solution to attain perfectly focused images throughout the distance range of the object. Helicon Remote software (Helicon Remote, 2017) was used to automate the focus stacking technique. The camera was connected to a laptop and was controlled by Helicon Remote (Figure 3). For each step the nearest and the farthest points were estimated with the use of the live view tool, in order to estimate the boundaries of the focusing distances. The DoF has been calculated with the use of the DoF calculator, an application for Android devices (DoF calculator, 2017).

Figure 3: Focus stacking acquisition set-up

The application calculates the DoF by using the camera model, the focal length, the focusing distance and the aperture value. The average focusing distance was calculated about *62cm* in the case that the object was photographed around and the aperture value was *f/16*, so the DoF was calculated *13cm*. The focusing distance changed in the session, in which the object was photographed from above and was set to *70cm*. In this case, the DoF was estimated to *18cm*. Helicon Remote software also estimates the DoF according to the same parameters, with the only difference that it is measured in focusing steps. The software has two options to take images in different focusing distances: automatically and manually. In auto mode, the software calculates the number of images that are necessary for the focus stacking procedure and the interval between two shots. In the manual operation, the user selects either the number of shots or the interval between them. It is preferable the interval between two shots to be smaller than the DoF, in order to attain the sharpest image. Thus, the interval between two shots was selected to be 80 focusing steps and the number of images were 5. Figure 4 illustrates an image with a shallow Depth of Field.
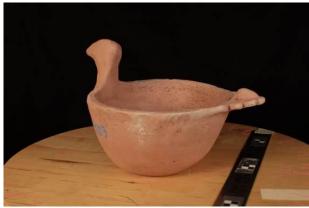

Figure 4: Image with shallow DoF

In this experiment, the photo set with the focus plane positioned in the middle of depth of the artefact was used to create another 3D model that served as the "Best-Focus" (BF) model for comparison. Figure 5 illustrated such an image.

Figure 5: An image used to create the BF model

5. DATA INTERGRATION

5.1 Focus stacking imaging

For the development of the focus stacking image the *pyramid approach* was used with the help of the Zerene Stacker software. Figure 6 represents an example of a focus stacked image with the pyramid approach. Having taken five images for each focus stacked image, the time necessary to produce each composite image was approx. 2-3 minutes. Hence about 90 minutes for the whole data set.
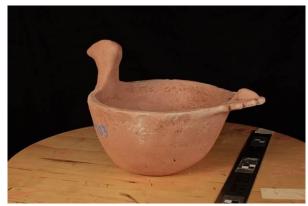

Figure 6: Focus stacked image

5.2 Image Based 3D Reconstruction

Agisoft Photoscan was used for the creation of the 3D models of the artefact. Two 3D models were created: the "Best focus" (BF) 3D model and one corresponding to the reconstruction from the focus stacking technique (FS). For each case a 3D triangular mesh was also created. Figure 7 depicts different views of the FS 3D model.

6. EVALUATION

An objective evaluation of the quality of the 3D models was performed. This evaluation was based on the number of the vertices produced during the alignment process, the dense point cloud generation and the number of the facets produced during the 3D mesh creation. Table 1 represents these values between the Best focus model and the FS model. The relative difference between the total number of vertices of the two sparse point clouds (BF and FS) is almost 10% (7.83). This is an indicator that the SFM algorithm can detect a higher number of matching points and thus result to a denser and hence more accurate

sparse point cloud. Note that both 3D models were produced using the same set of parameters during the SFM processing.
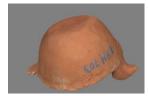

Figure 7: Different views of the FS 3D model of the artefact

Table 1 shows that the 3D model created from BF images has a greater number of vertices produced during the image alignment process than the model created by focus stacking images. On the other hand, Table 1 shows that in the dense point cloud generation both 3D models have produced similar number of vertices. This seems that the focus stacking technique affects the process of the dense image matching and 3D triangulation mesh.

	Vertices in Sparse Point Cloud	Vertices in Dense Point Cloud
BF	63,706	7,932,385
FS	58,720	7,957,583

Table 1: Results from Agisoft Photoscan

A subjective comparison by optical inspection of the 3D models was performed. The BF 3D model exhibits a higher level of noise when compared to the FS 3D model. Figure 8 illustrates the different surface noise levels between the FS, the reference, and the BF 3D models. A radiance scaling rendering shader has been used to accentuate the differences. Notice in the zoomed areas depicted in Figure 8 that the FS model is more similar to the reference one as surface features are more visible than on the BF model. It obvious that these details are lost due to the higher level of noise found in the BF 3D model.
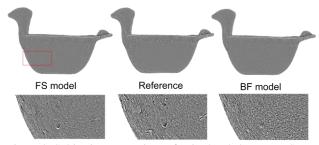

Figure 8: Subjective comparison of noise levels between FS, BF and reference models

6.1 Alignment and Registration

For the comparison of the SfM-MVS models with the reference model produced by the structured light scanner it was necessary for all of them to be at the same reference coordinate system. The CloudCompare (CC) software (CloudCompare) was used for this procedure. The alignment process was carried out within CC, in which common points were manually selected

between each pair of models. After the alignment, the registration between each pair of models was performed based on the ICP (Iterative Closest Point) algorithm. Table 2 presents the errors that were calculated during the alignment and registration procedure.

Procedure	BF-reference	FS- reference
Alignment	0.43	0.40
Registration	0.07	0.06

Table 2: Errors (in mm) in alignment and registration procedure

6.2 3D surface deviation comparisons

For the comparison of the surface deviation between each model pair the function of the cloud-to-mesh distance in CloudCompare software was used. The function computes the distances between each vertex of the point cloud to the nearest triangle of the mesh surface. This distance is calculated as follows: if the orthogonal projection of the vertex lies inside the surface defined by a triangle the distance between the vertex and its point-of-intersection on the surface is calculated; else the algorithms estimates the distance between the vertex and its projection to the nearest edge or vertex of the triangle (Koutsoudis et al. 2013). The BF and FS 3D models were used as point clouds whereas the reference 3D model was used as triangular mesh. Table 3 summarizes the surface deviations computed for the BF and FS 3D models in comparison to the reference model.

Procedure	BF-reference	FS- reference
Mean distance (mm)	0.003364	-0.002669
Std deviation (mm)	0.066985	0.074522

Table 3: Surface deviations of BF, FS 3D models in relation to the reference 3D model.

7. CONCLUDING REMARKS

Focus stacking is a technique that is mainly used in macro photography, in which the problem of the shallow DoF is obvious. In this paper, an attempt was presented to examine if the focus stacking technique affects the image-based 3D reconstruction in the case of the use of a normal lens. The focus stacking technique affects the process in that it results in a less noisy 3D model with more apparent surface details. As this paper presents a first attempt to identify benefits of focus stacking in artefact digitization, it is apparent that further investigation is needed.

REFERENCES

Brecko, J., Mathys, A., Dekoninck, W., Leponce, M., VandenSpiegel, D. and Semal, P., 2014. Focus stacking: Comparing commercial top-end set-ups with a semi-automatic low budget approach. A possible solution for mass digitization of type specimens. *ZooKeys*, *464*, p.1.

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,* Volume XLI-B5, pp. 229-236.

Gallo, A., Bruno, F., Muzzupappa, M. & La, M. F., 2012. Multi-view 3D reconstruction of small stone samples deteriorated by Marine organisms. *18th International Conference on Virtual Systems and Multimedia (VSMM)*, IEEE, pp. 181-187.

Gallo, A., Muzzupappa, M. & Bruno, F., 2014. 3D reconstruction of small sized objects from a sequence of multi-focused images. *Journal of Cultural Heritage,* 15(2), pp. 173-182.

Kontogianni, G., Chliverou, R., Koutsoudis, A., Pavlidis, G., and Georgopoulos, A., 2017, INVESTIGATING THE EFFECT OF FOCUS STACKING ON SFM-MVS ALGORITHMS, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-2/W3, 385-389, doi: 10.5194/isprs-archives-XLII-2-W3-385-2017

Koutsoudis, A., Vidmar, B., Arnaoutoglou F.,2013, Performance evaluation of a multi-image 3D reconstruction software on a low-feature artefact. *Journal of Archaeological Science,* 40(12), pp. 4450-4456.

Plisson, H. & Zotkina, L. V., 2015. From 2D to 3D at macro- and microscopic scale in rock art studies. *Digital Applications in Archaeology and Cultural Heritage,* 2(2), pp. 102-119.

Bowens Lights, http://www.bowens.co.uk/, (last accessed 06/2017)

Breuckmann, *smartSCAN 3D-HE*, http://www.eotech.fr/Fiches/produits/36_smartScan_HE_e_web.pdf, (last accessed 05/2016)

Cambridge in Colour, *Focus stacking & depth of field*, http://www.cambridgeincolour.com/tutorials/focus-stacking.htm, (last accessed 05/2017)

CloudCompare, http://www.danielgm.net/cc/, (last accessed 06/2017).

DoF application for Android devices, https://play.google.com/store/apps/details?id=com.aimenrg.dof (last accessed 06/2017)

Dpreview, https://www.dpreview.com/forums/post/58754021 (last accessed 06/2017)

Hampson Archaeological Museum, http://hampson.cast.uark.edu/, (last accessed 05/2016)

MeshLab, http://www.meshlab.net/ (last accessed 06/2017)

PRESIOUS - Predictive digitization, restoration and degradation assessment of cultural heritage objects, (FP7 - STREP 600533), http://www.presious.eu/, (last accessed 05/2017)

Zerene stacker software, http://zerenesystems.com/cms/home (last accessed 06/2017)