

QUANTIFYING DEPTH OF FIELD AND SHARPNESS FOR IMAGE-BASED 3D RECONSTRUCTION OF HERITAGE OBJECTS

E. Keats Webb^{1,2*}, S. Robson³, R. Evans²

¹ Smithsonian's Museum Conservation Institute, Suitland, Maryland, USA – webbekeats@si.edu

² University of Brighton, Brighton, UK – R.P.Evans@brighton.ac.uk

³ Dept. of Civil, Environmental and Geomatic Engineering, University College London, London, UK – s.robson@ucl.ac.uk

KEY WORDS: Image-based 3D reconstruction, cultural heritage, depth of field, sharpness

ABSTRACT:

Image-based 3D reconstruction processing tools assume sharp focus across the entire object being imaged, but depth of field (DOF) can be a limitation when imaging small to medium sized objects resulting in variation in image sharpness with range from the camera. While DOF is well understood in the context of photographic imaging and it is considered with the acquisition for image-based 3D reconstruction, an “acceptable” level of sharpness and associated “circle of confusion” has not yet been quantified for the 3D case. The work described in this paper contributes to the understanding and quantification of acceptable sharpness by providing evidence of the influence of DOF on the 3D reconstruction of small to medium sized museum objects. Spatial frequency analysis using established collections photography imaging guidelines and targets is used to connect input image quality with 3D reconstruction output quality. Combining quantitative spatial frequency analysis with metrics from a series of comparative 3D reconstructions provides insights into the connection between DOF and output model quality. Lab-based quantification of DOF is used to investigate the influence of sharpness on the output 3D reconstruction to better understand the effects of lens aperture, camera to object surface angle, and taking distance. The outcome provides evidence of the role of DOF in image-based 3D reconstruction and it is briefly presented how masks derived from image content and depth maps can be used to remove unsharp image content and optimise structure from motion (SfM) and multiview stereo (MVS) workflows.

1. INTRODUCTION

3D reconstruction processes assume that the object being recorded is “acceptably” sharp throughout the input image set. However, with small objects requiring close-up imagery limited depth of field (DOF) can be an issue. DOF describes the range of acceptable image sharpness both in front of and behind the plane of sharp focus. While DOF can be quantified for photographic imaging, an “acceptable” value for the image sharpness has not yet been quantified in the photogrammetric and computer vision communities where images are to be used for 3D reconstruction (Verhoeven, 2018).

1.1 Depth of Field (DOF) for 2D and 3D Imaging

DOF can be calculated from the focused distance, lens focal length, aperture, and the diameter of the circle of confusion (Ray, 2002). With a practical imaging system, light does not focus into a point but into a spot that is referred to as the circle of confusion, or blur circle. In photography, the circle of confusion is considered the parameter of acceptable blurriness or the criterion of permissible unsharpness. An observer views an object point in an image as sharp if the diameter of the circle of confusion is under the resolution limit (Luhmann et al., 2014). The diameter of the circle of confusion relates to the point spread function of the imaging system and can be regarded as the smallest element that a digital imaging system can resolve. A widely used circle of confusion diameter is 0.03 mm for the full frame 24 x 36 mm image format.

DOF is at its most problematic when photographing fine detail on small objects. Often the camera must be moved close to the object or the lens focal length increased in order to maximise magnification and fill the frame with the view of the small object. Both longer focal lengths and close working decrease the

DOF. Macro lenses offer an optimised lens option for close-up photography allowing higher magnification that can be achieved with conventional lenses, but they do not alter the DOF. Whilst decreasing lens aperture increases the DOF, small apertures reduce image quality due to optical diffraction. Such effects are widely known in the field of heritage imaging (Menna et al., 2012; Percoco et al., 2017; Sapirstein, 2018).

Image-based 3D reconstruction relies upon local image brightness variations for feature detection and dense matching. The visual and scientific value of output 3D reconstructions also benefit from low image noise and consistent textures for image draping. Whilst the need for consistent DOF across an imaged object is well recognized, the DOF requirement for 3D reconstruction given its definition from perceptual quantity and the idea of “acceptable” sharpness embodied in the circle of confusion has yet to be quantified (Verhoeven, 2018).

1.2 Spatial Frequency Response

Spatial Frequency Response (SFR) provides a measure of image contrast loss as a function of spatial frequency. It is important in the context of this work as it provides information about an imaging system's ability to maintain contrast as image details get smaller. The method is based on slanted-edge features in a target, and the SFR is derived from the Fourier transform of the line spread function (ISO 12233:2017). SFR results are reported by plotting the modulation level versus spatial frequency. The SFR at the 10% modulation provides a measure for the limiting resolution of the system, and the SFR at the 50% modulation provides a threshold as a sharpness indicator (ISO 19264-1:2017). The limiting resolution is the smallest distance between image points that can still be resolved (Burns and Williams, 2008). With both the 10% SFR and 50% SFR, the aim is to achieve the highest frequency but to not exceed the

* Corresponding author

Nyquist limit. The Nyquist frequency is the highest frequency that can be reliably reproduced without aliasing, and it is the half-sampling frequency or 0.5 cycles/pixel. In turn, sampling efficiency provides a convenient single value measure for comparing multiple SFR results (Burns and Williams, 2008; ISO 19264-1:2017). The sampling efficiency is the ratio of the limiting resolution to the sampling resolution. The limiting resolution is the frequency at which the SFR falls to 10% which can be calculated from the SFR of a slanted edge.

1.3 Improved DOF for 3D Imaging

A range of solutions have been developed to address DOF challenges such as extending DOF and increasing overall image sharpness. Outcomes include processing algorithms, hardware solutions and increasingly accessible commercial software.

Focus stacking extends DOF by combining a set of images acquired with varying focal positions into a single image with an increased DOF. Focusing stacking examples include 3D reconstruction of small archaeological objects and samples of encrustation from a marble statue (Gallo et al., 2014; Clini et al., 2016). Techniques often use commercial software like Zerene Stacker and Helicon Soft. However, the composite image contains zones of differing magnification and viewpoint making it problematic for photogrammetric self-calibration and the interior and exterior orientation determination. Focus stacking also increases the acquisition and processing times due to the increased number of images, often being cited as reasons that focus stacking was not selected (Gallo et al., 2014; Marziali and Dionisio, 2017; Verhoeven and Missinne, 2017; Sapirstein, 2018).

Hardware solutions are available for extending DOF including focus stacking rails automating camera movement to different focal positions (Nobel, 2017) and light-field camera systems (Levoy, 2006). Most accessible in the consumer market are digital camera modes for acquiring image sets with different focal positions. Current examples include live composite and focus stacking imaging modes, post focus simulation and focus stacking capabilities and focus shift capabilities.

DOF processing algorithms are an active research area in computer vision, most notable is Shape From Focus (SFF) (Nayar and Nakagawa, 1994). Local focus variations are used as depth cues while focus measure operators compute the focus level for each pixel in the image. The method derives shape from an image sequence of the same scene with variation in the focus. SFF has been used for a single view, micro imaging; however, Pertuz et al. (2013) and Billiot et al., (2013) worked to extend its use from the well-controlled scenarios of microscopy to complex, real scenes using conventional cameras.

1.4 Masking

Image pre-processing methods have been implemented into the image-based 3D reconstruction workflows to optimise images and increase the processing performance (Barazzetti et al., 2010; Remondino et al., 2016). Pre-processing can enhance image features that are important for the 3D reconstruction by improving local image feature contrast using a Wallis filter (Barazzetti et al., 2010; Remondino et al., 2016) or reducing noise with an adaptive smoothing filter (Remondino and El-Hakim, 2006; Barazzetti et al., 2010). Another widely explored pre-process is to mask out the background or non-essential features (Barazzetti et al., 2010; Koutsoudis et al., 2013; Gallo et al., 2014; Guidi et al., 2014; Troisi et al., 2015; Abate et al., 2016; Marziali and Dionisio, 2017; Sapirstein and Murray,

2017). These studies report improvements in alignment quality (Abate et al., 2016) and decreases in reconstruction processing times up to 75% (Koutsoudis et al., 2013; Gallo et al., 2014; Troisi et al., 2015). Background masking methods have included masking directly in the Agisoft 3D reconstruction software or using external image processing software like Photoshop (Porter et al., 2016; Marziali and Dionisio, 2017). Building on an initial step of background masking from Porter et al. (2016), Sapirstein (2018) presented a technique that using a low-resolution mesh to create masks in PhotoScan for a more precise object mask. Most studies have masked out the background and non-essential features; very few have discussed masking related to sharpness and DOF.

Verhoeven (2018) described masking out areas that are unsharp as fairly standard practice and detailed defocus estimating algorithms for automatic masking based on sharpness to speed up the reconstruction process. The work described a Matlab toolbox with fifteen working methods for mapping defocus blur and reported on three edge-based methods assessing accuracy, running time, and robustness. Verhoeven noted limitations on these available methods. Verhoeven recognised that there was a potential for some of the edge-based methods for masking out homogenous areas and those without edges (skies or studio backgrounds) and concluded that additional improvements were required for future implementation.

2. MATERIALS AND METHODS

This paper seeks to better understand the influence of lens aperture, viewing angle, object distance and image sharpness to evidence the limitations of DOF on the quality of 3D reconstruction. The work is carried out in the context of steps in a SfM-MVS workflow looking for simple changes in approach for example including automatic sharpness-based image masking as an accessible practical process.

2.1 Test Objects

Test objects and reference data are essential for reliable and repeatable quantification of both input 2D images and output 3D reconstructions. Test objects can be specifically designed to control variables and assess quality and can be replicated in a more reliable way than relying on heritage objects. Three test objects were used as part of this study: the Panel target, the Mango Vase, and the DICE target (Figure 1).



Figure 1. Panel target (left); Mango Vase (centre); DICE target (right).

The Panel target, is an aluminium plate (30.5 x 30 x 0.8 cm) coated with a pseudo-random pattern optimised for the optical detection of surface strain in engineered surfaces with Digital Image Correlation (Sargeant et al., 2016). The pseudo-random pattern has a high local contrast and small features, which are optimised to increase the sensitivity and practicality of image matching processes.

A planar surface should provide a simple geometric shape that can be assessed through a best-fit plane, or a flatness measure. However, the physical surface must be flat to better than the order of 0.01 mm detectable through 3D reconstruction at the imaging scales used. Repeated photogrammetric measurement confirmed that the Panel was not flat with a systematic saddle shaped pattern with maxima of the order of +/- 0.15 mm. An alternative to plane fitting was therefore required. One low-cost approach was to use an average of four 3D reconstructions as a reference surface. Assessments using the average mesh revealed a ripple pattern with an amplitude of about +/- 0.02 mm (Figure 2, left). An independent measurement was therefore conducted using a laser tracker and probe with six degrees of freedom to assess the reliability of the average mesh. Probing confirmed the plate's overall unflatness. The ripple pattern was absent from the probe data (Figure 2, right) highlighting that it was most likely attributable to unmodeled lens distortion in the 3D reconstruction and represents a fundamental limitation of the 3D reconstruction approach.

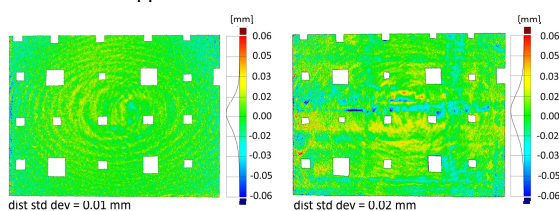


Figure 2. Surface discrepancy maps. 3D reconstruction with average mesh (left) and probe measurements (right).

Instrumentation to provide an independent check is not always available for heritage projects and the resulting measures may not provide the information necessary. For example, a touch probe includes discrete points and is unlikely to give the same surface sampling density to highlight local detail, but it provided an independent check on the overall shape that did not have the same systematic errors of a photogrammetric measurement. To reflect a low-cost heritage approach, the average mesh was used for assessment in this paper, with the caveat that it is not an independent reference and accumulated the systematic errors of the reconstructions used to create it.

Whilst the continuous surface, high contrast features make the Panel ideal for metric testing, the lack of sharp edge changes, discontinuities and 2D shape of the Panel target cannot provide full system performance information when documenting a similarly high contrast museum object in the round.

A second custom test object, the Mango Vase, includes line patterns, pigment patches, and varnished areas on a wood vase (19 x 13 cm) (Webb, 2015, 2020). Linking more closely with heritage objects and materials, the Mango Vase was used to provide evidence for the limitations of DOF when recording a small detailed object in the round. The reference surface for the Mango Vase was a 3D scan made with an AICON3D smartSCAN-HE structured light scanner provided by the Smithsonian Digitization Program Office. Once configured with an S-150 field of view, 240 mm base length and 370 mm working distance, the scanner was calibrated with a calibration plate following the AICON3D scanning procedure. An AICON3D automated turntable was used to automate the recording of the object in the round. The rotational symmetry of the vase made alignment of the reference data with the 3D reconstruction results challenging and any offset in the alignment impacted the results of the comparison. Comparisons between point clouds showed localized systematic discrepancies

in the areas around the neck of the vase and in looking at solutions in the round. The maximum magnitude of 0.6 mm can significantly impact alignment accuracy, limiting critical activities such as assessing change in object condition from 3D reconstruction results.

The third object includes a test pattern designed to quantify sampling efficiency as part of a camera characterization and image quality assessment process at each lens aperture, viewing angle, and camera to subject range used. The target is an output from the US-based Federal Agencies Digitization Guidelines Initiative (FADGI) (Rieger, 2016) and is available commercially as the Digital Imaging Conformance Evaluation (DICE) target. Spatial frequency analysis was conducted using the accompanying GoldenThread software (Image Science Associates, Rochester, NY, USA) and *sfrmat3* (Peter D. Burns, LosBurns Imaging Software). Quantitative analysis of SFR and sampling efficiency computed from each DICE target image were combined with metrics from a series of comparative 3D reconstructions to provide insights into the connection between DOF and output model quality.

2.2 Imaging System, Processing and Assessment

The camera used for this research was a Canon 5D Mark II with a Coastal Optics 60mm UV-VIS-IR apochromatic macro lens. This camera has a full-frame CMOS sensor (36 x 24 mm) with a maximum resolution of 21.1 MP (5,616 x 3,744 pixels) and a pixel pitch of 6.4 μm . The lens has no focus shift from UV through IR and is specified for low aberration, low distortion spectral imaging in forensics, scientific and fine art imaging. Two Bowers Gemini GM400Rx studio strobes with umbrellas were used to illuminate the targets. RAW images were acquired and processed using an Adobe Camera RAW (ACR) RAW processing workflow.

Image-based 3D reconstructions were processed using SfM-MVS through Agisoft PhotoScan Pro version 1.3.3. The processing followed the *error minimisation workflow* developed by Cultural Heritage Imaging (CHI) and the US Bureau of Land Management (BLM) (Schroer et al., 2017). Processing includes an initial alignment of the images and the creation of a sparse point cloud. The optimised workflow is an iterative process of gradually selecting and removing points while performing bundle adjustments to refine the alignment of the images and sparse point cloud and to optimise camera calibration. The final steps include building a dense point cloud, generating a mesh and exporting a model. For this research, a limited parameter set was selected (for two radial distortion coefficients and two tangential distortion coefficients) for the camera model to avoid over-parameterisation (James et al., 2017).

Input images were assessed through the FADGI star rating which provides an indication of the acceptable level of sharpness relating to paintings and other two dimension art (Rieger, 2016, p. 47). 3D reconstructions were assessed using the free version of the GOM Inspect software. Certified by both PTB and NIST, this provides an accessible and traceable tool for heritage recording professionals allowing surface deviations from a reference or comparative surface to be visualised as coloured discrepancy maps with associated histograms.

2.3 DOF Imaging Tests

2.3.1 Panel target

Two tests were conducted using the Panel. "DOF-3D-Plane" investigated the effect of aperture with incremental changes in

the camera-object angle, whilst “DOF-Sharp-SFR” assessed change in aperture with incremental variation in camera-object distance. Both tests imaged the DICE target at the middle camera position at each of the camera-object angles and distances to link 2D image quality to the 3D reconstruction. Tests were carried out at apertures of $f/5.6$, $f/11$ and $f/32$ in order to understand image quality influences in terms of DOF, the range of “acceptable” sharpness, and diffraction.

All images were acquired at ISO 100 with a 1/100 sec shutter speed. The flash output power of the GM400Rx strobe was increased from 1 through 3 to 6 as the aperture diameter decreased from $f/5.6$ through $f/11$ to $f/32$ to maintain consistent illumination at the sensor surface. Optimum focus was set using Live View at 200% magnification, viewing the central feature of the DICE target or the pattern at the centre of the Panel target. Focus was held fixed for each lens aperture setting.

An experimental imaging geometry was required that could reproducibly position the camera to capture convergent image networks while maintaining the consistent relationships between the Panel target and illumination. This was achieved by modifying components of a *camera positioning robot* (Sargeant et al., 2013). The Panel target was mounted on a fixture that could be moved up and down and be locked in place along the section with an initial camera-object distance of 600 mm (Figure 3). Two Newport LMS linear stages were stacked allowing the camera to be positioned in an area 300 x 600 mm with the camera height staying constant. The rotating stage was stacked onto the linear stages with the camera was mounted on top with a locked tripod head allowing for the camera angle to change in relation to the target. The setup allowed the camera-object angle to be incrementally changed by 10° starting from 0° and rotating to 30° .

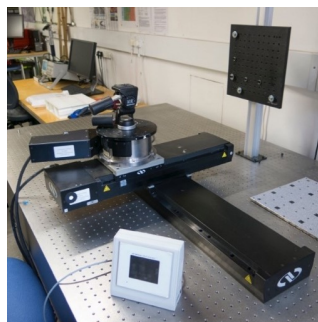


Figure 3. Camera positioning setup.

For each Panel imaging geometry, one central and six convergent images were acquired from three target heights and three camera positions. In each case the camera pointed towards the centre of the target such that optical axes converged behind the object plane. Image networks were acquired at $f/5.6$, $f/11$ and $f/32$ over four Panel angles ($a_1 = 0^\circ$, $a_2 = 10^\circ$, $a_3 = 20^\circ$, $a_4 = 30^\circ$). SFR measures were made using *sfrmat3* as only the images from the 0° angles were readable in GoldenThread software, likely due to target image geometry (Webb 2020).

For the DOF-Sharp-SFR test the procedure for the 0° viewing angle was repeated with the addition that the camera-object distance was incrementally changed in 5 mm increments moving through and beyond the range of computed DOF, to assess sharpness and resulting 3D reconstructions. The acquisition for this test included a range of -15 mm to +50 mm (from sharp focus position) at $f/5.6$ and $f/11$. The DOF range at $f/32$ (197.04 mm) was too big for the available laboratory space.

2.3.2 Mango Vase Test Object

Tests with the Mango Vase investigated the limitations of DOF with considerations of the direct applicability to museum objects. Image sets were acquired with five apertures ($f/5.6$, $f/8$, $f/11$, $f/16$, $f/32$) at ISO 100 with a 1/100 sec shutter speed. The following list includes the aperture and corresponding GM400Rx flash power: $f/5.6$ (1); $f/8$ (2); $f/11$ (3); $f/16$ (4); $f/32$ (6) needed for consistent illumination.

With the camera mounted on a tripod, a central turntable setup suited to imaging small to medium size heritage objects in the round was used to record the Mango Vase. The turntable allowed the object to be rotated at a constant 600 mm camera-object distance with minimal object handling whilst lighting with GM400Rx strobes and umbrellas allowed consistent shadow free illumination (Webb et al. 2015).

2.3.3 Masking

SfM-MVS workflow allows image content to be masked or blocked out if it is to be ignored in the 3D reconstruction computation. This procedure is designed to remove background from objects, but if unsharp image regions are detectable they can in principle be masked. Two potential automatic masking methods, the first based on image content and the second on depth maps, were explored.

Image content masking used the Adobe Photoshop CC 2019 focus area selection tool to mask out regions of the image that were considered out-of-focus. Software selection is based on unit-less user parameters: “In-Focus Range” and “Image Noise Level” with little technical information concerning the implementation. Settings were established by masking the DICE target and the Panel with the same settings to correlate the 2D image quality measured from the DICE target with the resulting masks of the Panel. Once the parameters were selected, a Photoshop Action script was recorded to batch process an image set with the image content masks.

The second method used 8 bit greyscale depth maps (or range images) in which pixel values denote camera-object distance. In PhotoScan, depth maps can be generated as output for each image during a first pass through the dense point cloud processing. The depth maps were used to create binary masks in Matlab with these masks correlating to sharpness levels based on camera-object distance and DOF.

3. RESULTS AND DISCUSSION

3.1 DOF-3D-Plane

The sampling efficiency results and FADGI star rating from the DICE target (Table 1) can be linked to the surface discrepancy maps of the 3D reconstruction results of the Panel (Figure 4).

At $f/32$ the complete target was within the calculated DOF at all angles, hence all image content was within the acceptable range of sharpness. Consistent 3D reconstructions were produced at all four angles (range 0.12 mm, standard deviation, σ , 0.2 mm). However, due to diffraction at this small aperture, the input image quality is well below the minimal FADGI 1-star rating and would not be acceptable as a museum record.

The $f/5.6$ datasets with the smallest DOF achieved 1-star rating at a_1 (0°) and had the poorest sampling efficiency with increasing angle. DOF is a clear limitation on the 3D reconstructions starting at a_2 (10°) (range 0.18 mm, σ 0.04 mm)

reducing further to (range 0.3 mm, σ 0.07 mm) by a4 (30°). A small DOF from a large aperture diameter should be avoided due to the overall reduction of image quality and the impact of the small DOF on 3D reconstructions.

f/5.6			
Horizontal			
	Left	Centre	Right
a1	68	65	61
a2	25	58	66
a3	14	54	27
a4	12	52	16

f/11			
Horizontal			
	Left	Centre	Right
a1	79	84	79
a2	51	84	79
a3	29	81	53
a4	22	76	33

f/32			
Horizontal			
	Left	Centre	Right
a1	51	55	52
a2	52	55	52
a3	50	54	51
a4	48	52	50

Vertical			
	Left	Centre	Right
a1	60	59	51
a2	21	58	74
a3	13	50	26
a4	11	44	15

Vertical			
	Left	Centre	Right
a1	80	84	77
a2	49	83	80
a3	27	83	48
a4	21	81	30

FADGI star rating			
★★★	★★	★	below 1 star

Table 1. DOF-3D-Plane sampling efficiency (%) and colour coded FADGI star rating. Results are provided for three areas of the DICE target (left, centre, right).

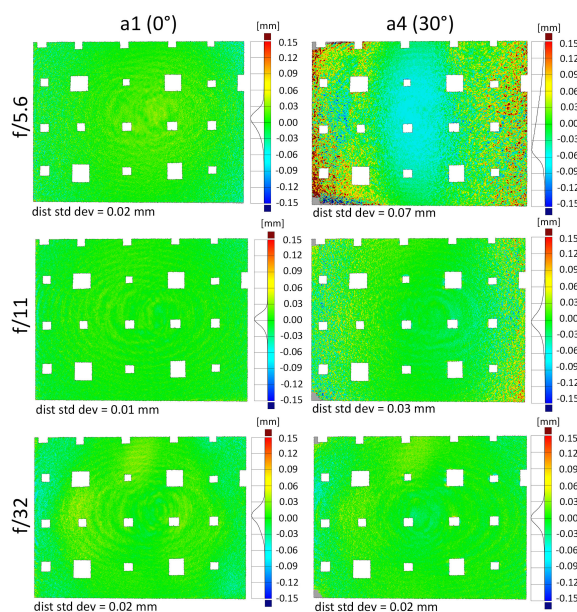


Figure 4. Surface discrepancy maps comparing 3D reconstructions with averaged reference mesh for a1 (0°) and a4 (30°) and three apertures (f/5.6, f/11, f/32).

At f/11 the data give the best image quality from this lens maintaining a FADGI 3-star rating at the centre of the target. Whilst the effect of DOF is observed in the f/11 datasets, there is evidence that high frequency features can give the best 3D reconstructions (range 0.6 mm, σ 0.01 mm) at moderate angles. The f/11 reconstruction data at a3 (20°) is similar to that of the f/32 data; however, in common with the f/8 images, f/11 maintains higher sampling efficiencies than f/32. The

conclusion is that an aperture that best balances DOF and image quality needs careful scrutiny, with each setting being checked for a given imaging configuration.

3.2 DOF-Sharp-SFR

SFR was calculated from the central feature of the DICE target averaging the two horizontal slanted edge features. SFR50, amplitude at 50% SFR, provides a useful way of presenting results and establishing how they relate to DOF.

Whilst the highest SFR50 frequency for the f/5.6 image sets is for the focus position, this is not the case for the f/11 image set and it appears that a better focus was some 10 mm behind the focus position (Figure 5). The 60 mm lens is a manual lens and it is likely that the focus was slightly behind the panel for this image set.

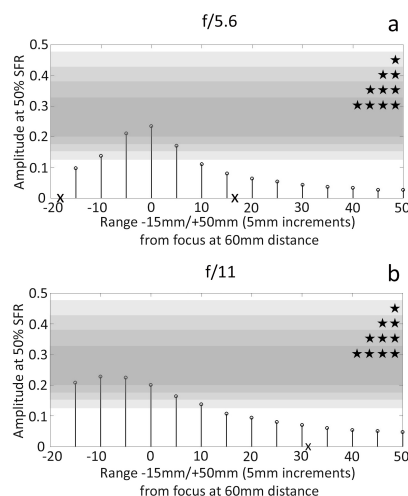


Figure 5. SFR50 for f/5.6 (a) and f/11(b) image sets. Grey range and stars indicate FADGI star ratings quality range and "x" on x-axis marks calculated design DOF.

The SFR50 plots for f/5.6 and f/11 indicate that image quality falls below the FADGI 1-star rating before reaching the near or far limit of the 0.03 mm circle of confusion used in the design DOF calculation. This suggests that a smaller diameter for the circle of confusion is necessary for achieving results within the FADGI star rating system.

3D reconstructions were assessed in GOM Inspect with comparisons to the averaged reference mesh. The f/5.6 results (Figure 6) showed an overall increase in noise (σ 0.03 mm to σ 0.09 mm) as the camera-object distance moves away from the plane of sharp focus. Beyond the DOF (indicated by the white box and starting at 20 mm), more random noise is observed although this is still within $\sim \pm 0.05$ mm.

Even at the maximum distance from the focus position (50 mm) with low image quality (sampling efficiency of 11.5%), the Panel is still reconstructed within ± 0.15 mm with the largest discrepancies along the edges of the model. The visual example of the central feature of the DICE target provides evidence of low image quality and an inability to resolve details on the target (Figure 6). Even though these images are used to produce a 3D reconstruction of the Panel, the images would not be useful as records of a heritage object surface and fall well below any FADGI star rating. The high contrast of the Panel proves to be resilient even when the image data is not sharp.

If the 2D input images were to fall into the FADGI star rating for both $f/5.6$ and $f/11$, the circle of confusion diameter would need to be around 0.01 mm giving a DOF of 11.20 mm for $f/5.6$ and 22.01 mm for $f/11$. This small diameter, one-third of the standard 0.03 mm and under the size of 2 pixels for the Canon 5D Mark II, would only prove beneficial for 2D image quality and remaining within the FADGI star rating guidance.

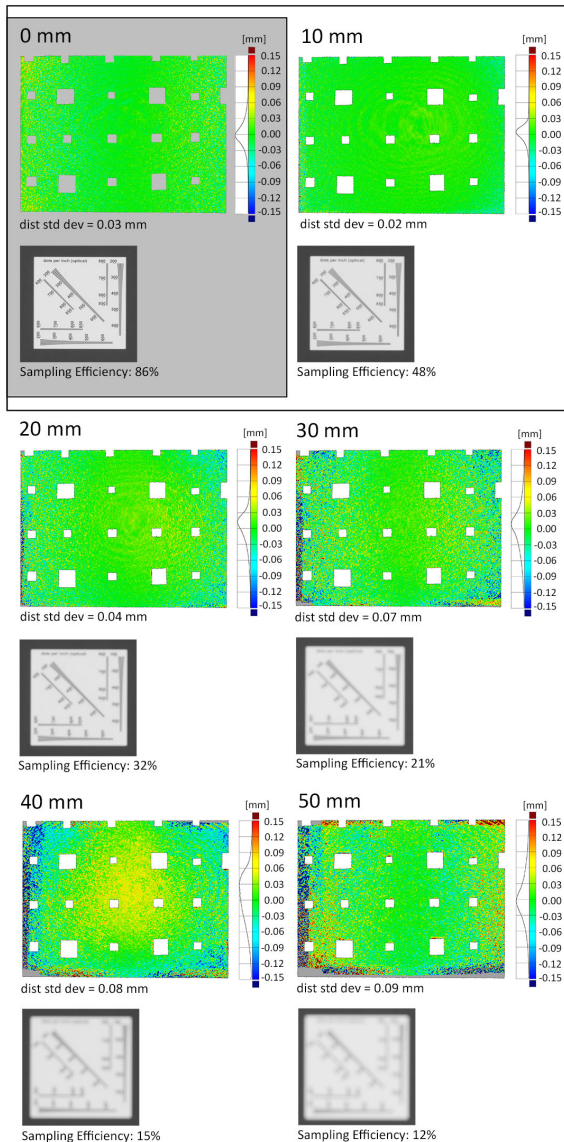


Figure 6. Surface discrepancy maps comparing 3D reconstructions ($f/5.6$ image sets) with averaged reference mesh at different camera-object distances. Image details of the DICE target central feature visualize decreasing sampling efficiency with changing camera-object distance. The grey box indicates result within the FADGI star rating and the white box indicates positions within the calculated DOF.

The DOF-Sharp-SFR 3D reconstruction results provide more flexibility for an increased diameter for the circle of confusion and DOF. Within the parameters of this experiment, a 3D reconstruction of the panel is always produced. At all distances for both $f/5.6$ and $f/11$, the image orientation and dense matching are successful and resulting models have less than +/-

0.15 mm deviation from the average mesh. If the maximum distance from the focus position for the test (50 mm) was used as the extreme near limit (not knowing how much further the camera-object distance could be reduced before significantly influencing the 3D reconstruction), the diameter of the circle of confusion would be about 0.1 mm for $f/5.6$ and about 0.05 mm for $f/11$ or about 15 pixels for $f/5.6$ and just under 8 pixels for $f/11$ for the Canon 5D Mark II.

The Panel proved to be resilient for 3D reconstructions with image data degraded from unsharpness, and it should be noted that 3D reconstruction circle of confusion estimates are potentially biased by the high local contrast of the pseudo-random pattern which are optimised for image matching. This resilience is unlikely to correspond with most heritage objects, which would likely have surfaces with lower local contrast and lower spatial frequency. Furthermore, an important consideration is the reuse of images, which would require the image content to be “acceptably sharp” for applications separate from 3D reconstruction.

3.3 DOF-3D-MangoVase

Based on spatial frequency analysis (Figure 7) the imaging system performed best around $f/11$ with $f/8$ and $f/16$ showing comparable performance mostly within the FADGI 3- and 4-star ratings. This aligned with what would be expected from practical use of this lens with the optimal aperture being stopped down a few stops from the largest lens aperture diameter. Significantly the $f/16$ results did not show the effect of diffraction, but the $f/32$ did, falling to a FADGI 1-star rating and below.

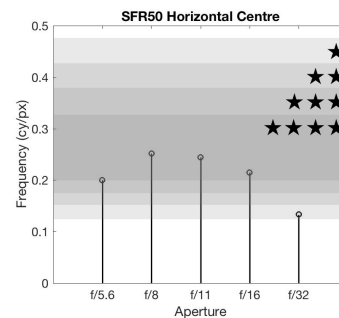


Figure 7. SFR50 for Mango Vase image sets. Grey gradients and stars indicate FADGI star ratings quality range.

Visualisations from the 3D reconstruction process provide evidence of the impact of aperture and DOF on the image matching and tie point identification (Figure 8, top row). This impact was most notable from the highest view of the camera network where the increased camera-object angle revealed the effect of DOF. The results for the identification of tie points for the Mango Vase showed the increased number of tie points and larger coverage area of the object as the DOF increases.

The vase is 190 mm high, so the largest DOF with an aperture of $f/32$ (197.04 mm) would include the full height of the object and this aperture resulted in points in the background and covering the full vase. The image set with the smallest DOF ($f/5.6$) resulted in points clustered around the top and shoulder of the vase where the image data is in focus. These results were expected and provided evidence that the detection of tie points corresponds with our expectations of DOF and related image sharpness needed for both feature detection and dense matching.

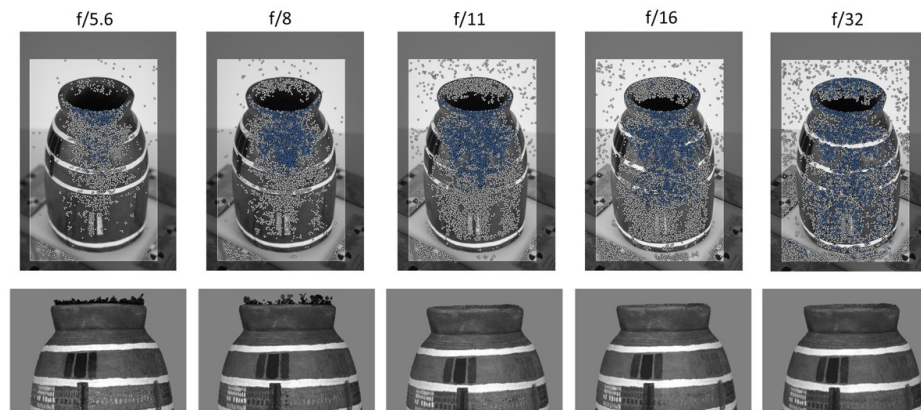


Figure 8. Top Row: PhotoScan matching points (i.e., tie points) results with blue points indicating used matches and correspond to points of the sparse point cloud and white points are unused matches. Bottom Row: Detailed views of resulting models with artefacts resulting from the smallest DOFs ($f/5.6$ and $f/8$).

Output 3D models highlight the effect of aperture selection (Figure 8, bottom row). The two models with the smallest DOF ($f/5.6$ and $f/8$) include artefacts reconstructed around the rim of the vase, whereas these are not present in the models with the larger DOF ($f/11$, $f/16$, $f/32$). The larger DOF means that the rim and the interior of the rim are better resolved in the input images and the model results in a more reliable reconstruction of these features.

The results of the DOF-3D-Mango experiment aligned with the results from the DOF-3D-Plane and showed that the larger aperture diameter (represented by $f/5.6$) should be avoided due to the overall reduced image quality and the impact of the small DOF on the resulting 3D reconstruction. An optimal aperture (represented by $f/11$) had a balance of DOF and image quality with the effects of DOF still be observed at the greater camera-object angles. The smaller aperture diameter (represented by $f/32$) showed similar performance to $f/8$, $f/11$ and $f/16$, but the SFR analysis showed a significant decrease in image quality from diffraction providing evidence that the 3D reconstruction process tolerates the reduced image quality from diffraction. Even though the diffraction is tolerated by the process, high spatial frequency features may not be resolved because of the effect of diffraction on the 2D image quality.

3.4 Masking

Both masking methods consistently increased the number of tie points and projections over the unmasked case (Table 2). This was unexpected with less input image information being available, however it appears that removing poor quality input image data prior to processing benefits the “black-box” PhotoScan workflow. However, for most purposes the 3D reconstruction benefit of DOF pre-masking is marginal and may not justify the additional time investment for the methods presented here.

Masking has proven beneficial for the 3D reconstruction process as evidenced by the internal masking tools in PhotoScan and the pre-processing workflows to mask out the background or non-essential features. Masking has been used to improve the quality of the alignment and decrease the reconstruction processing time, so by streamlining methods masking from sharpness would be beneficial. Future research could include testing with image stacking to improve the reliability of the camera calibration with the stacking workflow, investigating focus measure operators from Shape From Focus (SFF) as a means for masking based on sharpness, and even testing the Matlab toolbox presented by Verhoeven (2018) despite stated

challenges and limitations with available defocus estimating algorithms.

	No Masks	Depth Map Masks	Image Content Masks
tie points	30,218	45,682	44,998
projections	84,719	126,855	125,504
reprojection error (pix)	0.21	0.21	0.22
dense pt cloud (pts)	13,141,477	13,005,868	11,846,533

Table 2. PhotoScan processing results for masking methods of Mango Vase $f/5.6$ image sets

4. CONCLUSIONS

The experimental design (test objects, image geometries and lens settings) ensured that DOF when recording small objects would impact input image sampling efficiency metrics to levels below 30% and even down to 12% SFR. Based on the FADGI star rating used to define acceptability for 2D heritage imaging, results demonstrate that 3D reconstructions are tolerant of lower SFR values such that reconstructions can be generated from images containing significantly more blur (sampling efficiency as low as 12%) than would be acceptable for 2D images acting as an object record (sampling efficiency above 80%). An optimal aperture (e.g., $f/11$) balances DOF and image quality.

With a planar test object, the gradual increase in range from the camera and commensurate decrease in SFR shows how the central part of the field holds up well whilst the edges of the images show increasing levels of reconstruction noise altering the shape of the surface discrepancy histograms. If the maximum distance from the focus position for the test (50 mm) was used as the DOF near limit, the diameter of the circle of confusion would be 0.1 mm for $f/5.6$. At this distance with a sampling efficiency of 12%, the image quality is unacceptable for an object record, but acceptable for 3D reconstruction.

Heritage records are often used for multiple purposes. If the 2D input images were to fall within the FADGI star rating ensuring an appropriate image quality as an object record, the diameter of the circle of confusion would need to be 0.01 mm for $f/5.6$, one-third of the routinely used 0.03 mm diameter.

Whilst we can suggest values that might be used for an acceptable circle of confusion in the DOF algorithms, the

increase in the usable range of image sharpness from 2D to 3D recording is influenced by the localised optical properties of the surfaces being recorded. This is unsurprising given the importance of local image gradients within dense matching algorithms. Challenges will increase as smaller objects are imaged, extending into macro photography. However, in combination with the automation step of sharpness-based masking which allows image content control under varying DOF and local image detail, these results point to the value of an iterative sharpness-based masking step in the 3D reconstruction workflow.

ACKNOWLEDGEMENTS

This work is part of a PhD at the University of Brighton and University College London through the EPSRC funded Science and Engineering Arts Heritage and Archaeology (SEAHA) Centre for Doctoral Training (Webb 2020). Project partners were Smithsonian's Museum Conservation Institute and Analytik Ltd.

REFERENCES

- ISO 12233:2017 Photography—Electronic still-picture cameras—Resolution measurements, Geneva: International Organization for Standardization.
- ISO 19264-1:2017 Photography—Archiving systems—Part 1: Reflective originals, Geneva: International Organization for Standardization.
- Abate, D., Hermon, S., Eliades, I., 2016. Virtual and Physical Recomposition of Fragmented Ecclesiastical Frescoes Using a Photogrammetric Approach. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B5, 12-19.
- Barazzetti, L., Scaioni, M., Remondino, F., 2010. Orientation and 3D modelling from markerless terrestrial images: Combining accuracy with automation. *Photogram. Rec.* 25, 356–381.
- Clini, P., Frapiccini, N., Mengoni, M., Nespeca, R., Ruggeri, L., 2016. SFM technique and focus stacking for digital documentation of archaeological artifacts. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B5, 229–236.
- Gallo, A., Muzzupappa, M., Bruno, F., 2014. 3D reconstruction of small sized objects from a sequence of multi-focused images. *J. Cult. Herit.* 15, 173–182.
- Guidi, G., Gonizzi, S., Micoli, L.L., 2014. Image pre-processing for optimizing automated photogrammetry performances, *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, II-5, 145–152.
- James, M.R., Robson, S., Oleire-oltmanns, S., Niethammer, U., 2017. Optimising UAV topographic surveys processed with structure-from-motion: Ground control quality, quantity and bundle adjustment. *Geomorphology* 280, 51–66.
- Koutsoudis, A., Vidmar, B., Armaoutoglou, F., 2013. Performance evaluation of a multi-image 3D reconstruction software on a low-feature artefact. *J. Archaeol. Sci.* 40, 4450–4456.
- Levoy, M., 2006. Light fields and computational imaging. *Computer* 39, 46–55.
- Luhmann, T., Robson, S., Kyle, S., Boehm, J., 2014. *Close-Range Photogrammetry and 3D Imaging*, 2nd ed. De Gruyter, Berlin.
- Marziali, S., Dionisio, G., 2017. Photogrammetry and macro photography. The experience of the MUSINT II Project in the 3D digitizing process of small size archaeological artifacts. *Stud. Digit. Herit.* 1, 298–309.
- Menna, F., Rizzi, A., Nocerino, E., Remondino, F., Gruen, A., 2012. High Resolution 3D Modeling of the Behaim Globe. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXIX-B5, 115–120.
- Nayar, S.K., Nakagawa, Y., 1994. Shape from Focus. *IEEE T. Pattern Anal.* 16, 824–831.
- Nobel, J.J.I.-L., 2017. Buying a Focus Stacking Rail?. The Extremist: Companion blog to extreme-macro.co.uk. URL <http://extreme-macro.co.uk/extremist/buying-a-focus-stacking-rail/> (accessed 1.2.18).
- Percoco, G., Guerra, M.G., Sanchez Salmeron, A.J., Galantucci, L.M., 2017. Experimental investigation on camera calibration for 3D photogrammetric scanning of micro-features for micrometric resolution. *Int. J. Adv. Manuf. Tech.* 1–13.
- Pertuz, S., Puig, D., Garcia, M.A., 2013. Analysis of focus measure operators for shape-from-focus. *Pattern Recogn.* 46, 1415–1432.
- Porter, S.T., Roussel, M., Soressi, M., 2016. A Simple Photogrammetry Rig for the Reliable Creation of 3D Artifact Models in the Field. *Adv. Archaeol. Practice* 4, 71–86.
- Ray, S.F., 2002. *Applied Photographic Optics*, 3rd ed. Focal Press, Oxford.
- Remondino, F., El-Hakim, S., 2006. Image-based 3D Modelling: A Review. *Photogram. Rec.* 21, 269–291.
- Remondino, F., Gaiani, M., Apollonio, F., Ballabeni, A., Ballabeni, M., Morabito, D., Heritage, C., 2016. 3D Documentation of 40 Kilometers of Historical Porticoes – the Challenge. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B5, 12–19.
- Rieger, T. (Ed.), 2016. *Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Files*, Federal Agencies Digitization Guidelines Initiative. Rochester.
- Sapirstein, P., 2018. A high-precision photogrammetric recording system for small artifacts. *J. Cult. Herit.* 31, 33–45.
- Sapirstein, P., Murray, S., 2017. Establishing Best Practices for Photogrammetric Recording During Archaeological Fieldwork. *J. Field Archaeol.* 42, 337–350.
- Sargeant, B., Hosseinaveh, A., Erfani, T., Robson, S., Boehm, J., 2013. A webcam photogrammetric method for robot calibration. *Proc. SPIE* 8791, 1–13.
- Sargeant, B., Robson, S., Szigeti, E., Richardson, P., El-Nounu, A., Rafla, M., 2016. A method to achieve large volume, high accuracy photogrammetric measurements through the use of an actively deformable sensor mounting platform. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B5, 123–129.
- Schroer, C., Mudge, M., Lum, M., 2017. Photogrammetry Training: Practical, Scientific Use of Photogrammetry in Cultural Heritage.
- Troisi, S., Del Pizzo, S., Gaglione, S., Miccio, A., Testa, R.L., 2015. 3D models comparison of complex shell in underwater and dry environments, in: 2015 Underwater 3D Recording and Modeling. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-5/W5, 215–222.
- Verhoeven, G.J., 2018. Focusing on Out-of-Focus: Assessing Defocus Estimation Algorithms for the Benefit. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2, 1149–1156.
- Verhoeven, G.J., Missinne, S.J., 2017. Unfolding Leonardo da Vinci's Globe (AD 1504) to Reveal its Historical World Map. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-2/W2, 303–310.
- Webb, E.K., 2020. Optimising Spectral and 3D Imaging for Cultural Heritage Documentation Using Consumer Imaging Systems. PhD Thesis. University of Brighton.