Focus-Stacking Technique in Macro-Photogrammetry

STEFANO MARZIALI, Academy of Fine Arts of Verona, Italy
ELEONORA MARZIALI, Uni Nettuno, Italy

Small objects are an essential part of archaeological collections. Structure from Motion scanning of Cultural Heritage architectures and collections is getting more and more common , the tiny objects are often omitted for the technical difficulties the digitalization process of such objects impose. This research focuses on the most common issue in macro-photogrammetry: the depth of field, which is the distance in a scene between the nearest and farthest objects that appear acceptably sharp in an image. At a high magnification levels, the small depth of field can compromise the picture alignment or create noisy point cloud. Moreover, it can significantly reduce the quality of the texture, creating blurred patches during the picture projection phase. The proposed solution is the Focus-Stacking technique, a digital image processing method which combines multiple macro-images taken at different focus distances to give a resulting image with a greater depth of field than any of the individual source images. The feasibility of this method for shape measurement in the SfM process has already been tested and proved to be very reliable, but the effect on texture reproduction is not been verified. This research focused on this aspect with an experimentation based on several archaeological test cases. To evaluate the texture and the source images quality, an algorithm was created, using the image gradient method to calculate the image sharpness.

Key words:
Photogrammetry; macrography; focus stacking; depth of field, texture.

CHNT Reference:
Stefano Marziali et al. 2019. The Focus-Stacking Technique in Macro Photogrammetry.

RESEARCH AIM

This paper presents a methodology for the 3D reconstruction of small objects with high quality textures with the Structure from Motion technique. The proposed approach uses an image fusion algorithm (focus-stacking) to overcome the major problems related to macro-photogrammetry. The feasibility of this method for shape measurement is already been tested and proved to be very reliable (with an accuracy that can reach the 0,05% of the bounding box) [Gallo 2014], but none has verified the effect on texture reproduction. This research focused on this aspect with an experimentation based on several test cases.

PROBLEM DEFINITION

To photograph very small objects for a SfM photogrammetry survey, an acquisition setup based on a digital camera equipped with a macro lens with 1:1 magnification rate is necessary. The camera should be placed very close to the specimen to maximize its magnification power, but this configuration opens a very common problem in macrophotography: depth of field.

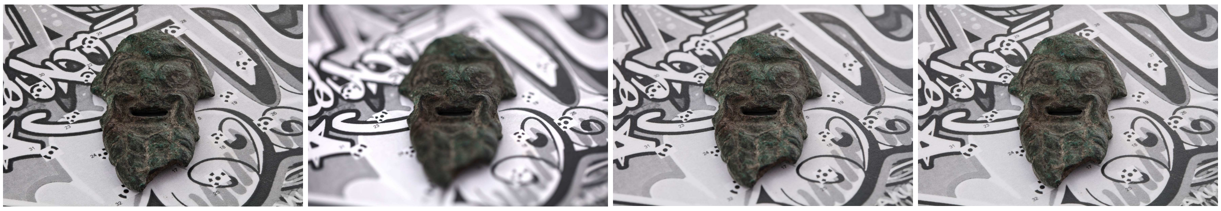
Depth of field (DoF) is the distance between the nearest and farthest objects in a scene that appear acceptably sharp in an image. It varies depending on focal length, diaphragm aperture and focusing distance according with the formula [Ray 2002]

$$DoF \approx \frac{2NcF^2d^2}{F^4 - N^2c^2d^2}$$

where N is diaphragm, F is focal length, c is circle of confusion and d is distance.

Authors addresses: Stefano Marziali, Restoration Department, Academy of Fine Arts of Verona, via Montanari 5, 37122, Verona, Italy; email: stefano.marziali@accademiabelleartiverona.it; Eleonora Marziali, UniNettuno, Italy; email: eleonora.marziali@gmail.com

To frame tiny objects in 1:1 scale it is necessary to employ lenses with very high focal length value (F). This causes the main technical problem of macrophotography: the higher the focal length, the smaller the depth of field, consequently, images will have a very shallow part on focus. As general rule, during the reconstruction process, blurred images should be avoided: they can cause difficulties in camera matching or dense point cloud generation. What more important, moreover, they will generate artifacts in the final texture: parts of the object will be covered with unfocused parts of the pictures.



Figs. 1-4. 1) Focus-stacking; 2) f/2.8 aperture; 3) f/11 aperture; 4) f/32 aperture.

EXISTING SOLUTION

To solve this issue, it is common to use a very narrow diaphragm aperture (N) to have a wider depth of field [Marziali et al. 2017], but extreme diaphragm values are too far from the optical optimum [Williams 1989] to have the sharpest image and it can cause difficulties, once again, in camera matching or texture generation.

PROPOSED SOLUTION: FOCUS-STACKING

Focus stacking (also known as focal plane merging and z-stacking or focus blending) is a digital image processing technique which combines multiple images taken at different focus distances (Figs. 1-4) to give a resulting image with a greater depth of field (DoF) than any of the individual source images [Ray 2002].

After setting the nearest and the farthest focus point in the scene, the software calculates the number of shots necessary to cover the distance at a specific diaphragm aperture. The camera shoots a sequence of pictures at different focus points. Once acquired the images, the focus-stacking algorithm analyzes every image finding the sharpest part. It selects this part, masking the less sharp part of the image so it will be not used in the process, and adds it to the sharpest part of the sequent image. In this way, a new image that is the sum of the sharpest part of the two images is obtained. The process is recursively repeated for every image in the sequence.

EXPERIMENTATION

Data collection methodology

The proposed methodology employs the setup with the fixed camera position, as described in [Marziali et al. 2017]. This setup simplifies the acquisition process. It permits, indeed, to keep frame and focus point fixed, to reduce considerably the acquisition time and to avoid camera movements and the consequent refocus. Despite this benefit, this setup can generate issues in pictures alignment for the change of the shadows position on the artifact. The problem can be partially solved with a very uniform illumination, casting no net shadows or specular reflections. This was ensured by placing the object inside a lightbox, a translucent box which diffuses the transmitted light. The artifact was lighted up with three halogen lights.

The camera used in this project was a DSLR Canon 5DS with full frame sensor of 50,6 MP, equipped with a Canon EF 100mm f/2,8 L Macro IS USM. This specific objective has very low distortion and supports high resolution camera. It, moreover, lets to frame at 1:1 magnification rate without having to approach excessively to the object, which, otherwise, could create operational problems like unwanted shadows or reflections.

For the acquisition process, it is recommended to use a tether camera setup to avoid camera shaking and control the focus process. For this research Helicon Remote was used. It is commercial camera control software designed for focus stacking applications. The acquisition process must be divided in steps and repeated at different angles. In every camera position, after setting the nearest and the farthest focus point in the scene, the software calculates the necessary number of shots to cover the distance at a specific diaphragm aperture. It is recommendable to control the overlapping level with a test burst, because in this experiment often the software underestimated the number of pictures necessary to guarantee a perfect overlapping (Figs. 7-8). After the setup, the camera is ready to take a sequence of pictures at different focus points.

Every sequence of pictures was taken at a specific diaphragm aperture, as shown in Table 2, and ISO 100, the smallest value for this camera model, to reduce the digital noise. The time value was set to obtain the right exposure.

Data set

To test the effect of focus stacking technique on texture generation, a group of five archaeological artifacts were chosen. Four of them are from the collection of Museo Archeologico di Bergamo, the other one is from a private collection. The complete list of the artifact can be found in Table 1.

Table 1. Data set

	Description	**Measure**
Artifact 1	It is a vitreous paste *unguentarium* (Roman period VI-III BC?), a type of a little jar to conserve oils, balms and perfumes. This *unguentarium*, is in black vitreous paste with white glass filament. It is a form of amphora and its dimension is cm. 4 about. It presents a circle nozzle, a short not distinct neck, a main body strongly elongated, two handles set vertically from hem to shoulder. The decoration is on the maximum expanding band to chevrons.	4 cm
Artifact 2944	This artefact is a little bronze bird, probably is a cap of *unguentarium* or same.	4 cm x 3 cm
Artifact 3246	This artifact is a bronze mask from unknown location (Roman period?). Probably, the mask represents Silenus.	5 cm x 4 cm
Artifact 3250	This artefact is a pendant (Roman period?)	2 cm x 2 cm
Artifact 493	This artefact is a stamp (Roman period?), its presents a handle on the long side, on the opposite side there are a symbol and on the top a written “CETI(?)” there are also traces of color.	5 cm x 3 cm

Every artifact has been acquired with three different diaphragm apertures, as shown in the next table:

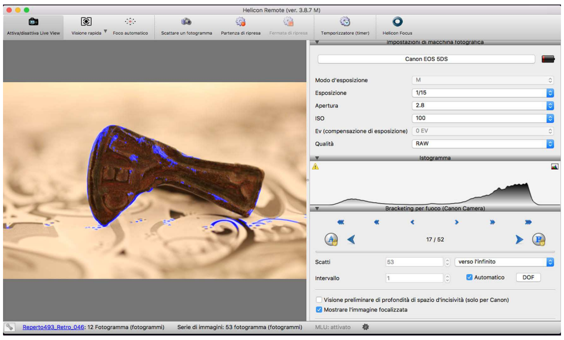
Table 2. Apertures

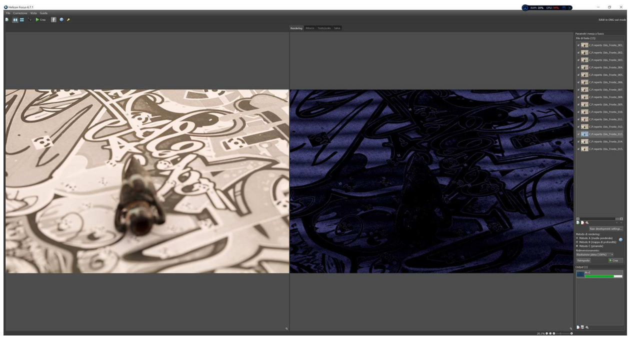
Aperture	**Description**	**Pictures for artifact**
f/11 (a)	Medium diaphragm (optical optimum) focusing the object center	120
f/11 (b)	Medium diaphragm (optical optimum) focusing at different steps for focus stacking merging	2200
f/32	Close diaphragm to maximize the depth of field	120
f/2,8	Open diaphragm with a narrow depth of field. This set of pictures has been used as control group	120

Every artifact was photographed by considering only one half at a time (Marziali, 2017). For each half were performed sixty (20+20+20) steps by turning the base of 18° between shots and varying the camera angle (20°, 30° and 45°), to ensure the observation of the object in all its parts.

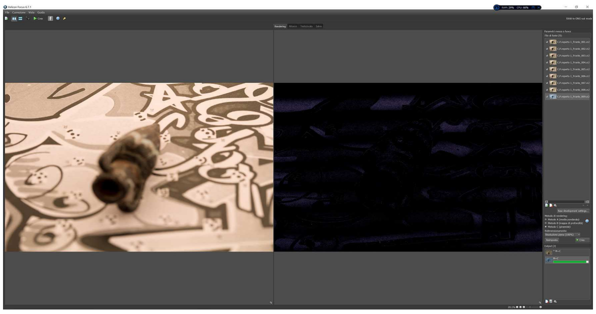
Each object was framed in the center of the view, to exploit the maximum lens sharpness and reduce the optical aberration.

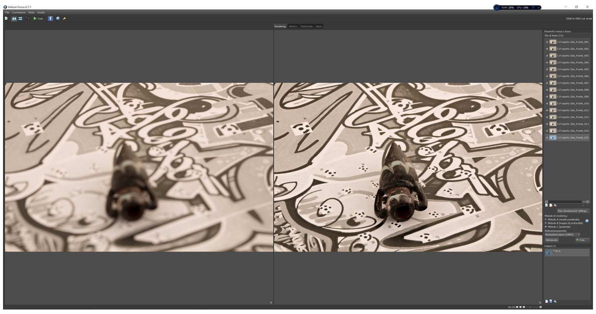
The photos were recorded in the native Canon CR2 digital negative. The transfer data was automatized tethering the camera to a laptop. The transfer was managed by the software Helicon Remote.





Figs. 5-6. 5) Helicon Remote during acquisition. The blue area defines the focus plane and the depth of field; 6) Overlapping good.





Figs. 7-8. 7) Overlapping bad; 8) Focus-stacking process.

Focus-stacking.

For focus stacking, in total 2.200 pictures have been acquired for every artifact. 15 pictures for every step at 45°, 20 pictures at 20° and 30°. All the pictures are automatically divided in folders, one for every step, by Helicon Remote.

Once acquired, the images are imported in software for photo-staking merging. For this purpose, Helicon Focus was chosen as it proved to be very reliable [Brecko et al. 2014]. Helicon Focus permits to choose between three different algorithms: *average* (method A), *depth map* (method B), *pyramid* (method C). Which method will work best depends on the image, the number of images in the stack, and whether the images were shot in random or consecutive order. Thought HeliconSoft, the producer software house, presents some recommendations regarding which method to choose, there is no strict rule as to which method will work better in any particular case. For this project, the method C was chosen for is better local contrast, that could be useful in the feature detection phase, but for completeness here's a brief explanation of each method[1]:

- Method A computes the weight for each pixel based on its contrast, after which all the pixels from all the source images are averaged according to their weights.
- Method B finds the source image where the sharpest pixel is located and creates a "depth map" from this information. This method requires that the images be shot in consecutive order from front to back or vice versa.

[1] http://www.heliconsoft.com/helicon-focus-main-parameters/

- Method C uses a pyramid approach to image representation. It gives good results in complex cases (intersecting objects, edges, deep stacks) but increases contrast and glare.

As described in the previous paragraph, all the photos were taken in RAW format, and imported directly in Helicon Focus for merging in their CR2 native format. It permitted to avoid a whole step in the data acquisition sequence, saving the time of the post-production and prevent the quality loss caused by a double JPEG conversion.

After the import phase, all the images were merged, as described with method C, in batch by the focus stacking algorithm in Helicon Focus and saved in Adobe DNG, the open source format developed by Adobe for storing[2] [Bennett and Wheeler 2010].

Every focus-stacked image was then pre-produced to be used in a SfM process for a 3D reconstruction.

Pre-processing.

The RAW files (CR2 files for f2,8, f11 and f32 acquisition and DNG for focus-stacked images) have been imported in Adobe Lightroom 6 for pre-processing before the SfM reconstruction. The CR2 files were converted in Adobe DNG storing.

The aim to the pre-processing phase of digital negative files was to facilitate the searching of homologous points by the photogrammetric software [Ballabeni et al. 2015] and to have radiometrically-calibrated images ensuring the consistency of surface colors along all the images.

With respect of color characterization, the color targets based technique [Hong et al. 2001] was adopted, using a set of differently colored samples measured with a spectrophotometer. The target XRite ColorChecker [McCamy et al. 1976] was employed during the image acquisitions, considering the measurements of each patch [Pascale 2006] and used as reference for exposure equalization and color balance. The use of the color target is necessary to have an accurate color acquisition, because, as in Hong et al. [2001], the RGB signals generated by a digital camera are device-dependent3. Furthermore, they are not colorimetric (i.e. the output RGB signals do not directly correspond to the device-independent tristimulus values based on the CIE standard colorimetric observer). To obtain a device-independent calibration, the XRite ColorChecker Camera Calibration software was used. It matches the photographed physical reference chart with a reference chart color space with ideal data values for the chart and creates a means converting the device color space to the reference chart color space The process produces a HSL

variation sheet saved in a color profile table (.dcp). It should be used at the start of the pre-processing phase. Unlike in the methodology described in (Marziali, 2017), clarity and sharpening value were set at value 0 to not falsify the evaluation of the image sharpening.

DNG files were adjusted, as shown in Table 3 and saved in JPEG (quality 12, sRGB color profile).

Table 3. Pre-processing in Adobe Lightroom 6.

Options	Details
Color balance – color temperature	Personal - set with the 18 % gray reference surface on the ColorChecker
Color balance - camera calibration	Process: 2012 Profile: generated with ColorChecker Camera Calibration software
Exposure equalization	Set with the 18 % gray reference surface on the ColorChecker
Clarity	+ 0
Sharpening	Amount: 0
Camera calibration	Process: 2012 Profile: generated with ColorChecker Camera Calibration software

[2] https://helpx.adobe.com/it/photoshop/digital-negative.html

[3] Different digital cameras produce different RGB responses for the same scene.

Data analysis.

For every artifact, the four sets of images were imported in Agisoft Photoscan in two chunks (one for the front of the object and one for the back) and aligned. To generate the dense point cloud and the mesh just the focus-stacked images were used. The resulting mesh was edited to simplify the UV. Four different textures were generated reprojecting the four-aperture image sets independently. This step was mandatory to create four different textures on four different models of the same object perfectly comparable, since they have the same overlapping shape.

Both source photographs and generated textures sharpening were tested. To evaluate the image quality an algorithm was created[4] in MATLAB (Fig.9). The algorithm calculates the image sharpness using the image gradient method [Zhu 2010]. This is the simplest and basic measure of sharpness and its main limit is the sensibility of the algorithm to the digital noise [Zhu 2010]. This flaw was considered irrelevant for this analysis, because the whole data set was uniform (all the pictures have been taken at the same ISO value) and almost noise-free (100 ISO), and the algorithm finds itself decent practical usage.

```
for z=1:size(files,1) % loop through the pictures in your input directory
    I=imread(files(z).name); % load the picture
    G=double(rgb2gray(I));

    % measure the sharpness of original image
    [Gx, Gy]=gradient(G);
    S=sqrt(Gx.*Gx+Gy.*Gy);
    sharpness=sum(sum(S))./(numel(Gx));
end
```

Fig. 9. The algorithm calculating the image sharpness using the image gradient method.

The Algorithm produces a numeric value in a range of 1-10. The complete analysis process is resumed in the next graphic (Fig. 10).

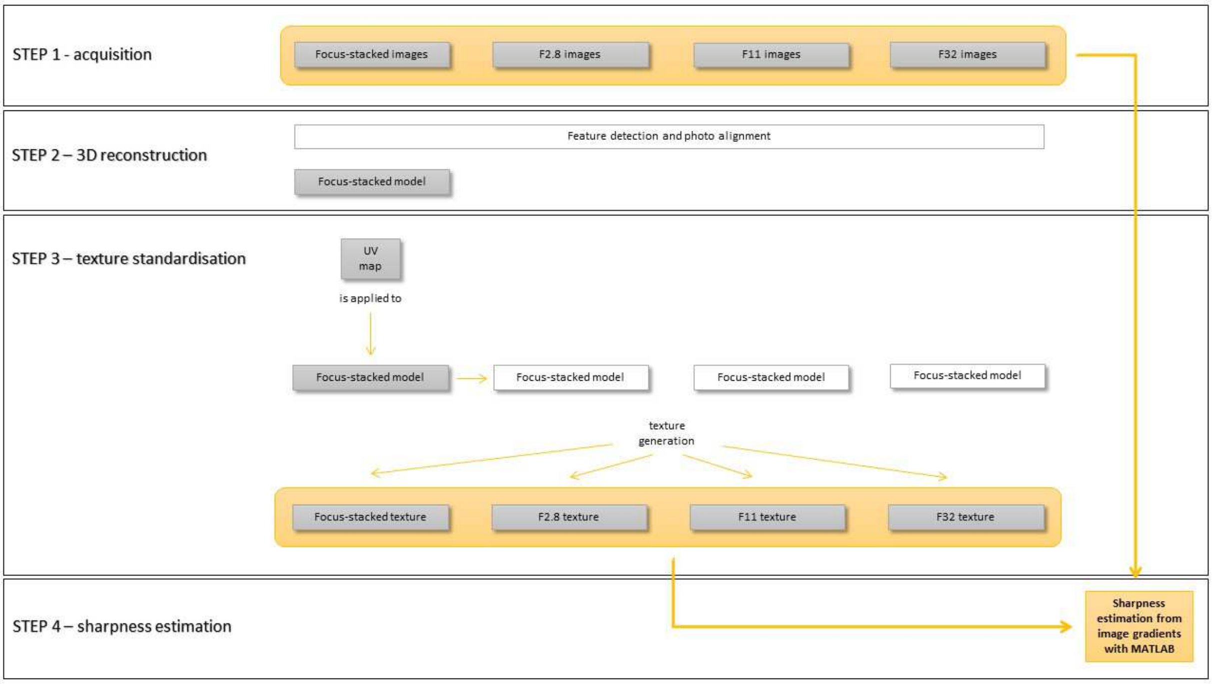


Fig. 10. Analysis process.

4 Based on https://it.mathworks.com/matlabcentral/fileexchange/32397-sharpness-estimation-from-image-gradients

RESULTS

Reconstruction

For SfM reconstruction, Agisoft Photoscan was set-up with the settings shown in Table 4. The focus-stacked images shown very good results in the process.

Table 4. Agisoft Photoscan Align Photos settings.

Options	Details
Accuracy	High
Preselection	Generic
Key point limit	40,000
Tie point limit	4,000

After the feature detection and alignment phase and prior of any alignment correction operation, some of the non-focus-stacked models presented some reconstruction errors, as shown in Table 5.

Table 5. Reconstruction results.

	1	493	2944	3246	3250
Focus-stacking	ok	ok	ok	ok	ok
f/2.8	ok with errors	ok with errors	ok	no	no
f/11	ok with errors	ok with errors	ok	ok with errors	ok with errors
f/32	ok with errors	ok	ok	ok with errors	ok with errors

Image sharpening.

The improvement in texture generation was quantified measuring the image sharpness in a value between 1-10. The texture generated with focus stacked images shows a sharpness value three times higher than the other images. The same difference was observed in the chunks of images used for the texture generation. The average results are listed in Table 6:

Table 6. Results.

Diaphragm	Sharpness (1-10 scale)
Focus-stacking	6,68
f/2.8	2,01
f/11	2,34
f/32	1,90

CONCLUSIONS

In the SfM pictures acquisition phase, if the object detected is large enough, the photographs can be taken with a good depth of field and sharpness, so most of the problems concerning the texture quality are solved by the software. If the object is very small, however, the use of macro lenses is necessary, and the photographs will be characterized by a shallow depth of field and sequent low-sharpness value. This entails a whole series of problems during the texture generation process.

This research tried to give a solution to this problem which gave excellent results in terms of quality of the generated texture. The proposed solution is the photographic technique called Focus-stacking. Through this technique is possible to create a single frame completely in-focus staking more than one partially in-focus image. Processing a chunk of images created with this technique, the software will be able to create a perfectly in-focus texture and a 3D model more like the original, as the focus-stacked image sharpness is been quantified to be three times higher the original set of photographs (Table 6).

The increased sharpness also entails a series of technical and aesthetic advantages: the finished model is more like the actual object and the reconstruction process is speeded up and more accurate (Tab. 5).

In addition, the Focus-stacking process does not slow down the reconstruction time in any way, as the number of frames used is the same as those of the fixed-focus models. Moreover, the focus-stacking software are low-cost or even freeware.

A disadvantage is the acquisition time and the focus-stacking processing, which is obviously greater than the time necessary for common the single-focus frames acquisition that requires no extra processing.

	Advantages	**Disadvantages**
Focus-stacking	Reduced modeling time (frame alignment, dense cloud creation, mesh generation and texture) Significant reduction of errors due to misalignment during frame alignment, creation of the dense cloud and the generation of meshes and textures Texture precision, all in focus and detailed Lower reconstruction errors Higher sharpness	Dilated acquisition times Expanded post-processing times You need to use more than one software as a result you need multiple steps in the workflow
Focus fixed	Reduced acquisition times Reduced post-processing time Fewer steps in the workflow	Expanding modeling times due to errors in frame alignment, creation of the dense cloud, and generation of meshes and textures Lower sharpness Details not easily visible Not precise and legible texture Frames not completely focused

REFERENCES

Alessandro Gallo et al. 2014. 3D reconstruction of small sized objects from a sequence of multi-focused images. Journal of Cultural Heritage, 15, 173–182

Sidney Ray. 2002. Applied photographyc optics. Focal Press

Stefano Marziali et al. 2017. Photogrammetry and macro photography. The experience of the MUSINT II Project in the 3D digitizing process of small size archaeological artifacts. Studies in Digital Heritage. 1. 298. 10.14434/sdh.v1i2.23250.

John B. Williams 1989. Image Clarity: High-Resolution Phoptography. Focal Press

M. J. Bennett and F. B. Wheeler. 2010. Raw as archival still image format: A consideration. In *Archiving 2010 - Preservation Strategies and Imaging Technologies for Cultural Heritage Institutions and Memory Organizations*, Final Program and Proceedings: 185-193.

J. Brecko, A. Mathys, W. Dekoninck, M. Leponce, D. Vandenspiegel and P. Semal. 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.

A. Ballabeni, F. I. Apollonio, M. Gaiani and F. Remondino. 2015. Advances in image pre-processing to improve automated 3d reconstruction. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XL-5/W4, 2015.

G. Hong, M. R. Luo and A. Rhodes. 2001. A study of digital camera colorimetric characterisation based on polynomial modelling. In *Color Research and Application*, 26(1), pp. 76-84.

C.S. McCamy, H. Marcus and J.G. Davidson. 1976. A color rendition chart. In *Journal of Applied Photographic Engineering*, Vol. 11(3): 95-99.

D. Pascale. 2006. RGB coordinates of the Macbeth ColorChecker. The BabelColor Company, Montreal, Canada.

Xiang Zhu. 2010. Automatic Parameter Selection for Denoising Algorithms Using a No-Reference Measure of Image Content. in *IEEE transactions on image* processing, vol. 19, no. 12, 2010.

Imprint:
Proceedings of the 22nd International Conference on Cultural Heritage and New Technologies 2017.
CHNT 22, 2017 (Vienna 2019). http://www.chnt.at/proceedings-chnt-22/
ISBN 978-3-200-06160-6

Editor/Publisher: Museen der Stadt Wien – Stadtarchäologie
Editorial Team: Wolfgang Börner, Susanne Uhlirz
The editor's office is not responsible for the linguistic correctness of the manuscripts.
Authors are responsible for the contents and copyrights of the illustrations/photographs.