Uses of 3D digitization techniques based on RGB images for the registration and sculptural creation of medals

Iria Groba-Martín

Within the world of digital design and manufacturing, 3D digitization systems based on RGB (red, green and blue) images[1] only pose a small fraction of the available techniques. Besides being contactless, these digitization systems are described as passive, as unlike laser or structured light scanners they don't project any kind of optical signal over the registered subjects (Chaudhuri, 1999).

One common way to classify these techniques is with respect to the number of input images. Multi-image systems would include photogrammetry, photometric stereo, or shape from focus, while examples of stereo systems could be found in binocular stereo or smartphone depth map generation. Within single-image systems we would find colour photometric stereo, shape from shading or monocular depth inference techniques, which use the latest AI (artificial intelligence) technologies to create three-dimensional shape from an individual photograph.

3D and 2.5D digitization systems

Although the term '3D' usually encompasses both 3D and 2.5D shapes (as in this project's title), there are significant distinctions between these two concepts (fig. 1). Whilst 3D refers to any kind of sculpture in the round, the term 2.5D is used to describe three-dimensional shape that can be fully seen from one single point of view and, thus, has no undercuts (Leithinger *et al.*, 2011). This implies that such a shape can be represented through a 'depth map', a digital monochrome image that contains information about the distance between the surface of objects or scenes and the camera or viewer.

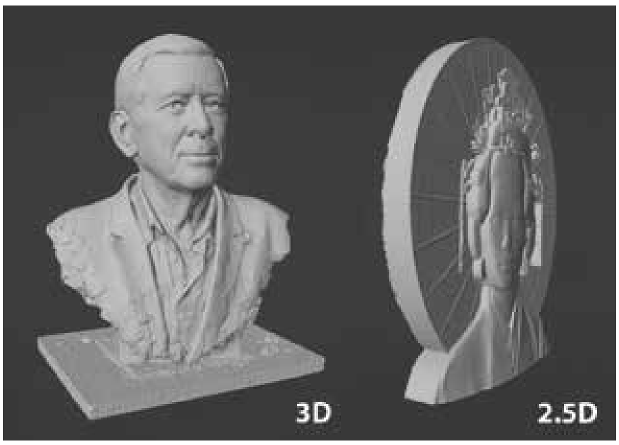

Fig 1. Examples of 3D and 2.5D digitized shapes of two pieces: *Miguel* (bust sculpture, 2023, acrylic resin, 540 mm, and *Umbra* (medal, 2023, PLA, 145 mm)

From the previously mentioned systems, only photogrammetry and monocular depth inference are capable of directly creating full 3D models. The other techniques generally capture shape in 2.5D or through depth maps, from which 2.5D shape can be retrieved. Although this might be problematic for registering sculpture in the round, it isn't so troublesome when it comes to capturing bas-reliefs such as medals, as they can usually be fully seen from one point of view and tend to have little or no undercuts.

Taking this into account, several inquiries arise. Which system would be more appropriate for medal digitization? Is the automatic generation of bas-reliefs through 2.5D digitization plausible? To answer these questions, it was decided to test the 3 RGB image-based 3D (and 2.5D) digitization systems that seemed most suitable: photogrammetry, photometric stereo, and shape from focus.

Photogrammetry

Photogrammetry is the most utilized image-based 3D digitization technique (Díaz Alemán, 2021). Using several pictures taken around an object, it can retrieve its complete 3D shape and colour, representing them through 3D models and texture maps (Schenk, 2005). Thanks to the available user-friendly software, it is easy to use for non-experts. It can be employed for registering small (about a few cm) to very large subjects (Luhmann *et al.*, 2013). When properly used, it can achieve great results and create models with a very good level of detail.

The sculpture community usually employs close-range digital photogrammetry, as digitized elements are typically

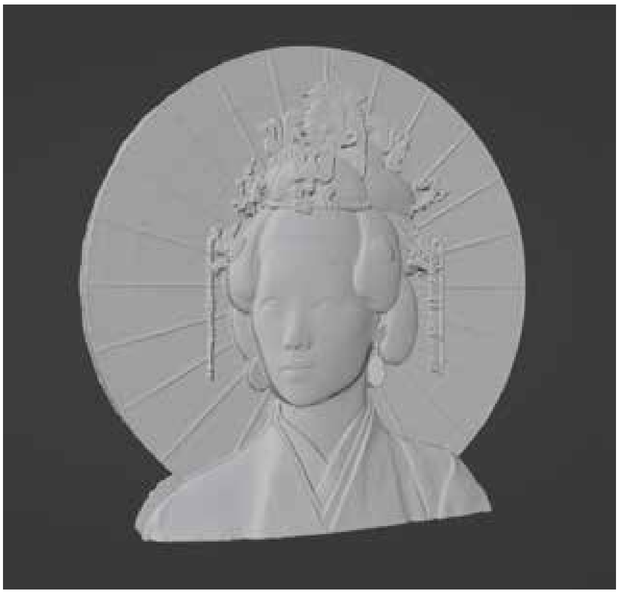

Fig 2: *Umbra*, photogrammetric model

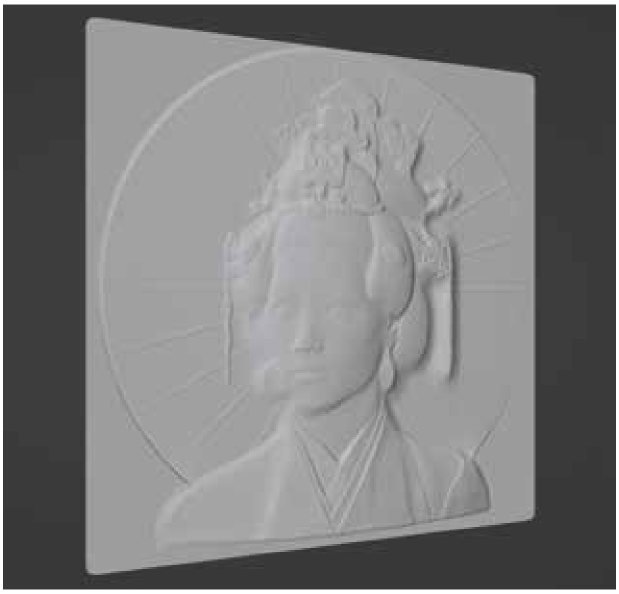

Fig 3: *Umbra*, photometric stereo model

under 200 m big (Luhmann *et al*., 2013). When digitizing objects, single-camera systems are generally used. Although multi-camera setups which take all the pictures simultaneously would be more appropriate for capturing people or objects in motion, the high cost associated with this approach limits its use to large companies, such as those from creative industries (Struck *et al*., 2019).

Photometric stereo

Photometric stereo works with pictures taken keeping the camera and the object static and changing only the direction of the incident light (Woodham, 1979). It allows to capture the shape in 2.5D and is optimal for registering very small details from objects that have little depth. However, it is not an appropriate system for digitizing large elements, as the mathematical model behind this technology assumes directional light and orthogonality in the images, both of which are rarely met in practice. Although adjustments such as distancing light sources and using cameras with long lenses can approximate these ideal conditions and reduce errors, achieving them is considerably easier when working with small subjects. As a result, overall digitized shape in larger subjects is usually slightly deformed (Karami *et al*., 2021).

The direct results of photometric stereo are ‘normal maps’, textures that represent the orientation of the surface of the captured object through an RGB colour representation system. When no undercuts or 90-degree jumps exist, these can be transformed into depth maps, from which 2.5D shape can be retrieved. Simple-to-use software exists both for the process of transforming images into a normal map, and for the step of converting the normal maps into depth maps.

Shape from focus

Shape from focus is a digitization system based on how the human eye analyses depth through its focus apparatus

Fig 4: *Bride*, 2017,
terracotta, 145 mm,
photogrammetric model

Fig 5: *Bride*, photometric stereo model

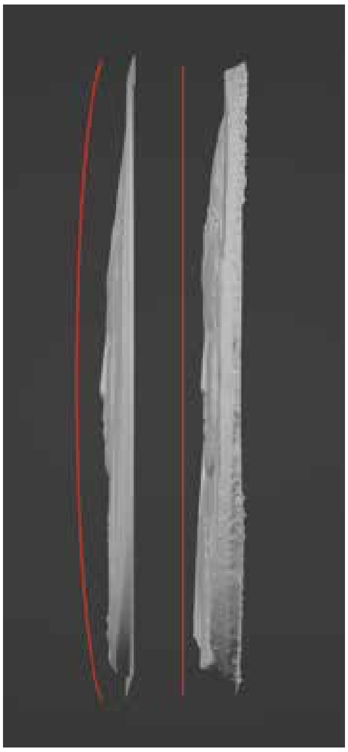

Fig 6: *Bride*, photometric stereo model (left) vs photogrammetric model (right)

(Nayar and Nakagawa, 1994). It uses a sequence of images taken with different focus planes to estimate local depth (Pavlidis and Koutsoudis, 2022). As best results are obtained with images that have a shallow depth of field, it is most appropriate for capturing very small subjects with a reasonable deepness. The varying focus plane can be achieved either by slowly changing the distance between the subject and the camera, or by keeping both static and varying the position of the lens by rotating the focus ring (Nayar, 1989). The result is a low-detail depth map, that can be transformed into a 2.5D model.

Although the results are not as detailed as with the previously proposed systems, shape from focus can be particularly useful when dealing with objects that are too small to be digitized through photogrammetry and have a shape that makes them unsuitable to be registered through photometric stereo (for having, for example, 90-degree jumps, or a big depth that would produce many cast shadows). Some common applications would be microscopic 2.5D digitization (Koch and Bruenger, 2021) or automatic focus for DSLR (Digital Single Lens Reflex) cameras (Nayar and Nakagawa, 1994).

Aspects to consider when digitizing

The three essential aspects must be considered when digitizing any kind of object are surface, size and shape:

1. The surface characteristics of the object will affect the results, as optic digitization systems are usually sensible to shininess, and tend to give worse results when dealing with darker elements. Best outcomes would be achieved when working with light-coloured subjects and fully diffuse surfaces with no specular components (these are referred to as Lambertian surfaces).
2. The dimensions of the digitized object will also influence the outcome, as not all techniques are suitable for all object sizes due to camera constrains, light falloff, etc.
3. Finally, the shape of the object should also be contemplated, as some techniques are sensible to 90 degree jumps in 2.5D surfaces, to undercuts or to complicated configurations.

From bas-relief to bas-relief: Methodology

Photogrammetry and photometric stereo were carried out with a shiny, 24 mm wide coin and with two 145 mm wide medals with Lambertian surfaces, one with 90-degree jumps and larger depth (*Umbra*), and another shallower one with no 90-degree steps (*Bride*). Shape from focus was only tested on these last two medals.

The methodology used was the following:

1. With photogrammetry, a high F number and automatic focus were employed. A very large number of images were taken with a hand-held camera by moving around the objects. The software chosen for the photogrammetric reconstruction, Reality Capture, gave as outputs a 3D model and an ‘albedo texture map’.
2. With photometric stereo, 16 images with varying incident light direction were taken for every object. To keep its position and viewing direction consistent, the camera was set up on a tripod facing the place in the floor where the digitized medal or coin was placed. A 55 mm lens was employed to get as close to orthogonality as possible, and a high F number was used to guarantee depth of field (although this wouldn’t have been necessary due to the shallowness of the digitized objects). A small chrome ball was also placed on the scene for automatic light calibration.

   The software employed for the photometric stereo reconstruction, Details Capture, used the pictures to retrieve an ‘albedo map’, a ‘global normal map’, a ‘flat normal map’ (a type of normal map flattened by the software) and a ‘height map’ (or depth map) created from this last one (inaccurate due to the flattening of the flat normal map). Next, the global normal map was transformed into a depth map using the software FriendlyShade Normalizer. The exposure curve of the resulting image was adjusted to increase the number of tones, and, finally, it was transformed into a 2.5D mesh in the software Blender by using it as a ‘displacement map’.

3. As with photometric stereo, shape from focus was performed with the camera placed on a tripod facing the floor. A low F number was employed to lower the depth of field, and a stack of images with different focus planes was taken by rotating the focus ring. The images were processed in the software Helicon Focus Pro. The outputs with this software were a fully focused image, a smoothed 2.5D model and a depth map that was transformed into an unsmoothed 2.5D model using the procedure described for photometric stereo.

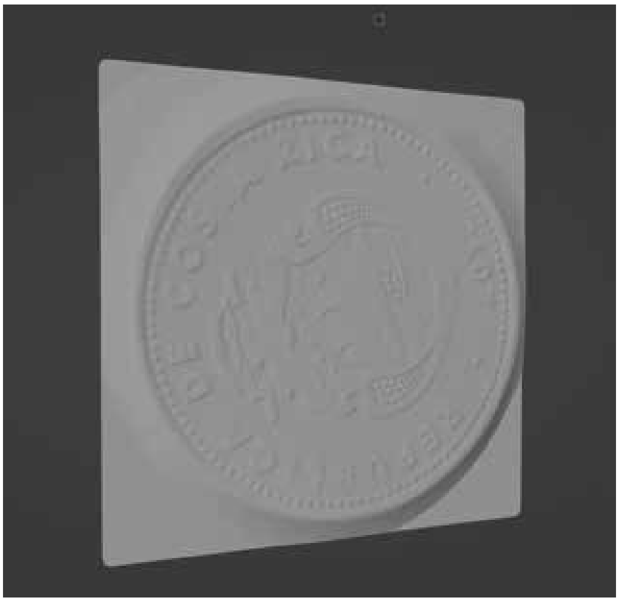

Fig 7: Shiny coin, photometric stereo model

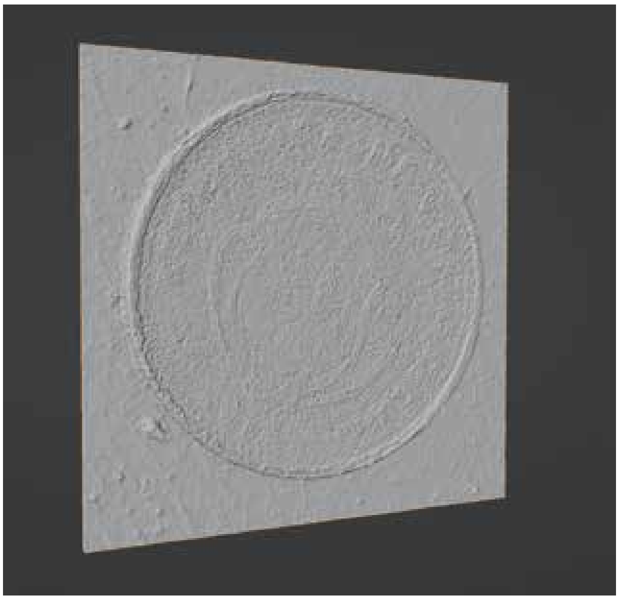

Fig 8: Shiny coin, photogrammetric model

In all the cases, the images were taken in RAW format, and their exposure curves were adjusted in the software Darktable before exporting them as JPEGs and processing them with the corresponding software.

From bas-relief to bas-relief: Results

When capturing the 145 mm wide medals, photogrammetry proved to be more appropriate than photometric stereo. This was especially evident with the *Umbra* medal: when comparing fig. 2 (photogrammetry) with fig. 3 (photometric stereo) we can see inaccuracies in the areas with 90-degree jumps that were caused by cast shadows projected during the image acquisition phase. The surface results in the photogrammetric registration of *Bride* (fig. 4) were quite similar to those obtained through photometric stereo (fig. 5). However, when closely examining their general shape, it is evident that light falloff caused a slight global deformation in the model generated through photometric stereo (fig. 6).

Nevertheless, the registration of the shiny coin was more successful with photometric stereo (fig. 7) than with photogrammetry (fig. 8). Photometric stereo stood out as the only system capable of accurately capturing the surface without introducing any noise. This suggests that it may be less sensitive to specular surfaces compared to photogrammetry. Furthermore, no evident global deformations appeared (fig. 9). This is because light falloff decreases exponentially when increasing the distance between the light source and the illuminated subject. With respect to their own sizes, the light source was much farther away in the case of the coin than in that of the medals, so light falloff didn't affect the results as much and no bulging appeared.

Fig 9: Shiny coin, cross-sectional view of photometric stereo model

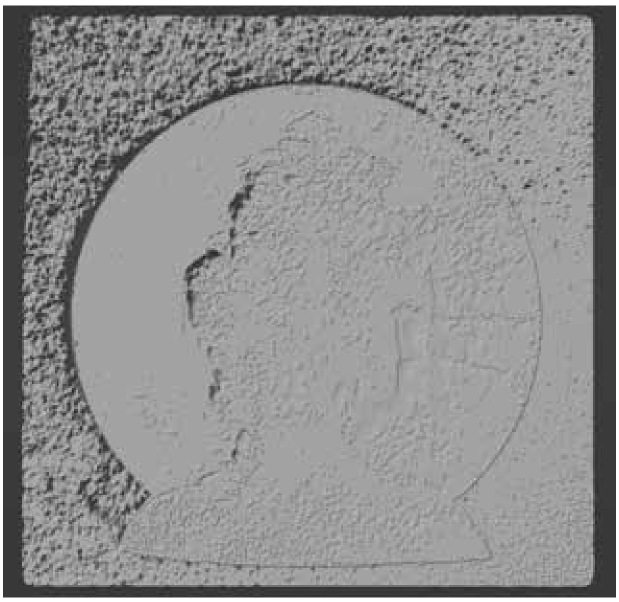

Fig 10: *Umbra*, shape from focus model

Fig 11: *Miguel*, automatic generation of bas-relief through photometric stereo

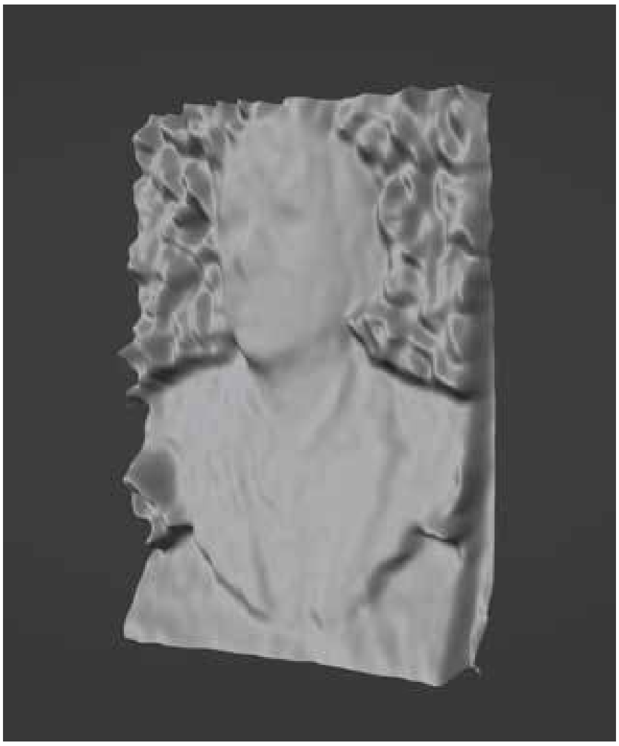

Fig 12: *Miguel*, automatic generation of bas-relief through shape from focus

Therefore, it can be said that photogrammetry gives better results than photometric stereo in the registration of larger subjects, as models generated with this last technique suffer from some global deformation due to light falloff. The digitizing of the coin showed two things: (1) that when dealing with specularity, photometric stereo seems to be far more robust than photogrammetry, and (2) that photometric stereo is most appropriate for digitizing smaller objects, which seem to be less affected by bulging do to the exponentiality of light falloff. However, it cannot be concluded that photometric stereo is better than photogrammetry at digitizing small subjects, as no Lambertian small objects were tested with both techniques. Another important characteristic of photometric stereo is that the system seems to work better with shallower surfaces that have no 90-degree jumps, as the presence of cast shadows during image acquisition gives inaccuracies in retrieved normal and depth maps. Depth from focus proved to be unsuitable for medal digitization (fig. 10).

From 3D shape to bas-relief: Methodology

When considering the creation of bas-reliefs from 3D shape, photogrammetry combined with 3D deformation and modelling can produce high quality results. However, the process of transforming a 3D model into a 2.5D one is not automatic and requires considerable amounts of time. Therefore, we will only consider photometric stereo and shape from focus for such a purpose.

Photometric stereo was employed to digitize a bust, *Miguel*. Shape from focus was used to digitize this same sculpture, several small objects, and various landscapes. The results for the digitization of the bust, a small shell and a close-range landscape will be shown.

The methodology employed with photometric stereo was similar to the one described in the previous section, with the slight distinction that the camera was placed on a tripod looking out-front, as the sculpture could not be laid on the floor looking upwards. The same thing applied to shape from focus. In this last case two different lens, a 55 mm and 18 mm one, were tested on the same subjects to see which of them would give better results, as shorter lens tend to have a shallower depth of field.

From 3D shape to bas-relief: Results

Photometric stereo showed good results in the automatic creation of bas-reliefs through the 2.5D capture of *Miguel* (fig. 11), especially when placing the illumination sources so that no or little cast shadows appeared during the image capture. Plus, it had an advantage: the software employed (Details Capture) retrieves two kinds of normal maps (the ‘flat and the ‘global’ one), from which two different types of depth maps can be obtained. By combining them, it was possible to easily alter the depth of the detail in the 2.5D model.

Although not very good, shape from focus results with the large figurative sculpture were better than the ones obtained with the medal digitization, especially when an 18 mm lens (fig. 12) was used instead of a 55 mm one. This is due to the smaller depth of field obtained with such an apparatus.

As expected, the results with the small object (fig. 13) were far better thanks to their high depth and to the fact that, when photographing small subjects, cameras usually capture them with a very shallow depth of field. The digitization of scenes was also proven possible only in the case of close-range setups (fig. 14), as the difficulty of

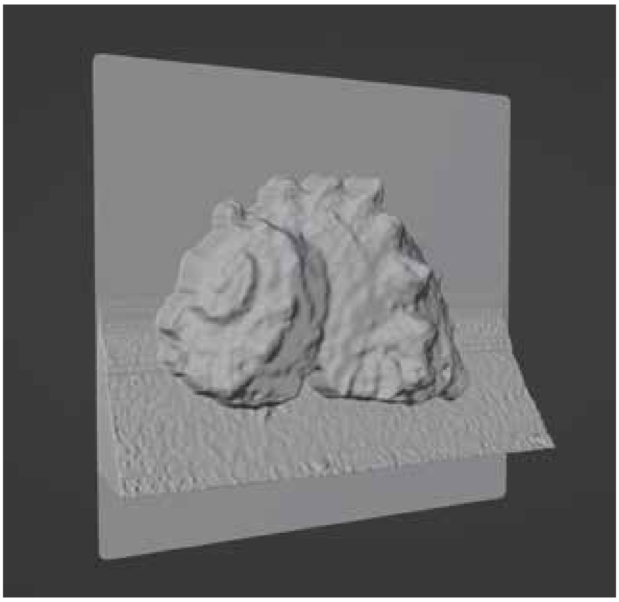

Fig 13: Shellfish, automatic generation of bas-relief through shape from focus

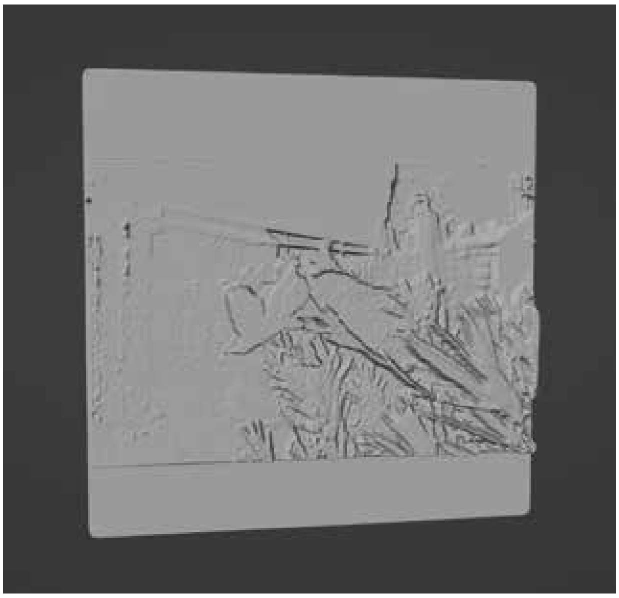

Fig 14: Landscape, automatic generation of bas-relief through shape from focus

getting blurry planes in longer-range ones prevented any good results from being obtained.

Overall, both systems demonstrated their suitability for the automatic generation of bas-reliefs from 3D shapes. When dealing with larger subjects, photometric stereo proved to be more appropriate. Shape from focus gave better results when working with smaller subjects with a reasonable depth and was also capable of creating automatic bas-reliefs from close-range landscapes. However, further experiments with photometric stereo would be necessary to see which system is most appropriate for these last two cases.

Applications

The applications of the described 3D digitization systems are numerous. Registering bas-reliefs can serve for conservation and restauration purposes, as well as for digitally studying art pieces. Digital exhibition of medals, their enlargement and reduction, and digital alteration through modelling or sculpting would be other possibilities. On the other hand, digitizing 3D shape in 2.5D would serve for the automatic and semiautomatic generation of bas-reliefs from objects, humans and even some scenes. This could have interesting creative applications in areas such as sculpture or medal-making.

Conclusions

Overall, we have seen that 3D (or 2.5D) digitization can be used both for the registration of medals and for the automatic creation of bas-reliefs from three-dimensional objects and some types of scenes. Although photogrammetry is the most used system, other techniques have shown interesting results and proven to be more useful when dealing with small, shallow shiny objects (photometric stereo), or when wanting to automatically turn objects or scenes into bas-reliefs (photometric stereo and shape from focus).

BIBLIOGRAPHY

Chaudhuri, S.: *Depth from defocus: A real aperture imaging approach*, New York, 1999.

Díaz Alemán, M. D.: Modelado 3D de precisión en procesos de digitalización de escultura construida, *AusArt Journal for Research in Art,* vol. 7, no. 2, 2021, pp. 113-125.

Karami, A., Menna, F. and Remondino, F.: Investigating 3d reconstruction of non-collaborative surfaces through photogrammetry and photometric stereo, *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B2-2021, XXIV ISPRS Congress,* 2021, pp. 519-526.

Koch, R. and Bruenger, J.: Depth estimation, in *Computer Vision: A Reference Guide* (ed. Katsushi Ikeuchi), Cham, 2021, pp. 290-294.

Leithinger, D., Lakatos, D., DeVincenzi, A., Blackshaw, M. and Ishii, H.: Direct and gestural interaction with relief: A 2.5D shape display, *Proceedings of the 24th annual ACM symposium on User interface software and technology*, Santa Barbara, 2011, pp. 541-548.

Luhmann, T., Robson, S., Kyle, S. and Boehm, J.: *Close-Range Photogrammetry and 3D Imaging*, London, 2013.

Nayar, S. K.: *Shape from focus*, Pittsburgh, 1989.

Nayar, S. K., and Nakagawa, Y.: Shape from focus, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 8, 1994, pp. 824-831.

Pavlidis, G. and Koutsoudis, A.: 3D digitization of tangible heritage, in *Handbook of Cultural Heritage Analysis* (ed. Sebastiano D'Amico and Valentina Venuti), Cham, 2022, pp. 1363-1404.

Schenk, T.: *Introduction to Photogrammetry*, Ohio, 2005.

Struck, R., Cordoni, S., Aliotta, S., Pérez-Pachón, L. and Gröning, F.: Application of photogrammetry in biomedical science, in *Biomedical Visualisation. Advances in Experimental Medicine and Biology*, vol. 1120 (ed. Paul M. Rea), Cham, 2019, pp. 121-130.

Woodham, R. J.: Photometric stereo: A reflectance map technique for determining surface orientation from image intensity, *Proceedings SPIE 0155, Image Understanding Systems and Industrial Applications I*, 1979, pp. 136-143.

NOTES

1. The term ‘RGB image’ refers to digital colour images composed of three primary colour channels: red, green, and blue. These channels combine to create the full spectrum of colours visible to the human eye. Most digital images used today, including those captured with smartphones and digital cameras, are RGB images.