The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-2/W8-2024
8th International ISPRS Workshop LowCost 3D - Sensors, Algorithms, Applications, 12–13 December 2024, Brescia, Italy

(R)-evolution in the implementation of photometric stereo – Application to archaeology

Antoine Laurent[1,4*], Benjamin Coupry[1], Jean Mélou[1],
Yvain Quéau[2], Carole Fritz[3], Jean-Denis Durou[1]

[1] IRIT, Université de Toulouse, France [2] GREYC, Université de Caen, France
[3] LAMS, Sorbonne Université, Paris, France [4] TRACES, Université de Toulouse, France

Keywords: Cultural heritage, 3D digitization, photometric stereo.

Abstract

There are two main families of photographic 3D reconstruction methods. On the one hand, so-called geometric methods (photogrammetry), which use images of the scene taken from different viewpoints, are based on the principle of triangulation. Photometric methods, on the other hand, relate the appearance of a 3D point (grey level or colour levels) to the orientation of the normal, relative to the direction of the incident light. While photogrammetry can be used to estimate the overall shape of a 3D scene, provided that the scene is sufficiently textured, photometric methods highlight the details of the relief, if the physical model linking the lighting to the relief of the scene and its reflectance is sufficiently realistic. In this article, we introduce a method of photographic 3D reconstruction that is both multi-view and multi-lighting, which combines the advantages of both approaches, and is particularly easy to implement in the archaeological context, since no object other than the scene is required to estimate the illumination.

1. Introduction

This article looks at 3D reconstruction methods for archaeology. The instruments most commonly used for 3D surveys in the field are various types of scanner and the camera, which can be used to acquire data for different 3D reconstruction methods. The price range for the initial investment in scanners is around several tens of thousands of euros, to which maintenance is generally added, and above all the software, which costs a few thousand euros. We therefore disqualify these instruments from the outset, from the point of view of low-cost 3D reconstruction.

The photographic method most commonly used in archaeology is *photogrammetry*, which consists of taking several shots of a scene from different angles, and deducing the relief from the distortions this causes in the images. When it comes to photogrammetry software, there is a clear dichotomy between licensed and open-source software. Licensed software has attractive, ergonomic interfaces, but the "black box" nature of the algorithms limits the accuracy of results.

The situation is often reversed for software of the second type, which sometimes has rudimentary interfaces. This is because of the development time, resources and skills required to make them competitive. This investment is generally used to improve performance and update algorithms. This highlights the difficulties of funding open-source solutions, and from an ethical point of view, the temptation for some researchers to use licensed solutions, instead of supporting solutions that are open to all.

Photometric stereo (PS) is another photographic method of 3D reconstruction, less well known than photogrammetry, which uses several photographs taken from the same point of view, but under different lighting conditions. Compared with photogrammetry, it estimates not only the relief of a 3D object, but also its reflectance, i.e. its true colour.

The value of PS for the 3D reconstruction of heritage is well established (Mélou et al., 2022). However, using this technique in a museum or on an archaeological site can be tedious. In this article, we present an easy-to-use acquisition and processing protocol based on the fusion of PS and photogrammetry.

The key to this protocol is to simplify the lighting calibration procedure, which is an essential step in PS. Although there are some variants of PS where lighting is added to the unknowns, this variant of PS, known as "uncalibrated PS", is an ill-posed problem.

An even more difficult PS problem, in which the illuminations can be arbitrary ("universal PS"), can be solved with the advent of trained neural networks, which provide very realistic results (Hardy et al., 2024). However, it turns out that ambiguities of the "bas-relief ambiguity" type (Belhumeur and Kriegman, 1998) are not eliminated for all that.

We therefore resolutely opt for the calibrated version of PS, where the lighting must be known. Calibrating the lighting in situ means positioning a calibration pattern in the scene at the time of shooting, which is usually either a reflecting sphere or a matte white sphere.

However, this poses a number of problems, making data acquisition more cumbersome and reducing the quality of the results. One solution is to use an RTI dome, which can be calibrated beforehand. In addition to the cost of such an instrument, new problems arise, in particular the reduction in the field of view.

In this article, we show how to make optimal use of multi-view and multi-lighting data, using the object itself as an illumination calibration pattern, without degrading the results. On the contrary, the PS implementation we recommend combines the advantages of the multi-view and multi-lighting approaches, while reducing the hardware to the bare minimum, i.e. a camera and a flash.

2. Classical formulation of calibrated PS

The physical model we adopt is the Lambertian model, which is well suited to archaeology, given that many objects are both opaque and diffusive. This model expresses the grey level $I(\mathbf{p})$ of an image point $\mathbf{p}$ in the following particularly simple form:

$$I(\mathbf{p}) = \rho(\mathbf{P})\,\mathbf{n}(\mathbf{P})^\top \mathbf{s}(\mathbf{P}) \tag{1}$$

In (1), $\rho(\mathbf{P}) \in \mathbb{R}^+$ denotes the *albedo* (the colour) of the point $\mathbf{P}$ in the scene which is projected at $\mathbf{p}$ into the image, $\mathbf{n}(\mathbf{P}) \in \mathbb{R}^3$ denotes the unit outgoing *normal* to the surface at point $\mathbf{P}$, and the vector $\mathbf{s}(\mathbf{P}) \in \mathbb{R}^3$, called the *illumination vector*, characterizes the direction and intensity of the luminous flux incident at this point.

This calibrated approach of PS (Woodham, 1980) assumes that the illumination vector $\mathbf{s}(\mathbf{P})$ is known, without necessarily assuming that it is uniform, i.e. independent of $\mathbf{P}$. By introducing as a new unknown $\mathbf{v}(\mathbf{P}) = \rho(\mathbf{P})\,\mathbf{n}(\mathbf{P})$, Equation (1) becomes linear. Moreover, the use of p lighting $\mathbf{s}^i(\mathbf{P})$, $i \in \{1, \dots, p\}$, can make the estimation of $\mathbf{v}(\mathbf{P})$ well-posed. Indeed, the p grey levels $I^i(\mathbf{p})$, $i \in \{1, \dots, p\}$, can be grouped together in matrix form:

$$\begin{bmatrix} I^1(\mathbf{p}) \\ \vdots \\ I^p(\mathbf{p}) \end{bmatrix} = \begin{bmatrix} \mathbf{s}^1(\mathbf{P})^\top \mathbf{v}(\mathbf{P}) \\ \vdots \\ \mathbf{s}^p(\mathbf{P})^\top \mathbf{v}(\mathbf{P}) \end{bmatrix} \tag{2}$$

which can be rewritten as:

$$\begin{bmatrix} I^1(\mathbf{p}) \\ \vdots \\ I^p(\mathbf{p}) \end{bmatrix} = \begin{bmatrix} \mathbf{s}^1(\mathbf{P})^\top \\ \vdots \\ \mathbf{s}^p(\mathbf{P})^\top \end{bmatrix} \mathbf{v}(\mathbf{P}) \tag{3}$$

As soon as $p \geq 3$ lighting vectors $\mathbf{s}^i(\mathbf{P})$ are used, and these vectors are not all coplanar, the matrix of the second member of (3), obtained by concatenation of these vectors, is of rank 3. The estimation of $\mathbf{v}(\mathbf{P})$ then becomes a well-posed problem, whose least squares solution is written:

$$\mathbf{v}(\mathbf{P}) = \begin{bmatrix} \mathbf{s}^1(\mathbf{P})^\top \\ \vdots \\ \mathbf{s}^p(\mathbf{P})^\top \end{bmatrix}^\dagger \begin{bmatrix} I^1(\mathbf{p}) \\ \vdots \\ I^p(\mathbf{p}) \end{bmatrix} \tag{4}$$

where the superscript † denotes the pseudo-inverse. Knowing that the normal $\mathbf{n}(\mathbf{P})$ is of norm 1, it is easy to deduce from (4):

$$\mathbf{n}(\mathbf{P}) = \frac{\mathbf{v}(\mathbf{P})}{\|\mathbf{v}(\mathbf{P})\|} \quad ; \quad \rho(\mathbf{P}) = \|\mathbf{v}(\mathbf{P})\| \tag{5}$$

3. Use of a pre-calibrated multi-lighting system

The simplest PS implementation is to use an RTI (reflectance transformation imaging) dome, which is a pre-calibrated multi-lighting system. RTI (Díaz-Guardamino et al., 2015) is an interactive lighting simulation technique that involves taking a large number of photographs of an object under different lighting, each illumination coming from one of the LEDs attached to a hemisphere, whose position is fixed, relative to the camera. Figure 1 shows two situations where such a dome is used in an archaeological context: a mosaic and an painted prehistoric cave. Apart from the problem of the field of view, which is limited because of the hemispherical structure of the dome, such an instrument does not fit into the framework of low-cost 3D reconstruction.

Figure 1. Installation of the RTI dome: (a) on the Seasons mosaic at Saint-Romain-en-Gal (Rhône, France); (b) in front of an painted wall in the prehistoric Chauvet cave (Ardèche, France).

4. Estimating the illumination using a calibration pattern

In the expression (1) of the grey level, the two vectors $\mathbf{n}(\mathbf{P})$ and $\mathbf{s}(\mathbf{P})$ can be interchanged, since they only intervene via their scalar product, which corresponds to the so-called *shading*. Knowing that $\rho(\mathbf{P})$ and $\mathbf{n}(\mathbf{P})$ can be estimated from the knowledge of $p \geq 3$ lighting vectors $\mathbf{s}^i(\mathbf{P})$ supposedly not all coplanar, is it possible, by analogous reasoning, to estimate $\rho(\mathbf{P})$ and $\mathbf{s}(\mathbf{P})$ from knowledge of the normals at $q \geq 3$ points $\mathbf{P}_j$, $j \in \{1, \dots, q\}$, if these normals are not all coplanar?

By entering a vector equivalent to $\mathbf{v}(\mathbf{P}) = \rho(\mathbf{P})\,\mathbf{n}(\mathbf{P})$, in this case $\mathbf{w}(\mathbf{P}_j) = \rho(\mathbf{P}_j)\,\mathbf{s}(\mathbf{P}_j)$, the q grey levels of q points $\mathbf{P}_j$ of known normal can be grouped together in a matrix form similar to (2):

$$\begin{bmatrix} I(\mathbf{p}_1) \\ \vdots \\ I(\mathbf{p}_q) \end{bmatrix} = \begin{bmatrix} \mathbf{n}(\mathbf{P}_1)^\top \mathbf{w}(\mathbf{P}_1) \\ \vdots \\ \mathbf{n}(\mathbf{P}_q)^\top \mathbf{w}(\mathbf{P}_q) \end{bmatrix} \tag{6}$$

but it is impossible to factor $\mathbf{w}$ in (6) as was done for $\mathbf{v}$ in (3), because this vector depends on the point $\mathbf{P}_j$. In other words, whereas p equations (1) at the same point $\mathbf{P}$, subjected to p different illuminations, constitute a system of p linear equations with 3 unknowns, it turns out that q equations (1) at q points of known normal, under the same illumination, constitute a system of q linear equations with $3\,q$ unknowns, which correspond to the coordinates of the q vectors $\mathbf{w}(\mathbf{P}_j)$. The problem is therefore ill-posed.

An admittedly very strong assumption, but sufficient to make this problem well-posed, consists in supposing not only that the normal is known at q points $\mathbf{P}_j$, but that these q points have the same albedo ρ and are subjected to the same illumination $\mathbf{s}$. In this case, the vector $\mathbf{w}$ is independent of $\mathbf{P}$ and can be estimated, in least squares, in a form strictly analogous to (4):

$$\mathbf{w} = \begin{bmatrix} \mathbf{n}(\mathbf{P}_1)^\top \\ \vdots \\ \mathbf{n}(\mathbf{P}_q)^\top \end{bmatrix}^\dagger \begin{bmatrix} I(\mathbf{p}_1) \\ \vdots \\ I(\mathbf{p}_q) \end{bmatrix} \tag{7}$$

To deduce $\mathbf{s}$ from $\mathbf{w}$, the fact that $\mathbf{s}$ is not a unit vector, unlike $\mathbf{n}$, is not a problem, because we can arbitrarily set $\rho = 1$, which implies $\mathbf{s} = \mathbf{w}$, as far as this is done for all the lighting, in order to guarantee the consistency of all the estimates.

5. Using a spherical calibration pattern

The hypotheses that make the previous estimation problem well-posed are more or less easy to justify (Hernández et al., 2008). It is possible to know the normal at a certain number of points from multi-view data, thanks to structure-from-motion and multi-view stereo. It is more difficult to ensure that the lighting $\mathbf{s}$ is independent of the point $\mathbf{P}$ on the surface, unless these points are close to each other. Finally, it is only possible to know q $\mathbf{P}_j$ points of the same albedo thanks to a prior knowledge on the object, which is the case for lighting calibration patterns, for which relief and colour are perfectly known.

If, in order to estimate the illumination $\mathbf{s}$, it is necessary to know the normal at $q \geq 3$ points $\mathbf{P}_j$ with the same albedo ρ, it should be remembered that these normals do not all have to be coplanar, so that the pseudo-inverse matrix which appears in (7) is defined. Consequently, this disqualifies planar patterns. The simplest solution is to use a sphere of uniform colour. The silhouette of the image of a sphere in perspective projection is an ellipse whose eccentricity increases with the distance from the center of the sphere to the optical axis of the camera. Not only does this object have the widest possible variety of normal orientations, but the normal is also perfectly known at any visible point on the silhouette.

This first version of PS was used on a "statue-menhir" in the Musée Fenaille in Rodez (see Figure 2)). The white sphere placed on a tripod is used to calibrate the lighting. Three precautions must be taken when positioning it: it must be in the camera's field of view, and it must neither mask the statue-menhir nor cast a shadow on it.

Figure 2. PS applied to the "Dame de Saint-Sernin" statue-menhir (Musée Fenaille, Rodez). Top line: two photographs taken under two different lighting conditions. The white sphere is used to calibrate the lighting. Bottom line: albedo map and normal field estimated from around twenty images. As expected, the estimated albedo of the white sphere is free of shading effects.

These constraints make it difficult to implement PS outside the laboratory, for example when digitizing archaeological artefacts in the field. Moreover, to estimate the illumination a posteriori we need to cut out the silhouette of the sphere. However, even if the albedo estimated on the result of Figure 2, which is uniformly white, stands out clearly against the background, it is more difficult to segment the sphere silhouette in one of the images, because of the shading and especially the self-shadow, which corresponds to the points $\mathbf{P}$ of the sphere where the scalar product $\mathbf{n}(\mathbf{P})^\top \mathbf{s}(\mathbf{P})$ is negative. Therefore, if the sphere is fixed in the scene when the lighting varies, it is relevant to trim the sphere on the estimated albedo.

A very special case is that of the Chauvet cave, which is a prehistoric painted cave where, for reasons of soil conservation, it is impossible to leave the highly constrained pathways. It is only possible to position the sphere close to certain walls at a distance from the pathways by holding it at arm's length at the end of a pole (see Figure 3). In order to limit motion blur, the exposure time must be reduced, which reduces the depth of field. In addition, it becomes impossible to crop the sphere on the estimated albedo, since it is not fixed in the scene. It is therefore necessary to trim the sphere in each image separately, which can become tedious. For this reason, we trained a neural network to identify spheres in an image under a wide range of lighting conditions (Fainsin et al., 2023).

Figure 3. Calibrated PS used in the Chauvet cave (Ardèche, France). Top: the calibration sphere is held at arm's length at the end of a pole, as it is impossible to get closer than six meters to certain walls. Bottom left: an exposure time that was too long, even though it was well suited to the lighting conditions, rendered the image of the sphere blurred and therefore unusable. Bottom right: the image of the sphere becomes clearer after reducing the exposure time and increasing the aperture of the lens, but this is at the expense of depth of field. Another problem: the pole and the sphere cast shadows on the wall.

The use of a calibration sphere does make it possible to use calibrated PS in a wide variety of contexts and at low cost, but care must be taken when positioning the sphere in the camera field. If it is possible to dispense with the use of such a calibration sphere, this will considerably improve the scope of application of PS. We will see how to use the scene itself as a calibration pattern.

6. Estimating lighting from the scene itself

The lighting estimation method we propose aims to estimate *simultaneously* the p lighting vectors $\mathbf{s}^i$, assumed to be uniform, and *not separately*, as in Section 5. We again assume that the normal is known at q points $\mathbf{P}_j$, but make no assumptions about the albedo at these points (Coupry et al., 2025).

We reformulate the Lambertian model (1) as follows, where $i \in \{1, \ldots, p\}$ and $j \in \{1, \ldots, q\}$:

$$\frac{I^i(\mathbf{p}_j)}{\rho(\mathbf{P}_j)} - \mathbf{n}(\mathbf{P}_j)^\top \mathbf{s}^i = 0 \tag{8}$$

This implicitly assumes that $\rho(\mathbf{P}_j) \neq 0$, i.e. that the $\mathbf{P}_j$ points used for lighting estimation are not black. It is in any case preferable to disqualify such points, if only to avoid using grey levels close to 0, which are subject to "thermal noise" from the sensor. Fortunately, these points are easy to identify, as their grey levels are close to zero under any lighting conditions. To simplify notation, we rewrite (8):

$$I^i(\mathbf{p}_j)\, \alpha_j - \mathbf{n}(\mathbf{P}_j)^\top \mathbf{s}^i = 0 \tag{9}$$

where $\alpha_j = 1/\rho(\mathbf{P}_j)$. If the normal $\mathbf{n}(\mathbf{P_j})$ is known, Equation (9) is linear with respect to the unknowns α_j and $\mathbf{s}^i$.

If we know the normal at q points $\mathbf{P}_j$, and the scene is photographed under p uniform lighting $\mathbf{s}^i$, the system formed by the $p\,q$ linear equations (9) has $3\,p+q$ unknowns. It is therefore over-determined whenever $p\,q \geq 3\,p+q$. Given that $p \geq 3$, this condition is satisfied as soon as $q \geq 5$, a condition that is easy to guarantee in practice.

Equations (9) are not only linear, they are also homogeneous. To avoid the trivial solution, i.e. $\alpha_j = 0$, $\forall\, j \in \{1, \ldots, q\}$, and $\mathbf{s}^i = [0, 0, 0]^\top$, $\forall\, i \in \{1, \ldots, p\}$, it is necessary to add constraints. An albedo represents a proportion of light energy. It must therefore be positive and less than 1. Consequently, the system of equations (9) must be completed by the following q constraints:

$$\alpha_j \geq 1, \quad j \in \{1, \ldots, q\} \tag{10}$$

To ensure a degree of robustness, the coupled estimation of albedo and lighting vectors can then be rewritten as the following optimisation problem:

$$\begin{aligned} \min_{\{\alpha_j\},\{\mathbf{s}^i\}} \quad & \sum_{i=1}^{p} \sum_{j=1}^{q} \left| I^i(\mathbf{p}_j)\, \alpha_j - \mathbf{n}(\mathbf{P}_j)^\top \mathbf{s}^i \right| \\ \text{s. t.} \quad & \alpha_j \geq 1, \quad j \in \{1, \ldots, q\} \end{aligned} \tag{11}$$

which can be solved, for example, by the reweighted least squares method.

How can we find out the normal $\mathbf{n}$ at $q \geq 3$ points $\mathbf{P}_j$ in the scene? We obtain this knowledge by observing that each illumination $\mathbf{s}^i$ varies at frequencies lower than the albedo ρ and the normal $\mathbf{n}$. In other words, if we can estimate the global relief of the scene, i.e. without bias in the low frequencies, using an appropriate method, solving Problem (11) allows us to obtain a good estimate of the illumination. To do this, we can use a 3D mesh obtained either by photogrammetry, profilometry or a scanner. We prefer the first solution, which does not require any registration between both 3D reference frames, if one of the photogrammetry shots is taken from the pose used for PS.

It should be noted that the illumination estimation method that we recommend, without a calibration sphere, also makes it possible to estimate the albedo $\rho(\mathbf{P}_j) = 1/\alpha_j$, $j \in \{1, \ldots, q\}$, but this estimate is only reliable at low frequencies, if this is the case for the normals $\mathbf{n}(\mathbf{P}_j)$. It therefore makes more sense to re-estimate the albedo a posteriori, at the same time as the normals, using the classical version of calibrated PS.

7. Demonstration of a low-frequency bias in the relief

We illustrate this new variant of PS, which does not use a calibration sphere, on an archaeological artefact located in the prehistoric cave of Mas d'Azil (Ariège, France). This artefact is located in a room that is difficult to access (see Figure 4), where it would have been very impractical to use a calibration sphere. Indeed, it is inconceivable that the sphere could come into contact with the artefact. The "Mask" is a sculpted representation of a face seen in profile, in which the artist has taken advantage of the natural shape of the wall.

Figure 4. Difficult acquisition conditions in the "Mask" gallery in the Mas d'Azil cave: the use of PS has to be adapted to the reality of the terrain.

Figure 5 shows the camera poses during multi-view acquisition, estimated by structure-from-motion. The next stage in the classic photogrammetry pipeline, i.e. multi-view stereo, provides a 3D reconstruction that is faithful to the original in the low frequencies, but which does not reproduce the relief details. The normal field at the top of Figure 7 is deduced from this 3D reconstruction. Figure 6 shows three photographs (from a dozen) used by calibrated PS, with the illumination estimated from the 3D reconstruction obtained by multi-view stereo.

Figure 7 shows the normal field obtained by multi-view stereo, and that obtained by our new PS method. For the comparison to be meaningful, an equivalent number of photographs taken at the same distance from the artefact were used by both methods. It is clear that the latter result is better resolved, but also includes a bias in the low frequencies, which stems from the lighting model, which was wrongly assumed to be directional.

8. Correction of low-frequency bias in the relief

Unless the incident luminous flux is perfectly parallel and uniform, which is difficult to guarantee in practice, the calibration method presented in the previous section causes a bias in the low frequencies of the estimated normal field, as shown in Figure 7. This is because each illumination $\mathbf{s}^i$ is characterized by a vector independent of $\mathbf{P}$, showing that we have used a directional illumination model.

Figure 5. Multi-view acquisition of the "Mask": camera poses were estimated by structure-from-motion from around twenty photographs taken under constant lighting.

Figure 6. Multi-lighting acquisition of the "Mask": three photographs (from a dozen or so) taken under different lighting conditions, with the camera positioned at a constant angle.

To correct for this bias, we could assume that each illumination vector varies spatially, both in direction and intensity, and therefore denote it by $\mathbf{s}^i(\mathbf{P})$. However, we know that the estimation of lighting $\mathbf{s}^i(\mathbf{P})$ depending on $\mathbf{P}$ constitutes an ill-posed problem, which requires introducing a prior. We have already assumed that lighting varies less quickly than the relief of the scene or its color. Rather than resorting to an ad hoc regularization term, we prefer to introduce a linear interpolator $\mathbf{s}^i(\mathbf{P})$, using a grid whose parameters $\theta^i \in \mathbb{R}^{3r}$ characterize r lighting vectors defined in r control points, as shown in Figure 8.

To validate this last variant of PS, which is both easy to implement and unbiased, we use an object from the Musée Saint-Raymond in Toulouse (France), which represents a ceramic Cypriot head with calathos (cf. Figure 9).

Figure 10 shows three 3D reconstructions obtained with different methods. The correction of bias in low frequencies by the variant of PS that we propose is well visible, at least qualitatively. Indeed, the normal field on the right, obtained with bias correction, is closer to the normal field obtained by photogrammetry, in comparison with the normal field obtained by calibrated PS from data acquired by an RTI dome. Since the LEDs of the dome being are located at short distance from the object, they cause non-directional lighting.

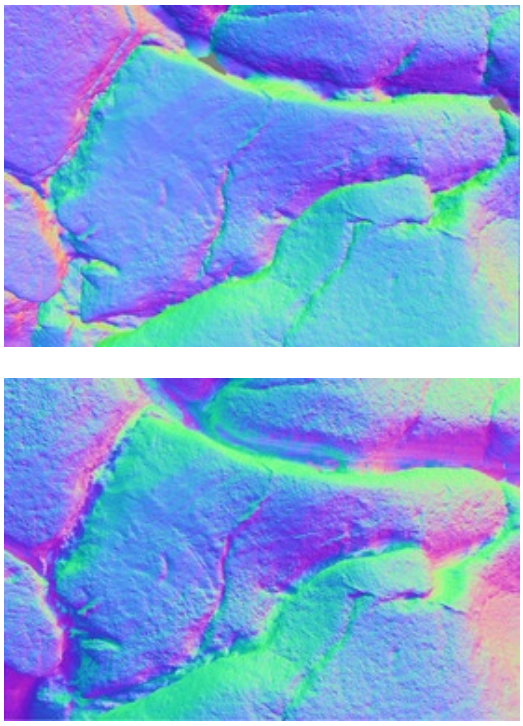

Figure 7. Normal fields estimated by photogrammetry (top) and by our new PS method (bottom), using the same acquisition distance. This second estimate clearly shows a bias in the low frequencies, due to an overly simplistic illumination model.

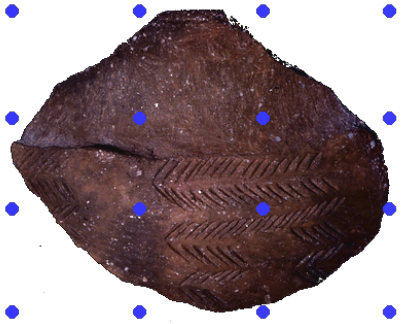

Figure 8. Interpolation grid used for local lighting estimation. Each blue dot represents one of the r control points.

Figure 9. Ceramic Cypriot head from the Musée Saint-Raymond (Toulouse, France).

Figure 10. Three normal fields of the Cypriot head represented in false colors. Left column: photogrammetry result obtained with the Metashape software. Center column: calibrated PS result obtained with an RTI dome. Right column: result obtained with the new PS method we propose. The bottom line shows enlargements of an area located at the mouth.

9. Conclusion and perspectives

In this article, we have presented several variants of calibrated PS, which differ in how the lighting is calibrated. To go in the direction of low-cost 3D reconstruction, we left out the RTI domes, which yet allow for pre-calibrating the lighting. We then showed that the use of a calibration sphere is not necessarily appropriate for the various contexts of heritage digitization. A 3D model of the scene that is coarse, but reliable in low frequencies, can be used as a calibration pattern. But to avoid the bias in low frequencies coming from a directional lighting model notoriously too simplistic, we finally proposed a new version of calibrated PS where each lighting is estimated in different control points, then interpolated over the whole object.

This version of calibrated PS has several advantages for archaeology. In addition to the ease of implementation of PS, since the scene itself is used to calibrate lighting, calculation time is reduced, because the acquisition of an initial 3D mesh by photogrammetry can be relatively coarse and therefore rely on a smaller number of acquisitions. Indeed, it is often the multi-view stereo step that constitutes the bottleneck in photogrammetry pipelines. In this regard, we regularly disseminate our advances in photometric 3D reconstruction in the open-source software AliceVision / Meshroom (Griwodz et al., 2021), as shown by the screenshot of Figure 11. Finally, it should be noted that the processing chain we recommend is frugal, since it requires only a camera and a flash in terms of equipment.

In addition to the end-to-end implementation of this work in AliceVision / Meshroom, we plan to improve a very recent reconstruction method called RNbNeuS (Brument et al., 2024), which consists of obtaining the complete 3D reconstruction of a scene, allowing the object to be seen from all angles, based on recent differentiable rendering techniques such as NeuS2 (Wang et al., 2023).

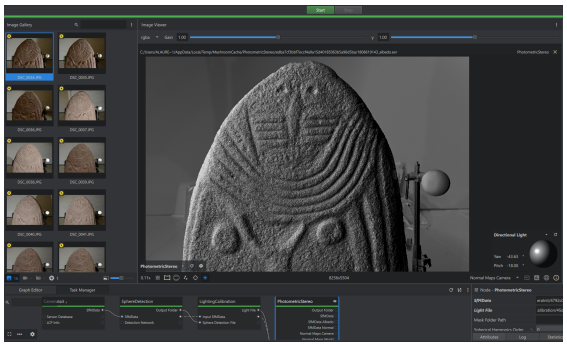

Figure 11. Implementation of PS in AliceVision / Meshroom software: example of relighting a 3D reconstructed scene.

References

Belhumeur, P., Kriegman, D., 1998. What is the set of images of an object under all possible illumination conditions? *International Journal of Computer Vision*, 28, 245–260.

Brument, B., Bruneau, R., Quéau, Y., Mélou, J., Lauze, F., Durou, J.-D., Calvet, L., 2024. RNb-NeuS: Reflectance and Normal-based Multi-View 3D Reconstruction. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5230–5239.

Coupry, B., Laurent, A., Brument, B., Mélou, J., Quéau, Y., Durou, J.-D., 2025. Assessing the Quality of 3D Reconstruction in the Absence of Ground Truth: Application to a Multimodal Archaeological Dataset. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. To appear.

Díaz-Guardamino, M., Sanjuán, L. G., Wheatley, D., Zamora, V. R., 2015. RTI and the study of engraved rock art: A re-examination of the Iberian south-western stelae of Setefilla and Almadén de la Plata 2 (Seville, Spain). *Digital Applications in Archaeology and Cultural Heritage*, 2(2–3), 41–54.

Fainsin, L., Mélou, J., Calvet, L., Carlier, A., Durou, J.-D., 2023. Neural detection of spheres in images for lighting calibration. *Proceedings of the 16th International Conference on Quality Control by Artificial Vision*.

Griwodz, C., Gasparini, S., Calvet, L., Gurdjos, P., Castan, F., Maujean, B., De Lillo, G., Lanthony, Y., 2021. AliceVision Meshroom: An open-source 3D reconstruction pipeline. *Proceedings of the 12th ACM Multimedia Systems Conference*, 241–247.

Hardy, C., Quéau, Y., Tschumperlé, D., 2024. Uni MS-PS: a Multi-Scale Encoder Decoder Transformer for Universal Photometric Stereo. *Computer Vision and Image Understanding*, 248.

Hernández, C., Vogiatzis, G., Cipolla, R., 2008. Multiview photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(3), 548–554.

Mélou, J., Laurent, A., Fritz, C., Durou, J.-D., 2022. 3D digitization of heritage: photometric stereo can help. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 48, 145–152.

Wang, Y., Han, Q., Habermann, M., Daniilidis, K., Theobalt, C., Liu, L., 2023. Neus2: Fast learning of neural implicit surfaces for multi-view reconstruction. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 3295–3306.

Woodham, R., 1980. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1), 139–144.