Multi-Light Energy Map

Florian Lardeux, Sylvain Marchand, Petra Gomez-Krämer

EUROGRAPHICS Workshop on Graphics and Cultural Heritage (2018)
R. Sablatnig and M. Wimmer (Editors)

Multi-Light Energy Map

F. Lardeux, S. Marchand & P. Gomez-Krämer

L3i, University of La Rochelle, France

Abstract

We propose a model to represent quasi-flat objects, such as coins or amphora stamps. These objects are flat surfaces, meaning their length and their width largely exceed their height, and feature a distinctive relief. This relief characterizes the object and its perception is directly influenced by the position of the object, the light direction and the viewer's direction. Our model is a single image representation containing the underlying structural variations of the object. This model, that we call 'Multi-Light Energy Map', is constructed out of several classic images taken with several illumination directions without computing the object's surface normals. We found a way to extract useful information out of this sequence of images and compile it into our map. We eventually explain how we can use this model in the case of image registration of ancient coins.

Categories and Subject Descriptors (according to ACM CCS): I.3.8 [Computer Graphics]: Applications—I.4.3 [Image Processing and Computer Vision]: Enhancement—Registration I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Shading,Shape I.4.10 [Image Processing and Computer Vision]: Image Representation—

1. Introduction

A quasi-flat object can be defined as an object which height is negligible with respect to its width and length, but has nonetheless a varying height that makes it recognizable. Many archaeological artifacts produced by some matrix are quasi-flat, e.g. seals, amphora stamps, medieval tiles... Coins are a typical example of this kind of objects. The major issue we encounter when analyzing these objects is that illumination conditions have a strong influence over the appearance of the object in the final image. Thus depending on the orientation of light we can end up with totally different images, even though as human we recognize them quickly. Because of this, comparison or registration can be tricky and unreliable.

Several methods already exist to make the object resilient to light conditions. One can for instance retrieve its surface, which inherently does not depend on light conditions. In the 1980s, Horn introduced shape-from-shading in [Hor89], which aims at reconstructing the three-dimensional shape of the surface from the brightness of a single gray scale image. This method is known to be an ill-posed problem and causes the surface to be actually hard to compute. Alternatively, photometric stereo in [Woo80] takes several images with different light orientations into account. Some techniques allow to integrate the normals to get the height map of the object (e.g., [FC88]) with the risk of integrating noise and render unreliable outputs. In [MGW01], Malzbender introduced the Polynomial Texture Maps to store light information in parametric pixels and render the object using light as a parameter. This technique, although interesting to render the surface under any illumination conditions, does not really allow us to make a light-resistant version of the object as it still depends on light. In [Mar14], a hybrid model-to-image model for the registration of ancient coins is used. The idea is to capture coins with several light orientations and use correlation to get a similarity factor and registration parameters. The viewing conditions and the rigid deformations are successively estimated and updated until convergence. This approach motivates us to create a unique stable representation that would be usable immediately for registration. We could even directly find our object's contours. For instance, in [RTF*04], the authors present a technique that uses multi-flash imaging to capture depth edges of a 3D scene. The goal is not specifically to find edges but rather to use them to create a non-photorealistic version of the scene. Four lights are placed just around the camera lens and light up the scene successively. This technique is interesting because it also forces light conditions but this time to directly retrieve strong depth variations of the scene which capture relevant information.

Our technique is close to a multi-view approach that stores several images, but we intend to keep only one image as a final model. Therefore, we want to prioritize a model that is close enough to the object without having the need to store too much data (as done in [MGW01] or [Woo80] for instance) and without risking to add noise when integrating normals to get a height map. On the other hand, since we would like to keep as much relevant information as possible, we do not want to directly compute contours as it is done in [RTF*04]. We take several images with the same view point by changing only the light orientations in order to calculate a representation that is independent of light and easily reproducible. We call this representation the "Multi-Light Energy Map". We do not

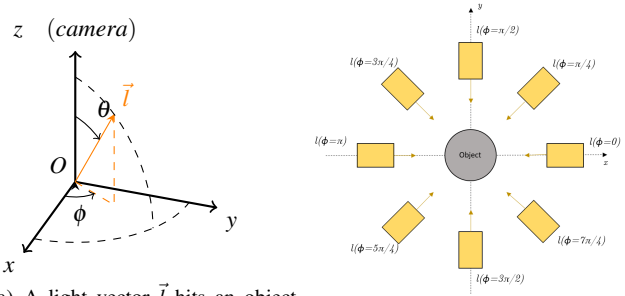
intend to perform a reconstruction of the object. Rather, this technique can be used a first step for image comparison.

The remainder of the paper is organized as follows. We first explain our setup to acquire the images in Section 2. We then detail the mathematics to derive our model and explain the procedure to compute our map in Section 3. In Section 4 we use our map for image registration and we compare it to the previous work of [Mar14]. We conclude in Section 5.

2. Image Acquisition

Ideally, we would like to set the illumination conditions so as to cover all different possible orientations within the hemisphere around the object. However, on top of being quite demanding, we actually do not need to care for all possible positions of the elevation angle, which is the angle formed by the z-axis and the light vector. We would like to get a set of images that underline nicely the main features of the object. Fig. 1(a) shows a light ray that hits the object (whose center is O) with an azimuth angle ϕ and an elevation angle θ. Ideally the best choice for our elevation angle is $\theta = \frac{\pi}{2}$ because the object is simply a flat surface with some characteristic relief. By using grazing light we exacerbate the steepest edges of the height map, which represent the most important features of our object. As shown in Fig. 1(b), the camera is placed on the z-axis and we move the light around the object at equally spaced positions.

Now that we have completed our acquisition, we get a set of m images, each at a different orientation. Ideally, the k^{th} image is taken with an incident light azimuth angle $\phi = \frac{2k\pi}{m}, k \in [\![0, m-1]\!]$. We would like to obtain a single image out of this set that will be our light resistant representation.



(a) A light vector $\vec{l}$ hits an object with elevation θ and azimuth ϕ.

(b) Images acquisition from the camera view: example with 8 different light azimuths.

Figure 1: *Geometry of the acquisition.*

To reduce the impact of noise, we apply anisotropic diffusion [PM90] to smooth our images because the relevant information of the object is contained in its physical edges. As stated in [PM90], Gaussian blurring does not respect the natural boundaries of objects. Only few iterations are enough depending on the quality of initial images.

3. Energy Map Computation

Consider the Lambertian image model:

$$I(x,y,\phi,\theta) = \vec{l}(\phi,\theta)\cdot\vec{n}(x,y) \tag{1}$$

where $\vec{l}$ is the incident light vector, characterized by its direction and $\vec{n}$ is the normal vector to the surface of the object at pixel (x,y). Thus we have:

$$\vec{l}(\phi,\theta) = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta)$$

$$\text{and } \vec{n} = (n_x, n_y, n_z)$$

It should be noted that we chose a Lambertian model to describe our image signal because it is fairly simple to use. This hypothesis, although formally wrong in most cases allows us to derive a model that gives good results nonetheless.

Intuitively, wherever we encounter a steep edge with a certain orientation, our image signal I should display strong variations, and where there is no relief at all, we should get no difference in the signal from one image to another. We will first explain the calculation in the continuous case and then derive a usable formula for our discrete case.

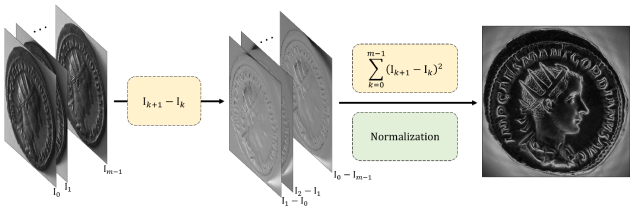


Figure 2: *The pipeline of our method: pairs of successive images are subtracted and then used into the variance calculation. We normalize the final map after the calculations.*

3.1. Mathematical Intuition in the Continuous Case

In order to find how strong the variations occur we calculate the derivative of the model described in Eq. 1 with respect to the azimuth ϕ:

$$\frac{\partial I}{\partial \phi} = \frac{\partial \vec{l}}{\partial \phi}\cdot\vec{n} = \sin\theta\cdot(n_y\cos\phi - n_x\sin\phi)$$

Our final map will be a function of x and y so that at each pixel (x,y) we get how wide are the variations over all orientations. We thus calculate the variance of our signal $\frac{\partial I}{\partial \phi}$:

$$\sigma^2 = \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{\partial I}{\partial\phi}\right)^2 \mathrm{d}\phi - \frac{1}{2\pi}\underbrace{\left(\int_0^{2\pi}\frac{\partial I}{\partial\phi}\,\mathrm{d}\phi\right)^2}_{=0 \text{ (mean of sine over its period)}}$$

$$= \frac{1}{2\pi}\sin^2\theta\left(n_x^2\underbrace{\int_0^{2\pi}\sin^2\phi\,\mathrm{d}\phi}_{=\pi} + n_y^2\underbrace{\int_0^{2\pi}\cos^2\phi\,\mathrm{d}\phi}_{=\pi} - 2n_xn_y\underbrace{\int_0^{2\pi}\sin\phi\cos\phi\,\mathrm{d}\phi}_{=\frac{1}{2}\int_0^{2\pi}\sin 2\phi\,\mathrm{d}\phi=0}\right)$$

$$\boxed{\sigma^2 = \frac{1}{2} \cdot \sin^2\theta \cdot \left\|\vec{n}_{xy}\right\|^2} \tag{2}$$

where we denote:

$$\left\|\vec{n}_{xy}\right\| = \sqrt{n_x^2 + n_y^2}$$

Because $\vec{n}_{xy}$ is the normal vector projected onto the (xOy) plane, we have shown above that the variance σ^2 is directly connected to the gradient of the height map h in the plane (xOy). The value $\left\|\vec{n}_{xy}(x,y)\right\|$ tells us how strong the variations are in the height map at a certain position (x,y) in the plane. Therefore this simple calculation gives us information about the variations in the object structure. Note that we could retrieve the norm of the height map gradient $\left\|\vec{\nabla}h\right\|$, but experiments showed equivalent results as for Eq. 2.

3.2. Discrete Computation

Let us consider the discrete formulation:

$$\frac{\partial I}{\partial \phi} \to \frac{\Delta I}{\Delta \phi} = \frac{I(\phi_{k+1}) - I(\phi_k)}{\Delta \phi} = \left(\frac{\vec{l}_{k+1} - \vec{l}_k}{\Delta \phi}\right) \cdot \vec{n}$$

where $\phi_k = \frac{2k\pi}{m}$, $\Delta\phi = \frac{2\pi}{m}$ and $\vec{l}_k$ is the light vector at position k. Of course if $k = m-1$, we compute $\frac{I(\phi_0) - I(\phi_{m-1})}{\Delta\phi}$.

Then,

$$\frac{\Delta I}{\Delta \phi} = \frac{1}{\Delta \phi}\left[(\cos\phi_{k+1} - \cos\phi_k)\sin\theta \cdot n_x + (\sin\phi_{k+1} - \sin\phi_k)\sin\theta \cdot n_y\right]$$

In the same way, we calculate the variance of our values. We denote $\Delta I_k = I_{k+1} - I_k$:

$$\begin{aligned}
\sigma^2 &= \frac{1}{m}\sum_{k=0}^{m-1}\left(\frac{\Delta I}{\Delta\phi}\right)^2 \\
&= \frac{1}{m\Delta\phi^2}\sum_{k=0}^{m-1}\left[(\cos\phi_{k+1} - \cos\phi_k)\sin\theta \cdot n_x + (\sin\phi_{k+1} - \sin\phi_k)\sin\theta \cdot n_y\right]^2 \\
&= \frac{1}{m}\left[\left(\frac{\sin\theta}{\Delta\phi}\right)^2 \left(n_x^2 \sum_{k=0}^{m-1}(\cos\phi_{k+1} - \cos\phi_k)^2 + n_y^2 \sum_{k=0}^{m-1}(\sin\phi_{k+1} - \sin\phi_k)^2\right)\right. \\
&\qquad \left. + 2n_x n_y \underbrace{\sum_{k=0}^{m-1}(\cos\phi_{k+1} - \cos\phi_k)(\sin\phi_{k+1} - \sin\phi_k)}_{=0}\right]
\end{aligned}$$

We use the trigonometric identities in our calculations:

$$\begin{cases}
\cos\phi_{k+1} - \cos\phi_k = -2\sin\left(\frac{\phi_{k+1} - \phi_k}{2}\right)\sin\left(\frac{\phi_{k+1} + \phi_k}{2}\right) \\
\sin\phi_{k+1} - \sin\phi_k = 2\sin\left(\frac{\phi_{k+1} - \phi_k}{2}\right)\cos\left(\frac{\phi_{k+1} + \phi_k}{2}\right)
\end{cases}$$

$$\begin{aligned}
\sigma^2 &= \frac{1}{m}\left(\frac{\sin\theta}{\Delta\phi}\right)^2 \left(n_x^2 \sum_{k=0}^{m-1} 4\sin^2\left(\frac{\phi_{k+1} + \phi_k}{2}\right)\sin^2\left(\frac{\Delta\phi}{2}\right) + \right. \\
&\qquad \left. n_y^2 \sum_{k=0}^{m-1} 4\cos^2\left(\frac{\phi_{k+1} + \phi_k}{2}\right)\sin^2\left(\frac{\Delta\phi}{2}\right)\right) \\
&= \frac{4}{m}\sin^2\left(\frac{\Delta\phi}{2}\right)\left(\frac{\sin\theta}{\Delta\phi}\right)^2 \\
&\times \left(n_x^2 \underbrace{\sum_{k=0}^{m-1}\sin^2\left(\frac{2(k+1)\pi}{m}\right)}_{=\frac{m}{2}} + n_y^2 \underbrace{\sum_{k=0}^{m-1}\cos^2\left(\frac{2(k+1)\pi}{m}\right)}_{=\frac{m}{2}}\right) \\
&= 2\cdot\sin^2\left(\frac{\Delta\phi}{2}\right)\cdot\left(\frac{\sin\theta}{\Delta\phi}\right)^2 \cdot \left(n_x^2 + n_y^2\right)
\end{aligned}$$

Finally,

$$\boxed{\sigma^2 = \frac{1}{2}\cdot\mathrm{sinc}^2\left(\frac{\pi}{m}\right)\cdot\sin^2\theta\cdot\left\|\vec{n}_{xy}\right\|^2}$$

where $\mathrm{sinc}(x) = \frac{sin(x)}{x}$. We can compute our map $\left\|\vec{n}_{xy}\right\|$ as:

$$\left\|\vec{n}_{xy}\right\| = \sigma\cdot\sqrt{\frac{2}{\mathrm{sinc}^2\left(\frac{\pi}{m}\right)\cdot\sin^2\theta}} \tag{3}$$

We actually do not need to calculate all of this, as we will normalize our map in the end. Therefore, the energy map $\mathscr{E}$ is calculated as:

$$\boxed{\mathscr{E} = \sqrt{\sum_{k=0}^{m-1}{\Delta I_k}^2}} \tag{4}$$

The whole process of computing the energy map is summarized in Fig. 2. This mathematical object describes how strong the variations are in the height map, since it is directly connected to its gradient norm. Fig. 3 is an example of such an image. This map synthesizes some of the main features of the analyzed quasi-flat object, which are the relief edges. The brighter the pixels, the steeper the edges in the surface relief. We found experimentally that even $m = 4$ gives convincing results so far. The differences between the use of 4, 8 or 16 images does not change significantly the quality of our final map.

Figure 3: *Energy map obtained from 16 images of a coin.*

4. Application: Image Registration with our Model

The goal of such model is to be used in the pipeline of techniques such as object classification or recognition. Here, we will demon-

strate the efficiency of our model in the simple case of image registration, as done in [Mar14, Mar15]. The idea is to use registration to compute the similarity between two coins. In these examples, different light conditions were also used, either by alternatively updating light and viewing conditions, or by creating a single image model which is resilient to the illumination conditions. We take an approach similar to the one presented in [Mar15], except that here we do not explicitly calculate the normals. We reproduced the experiments introduced in [Mar14] (classic image registration using the Fourier-Mellin transform), published at GCH 2014, with two new sequences of images.

Figure 4: *Registration process with the energy map. Left and center: energy maps of the same coin taken with different sets of images. Right: result of registration (i.e. the transformed first image).*

If we perform a registration directly on the images of the coins from 2 different series of 16 images, we get varying results. The similarity factor gets up to 0.98 and down to 0.26, with a mean value of **0.64** (standard deviation 0.17). Thus registration is highly unreliable using only image captures as is. In Fig. 4 we see the result of registration performed on two different energy maps of the same coin using the same image sequences. The similarity factor obtained here is **0.94**. On Fig. 5 we show the efficiency of our map to perform registration on images of the same coin from two different series of captures. The use of the two images yields a poor registration as the effect of light orientation impacts on the image appearance. On the other hand, using energy maps of the coin to evaluate the appropriate transformation gives fairly good results when applied to any of the images from the second series. We also show the image difference, defined as $diff = T(I_2) - I_1 + 0.5$, where T denotes the appropriate rigid transformation and I_1, I_2 are the normalized images in $[0,1]$. Medium gray areas denote equal parts between images whereas dark and bright parts emphasize differences. When using directly the images, the differences are noticeable where the coins are not properly superimposed. Using the energy maps we only see differences in shading.

5. Conclusions and Future Work

In this short paper, we have introduced the Multi-Light Energy Map to represent a quasi-flat object, such as ancient coins or amphora stamps. The advantage of this model is threefold: it is independent of light conditions because it stems from a series of images with controlled light directions. The final model is very simple and light: it is one image that captures an equivalent of the norm of the height map gradient of the object. Finally, the model is easily reproducible: the same object will yield the same energy map. This model is computed as a mean to perform coin recognition, registration, comparison or even contour extraction. We intend to run more tests when we will have the chance to extend our database and we will investigate the use of such model for coin classification.

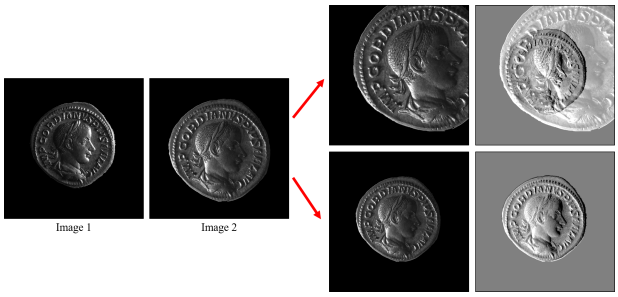


Figure 5: *Registration without (top right) and with (bottom right) energy maps. Left: two images of the same coin from different series. Results on the right: first column shows the result using the images as is (top) or using the energy maps and then using the transformation on image 2; second column shows the differences between image 1 and the transformed image 2.*

References

[FC88] FRANKOT R., CHELLAPPA R.: A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence 10*, 4 (jul 1988), 439–451. URL: https://doi.org/10.1109/34.3909. 1

[Hor89] HORN B. K. P.: Shape from shading. MIT Press, Cambridge, MA, USA, 1989, ch. Obtaining Shape from Shading Information, pp. 123–171. URL: http://dl.acm.org/citation.cfm?id=93871.93877. 1

[Mar14] MARCHAND S.: IBISA: Making Image-Based Identification of Ancient Coins Robust to Lighting Conditions. In *EUROGRAPHICS Workshop on Graphics and Cultural Heritage (GCH)* (Darmstadt, Germany, Oct. 2014), pp. 13–16. URL: https://hal.archives-ouvertes.fr/hal-01073329. 1, 2, 4

[Mar15] MARCHAND S.: The IBISA System for Computer-Assisted Coin Identification and Hoards Study. In *Proceedings of the XV International Numismatic Congress (INC), volume I* (Taormina, 2015), pp. 204–206. 4

[MGW01] MALZBENDER T., GELB D., WOLTERS H.: Polynomial texture maps. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 2001), SIGGRAPH '01, ACM, pp. 519–528. URL: http://doi.acm.org/10.1145/383259.383320. 1

[PM90] PERONA P., MALIK J.: Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence 12*, 7 (jul 1990), 629–639. URL: https://doi.org/10.1109/34.56205. 2

[RTF*04] RASKAR R., TAN K.-H., FERIS R., YU J., TURK M.: Non-photorealistic camera: Depth edge detection and stylized rendering using multi-flash imaging. In *ACM SIGGRAPH 2004 Papers* (New York, NY, USA, 2004), SIGGRAPH '04, ACM, pp. 679–688. URL: http://doi.acm.org/10.1145/1186562.1015779. 1

[Woo80] WOODHAM R. J.: Photometric method for determining surface orientation from multiple images. *Optical Engineering 19*, 1 (feb 1980). URL: https://doi.org/10.1117/12.7972479. 1