Combination method: Photometric stereo with Shadows

Amina Dulac, Sandy Martedi, Hideo Saito
Graduate School of Science and Technology
Keio University
3-14-1 Hiyoshi Kohoku-ku, Yokohama 223-8522, Japan
amina.dulac@centrale-marseille.fr, sandy@hvrl.ics.keio.ac.jp, saito@hvrl.ics.keio.ac.jp

Kouichi Tezuka, Masayoshi Shimizu
Fujitsu Laboratories Ltd.
4-1-1 Kamikodanaka, Nakahara-ku, Kawasaki, Kanagawa 211-8588, Japan
ktezuka@jp.fujitsu.com, shimizu.masa@jp.fujitsu.com

Abstract

This paper presents an algorithm allowing to perform photometric stereo with high accuracy results in the presence of shadows. First, we choose a combination of three images suited to reveal the presence of shadow and use it to calculate the normal vector on each pixel. This normal vector is then used to recalculate the intensity in order to evaluate the distance with the real intensities and detect the shadow pixels on each image. If the error between the calculated intensities and the real intensities are superior to a certain threshold, the pixel is determined to be in the region of a shadow. The algorithm runs in loop to identify a valid set of images, that is to say without shadow, for each pixel. The method is designed to limit the computational cost. The results show that the accuracy is maintained as compared to much heavier computational cost algorithms.

Keywords: photometric stereo, shadows, 3D reconstruction.

1 Introduction

Photometric Stereo is a method used to recover the 3D shape of an object with high detail accuracy. For that it uses a set of images of the object from the same point of view but under different illuminations. In case of calibrated photometric stereo, which is the one being presented in this paper, the light sources directions are known. The approximation of Lambertian surface is also made. One remaining issue in photometric stereo is how to handle shadow areas. Some papers proposed algorithms to deal with the shadows in photometric stereo, however all the proposed methods require heavy calculations and are not suitable for real time processing. In this paper we describe a new algorithm with reduced time of calculation and producing results with as good accuracy as heavier methods. This method was developed in order to be used for 3D shape reconstruction of the human skin by photometric stereo technique for medical control.

This paper is separated into four sections. In the first section, we provide an overview of previous work. In section 2 we introduce our methods and in section 4 we conclude with results.

2 Related work

Photometric stereo is considered as an interesting intensity-based 3D shape recovering technique because of the precision of the results it can achieve on objects with Lambertian reflectance. Methods have also been developed to handle specular surfaces [7].The theory called calibrated photometric stereo requires known the light sources directions. Additionally, the principles of this method have now been extended to uncalibrated photometric stereo where the light sources directions are unknown [3]. This paper however uses calibrated photometric stereo since our application uses known light sources and focuses on the accuracy of the results as opposed to overcoming such constraints.

The main difficulty in the accuracy of the results lies in the presence of shadows. In order to have results reliable enough, especially for a medical use, the shadow areas need to be detected and handled properly. The traditional way to remove the shadows of an image is to apply an intensity threshold. However this simple technique only works on objects with constant surface albedo coefficient. Chandraker et al. [6] proposed a method allowing shadow labeling in photometric stereo with changing albedo and mutli-light sources using energy minimization where the "data term" is based on photometric stereo and the "smoothness term" supports the spatial continuity. Due to the exponential number of possible label configurations, minimizing this type of energy requires the use of a fast graph cuts algorithm from Boykov et al.[10]. However, even with this type of approximation the computational cost stays very high. Sunkavalli et al. [8] also proposed a method dealing with shadows in uncalibrated photometric stereo. Instead of reasoning about per-pixel intensity, their approach is reasoning about illumination subspaces using a RANSAC type algorithm, which needs about 1000 iterations thus requiring a large amount of time to compute.

Our work is a per-pixel approach but using targeted combinations to decrease the amount of calculation required. We show that by not considering the totality of the possible labels, by choosing the right combination for each pixel, we can achieve a similar degree of accuracy and allow for high-speed execution.

3 Identifying shadows

For clarity we will briefly describe notations. We consider a set of n images of a same object from the same point of view. Each image is illuminated by a light source j and has m surface points. This section begins by describing the photometric stereo theory. We then explicitly show the impact of the shadows in the equations and propose an algorithm to detect pixels corrupted by shadow in each image.

3.1 Calibrated Photometric Stereo

Photometric Stereo is a technique used to recover the 3D shape of an object from 2D images of it. The objects considered in this paper are assumed to have Lambertian reflectance. Given a set of images of the same object, from the same point of view but illuminated by different known light sources, the intensity of the pixel i in the image j is expressed as:

$$c_{ij} = \rho l_j^T n_i \,, \qquad (1)$$

where l_j is the light source direction vector, n_i the normal vector at the pixel i and ρ the albedo coefficient.

For *m* surface points and *n*>3 light sources, the concatenation of all pixels on all images leads to the following system:

$$I = L^T N \,, \qquad (2)$$

where I indicates a $n \times m$ intensity matrix, L denotes a $3 \times n$ light source matrix and N is a $3 \times m$ matrix containing the product of the albedo coefficients and normal vectors. For a number of linear independant light sources superior to 3, L is at least of rank 3. The normal matrix N can be recovered using the pseudo inverse of L as:

$$N = (L^T)^{+} I \,. \qquad (3)$$

The relative depth of each point can then be calculated by integrating the normal vectors.

3.2 Shadows

One limitation of this method comes from the presence of shadows. When a point of the surface of an object is not reached by the light the Photometry stereo theory is not applicable anymore. Areas of shadows can produce corrupted results in the normal map and thus lead to an inaccurate recovery of the depth. If we write equation (3) for one pixel we have:

$$\begin{cases} N_x = \sum_{j=1\ldots n} c_j R_x(L) \\ N_y = \sum_{j=1\ldots n} c_j R_y(L) \,. \\ N_z = \sum_{j=1\ldots n} c_j R_z(L) \end{cases} \qquad (4)$$

The coefficients $R_x(L), R_y(L)$ and $R_z(L)$ only depends on the lights source directions, which are known in the case of calibrated photometric stereo. When a pixel is in the shadow, only the term c_j representing its intensity is corrupted in the equation: the intensity value will be lower than it should be because of the shadow. Thus we can naturally infer that when the percentage of images touched by shadow for a pixel is low, the influence on the calculated normal vector for this pixel is almost undetectable. If the percentage of shadow images is however too important the calculation leads to a highly inaccurate normal vector. Based on this simple observation we can deduce that if we apply photometric stereo on three images among which one contains shadows, the normal vector of the pixels in the shadow will be highly corrupted, making it easier to detect. Consequently, a method based on combinations of three images to detect the shadows can be used.

3.3 Algorithm

We consider a set of n images containing some shadows. We also make the hypothesis for each pixel that there is at least three images without shadow. If that is not the case, Photometric Stereo theory is not applicable. We assume that the hypothesis is respected by choosing a number of images important enough and light source directions offering a good coverage of the object. The previous observations lead us to the following per-pixel algorithm:

1. Sort the images by intensity. Notice that we first convert the images into grayscale images. We organize this group of sorted images from the brightest to the darkest $\{im_1, im_2 \ldots im_n\}$.
2. Apply three sources photometric stereo using the darkest image im_n and two of the three brightest images in order to recover the normal vector N. The image choice relies on the fact that we consider that the three brightest images do not contain shadows and that if the pixel is in the shadow in some images then the darkest image will contain shadows. Thus we get a combination of three images, two of which without shadows and one potentially containing shadows.

3. Use the recovered normal vector N to recalculate the intensity $I_j = L_j{}^T N$ on each image j and calculate the global relative error with the real intensities $E_{tot} = \sum_{j=1}^{n} \left| \frac{I_j^{real} - I_j}{I_j^{real}} \right|$.
4. If $E_{tot} < \alpha$, we can regard the pixel as non-affected by shadow because the calculated intensities I_j are almost equal to the real intensities I_j^{real}. Then we attribute the set of images as a valid set to this pixel. If $E_{tot} > \alpha$, which mean that the calculated intensities I_j are too different from I_j^{real}, then we label this pixel as a shadow pixel in image n and repeat step one to three on the set of images $\{im_1, im_2 \ldots im_{n-1}\}$. The threshold α is chosen according to experimental results.

The loop runs until every pixel is attributed a valid set of images or until there is only three images left in the set of images, so there is maximum $n - 3$ loops and the number of combinations calculated is a $O(\binom{n}{3})$. As a comparison, the method from [6] requires the calculation of a $O(\sum_{k=3}^{n} \binom{n}{k})$ combinations and the fast graph cut algorithm needs about 100 iterations to reach a good minimum. The method developed in [8] does not operates on a per-pixel basis but still has a high computational cost because of its RANSAC type procedure. If we consider that the image contains h visibility subspaces, each occupying the same proportion of space then for m pixels the number of operations is $\sum_{k=1}^{h} 1000 \times (m - (k-1)\frac{m}{h})$.

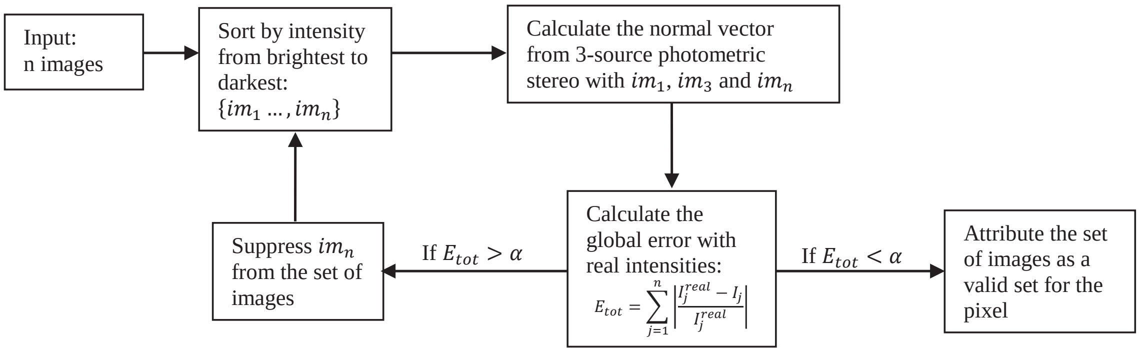


Figure 1. schema of the algorithm.

4 Results

In this section we present the results obtained with our method on synthetic data and real data. The first example we show is a synthetic set of images of a simple sphere on a plane (Figure 2). This dataset contains both attached and cast shadows and the albedo coefficient differs from the sphere to the plan. The multiple shadows are also overlapping, creating a more complex map of valid sets but the symmetry of the problem makes it easy to judge on the accuracy of the results. We can notice the difference of the normal map obtained by using classic photometric stereo where the influence of shadows is explicitly corrupting the result and with our shadow handling method where the marks left by the shadows do not appear anymore.

The second dataset presented contains eight images of a real object with a complex shadow map (Figure 3). Each color on the shadow map represents a different valid set of images, this means pixels sharing the same color are in the shadow in the same images. Please note that the color is decided arbitrary for each valid set of images. We can notice the produced shadow map seems very close to the true shadow map of the object. For comparison we show the shadow map obtained using the algorithm from [8], the accuracy of the shadows area is about as high as the one we achieve. Additionally, this is performed with a smaller computation time. The last example is a set of twelve pictures of a horse head shaped object (Figure 4). If we first look at the horse's ear we can see that the area is highly affected by shadows. As a consequence, the 3D reconstruction obtained from classic photometric stereo contains some artifacts. In this case a peak along the ear area is formed. Our method however avoids the formation of such inconsistencies by effectively detecting the shadows projected by the ear.

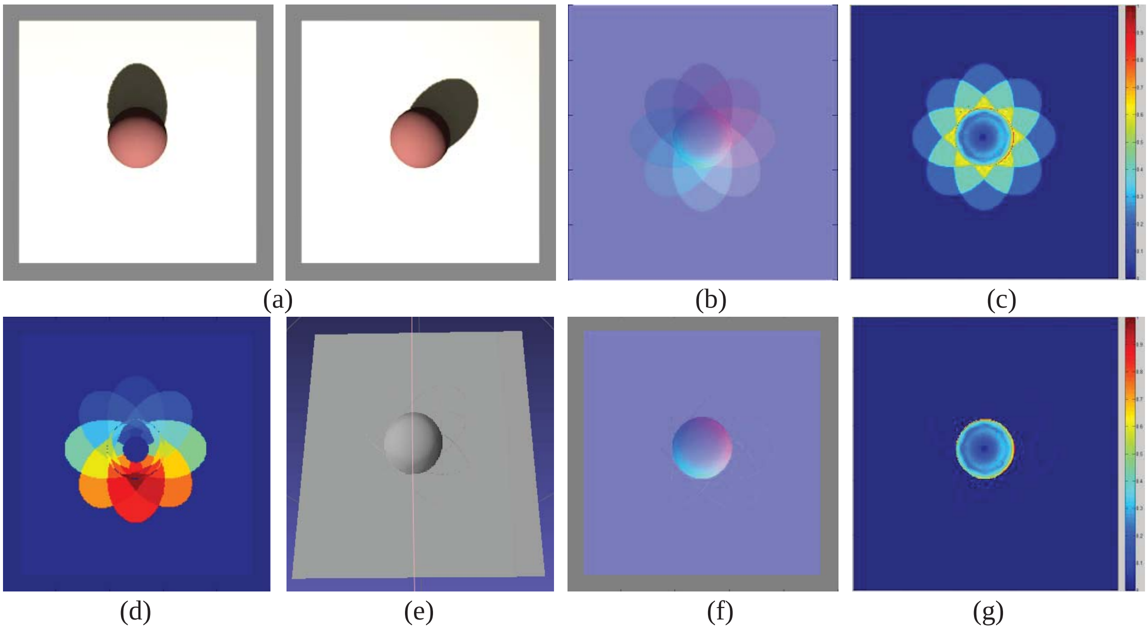


Figure 2. (a) 2 of the 8 input images of synthetic sphere on a plan. (b) normal map without applying our method. We can see the impact of the shadows. (c) error between the true normal map and the normal map (b) . (d) shadow map obtained with our method. (e) 3D reconstruction of the object after applying our method. (f) normal map obtained with our method, you can notice the disappearance of the marks previously made by the shadow areas. (g) error between the true normal map and the normal map obtained with our method (f).

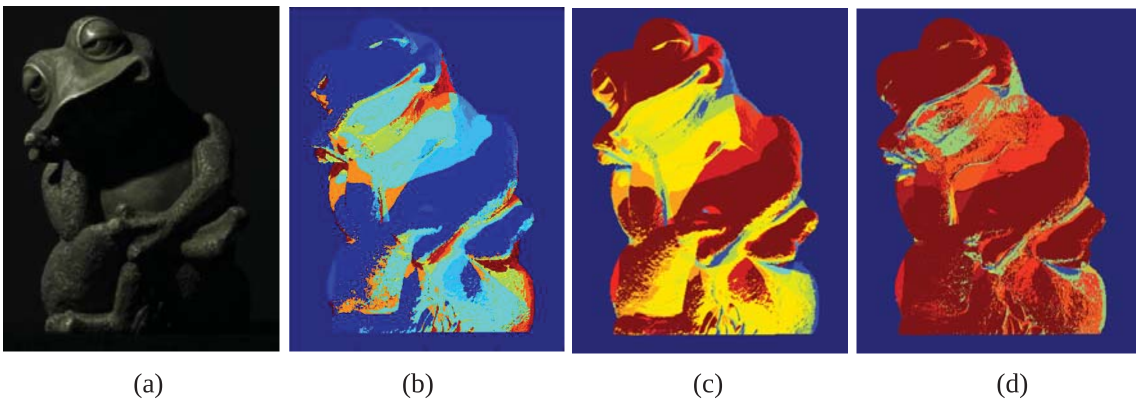
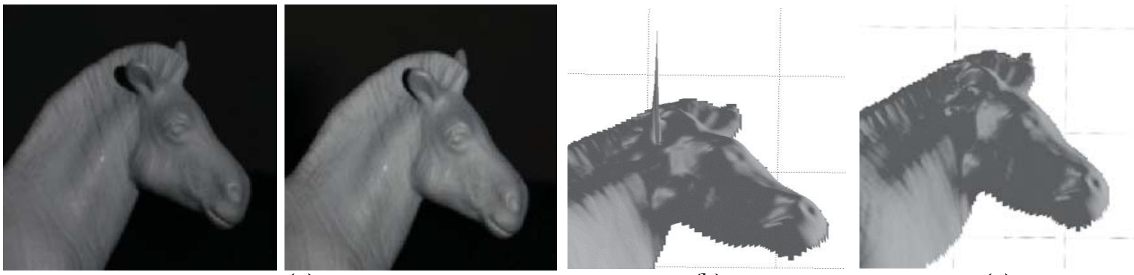


Figure 3. (a) sample input images of the statue of a frog. (b) shadow map obtained with our method. Pixels sharing the same color are in the shadow in the same images. Please note that the color of each valid set of images is chosen arbitrary. We can see how similar it is to the true shadow map (c) true shadow map taken from [8]. (d) shadow map obtained with method from [8].



(a) (b) (c)

Figure 4. (a) sample input images of the statue of a head of a horse. (b) 3D reconstruction obtained from classic photometric stereo. We can notice the pic the area of the ear due to cast shadows. (c) 3D reconstruction obtained from our method. The artifact created by the shadows of the ear have been removed. Note that image (b) and (c) look more pixelated than the input images due to the lost of definition caused by the passage from 2D to 3D.

5 Conclusion

We have presented in this paper a new method to detect both cast and attached shadows in photometric stereo with albedo variation and complex shapes. Using well-chosen combinations of three images to calculate the intensity error on a per-pixel basis allow us to avoid the heavy computational cost needed in the previous existing methods while conserving the high degree of accuracy in the results.

This paper only focuses on calibrated photometric stereo because it offers the best accuracy for medical applications such as human skin reconstruction, but one possible direction of work could be to extend the method to uncalibrated photometric stereo in order to use it in a wider range of applications.

References

[1] Horn, B. (1986) *Robot Vision*.

[2] Belhumeur, P.N., Kriegman, D.J., Yuille, A.L. (1999). The bas-relief ambiguity. Int. Journal of Computer Vision 35(1), 33-44.

[3] Basri, R., Jacobs, D., Kemelmacher, I. (2007). Photometric Stereo with general unknown lighting.Int. Journal of Computer Vision 72(3), 239-257.

[4] Barsky, S., Petrou, M. (2003). The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. PAMI, 25(10), 1239-1252.

[5] Hertzmann, A., Seitz, S. (2005). Example-based photometric stereo: Shape reconstruction with general, varying BRDFs. PAMI, 27(8), 1254-1264.

[6] Chandraker, M., Kahl, F., Kriegman, D. (2007). Shadowcuts: Photometric stereo with shadows. Inc: Proc. IEEE conf. Computer Vision and Pattern recognition.

[7] Ikeushi, K. (1981). Determining surface orientations of specular surfaces by using the photometric stereo method. IEEE Trans. Pattern Anal. Mach, Intell, 3(6), 661-669.

[8] Sunkavalli, K., Zickler, T., Pfister, H., (2010). Visibility Subspaces: Uncalibrated Photometric Stereo with Shadows. Inc: ECCV 2, volume 6312 of Lecture Notes in Computer Science, page 251-264. Springer.

[9] Woodham, R. (1978). Photometric stereo: A reflectance map technique for determining surface orientation from image intensity. In SPIE, volume 155.

[10] Boykov, Y., Vexler, O., Zabih, R. (2001). Fast approximate energy minimization via graph cuts. PAMI, 20(12), 1222-1239.