PHOTOMETRIC STEREO THROUGH AN ADAPTED ALTERNATION APPROACH

Carme Julià, Angel D. Sappa, Felipe Lumbreras, Joan Serrat and Antonio López

Computer Vision Center and Computer Science Department,
Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

ABSTRACT

Photometric stereo aims at finding the surface normal and reflectance at every point of an object from a set of images obtained under different lighting conditions. The obtained intensity image data are stacked into a matrix that can be approximated by a low-dimensional linear subspace, under the Lambertian model. The current paper proposes to use an adaptation of the *Alternation* technique to tackle this problem when the images contain missing data, which correspond to pixels in shadow and saturated regions. Experimental results considering both synthetic and real images show the good performance of the proposed *Alternation*-based strategy.

Index Terms— Photometric stereo, alternation technique, missing data

1. INTRODUCTION

The image of an object depends on many factors such as lighting conditions, viewpoint, geometry and albedo of the object. Hence, if a general description of the object is required, it is necessary to design a representation that capture all the image variation caused by those factors. Then, it could be used for object detection and recognition, for instance.

Photometric stereo aims at estimating the surface normals and reflectance of an object by using several intensity images obtained under different lighting conditions. The general assumptions are that the projection is orthographic, the camera and objects are fixed and the moving light source is distant from the objects. Hence, it can be assumed that the light shines on each point in the scene from the same angle and with the same intensity.

The starting point (as pointed out in [1]) is that the set of images produced by a convex Lambertian object, under arbitrary lighting, can be well approximated by a low-dimensional linear subspace of images. In [2], it is shown that, without shadows, a Lambertian object produces a 3D subspace of images. This linear property suggests to use factorization techniques to model the image formation and obtain each of the factors that contribute to it. The intensity of the pixels of the images are stacked into a *measurement matrix*, whose rows and columns correspond to each of the pixels and images, respectively.

One common assumption in most photometric stereo approaches is that the images do not have shadows nor saturated regions (e.g., [3]), which correspond to points with very low and high intensities values, respectively. This is due to the fact that these points do not follow a Lambertian model. Although if there are only a few of them the Lambertian model is a good approximation, their presence can bias the obtained results. Hence, some approaches propose methods to reject them or tend to reduce their influence on the results.

Hayakawa [4] presents a photometric stereo approach for estimating the surface normals and reflectance of objects, which is similar to the factorization method presented in [5] for the shape and motion estimation. This approach is based on SVD and, in order to obtain a unique decomposition, it uses one of the following constraints: 1) at least six known pixels have constant or known reflectance; 2) the light source intensity value is constant or known in at least six known images. Hayakawa proposes a strategy to deal with shadows. The idea is to select an initial submatrix, whose entries do not correspond to pixels in shadow. Then, the surface normal and reflectance of pixels in shadow are estimated by growing a partial solution obtained from the initial submatrix. Unfortunately, to find a submatrix without shadows is in general a quite expensive task. Furthermore, the SVD has a high computational cost when dealing with big matrices, which are common in this application.

Epstein et al. [6] present an approach based on [4] for learning models of the surface geometry and albedo of objects. They point out that in [4] the obtained reflection and light directions are recovered up to a rotation. In order to solve that ambiguity, they propose the *surface integrability*. Yuille et al. [7] propose and extension of [6] in which the SVD is applied to obtain the shape and albedo of objects under different unknown illuminations. They also propose a method to locate and reject shadows. It consists of an iterative process whose initialization is given by the SVD.

Basri et al. [1] also use the photometric stereo to recover the shape and reflectance properties of an object. The main advantage of this recent approach is that it allows arbitrary

This work has been supported by the Government of Spain under MEC project TRA2007-62526/AUT and research programme Consolider Ingenio 2010: MIPRCV (CSD2007-00018). The second author has been supported by The Ramón y Cajal Program.

lightings, including any combination of point sources and diffuse lightings. They use spherical harmonics [8], which form an orthonormal basis for describing functions on the surface of a sphere. Basri et al. propose to remove unreliable pixels, such as saturated pixels, and to fill in the missing data by using Wiberg's algorithm [9].

In this paper, the Alternation technique [10] is proposed to decompose the *measurement matrix*, which contains the intensity images, into the surface and light source matrices. The novelty of our proposal is that shadows and saturated regions are considered as missing data. Thus, the results do not depend on those pixels, which in general influence the results. This paper is structured as follows. First of all, the formulation used along the paper is introduced. Then, the Alternation technique adapted to the photometric stereo is presented. Experimental results with both synthetic and real images are given. Finally, concluding remarks are summarized.

2. FORMULATION

A measurement matrix I contains the grey-level intensity image data at p pixels through f frames in which only the light source is moving. In particular, the kth-row of I corresponds to the intensities of the kth-pixel in every image. At the same time, the jth-column of I corresponds to the intensities of all the pixels of the jth-frame. Hence, the matrix I is defined as:

$$I_{p\times f} = \begin{bmatrix} i_{11} & \dots & i_{1f} \\ \vdots & & \vdots \\ i_{p1} & \dots & i_{pf} \end{bmatrix} \qquad (1)$$

The space of images of the object obtained by varying the light source direction spans a three dimensional space [2], if there are not shadows or saturated regions. Therefore, it can be assumed that the rank of I is 3. Assuming a Lambertian reflectance model, this matrix can be factorized as:

$$I = RNMT \qquad (2)$$

where

$$R_{p\times p} = \begin{bmatrix} r_1 & & 0 \\ & \ddots & \\ 0 & & r_p \end{bmatrix} \qquad (3)$$

is the surface reflectance matrix (r represents the surface reflectance at each pixel),

$$N_{p\times 3} = \begin{bmatrix} \mathbf{n}_1 \dots \mathbf{n}_p \end{bmatrix}^t = \begin{bmatrix} n_{1x} & n_{1y} & n_{1z} \\ \vdots & \vdots & \vdots \\ n_{px} & n_{py} & n_{pz} \end{bmatrix} \qquad (4)$$

is the surface normal matrix (**n** represents the surface normal at each pixel),

$$M_{3\times f} = \begin{bmatrix} \mathbf{m}_1 \dots \mathbf{m}_f \end{bmatrix} = \begin{bmatrix} m_{x1} & \dots & m_{xf} \\ m_{y1} & \dots & m_{yf} \\ m_{z1} & \dots & m_{zf} \end{bmatrix} \qquad (5)$$

is the light-source direction matrix (**m** represents the light-source direction at each frame), and

$$T_{f\times f} = \begin{bmatrix} t_1 & & 0 \\ & \ddots & \\ 0 & & t_f \end{bmatrix} \qquad (6)$$

is the light-source intensity matrix (t represents the light-source intensity at each frame).

Using the above definitions, the surface matrix S and the light-source matrix L are defined as follows:

$$S_{p\times 3} = \begin{bmatrix} \mathbf{s}_1 \dots \mathbf{s}_p \end{bmatrix}^t = \begin{bmatrix} s_{1x} & s_{1y} & s_{1z} \\ \vdots & \vdots & \vdots \\ s_{px} & s_{py} & s_{pz} \end{bmatrix} = RN \qquad (7)$$

$$L_{3\times f} = \begin{bmatrix} \mathbf{l}_1 \dots \mathbf{l}_f \end{bmatrix}^t = \begin{bmatrix} l_{x1} & \dots & l_{xf} \\ l_{y1} & \dots & l_{yf} \\ l_{z1} & \dots & l_{zf} \end{bmatrix} = MT \qquad (8)$$

Hence, the *measurement* matrix can be decomposed as:

$$I = SL \qquad (9)$$

Therefore, the surface matrix S and the light-source matrix L can be recovered from intensity images obtained under varying illumination. Furthermore, synthetic images can be generated by considering arbitrarily light positions and substituting them to the expression (9).

3. ADAPTED ALTERNATION TO THE PHOTOMETRIC STEREO PROBLEM

An adaptation of Alternation [10] is proposed to factorize the matrix I, instead of using SVD. Entries of the matrix corresponding to pixels in shadow and saturated regions (also denoted as specularties) are considered as missing data. The algorithm is summarized below.

Algorithm:

1. Set a lower and an upper threshold to define the shadows and saturated regions, respectively. The lower threshold depends on the intensity values in each set of images, while the upper threshold is, in general, 255.

2. Consider the entries corresponding to shadows and saturated regions as missing data in I.

3. Apply the Alternation technique to I. The algorithm starts with an initial random $p \times 3$ matrix S_0 (analogously with a $3 \times f$ random L_0) and repeats the next two steps until the product $S_k L_k$ converges to I:

   - Compute L[1]: $L_k = (S_{k-1}^t S_{k-1})^{-1}(S_{k-1}^t I)$
   - Compute S[1]: $S_k = I L_k^t (L_k L_k^t)^{-1}$

Solution: S contains the surface normals and reflectance, L contains the light source direction and intensities and the product SL is the best rank-3 approximation to I.

[1]These products are computed only considering known entries in I.

However, as in the SVD case [4], the obtained decomposition is not unique, since any invertible matrix Q with size 3×3 gives the following valid decomposition:

$$I = SL = \hat{S}QQ^{-1}\hat{L} \tag{10}$$

Therefore, at the end of the algorithm, one of the constraints proposed in [4] is used to determinate the matrix Q:

1. The relative value of the surface reflectance is constant or known in at least six pixels. The matrix Q can be computed with the following system of p equations:
$$\hat{s}_k QQ^t \hat{s}_k^t = 1, \quad k = 1, \cdots, p \tag{11}$$
where $\hat{s}_k$ is the kth-vector of $\hat{S}$.

2. The relative value of the light-source intensity is constant or known in at least six frames. Here Q can be obtained by solving the following system:
$$\hat{l}_k^t QQ^t \hat{l}_k = 1, \quad k = 1, \cdots, f \tag{12}$$
where $\hat{l}_k$ is the kth-vector of $\hat{L}$.

If the value of the reflectance or the value of the light intensity is known, it is substituted to the corresponding above equation. Actually, if the value is not known, the reflectance and the light intensity are recovered only up to scale. In our experiments, the second constraint is used and a total of f equations (the number of available images) are considered.

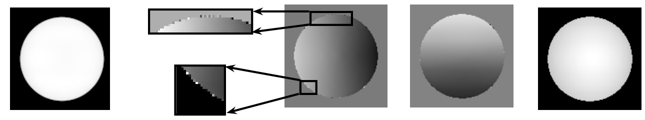
4. EXPERIMENTAL RESULTS

The aim at this Section is to show that results are improved when pixels in shadow and saturated regions are considered as missing entries in I. Hence, results obtained taking the full image intensity matrix are compared with the ones obtained when those particular entries are considered as missing data.

4.1. Synthetic Images

In this first experiment, the *POV-Ray* software is used to generate images that contain a sphere. The light source directions are generated by simulating a trajectory on a sphere and avoiding positions of the light source behind the object. A total of 66 images with a size of 120×120 pixels are generated, given rise to a measurement matrix of 14,400$\times$66 elements. Since only the entries corresponding to non-background pixels are considered, the final resulting matrix contains 7,499$\times$66 elements.

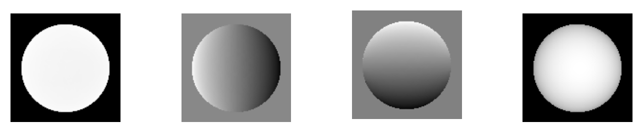
Fig. 1 shows the recovered factors in the case of non-missing data. It can be seen in Fig. 1 (a) that the reflectance is constant in all the points of the surface. Fig. 1 (b), Fig. 1 (c) and Fig. 1 (d) show each of the coordinates of the recovered surface normals. Notice that there are points on the boundaries of the visible surface in which the normal values are not properly recovered (see enlargements on Fig. 1 (b)). This is due to the fact that these points correspond to pixels in shadow in several images.



(a) R (b) N_x (c) N_y (d) N_z

Fig. 1. Synthetic images; (a) reflectance; (b), (c) and (d) coordinates of the surface normals.

In order to correct the wrongly recovered values in the normals, shadows are considered as missing data. Hence, they are not used for computing the factors in the third step of the adapted-Alternation algorithm (Section 3). Concretely, pixels whose intensity is lower than 120 are considered as missing data. With such a high threshold, a percentage of missing data of 57% is obtained. Fig. 2 shows the results corresponding to this missing data case. Here, the normal components do not present incorrect values in the boundaries.



(a) R (b) N_x (c) N_y (d) N_z

Fig. 2. Synthetic images, 57% of missing data; (a) reflectance; (b), (c) and (d) coordinates of the surface normals.

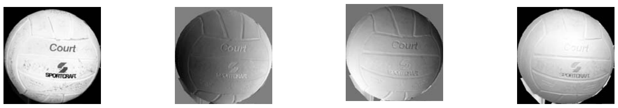
4.2. Real Images

Images from the Yale data base (http://cvc.yale.edu) are used. In particular, a scene containing a ball is presented here. Images are captured using a purpose-built illumination rig, which is fitted with 64 computer controlled strobes. The extreme cases, in which almost all the pixels of the image are in shadow, are not considered in this experiment, only 49 images are taken to generate the *measurement matrix*.

The images contain many regions of specular reflection, that is saturated pixels with an intensity of 255 (see Fig. 4 (top)). The images have a size of 294$\times$294 pixels, which give a measurement matrix of 66,921 rows and 49 columns. If the background pixels were considered, the matrix would have 86,436 rows. Fig. 3 shows the reflectance and the coordinates of the surface normals obtained taking all the intensity values.

Fig. 4 gives a comparison between the initial images (top) and the recovered ones with the product of the obtained factors (bottom). It can be seen that the saturated regions keep quite saturated in the recovered images.

If the saturated pixels (those for which the image intensity is equal to 255) are considered as missing data, the *measurement matrix I* has a percentage of missing data of about 28%.



(a) R (b) N_x (c) N_y (d) N_z

Fig. 3. Real images; (a) reflectance; (b), (c) and (d) coordinates of the surface normals.

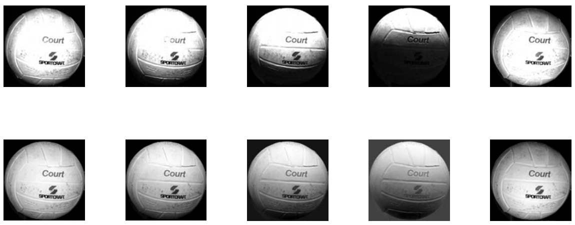


Fig. 4. (top) A set of the original images of the ball; (bottom) images recovered by the product of the recovered factors.

Fig. 5 shows the results obtained in this case. It can be seen that the reflectance (Fig. 5 (a)) is considerably less saturated than the recovered one when considering 100% of data.

(a) R (b) N_x (c) N_y (d) N_z

Fig. 5. Real images, 28% of missing data; (a) reflectance; (b), (c) and (d) coordinates of the surface normals.

Fig. 6 shows some initial images (top) and the recovered ones (bottom) in the case of missing data. Notice that the recovered images are not as saturated as the obtained in the full data case (Fig. 4 (bottom)).

5. CONCLUSION

This paper proposes the use of the Alternation technique, together with the corresponding adaptation, to tackle the photometric stereo problem. The goal is to obtain the normals and reflectance surface of an object from a given set of images obtained under varying illumination. Entries of the measurement matrix that correspond to saturated points or pixels in shadow in the image are considered as missing data. Experimental results with synthetic and real images show the viability of the proposed adapted Alternation approach. Furthermore, results are improved when shadows and specularities are considered as missing data.

Fig. 6. Case 28% of missing data: (top) a set of the original images of the ball; (bottom) images recovered by the product of the recovered factors.

6. REFERENCES

[1] R. Basri, D. Jacobs, and I. Kemelmacher, "Photometric stereo with general, unknown lighting," *International Journal of Computer Vision*, vol. 72, pp. 239–257, 2007.

[2] A. Shashua, "On photometric issues in 3D visual recognition from a single 2D image," *International Journal of Computer Vision*, vol. 21, pp. 99–122, 1997.

[3] L. Zhang, C. Brian, A. Hertzmann, and S.M. Seitz, "Shape and motion under varying illumination: unifying structure from motion, photometric stereo, and multi-view stereo," in *IEEE International Conference on Computer Vision and Pattern Recognition*, 2003.

[4] H. Hayakawa, "Photometric stereo under a light source with arbitrary motion," *Optical Society of America*, vol. 11, pp. 3079–3089, 1994.

[5] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: a factorization method," *International Journal of Computer Vision*, vol. 9, no. 2, pp. 137–154, November 1992.

[6] R. Epstein, A.L. Yuille, and P.N. Belhumeur, "Learning object representations from lighting variations," in *Object recognition workshop, ECCV*, 1996.

[7] A.L. Yuille, D. Snow, R. Epstein, and P.N. Belhumeur, "Determining generative models of objects under vaying illumination: shape and albedo from multiple images using SVD and integrability," *International Journal of Computer Vision*, vol. 35, pp. 203–222, 1999.

[8] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 218–233, 2003.

[9] T. Wiberg, "Computation of principal components when data is missing," in *Second Symposium of Computational Statistics*, 1976.

[10] R. Hartley and F. Schaffalitzky, "Powerfactorization: 3D reconstruction with missing or uncertain data," in *Australian-Japan advanced workshop on Computer Vision*, 2003.