Recovery of Surface Normals and Reflectance from Different Lighting Conditions

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
{cjulia,asappa,felipe,joans,antonio}@cvc.uab.es

Abstract. This paper presents a method for finding the surface normals and reflectance of an object from a set of images obtained under different lighting conditions. This set of images, assuming a Lambertian object, can be approximated by a three dimensional linear subspace, under an orthographic camera model and without shadows and specularities. However, a higher dimensional subspace is needed when images present pixels in shadow, specularities or ambient illumination. This paper proposes on the one hand to consider pixels in shadow and specularities as missing data; and on the other hand a rank-four formulation to recover the ambient illumination. An adaptation of the *Alternation* technique is introduced to compute the sought surface normals and light-source matrices. Experimental results show the good performance of the proposed *Alternation*-based strategy.

Keywords: Photometric stereo, Alternation technique, missing data.

1 Introduction

Photometric stereo aims at estimating the surface normal and reflectance at every point 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 is that the set of images produced by a convex Lambertian object, under arbitrary lighting, can be approximated by a low-dimensional linear subspace of images [1]. Concretely, without shadows and specularities, a Lambertian object produces a 3D subspace of images [2]. 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 image data are stacked into a *measurement matrix*, whose rows and columns correspond to each of the pixels and images, respectively. The Singular Value Decomposition (SVD) [3] is in general used to decompose this matrix into the surface and light-source matrices.

Most photometric stereo approaches assumes that images do not have shadows nor specularities (e.g., [4]), which correspond to points with very low and high

A. Campilho and M. Kamel (Eds.): ICIAR 2008, LNCS 5112, pp. 315–325, 2008.

intensities values, respectively. These points do not follow a Lambertian model. Although if there are only a few of them the Lambertian model could be used as 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 [5] presents a photometric stereo approach for estimating the surface normals and reflectance of objects, which is similar to the factorization method presented in [6] for the shape and motion estimation. This approach factorize the measurement matrix by using the SVD, assuming rank 3. Furthermore, Hayakawa proposes a strategy to deal with shadows. First of all, shadows and illuminated data are classified, by using an intensity threshold. The idea is to select an initial submatrix, whose entries do not correspond to pixels in shadow. Then, the surface normals 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. In addition, the SVD has a high computational cost when dealing with large matrices, which are common in this application. Epstein et al. [7] present an approach based on [5] for learning models of the surface geometry and albedo of objects. It is based on the SVD and also assumes rank 3. They point out that in [5] the obtained reflection and light directions are recovered up to a rotation. In order to solve that ambiguity, they introduce the *surface integrability*.

In real images, the presence of shadows, specularities or ambient illumination is quite common. In those cases, a subspace with a dimension higher than three is needed to approximate properly the measurement matrix [8]. Yuille et al. [9] propose an iterative method to locate and reject shadows. In addition, they propose an extension of [7] to a rank-four formulation that allows to recover the ambient illumination. In a recent paper, Basri et al. [1] proposes an approach that allows arbitrary 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. In particular, they present two methods, the first one uses a first order harmonic approximation (a 4D space), while the second one uses a second order harmonic approximation (a 9D space). They propose to remove unreliable pixels, such that saturated pixels, and fill in missing data by using Wiberg's algorithm [10].

This paper proposes on the one hand, to consider pixels in shadow and specularities as missing data, in order to reduce their influence to the results; and on the other hand, a rank-4 formulation that includes an ambient illumination term. Since the SVD can not be applied to a missing data matrix, an adaptation of the Alternation technique [11], which can deal with missing data, is introduced to factorize the measurement matrix into the surface and light-source matrices. Hence, not only the surface normals and reflectance are recovered, but also the ambient illumination. The rest of the paper is organized as follows. Section 2 introduces the classical rank 3 formulation. The Alternation technique adapted to the photometric stereo is presented in Section 3. Section 4 proposes a generalization to the rank-4 case that allows to recover the ambient illumination.

Experimental results with real images are given in Section 5. Finally, concluding remarks are summarized in Section 6.

2 Rank 3 Formulation

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

$$I_{p\times f} = \begin{bmatrix} i_{11} \dots i_{1f} \\ \vdots \quad\quad \vdots \\ i_{p1} \dots i_{pf} \end{bmatrix} \tag{1}$$

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 specularities. Therefore, it can be assumed that the rank of I is 3. Assuming a Lambertian reflectance model, this matrix can be factorized as:

$$I = RNMT \tag{2}$$

where

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

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

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

is the surface matrix ($\mathbf{n}$ represents the surface normal at each pixel),

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

is the light-source direction matrix ($\mathbf{m}$ represents the light-source direction at each frame), and

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

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:

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

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

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

$$I = SL \tag{9}$$

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

3 Adapted Alternation to Photometric Stereo

The problem presented above could be tackled by any factorization technique. In general, Singular Value Decomposition (SVD) is used to compute the S and L factors from a measurement matrix I; however, if entries of I corresponding to pixels in shadow or saturated regions (also denoted as specularities) are considered as missing data, SVD can not be applied. In this paper, an adaptation of Alternation [11], which is able to deal with missing data, is proposed to factorize the matrix I. The algorithm is summarized below for the rank 3 case, the extension to the rank 4 case is presented in the next Section.

Algorithm

1. Set a lower and an upper threshold to define the shadows and specularities, 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 specularities 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.

However, as in the SVD case [5], the obtained decomposition is not unique, since any 3×3 invertible matrix Q gives the following valid decomposition:

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

Therefore, at the end of the algorithm, one of the constraints proposed in [5] is used to determinate the matrix Q:

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

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:
$$\hat{s}_k Q Q^t \hat{s}_k^t = 1, \quad k = 1, \cdots, p \tag{11}$$
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:
$$\hat{l}_k^t Q Q^t \hat{l}_k = 1, \quad k = 1, \cdots, f \tag{12}$$
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 Generalization to the Rank 4 Case

This Section proposes a generalization of the previously presented formulation to the rank 4 case. It allows to include a term corresponding to the ambient illumination. With this new formulation, the equation (9) is transformed as:
$$I_{p\times f} = S_{p\times 3} L_{3\times f} + \mathbf{a}_{p\times 1} \tag{13}$$
where $\mathbf{a}_{p\times 1}$ is the ambient illumination, which does not depend on the light source direction. It could take a different value at each pixel. In matrix formulation, this equation can be expressed as:
$$I_{p\times f} = \begin{bmatrix} S \; \mathbf{a} \end{bmatrix} \begin{bmatrix} L \\ \mathbf{1} \end{bmatrix} \tag{14}$$
Notice that each of the factors can be of rank 4 at most. Therefore, in this case, the Alternation technique is applied considering a rank 4 value for I (step 3, Section 3) and the following decomposition is obtained:
$$I_{p\times f} = \widetilde{A}_{p\times 4} \widetilde{B}_{4\times f} \tag{15}$$
At each step of the Alternation, the last row of $\widetilde{B}$ is set to be a vector of ones. As in the rank-3 case, any 4×4 matrix Q gives the following valid decomposition:
$$I = \begin{bmatrix} S \; \mathbf{a} \end{bmatrix} \begin{bmatrix} L \\ \mathbf{1} \end{bmatrix} = \tilde{A}_{p\times 4} \tilde{B}_{4\times f} = A_{p\times 4} Q_{4\times 4} Q_{4\times 4}^t B_{4\times f} \tag{16}$$
The linear transformation $Q_{4\times 4}$ can be computed by using one of the aforementioned constraints. In order to compute it more easily, this matrix is separated into two different matrices: Q_1 and Q_2 with dimensions 4×3 and 4×1, respectively. That is,
$$I = AQQ^tB = A \begin{bmatrix} Q_1 \; Q_2 \end{bmatrix} \begin{bmatrix} Q_1^t \\ Q_2^t \end{bmatrix} B = \begin{bmatrix} AQ_1 \; AQ_2 \end{bmatrix} \begin{bmatrix} Q_1^t B \\ Q_2^t B \end{bmatrix} \tag{17}$$

If the surface reflectance is constant or known in every pixel, the matrix Q is computed by solving the linear systems defined by the equations:

$$a_k Q_1 Q_1^t a_k^t = 1, k = 1, \cdots, p \tag{18}$$

where a_k is the kth-vector of the first factor A, and

$$\mathbf{a} = AQ_2 \tag{19}$$

On the contrary, if the intensity of the light source is constant or known in every image, the matrix Q is computed by solving the linear systems defined by:

$$b_k^t Q_1 Q_1^t b_k = 1, k = 1, \cdots, p \tag{20}$$

where b_k is the kth-vector of the second factor B, and

$$\mathbf{1} = Q_2^t B \tag{21}$$

Finally, the Q matrix is used to obtain the final factors:

$$\tilde{A} = AQ, \quad \tilde{B} = Q^t B \tag{22}$$

5 Experimental Results

Images from the Yale data base (http://cvc.yale.edu) are used to validate the proposed approach. These images were captured using a purpose-built illumination rig, which is fitted with 64 computer controlled strobes. Extreme cases, in which almost all pixels of the image are in shadow, are not considered in these experiments. In particular, two different data sets are presented here; a scene containing: i) a ball; ii) a sculptured bust.

The objective of this Section is to show the improvements of the obtained results when pixels in shadow and specularities are considered as missing entries in I. Hence, results obtained taking the full image matrix I are compared with the ones obtained when those particular entries are considered as missing data.

5.1 Ball Images

These images contain regions of specular reflection, that is pixels with an intensity image of 255 (see Fig. 2 (top)). They have a size of 294×294 pixels and

Fig. 1. Ball images, recovered factors: (a) reflectance; (b) ambient illumination; (c), (d) and (e) x, y and z components of the surface normals

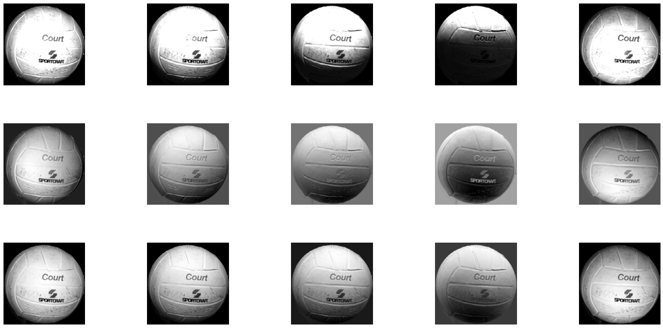

Fig. 2. (top) A set of the original images of the ball; (middle) images recovered by projecting the original ones onto a four-dimensional subspace; (bottom) images of the middle, adding the ambient component

a b c d e

Fig. 3. Ball images, 28% of missing data, recovered factors: (a) reflectance; (b) ambient illumination; (c), (d) and (e) x, y and z components of the surface normals

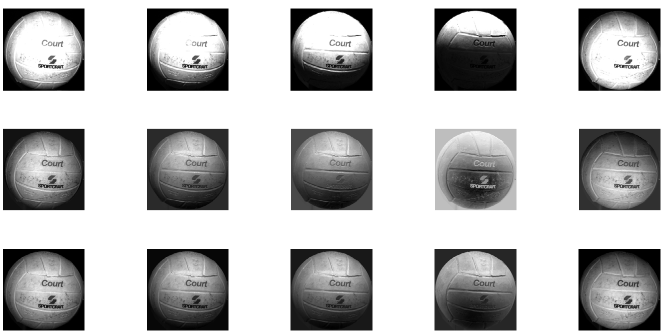

Fig. 4. Case 28% of missing data: (top) a set of the original images of the ball; (middle) images recovered by projecting the original ones onto a four-dimensional subspace; (bottom) images of the middle, adding the ambient component

only 49 images are considered, given a measurement matrix I with a size of 66,921×49 (background pixels are not considered). Fig. 1 shows the reflectance, ambient illumination and coordinates of the recovered surface normals, obtained taking the full I. Fig. 2 gives a comparison between the original images (top) and the recovered ones with the product of the obtained factors, without the ambient term (middle) and with the ambient term (bottom). It can be seen

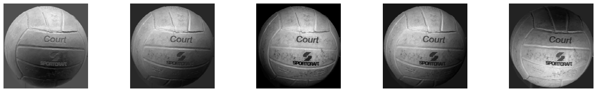

Fig. 5. Images synthesized by considering random light source positions

a b c d e

Fig. 6. Sculptured bust images, recovered factors: (a) reflectance; (b) ambient illumination; (c), (d) and (e) x, y and z components of the surface normals

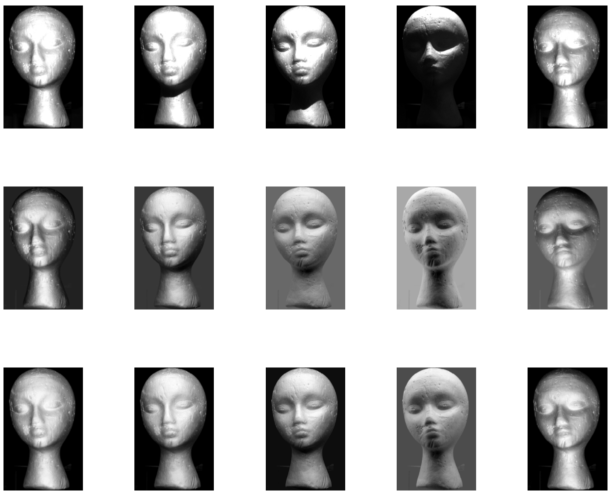

Fig. 7. (top) A set of the original images of the sculpture; (middle) images recovered by projecting the original ones onto a four-dimensional subspace; (bottom) images of the middle, adding the ambient component

that specular regions in the original images keep quite specular in the recovered ones.

The measurement matrix I has a percentage of missing data of 28% when specular pixels are considered as missing data. These missing data are not used for computing the factors in the third step of the adapted-Alternation algorithm (Section 3). Fig. 3 shows the results obtained in this case. It can be seen that the reflectance and ambient are considerably less specular than the recovered ones in the full data case (Fig. 1). Fig. 4 shows some original images (top) and the recovered ones in the case of 28% of missing data, without the ambient term (middle) and adding the ambient term (bottom). Notice that the recovered images are not as specular as the obtained in the full data case (Fig. 2(bottom)).

Just as an illustration, Fig. 5 shows five synthetic images obtained by taking random positions of the light source. The surface matrix (S) obtained in the previous case (28% of missing data) is multiplied by each of the new light positions. The ambient term is added, in order to give more realistic images. This can be very useful, for instance in object recognition or industrial inspection. Any light position can be considered, providing thus a wide range of different images.

5.2 Sculptured Bust Images

The images of this second real data set also contains saturated pixels (see Fig. 7 (top)). They have a size of 404×260 pixels and the obtained measurement matrix has a size of 65,246×49 (background pixels are not considered). The reflectance, ambient and surface normals obtained in this case are shown in Fig. 6.

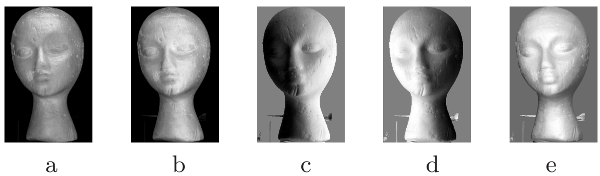

Fig. 8. Sculptured bust images, 16% of missing data, recovered factors: (a) reflectance; (b) ambient illumination; (c), (d) and (e) x, y and z components of the surface normals.

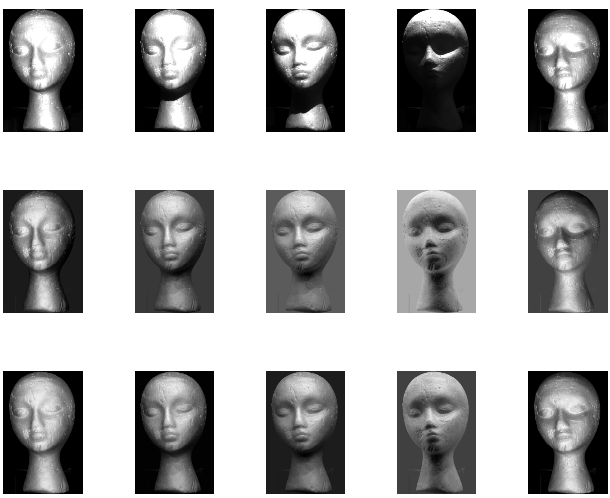

Fig. 9. Case 16% of missing data: (top) a set of the original images of the sculptured bust; (middle) images recovered by projecting the original ones onto a four-dimensional subspace; (bottom) images of the middle, adding the ambient component

Fig. 7 shows some original (top) and recovered images, without the ambient term (middle) and adding the ambient term (bottom). The original specular regions are a little specular in the recovered images (see Fig. 7 (bottom)).

The measurement matrix has a percentage of missing data of 16% when saturated points are considered as missing data. Fig. 8 shows the obtained results in this case. The ambient and reflectance are less saturated than in the full data case (Fig. 6). Finally, some original and recovered images are shown in Fig. 9. It can be seen that the recovered images ((middle) and (bottom)) are considerably less saturated than in previous case (see Fig. 7 (bottom)).

6 Conclusion

This paper presents a method to recover the surface normals and reflectance of an object from a set of images obtained under different lighting conditions. Intensity image data are stacked into a measurement matrix, which can be approximated by a rank-3 matrix, assuming a Lambertian object and an orthographic camera model. This paper proposes to consider pixels in shadow and specular regions as missing data. In addition, a rank-4 formulation that allows to recover the ambient illumination is introduced. An adaptation of the Alternation technique is used to factorize the measurement matrix into the surface and light-source matrices. Experimental results with real images show the viability of the proposed adapted Alternation approach. Furthermore, results are improved when specularities are considered as missing data. Analogous results were obtained when shadows were considered as missing data; due to the lack of space they are not presented here.

Acknowledgments. 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.

References

1. Basri, R., Jacobs, D., Kemelmacher, I.: Photometric stereo with general, unknown lighting. International Journal of Computer Vision 72, 239–257 (2007)
2. Shashua, A.: On photometric issues in 3D visual recognition from a single 2D image. International Journal of Computer Vision 21, 99–122 (1997)
3. Golub, G., Van Loan, C. (eds.): Matrix Computations. The Johns Hopkins Univ. Press (1989)
4. Zhang, L., Brian, C., Hertzmann, A., Seitz, S.: Shape and motion under varying illumination: unifying structure from motion, photometric stereo, and multiview stereo. In: IEEE CVPR, pp. 618–625 (2003)
5. Hayakawa, H.: Photometric stereo under a light source with arbitrary motion. Optical Society of America 11, 3079–3089 (1994)
6. Tomasi, C., Kanade, T.: Shape and motion from image streams under orthography: a factorization method. International Journal of Computer Vision 9(2), 137–154 (1992)

7. Epstein, R., Yuille, A., Belhumeur, P.: Learning object representations from lighting variations. In: Object recognition workshop, ECCV (1996)
8. Basri, R., Jacobs, D.: Lambertian reflectance and linear subspaces. IEEE Transactions on Pattern Analysis and Machine Intelligence 25, 218–233 (2003)
9. Yuille, A., Snow, D., Epstein, R., Belhumeur, P.: Determining generative models of objects under vaying illumination: shape and albedo from multiple images using SVD and integrability. Int. Journal of Computer Vision 35, 203–222 (1999)
10. Wiberg, T.: Computation of principal components when data is missing. In: Second Symposium of Computational Statistics, pp. 229–326 (1976)
11. Hartley, R., Schaffalitzky, F.: Powerfactorization: 3D reconstruction with missing or uncertain data. In: Australian-Japan advanced workshop on Computer Vision (2003)