

# A Factorization-Based Approach to Photometric Stereo

Carme Julià,<sup>1</sup> Felipe Lumbreras,<sup>2</sup> Angel D. Sappa<sup>3</sup>

<sup>1</sup> Intelligent Robotics and Computer Vision Group, Department of Computer Science and Mathematics, Universitat Rovira i Virgili, Spain

<sup>2</sup> Computer Science Department, Computer Vision Center (Universitat Autònoma de Barcelona)

<sup>3</sup> Computer Vision Center (Universitat Autònoma de Barcelona)

Received 1 April 2009; accepted 25 October 2010

**ABSTRACT:** This article presents an adaptation of a factorization technique to tackle the photometric stereo problem. That is to recover the surface normals and reflectance of an object from a set of images obtained under different lighting conditions. The main contribution of the proposed approach is to consider pixels in shadow and saturated regions as missing data, in order to reduce their influence to the result. Concretely, an adapted Alternation technique is used to deal with missing data. Experimental results considering both synthetic and real images show the viability of the proposed factorization-based strategy. © 2011 Wiley Periodicals, Inc. *Int J Imaging Syst Technol*, 21, 115–119, 2011; Published online in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/ima.20273

**Key words:** factorization technique; photometric stereo; missing data

## I. INTRODUCTION

Photometric stereo (Woodham, 1980) 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 of the scene from the same angle and with the same intensity.

The starting point of the photometric stereo problem is that the set of images produced by a Lambertian object, under arbitrary lighting, can be approximated by a low-dimensional linear subspace of images (Basri et al., 2007). Concretely, a Lambertian object produces a 3D subspace of images (Shashua, 1997). This linear property suggests to use factorization techniques to model the image formation and to recover each of the factors that contribute to it. Images are reshaped to columns, and the grey-level intensity image at each pixel is stacked into each row, generating thus a measurement matrix. The singular value decomposition (SVD) (Golub and Van Loan, 1989) is in general used to factorize the matrix and obtain thus a low-rank matrix approximation of this measurement matrix. Most photometric stereo

approaches assume that images do not have shadows nor saturated regions (e.g., Zhang et al., 2003), which correspond to points with very low and high intensity values, respectively. This assumption is enforced since 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 (1994) presents a photometric stereo approach for estimating the surface normals and reflectance of objects, which is similar to the factorization method presented in (Tomasi and Kanade, 1992) for the shape and motion estimation. This approach factorizes the measurement matrix with the SVD. Furthermore, Hayakawa proposes to classify shadows and illuminated data, 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, this 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., (1996) present an approach based on Hayakawa (1994) for learning models of the surface geometry and albedo of objects. It is also based on the SVD. They point out that in (Hayakawa, 1994) the linear ambiguity obtained by using the SVD is not always properly solved. To amend that ambiguity, they introduce the integrability constraints, which are used to ensure that the set of recovered surface normals forms a consistent surface. Yuille et al. (1999) propose another SVD-based method to recover the surface normals and reflectance of an object. In addition, they present an iterative method to reject shadows.

In a recent paper, Shim et al. (2008) construct a subspace model of the reflectance functions from an existing face database. Their approach is not based on factorization, but on the expectation-maximization (EM) algorithm, which is used to solve the lighting, pose and reflectance functions.

Argyriou and Petrou presents a recursive photometric stereo algorithm that allows to deal with shadows and highlights in (Argyriou and Petrou, 2008). The idea of their algorithm is to identify areas where most of the light directions give unreliable pixels (shadows and highlights) and to apply it only to reliable pixels.

Correspondence to: Carme Julià; e-mail: carme.julia@urv.cat

Contract grant sponsors: This work has been supported by the Government of Spain under projects TRA2007-62526/AUT and DPI2007-66556-C03-03, and research programme Consolider-Ingénio 2010 (CSD2007-00018).

Results show that obtained surfaces are better when shadows and highlights are identified. They assume the light positions known, which is not always a realistic assumption.

This article presents an adaptation of a factorization technique to decompose the matrix that contains the grey-level intensity image data, into the surface and light source matrices. Concretely, the Alternation technique, which has been widely used in the computer vision framework (e.g., Guerreiro and Aguiar, 2003; Hartley and Schaffalitzky, 2003), is used. The novelty of our proposal is that pixels in shadow and saturated regions are considered as missing data. Thus, results are not influenced by them.

The rest of the article is organized as follows. Section II introduces the Lambertian reflectance model and the formulation of the photometric stereo problem. The factorization-based technique adapted to the photometric stereo problem is presented in Section III. Experimental results with synthetic and real images are given in Section IV. Finally, concluding remarks are summarized in Section V.

## II. FORMULATION

The image intensity at the pixel  $(u, v)$  depends on the optical properties of the surface material (albedo), the surface shape, and the spectral distribution of the incident illumination. The reflectance characteristics of a given surface can be represented by a reflectance function  $\phi$  of three unit vectors: surface normal  $n = (n_x, n_y, n_z)^t$ , light-source direction  $m = (m_x, m_y, m_z)^t$ , and viewer direction  $v = (v_x, v_y, v_z)^t$ . Figure 1 shows the geometric reflectance model for image generation in the viewer-oriented coordinate system. Using the reflectance function  $\phi$ , the following equation describes the image-generation process:

$$i = t\phi(n, m, v), \quad (1)$$

where  $t$  contains the light source intensity at each image.

The most used reflectance model is the Lambertian model, which states that materials absorb and reflect light uniformly in all directions. Assuming that the image projection is orthographic and that there is only a distant point light source, the viewer direction and the light source direction can be considered to be constant over the image plane. The Lambertian model is given by the following equation:

$$i(u, v) = t\phi(u, v) = tr(u, v)n(u, v)m, \quad (2)$$

where  $r(u, v)$  is the albedo at the  $(u, v)$  pixel,  $n(u, v)$  is its surface normal, and  $m$  contains the light direction at each image. The albedo at each point on the object  $r(u, v)$  describes the fraction of light reflected at that point.

Let  $I$  be a measurement matrix containing the grey-level intensity image data at  $p$  pixels through  $f$  frames in which only the light source is moving. Assuming a Lambertian reflectance model, this matrix can be factorized as:

$$I = RNMT, \quad (3)$$

where

$$R_{p \times p} = \begin{bmatrix} r_1 & & 0 \\ & \ddots & \\ 0 & & r_p \end{bmatrix} \quad (4)$$

is the surface reflectance matrix that contains the surface reflectance at each pixel  $p$ ,

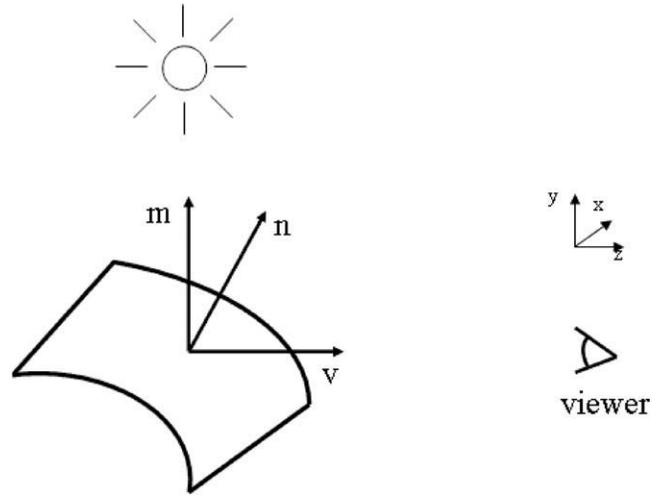


Figure 1. Geometric reflectance system model for image generation in the viewer-oriented coordinate system.

$$N_{p \times 3} = [n_1 \quad n_2 \quad n_3]^t = \begin{bmatrix} n_{1x} & n_{1y} & n_{1z} \\ \vdots & \vdots & \vdots \\ n_{px} & n_{py} & n_{pz} \end{bmatrix} \quad (5)$$

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

$$M_{3 \times f} = [m_1 \quad m_2 \quad m_3] = \begin{bmatrix} m_{x1} & \cdots & m_{xf} \\ m_{y1} & \cdots & m_{yf} \\ m_{z1} & \cdots & m_{zf} \end{bmatrix} \quad (6)$$

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

$$T_{f \times f} = \begin{bmatrix} t_1 & & 0 \\ & \ddots & \\ 0 & & t_f \end{bmatrix} \quad (7)$$

is the light-source intensity matrix that contains the light-source intensity at each frame  $f$ . Using these definitions, the surface matrix  $S$  and the light-source matrix  $L$  are defined as follows:

$$S_{p \times 3} = RN, L_{3 \times f} = MT \quad (8)$$

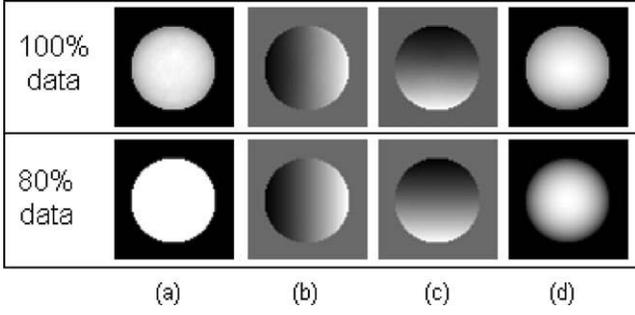
Hence, the measurement matrix can be decomposed as:

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

This decomposition can be obtained by using factorization techniques. In general, the SVD is used to recover the surface matrix  $S$  and the light-source matrix  $L$ , from intensity images obtained under varying illumination.

## III. FACTORIZATION TECHNIQUE ADAPTED TO PHOTOMETRIC STEREO PROBLEM

A common assumption in most photometric stereo approaches is that images do not contain shadows nor saturated regions, which correspond to points with very low and high intensity values, respectively.



**Figure 2.** Synthetic images: (a) recovered reflectance; (b), (c), and (d)  $x$ ,  $y$ , and  $z$  coordinates of the recovered surface normals; (top) full data case; (bottom) missing data case.

This is due to the fact that these points do not follow a Lambertian model. This article proposes to consider those points as missing entries in  $I$ , to reduce their influence to the results. Because the SVD cannot be used with missing data, an adaptation of the Alternation technique (e.g., Guerreiro and Aguiar, 2003; Hartley and Schaffalitzky, 2003) is used. The algorithm is summarized below:

**Algorithm.**

- 1). Set a lower and an upper threshold to define the shadows and saturated regions, respectively. Namely,  $\sigma_l$  and  $\sigma_u$ .
- 2). Define the following set

$$\Omega = \{(i, j) | \sigma_l < I(i, j) < \sigma_u\} \quad (10)$$

Only entries in  $I$  corresponding to pairs  $(i, j) \in \Omega$  are used during matrix factorization. That is, shadows and saturated regions, which correspond to pairs  $(i, j) \notin \Omega$ , are considered as missing data in  $I$ .

- 3). Apply the Alternation technique to  $I$ : the algorithm starts with a  $p \times 3$  random 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  $I$ :
  - (i) compute  $L_k$ ,
  - (ii) compute  $S_k$  from  $L_k$ :  $S_k = IL_k'(L_k L_k')^{-1}$

Remark that these products are computed only considering entries  $I(i, j)$  such that  $(i, j) \in \Omega$ .

**Solution.**  $S$  contains the surface normal 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 (Hayakawa, 1994), the obtained decomposition is not unique, since any invertible matrix  $Q$  with size  $3 \times 3$  gives the following valid decomposition:

$$I = SL = \hat{S}QQ^{-1}\hat{L} \quad (11)$$

Therefore, at the end of the algorithm, one of the two constraints proposed in (Hayakawa, 1994) is used to determinate the matrix  $Q$ :

1. The relative value of the surface reflectance ( $C_1$ ) is constant or known in at least six pixels. The matrix  $Q$  can be computed with the following system of  $p$  equations:

$$\hat{s}_k Q Q' \hat{s}_k' = C_1, k = 1, \dots, p \quad (12)$$

where  $\hat{s}_k$  is the  $k$ th-vector of  $\hat{S}$  and  $C_1$  is the value of the surface reflectance.

2. The relative value of the light-source intensity ( $C_2$ ) is constant or known in at least six frames. Here,  $P = Q^{-1}$  can be obtained by solving the following system:

$$\hat{l}_k P' P \hat{l}_k = C_2, k = 1, \dots, f \quad (13)$$

where  $\hat{l}_k$  is the  $k$ th-vector of  $\hat{L}$ , and  $C_2$  is the value of the light-source intensity.

If the values  $C_1$  or  $C_2$  are not known a priori, they are assumed to be 1. That is, the constraints (12) and (13) impose a constant reflectance at every pixel and a constant light-source intensity at every image, respectively. Therefore, in these situations, the reflectance and light-source intensity can be recovered only up to a constant.

## IV. 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 by taking the full matrix  $I$  are compared with the ones obtained when those particular entries are considered as missing data.

**A. Synthetic Data.** For the experiments with synthetic data, a sphere is generated in Matlab by assuming a center and a radius value. The surface normals at each point of the sphere can be easily computed and, by multiplying them with different light source directions, a matrix of intensities  $I$  is obtained. Different light source directions are obtained by simulating a trajectory and avoiding positions of the light source behind the object. A total of 46 images,  $71 \times 71$  pixels each, are generated, given rise to a measurement matrix of  $5041 \times 46$  elements. Since only the entries corresponding to nonbackground pixels are considered, the final matrix contains  $1941 \times 46$  elements.

Figure 2 (top) shows the recovered reflectance and surface normal at each point, if all data are considered. In a second experiment, points in saturated regions and shadow are considered as missing data and removed from the input matrix ( $I$ ). Saturated regions are defined by pixels whose intensity is higher than  $\sigma_u = 250$ , while shadows correspond to non illuminated pixels (the product  $SL$  is negative or 0, i.e.,  $\sigma_l = 0$ ). With such thresholds, a matrix with 20% of missing data is obtained. The recovered reflectance and surface normal at each point are shown in Figure 2 (bottom).

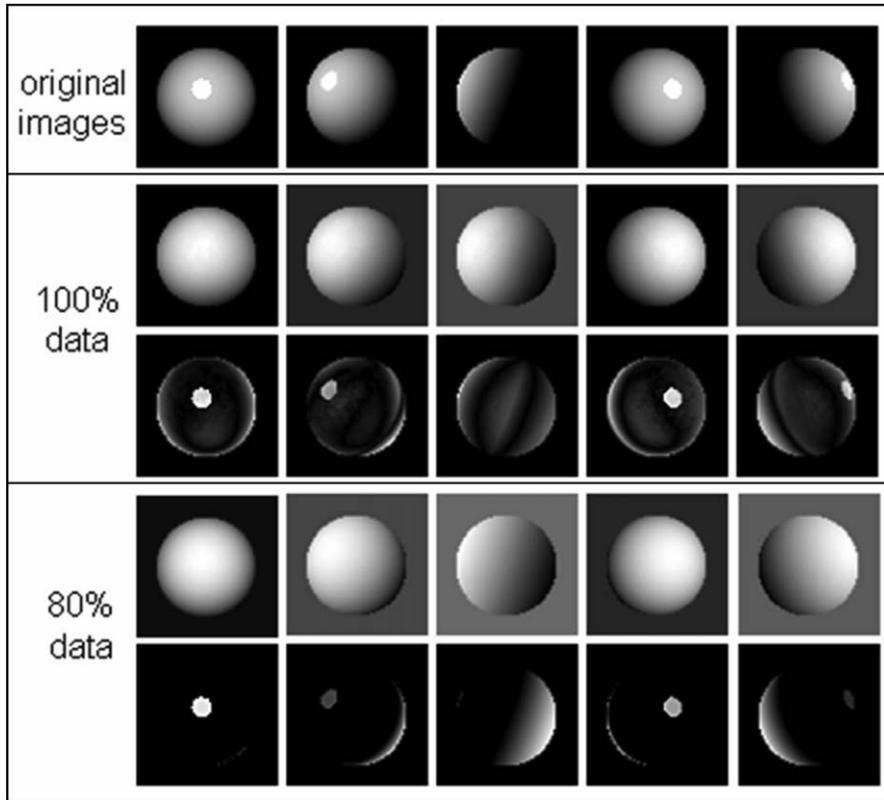
As ground truth of all factors are known in these experiments with synthetic data, the obtained error in each case can be computed (i.e., reflectance ( $e_R$ ), normals ( $e_N$ ), light direction ( $e_M$ ), and intensity ( $e_T$ )). Concretely, the root mean square error (rms) is used as a measure of goodness:

$$e = \frac{\|GT - R\|_F}{\sqrt{n}} \quad (14)$$

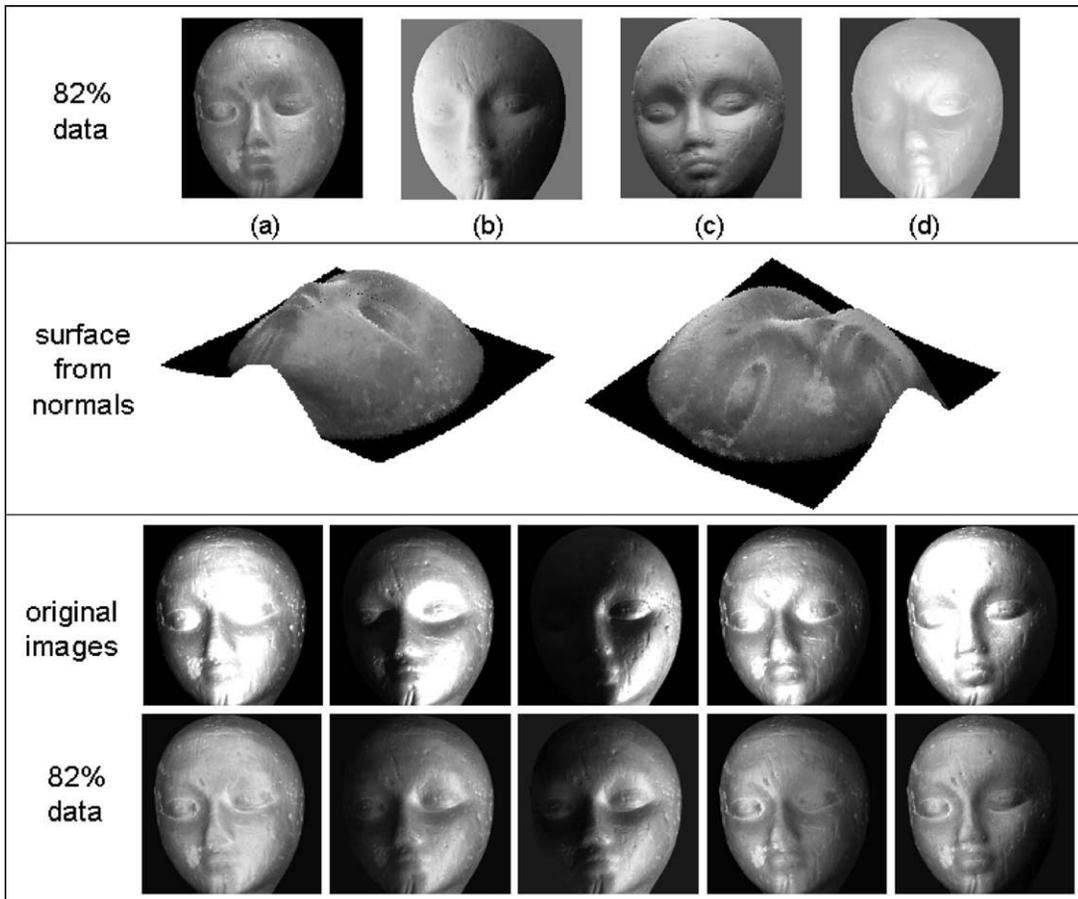
where GT and  $R$  are the ground truth and recovered factors, respectively,  $n$  is the number of elements in  $R$ , and  $\|\cdot\|_F$  is the Frobenius

**Table I.** Synthetic images: obtained errors for all recovered factors in both full and missing data cases.

Known data	$e_R$	$e_N$	$e_M$	$e_T$
100% data	0.079	0.096	0.075	20.677
80% data	7.547e-016	7.493e-016	5.094e-016	3.139e-013



**Figure 3.** (top) A set of original images; (middle) and (bottom) images recovered by the product of the resulting factors and images obtained by subtracting the recovered images from the original ones in the full and missing data case, respectively.



**Figure 4.** Real images: (top) (a) recovered reflectance; (b), (c), and (d) x, y, and z coordinates of the recovered surface normals; (middle) two views of the surface obtained by using the recovered normals; (bottom) a set of original images containing saturated points (top row) and images recovered by the product of the resulting factors (bottom row).

norm (Golub and Van Loan, 1989). As the Alternation depends on the initialization, 100 attempts are carried out for each hypothesis and the mean of the obtained errors is given.

Table I presents the errors obtained in both experiments and for all recovered factors. It can be seen that all factors are better recovered when pixels in shadow, and saturated regions are considered as missing entries and removed from the given input matrix ( $I$ ).

Figure 3 gives a comparison between the original images (Fig. 3 (top)) and those recovered by the product of resulting factors (middle row, top side: full data case; bottom row, top side: missing data case). Images obtained by subtracting the recovered images from the original ones are presented in the bottom side of middle and bottom row. It can be seen how pixels in shadows and saturated ones bias the results (full data case). This problem is solved when these pixels are not considered (missing data case) for computing factors  $L$  and  $S$  by the Alternation technique.

**B. Real Data.** To validate results obtained with synthetic data, a single experiment considering real data is presented. In particular, images from the Yale data base (<http://cvc.yale.edu>) are used in this experiment. Images are captured using an 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 this experiment, only 46 images are used to build the matrix  $I$ .

The images have a size of  $404 \times 260$  pixels each, which give a measurement matrix of  $65,246 \times 46$  elements. In this particular sequence, only saturated pixels ( $\sigma_u = 255$ ) are considered as missing entries.

This is due to the fact there is a large number of points in shadows. Hence, if both shadows and saturated regions were considered as missing entries (and removed from the matrix  $I$ ), some rows of  $I$  would contain only few known entries (less than three in most cases), and the factorization technique could not be applied to those particular points.

The matrix  $I$  contains 18% of missing data when saturated pixels ( $\sigma_u = 255$ ) are considered as missing entries. Figure 4 (top) shows the recovered reflectance and surface normal at each point. The surface of the studied object can be obtained by using the recovered normals (Kovesi, 2005) (Fig. 4 (middle)). Finally, Figure 4 (bottom) gives a comparison between the original images (top row) and those recovered by the product of resulting factors, when saturated pixels are considered as missing data (bottom row). It can be seen that saturated regions are considerably improved in the recovered images.

## V. CONCLUSIONS

This article proposes an adaptation of the Alternation technique to tackle the photometric stereo problem. The goal is to obtain the nor-

mal and reflectance surface of an object from a given set of images obtained under varying illumination.

The main contribution of this article is a technique that avoids biased results by removing entries (i.e., points in shadows and saturated regions) from the given measurement matrix. Experimental results with synthetic and real images show the viability of the proposed approach, as well as improvements on the results.

## REFERENCES

- V. Argyriou and M. Petrou, Recursive photometric stereo when multiple shadows and highlights are present, IEEE Int Conf Comput Vision Pattern Recognit (CVPR), Alaska, USA, 2008.
- R. Basri, D. Jacobs, and I. Kemelmacher, Photometric stereo with general, unknown lighting, Int J Comput Vision 72 (2007), 239–257.
- R. Epstein, A.L. Yuille, and P.N. Belhumeur, Learning object representations from lighting variations, Object Recognition Workshop, ECCV. Cambridge, UK, 1996.
- G.H. Golub and C.F. Van Loan (Editors), Matrix computations. The Johns Hopkins Univ. Press, Baltimore, Maryland, USA, 1989.
- R. Guerreiro and P. Aguiar, Estimation of rank deficient matrices from partial observations: Two-step iterative algorithms, EMMCVPR, Lisbon, Portugal, 2003, pp. 450–466.
- R. Hartley and F. Schaffalitzky, Powerfactorization: 3D reconstruction with missing or uncertain data, Australian-Japan Adv Workshop Comput Vision, Adelaide, Australia, 2003.
- H. Hayakawa, Photometric stereo under a light source with arbitrary motion, Opt Soc Am 11 (1994), 3079–3089.
- P. Kovesi, Shapelets correlated with surface normals produce surfaces, IEEE Int Conf Comput Vision, Beijing, China, 2005, pp. 994–1001.
- A. Shashua, On photometric issues in 3D visual recognition from a single 2D image, Int J Comput Vision 21 (1997), 99–122.
- H. Shim, J. Luo, and T. Chen, A subspace model-based approach to face relighting under unknown lighting and poses, IEEE Trans Image Process 17 (2008), 1331–1341.
- C. Tomasi and T. Kanade, Shape and motion from image streams under orthography: A factorization method, Int J Comput Vision 9 (1992), 137–154.
- J. Woodham, Photometric method for determining surface orientation from multiple images, Opt Eng 19 (1980), 139–144.
- A.L. Yuille, D. Snow, R. Epstein, and P.N. Belhumeur, Determining generative models of objects under varying illumination: Shape and albedo from multiple images using SVD and integrability, Int J Comput Vision 35 (1999), 203–222.
- 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, IEEE Int Conf Comput Vision Pattern Recognit (CVPR), Madison, USA, 2003.