

Photometric stereo under unknown light sources using robust SVD with missing data
Daisuke Miyazaki, Katsushi Ikeuchi,
in Proceedings of International Conference on Image Processing,
ICIP 2010, 2010.09

This file is a draft version. It may be different from the published version.

PHOTOMETRIC STEREO UNDER UNKNOWN LIGHT SOURCES USING ROBUST SVD WITH MISSING DATA

Daisuke Miyazaki and Katsushi Ikeuchi

Institute of Industrial Science, The University of Tokyo, Japan.

ABSTRACT

In this paper, we propose a novel photometric stereo method that uses singular value decomposition. Singular value decomposition can solve the photometric stereo problem when the light source direction is unknown; however, it has the critical problem of being sensitive to outliers. We therefore propose a novel singular value decomposition method that is robust to outliers. We also show some results of our photometric stereo method when applied to objects that involve not only diffuse reflection but also specular reflection.

1. INTRODUCTION

In this paper, we propose a photometric stereo method that is effective when the light source direction is unknown. Our method can be applied to objects that have both diffuse reflection and specular reflection. Our method is based on a novel singular value decomposition (SVD) method that is robust to outliers. In addition, we use graph cut for detecting shadow and specular reflection.

The uncalibrated photometric stereo estimates the surface normal of the object from a large number of images but without knowing the direction of light sources. Based on the theory given by Woodham et al. [14], Hayakawa [7] proposed the uncalibrated photometric stereo using singular value decomposition (SVD). Later, many methods [1–3, 6, 8, 10, 12, 15] have been proposed which are the extension to Hayakawa’s method. Generally, uncalibrated photometric stereo has a problem that it is sensitive to outliers such as specular reflection and shadow. In this paper, we propose a robust SVD method in order to overcome this problem. Mukaigawa et al. [9] proposed the RANSAC-based method, while our method, which is SVD-based, can be applied to many application fields that use singular value (*or* eigen value). The paper presented by Tomasi and Kanade [13] explains how to calculate the motion and shape of the object when some part of the data is missing, and our method is a generalized version of their method which can be applied not only for the structure-from-motion problem but also for the uncalibrated photometric stereo problem. PCAMD/SVDMD (PCA/SVD

with missing data) introduced by Shum et al. [11] also deals with the problem when the partial data are missing; however, the low-rank constraint is not incorporated in the PCAMD itself. The incremental SVD proposed by Brand [5] deals with both the low-rank constraint and the missing data, which are also dealt with in our method. Brand’s method applies the SVD incrementally for each iteration, while our method applies the SVD to whole data for each iteration in order to avoid the cumulative error intrinsic to an incremental approach.

2. FACTORIZATION USING SVD

If the number of the image is f and the number of the pixel is p , all input data can be represented by the following matrix:

$$I = SL = \begin{pmatrix} i_{11} & \cdots & i_{1f} \\ \vdots & \ddots & \vdots \\ i_{p1} & \cdots & i_{pf} \end{pmatrix}, \quad (1)$$

$$S = \begin{pmatrix} s_{1x} & s_{1y} & s_{1z} \\ \vdots & \vdots & \vdots \\ s_{px} & s_{py} & s_{pz} \end{pmatrix}, L = \begin{pmatrix} l_{x1} & \cdots & l_{xf} \\ l_{y1} & \cdots & l_{yf} \\ l_{z1} & \cdots & l_{zf} \end{pmatrix}. \quad (2)$$

We call I the image matrix, and its rank is 3. Here, we call $\mathbf{s} = (s_x, s_y, s_z)^\top$ the surface vector, and we call $\mathbf{l} = (l_x, l_y, l_z)^\top$ the light vector. The surface vector is the product of the unit surface normal vector and the albedo. The light vector is the product of the unit vector representing the light source direction and the light source intensity.

The image matrix can be decomposed by SVD as follows:

$$I = U W V^\top, \quad (3)$$

where the size of U , W , and V^\top are $p \times 3$, 3×3 , and $3 \times f$, respectively. The surface matrix S and light matrix L can be estimated as follows:

$$S = S' A, \quad L = A^{-1} L', \quad (4)$$

$$S' = U W^{1/2}, \quad L' = W^{1/2} V^\top, \quad (5)$$

where, S' is called pseudo surface matrix and L' is called pseudo light matrix. A is a 3×3 invertible matrix, and represents an ambiguity. We will solve this ambiguity using the constant albedo assumption [7, 14] and the occluding boundary constraint [10]. The above solution obliges the surface to obey Lambert’s law, and will be affected by outliers such as

D. Miyazaki is currently with Graduate School of Information Sciences, Hiroshima City University, Japan.

shadows and specular reflections. We propose a novel SVD method in Section 3, which is robust to outliers. We utilize the graph cut method as described in Section 4 in order to detect shadows and specular reflections. Our algorithm calculates this SVD method and this graph cut method alternatively for each iteration until convergence.

3. HOLE-FILLING SVD

In this section, we propose a robust SVD method, which can deal with the outliers included in the input data. The key idea of the proposed method, which we call ‘‘hole-filling SVD,’’ is to (1) Scrape the vector containing the outlier from the matrix in order to obtain the sub-matrix not containing outlier (2) Apply the SVD to the sub-matrix (3) Clean up the contaminated vector from the SVD result (4) Stick the refreshed vector to the sub-matrix.

First, we explain the fine-to-coarse and coarse-to-fine approach. We denote the input matrix that we want to apply the SVD to as X . First, we choose the column where the number of outliers is maximum in the input matrix, and remove it. This process, which we call ‘‘column scraping,’’ can be expressed as follows without loss of generality.

$$X_{k-1} = [X_k \mid \mathbf{x}_k], \quad (6)$$

where X_{k-1} is the $M \times N$ matrix, X_k is the $M \times (N-1)$ matrix, and \mathbf{x}_k is the $M \times 1$ vector. If we apply SVD to X_{k-1} , the result will be bad since the column vector \mathbf{x}_k in X_{k-1} contains many outliers; however, if we apply SVD to X_k , the result will be much better. Starting from the input matrix X_0 , we make the matrix smaller and smaller by scraping the noisy columns and rows. Finally, we obtain a small matrix with only a small amount of noise. If we apply SVD to such a matrix, we will obtain a good result. Using the SVD result, we want to clean up the contaminated vector in order to suppress the interference of the outlier, and stick it to the matrix. We use the term ‘‘column sticking’’ for this process. Starting from the SVD result of such small matrix, we make the matrix bigger and bigger by sticking the revised columns and rows. Consequently, we obtain a better SVD result of whole input matrix X . Row scraping and sticking can also be similarly defined. How to clean up the contaminated vector before sticking it to the matrix is explained as follows.

The key process of hole-filling SVD is to clean up the vector \mathbf{x} . The application of SVD can be represented as follows:

$$\begin{bmatrix} X \\ \mathbf{x}^\top \end{bmatrix} = \begin{bmatrix} U \\ \mathbf{u}^\top \end{bmatrix} W V^\top, \quad (7)$$

where X and U are $M \times N$ matrices, and W and V^\top are $N \times N$ matrices. W is the diagonal matrix where the singular value is in descending order. We will only explain the process for row scraping and sticking; however, the process for column scraping and sticking is also the same. Our algorithm assumes that the input matrix in an ideal case would be rank deficient. Suppose that there are g singular values that are

non-zero in an ideal case (e.g., $g = 3$ in photometric stereo (Eq. (3))). In this case, we can express the bottom part of Eq. (7) as follows.

$$\mathbf{x}^\top = \hat{\mathbf{u}}^\top \hat{V}^\top, \quad (8)$$

where \mathbf{x}^\top is $1 \times N$ vector, $\hat{\mathbf{u}}^\top$ is $1 \times g$ vector, and \hat{V}^\top is $g \times N$ matrix. Here, singular values are included in \hat{V}^\top . Since the input vector \mathbf{x}^\top contains outliers, we do not solve Eq. (8). First, we define the vector $\hat{\mathbf{x}}^\top$ and the matrix \hat{V}^\top as follows.

$$(\hat{\mathbf{x}}^\top)_j = \begin{cases} (\mathbf{x}^\top)_j & \text{(if } (\mathbf{x}^\top)_j \text{ is not outlier)} \\ 0 & \text{(if } (\mathbf{x}^\top)_j \text{ is outlier)} \end{cases}, \quad (9)$$

$$(\hat{V}^\top)_{ij} = \begin{cases} (\hat{V}^\top)_{ij} & \text{(if } (\mathbf{x}^\top)_j \text{ is not outlier)} \\ 0 & \text{(if } (\mathbf{x}^\top)_j \text{ is outlier)} \end{cases}. \quad (10)$$

We assume that we can specify which elements in the input matrix are outliers. The zero value is filled in for the elements which are specified as outliers. We obtain a much more reliable coefficient vector $\hat{\mathbf{u}}^\top$ as follows, since the outliers are not used for estimating it.

$$\hat{\mathbf{u}}^\top = \hat{\mathbf{x}}^\top \hat{V}^{+\top}, \quad (11)$$

where superscript ‘‘+’’ represents the pseudo-inverse matrix calculated by SVD. Now, we can revise the noisy vector \mathbf{x}^\top as follows:

$$(\hat{\mathbf{x}}^\top)_j = \begin{cases} (\mathbf{x}^\top)_j & \text{(if } (\mathbf{x}^\top)_j \text{ is not outlier)} \\ (\hat{\mathbf{u}}^\top \hat{V}^\top)_j & \text{(if } (\mathbf{x}^\top)_j \text{ is outlier)} \end{cases}. \quad (12)$$

The outlier elements are filled in with the ideal value, and we call this procedure ‘‘hole-filling.’’ Here, we only explained the process for row scraping and sticking; however, the process for column scraping and sticking is also the same. We stick this refreshed vector $\hat{\mathbf{x}}_k^\top$ to the matrix X_k and make the bigger matrix $X_{k-1} = \begin{bmatrix} X_k \\ \hat{\mathbf{x}}_k^\top \end{bmatrix}$. We make the matrix bigger and bigger sticking the vector cleaned up by the hole-filling process. Finally, we obtain a better SVD result for the whole input matrix X , namely, X_0 .

4. DETECTING OUTLIERS BY GRAPH CUT

In order to apply the hole-filling SVD (Section 3), we have to detect the outliers. For detecting the shadow and specular pixels, we use the graph cut (min-cut/max-flow) algorithm [4].

We assume that the ambient light is already removed from the input images; thus, the pixel brightness of the shadow is zero. We define the data cost term δ as follows:

$$\delta(1) = t_1, \quad \delta(0) = \begin{cases} 0 & (i_{pf} \geq t_3) \\ t_3 - i_{pf} & (t_2 < i_{pf} < t_3) \\ t_3 - t_2 & (i_{pf} \leq t_2) \\ 1 & \end{cases}. \quad (13)$$

Here, $\delta(0)$ represents the cost for a non-shadowed pixel, and $\delta(1)$ represents the cost for a shadowed pixel. If the pixel brightness i_{pf} is less than a threshold, the cost for a non-shadowed pixel $\delta(0)$ becomes maximum; thus, such a pixel

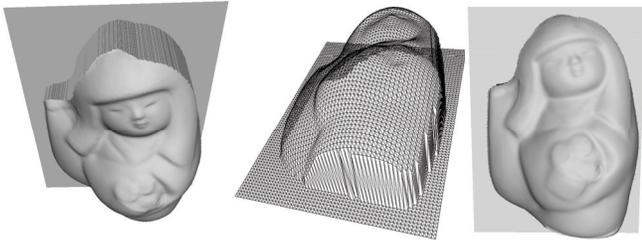


Fig. 1. Estimated shape.

can be considered as a shadowed pixel. However, such simple thresholding is not reliable due to the noise in the image; thus, we also add the smoothness cost term ν as follows:

$$\nu_{p,q}(\beta_p, \beta_q) = t_0 |\beta_p - \beta_q|, \quad (14)$$

where p and q represent the neighboring pixels, and β is a label that takes 0 or 1. In our experiments, we use $t_0 = 0.25, t_1 = 0.5, t_2 = 8, t_3 = 24$ determined empirically, where $i_{p,f}$ varies from 0 to 255. If the image size is 640×480 and the number of images is 100, the number of whole input pixels will be 30 millions; thus, the parameters t_0, t_1, t_2 , and t_3 are less sensitive to the final result since even if several hundreds of diffuse pixels are wrongly detected as shadow pixels or specular pixels, the detection error is quite small considering the 30 millions of input pixels. The graph cut is applied for each input image. The detection of the specular pixel is similar; thus, we will skip explaining it.

5. EXPERIMENTAL RESULT

5.1. Qualitative evaluation

We apply our SVD photometric stereo to a ceramics object which has specularity. We took 100 images with freely moving the light source, and we estimated the shape (Fig. 1).

If we apply the SVD without removing the outliers such as specular reflections and shadows, the whole surface shape will distort as shown in Fig. 2 (b). We detect the shadowed pixels and the specular pixels by graph cut robustly, and we use the hole-filling SVD that is not influenced by such outliers; thus, we obtain the correct shape of the object as is shown in Fig. 2 (a). If we use a simple thresholding to detect shadows and speculars as is done by Hayakawa [7], clumsy defects will appear at the estimated shape as is shown in Fig. 3 (b). Thanks to the fine-to-coarse and coarse-to-fine approach of hole-filling SVD, the estimated shape is smooth enough that the algorithm is insensitive to outliers (Fig. 3 (a)).

5.2. Quantitative evaluation

Fig. 4 (b) shows the results of our method. Fig. 4 (a) is ground truth. Fig. 4 (c) is the result of Hayakawa's method. Fig. 4 (3) represents the height error, and Fig. 4 (4) represents the surface normal error. The error of our method (Fig. 4 (d)) is smaller than the error of Hayakawa's method (Fig. 4 (e)). The

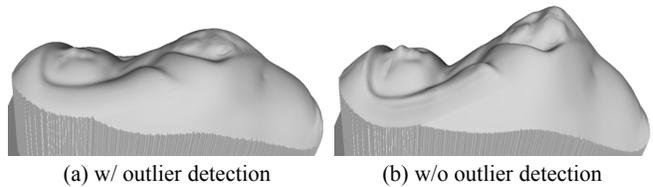


Fig. 2. The distortion caused by shadow and specular reflection: (a) Result of our SVD photometric stereo, which detects shadow and specular pixels robustly, (b) Result if we apply SVD to whole data, which contain shadow and specular reflection.

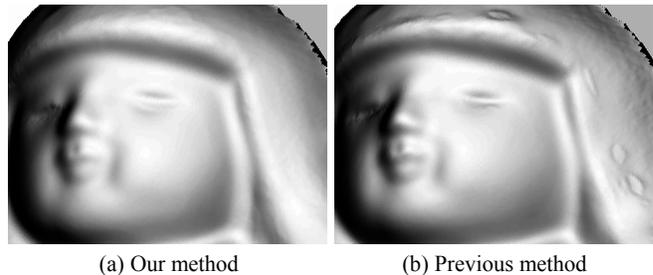


Fig. 3. The fraction caused by shadow and specular reflection: (a) Result of our SVD photometric stereo, which detects shadow and specular pixels with graph cut, (b) Result of Hayakawa's SVD photometric stereo, which detects shadow and specular pixels with simple thresholding.

numerical error is shown in Table 1. The computation time of the hole-filling SVD dominates almost the whole computation time of the system since it uses the SVD method multiple times for each iteration until convergence. Its computation time depends on the input data size and the performance of the computer, and the typical computation time of our experiments in several situations was around more than one hour and less than one day.

6. CONCLUSION

In this paper, we propose a photometric stereo that does not require the light direction to be known. We solve the ambiguity not by the users' operation but by the property of the target object itself. Thanks to our hole-filling SVD, our photometric stereo is not affected by specular reflection and shadow. Our

Table 1. Evaluation. Our method estimates the surface normal with less RMSE (root mean square error) than conventional method. The percentage in the table represents the relative ratio of the error compared to our result.

	Our PS	Hayakawa's PS
Height RMSE	0.16 cm (100%)	0.18 cm (108%)
Normal RMSE	23.8° (100%)	25.9° (109%)

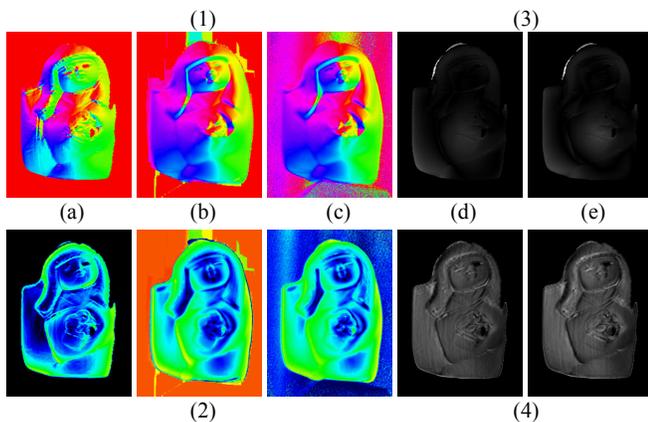


Fig. 4. Comparison. (a) True value, (b) The result of our SVD photometric stereo (PS), (c) The result of Hayakawa's SVD PS, (d) The error of our SVD PS, (e) The error of Hayakawa's SVD PS; (1) Azimuth angle (red: upper direction, blue: left-bottom direction, green: right-bottom direction), (2) Zenith angle (blue: 0° , red: 90°), (3) The difference of the height (the brighter the noisier), (4) The difference of the surface normal (the brighter the noisier).

hole-filling SVD can be applied to any other methods that use SVD.

7. REFERENCES

- [1] R. Basri, D. Jacobs, and I. Kemelmacher, "Photometric stereo with general, unknown lighting," *International Journal of Computer Vision*, vol. 72, no. 3, pp. 239–257, 2007.
- [2] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille, "The bas-relief ambiguity," *International Journal of Computer Vision*, vol. 35, no. 1, pp. 33–44 (1999)
- [3] N. Birkbeck, D. Cobzas, P. Sturm, and M. Jagersand, "Variational shape and reflectance estimation under changing light and viewpoints," in *Proceedings of European Conference on Computer Vision*, pp. 536–549 (2006)
- [4] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1124–1137, 2004.
- [5] M. Brand, "Incremental singular value decomposition of uncertain data with missing values," in *Proc. European Conference on Computer Vision*, pp. 707–720, 2002.
- [6] A. S. Georghiadis, "Incorporating the Torrance and Sparrow model of reflectance in uncalibrated photometric stereo," in *Proceedings of IEEE International Conference on Computer Vision*, pp. 816–825 (2003)
- [7] H. Hayakawa, "Photometric stereo under a light source with arbitrary motion," *Journal of the Optical Society of America A*, vol. 11, no. 11, pp. 3079–3089, 1994.
- [8] B. Kim and P. Burger, "Depth and shape from shading using the photometric stereo method," *CVGIP: Image Understanding*, vol. 54, no. 3, pp. 416–427 (1991)
- [9] Y. Mukaigawa, Y. Ishii, and T. Shakunaga, "Analysis of photometric factors based on photometric linearization," *Journal of the Optical Society of America A*, vol. 24, no. 10, pp. 3326–3334, 2007.
- [10] I. Sato, T. Okabe, Q. Yu, and Y. Sato, "Shape reconstruction based on similarity in radiance changes under varying illumination," in *Proc. International Conference on Computer Vision*, 2007.
- [11] H.-Y. Shum, K. Ikeuchi, and R. Reddy, "Principal component analysis with missing data and its application to polyhedral object modeling," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 17, no. 9, pp. 854–867, 1995.
- [12] D. Simakov, D. Frolova, and R. Basri, "Dense shape reconstruction of a moving object under arbitrary, unknown lighting," in *Proceedings of IEEE International Conference on Computer Vision*, pp. 1202–1209 (2003)
- [13] 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, 1992.
- [14] R. J. Woodham, Y. Iwahori, and R. A. Barman, "Photometric stereo: Lambertian reflectance and light sources with unknown direction and strength," in *Technical Report* (1991)
- [15] 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," *International Journal of Computer Vision*, vol. 35, no. 3, pp. 203–222, 1999.
- [16] L. Zhang, B. Curless, A. Hertzmann, and S. M. Seitz, "Shape and motion under varying illumination: Unifying structure from motion, photometric stereo, and multi-view stereo," in *Proceedings of IEEE International Conference on Computer Vision*, pp. 618–625 (2003)