Generalisation of Photometric Stereo technique to Q-illuminants

Vasileios Argyriou, Svetlana Barsky and Maria Petrou
Imperial College London
v.argyriou@, s.barsky@, maria.petrou@imperial.ac.uk

Abstract

We present a generalisation of the 4-light photometric stereo technique to an arbitrary number Q of illuminants. We assume that the surface reflectance can be approximated by the Lambertian model plus a reflectance component. The algorithm works in a recursive manner eliminating the pixel intensities affected by shadows or highlights based on a least squares error technique, retaining only the information coming from illumination directions that can be used for photometric stereo reconstruction of the normal of the corresponding surface patch. We report results for both simulated and real surfaces and compare them with the results of other state of the art photometric stereo algorithms.

1 Introduction

Photometric stereo (PS) methods can derive orientation and reflectance values for each pixel, using multiple images captured from the same view point under different illumination directions [18]. In order to recover gradient information of a surface patch, photometric stereo assumes that neither highlights nor shadows are present in the images used. In fact under orthographic projection, the only light source direction illuminating a scene which can result in an entirely shadowless image is the one aligned with the viewing direction. Therefore, shadowing is nearly unavoidable in most images. Most reconstruction algorithms either ignore the problematic pixels [3] or explicitly exclude data with shadows and highlights [12, 14].

Since at least three images, corresponding to three different lighting directions, are needed to recover the surface normal, it is a common practice to use more than three illumination directions, to increase the possibility of all surface points being illuminated by at least three sources. Thus, the challenges are first to identify which locations are problematic (i.e. correspond to highlights or shadows), and then to exploit the implied constraints to the fullest extent to solve the PS problem using for each pixel the three most appropriate images.

A common heuristic for dealing with highlights and shadows is to simply compare the intensity values of the pixels with a global threshold in order to decide which values are affected by highlights and shadows and which not, and consequently either discard the information from the corresponding light source or use it for estimating the surface normal and surface albedo. Coleman and Jain [6], and Solomon and Ikenchi, [16] proposed a

BMVC 2008 doi:10.5244/C.22.21

method for determining highlights in the absence of shadows utilising four images of the same surface. Combining all the recovered albedos from all four possible triplets of pixels, the surface patch was regarded as a highlight if the four recovered albedos differed significantly. Barsky and Petrou [2, 3, 4] presented an algorithm for estimating the local surface gradient by using four source colour photometric stereo in the presence of highlights and shadows. Chandraker et al. in [5] proposed an algorithm for performing Lambertian photometric stereo in the presence of shadows. Estimating the per pixel light source visibility, shadow maps were obtained and used to constrain the surface normal integration.

Levine and Bhattacharyya [11] used support vector machines to identify shadow boundaries and, based on this information, shadowed regions, without requiring any prior information regarding the scene. Finlayson et al. [7] proposed a method to recover intrinsic images that are free of shadows based on entropy minimisation. An approach for cast shadow estimation based on the simultaneous surface shape reconstruction using shape-from-shading and class-based surface completion was introduced by Smith and Hancock in [15]. Rushmeier et al. [14] proposed a five light source photometric stereo system where the highest and lowest values in five components are discarded to avoid highlight and shadow. A six light source photometric stereo technique was suggested by Sun et al. [17] that employs a slightly more sophisticated decision criterion so as to discard pixels with doubtful values. Georghiades [8, 9] proposes a shape recovery method for non-Lambertian surfaces. This work includes an approach to excluding shadow and saturated pixels using a more complex model of reflectance (Torrance and Sparrow).

In order to separate shadows and highlights, the use of colour information is suggested by Barsky and Petrou in [2, 3, 4]. The chromaticity of the brightest pixel is compared with the chromaticity of the darker pixels and, if the difference exceeds a certain threshold, the pixel is labelled as a highlight. This method cannot provide reliable classification if the chromaticity of the surface colour is close to the chromaticity of the incident light. In this case, Barsky and Petrou suggested to discard the brightest pixel and reconstruct the normal using only the three darkest pixels; the pixel will be regarded as a highlight if the recovered normal is close to the specular direction of the corresponding imaging configuration. In the rest of the problematic quadruples, the darkest pixel value is discarded as a shadow, and the colour and the normal are recovered using the three brightest pixel values. The main problem of this method is that it cannot be used for more than four illumination sources, since it would not be possible to determine the reliable sources in case they are not the majority. Furthermore, this technique is not able to identify complicated cases such as cases when two of the four light sources produce shadows or highlights, and so it would lead to erroneous reconstruction in such cases. In this paper we shall generalise this method to work with Q lighting directions. Once highlights and shadows have been excluded, the problem we have to solve is that of PS applied to a Lambertian surface using $\tilde{Q}$ lighting directions, where $\tilde{Q} \leq Q$.

2 Photometric Stereo for Lambertian Surfaces

For a Lambertian object illuminated by a light source of parallel rays, the observed image intensity I at each pixel is given by the product of albedo ρ and the cosine of the incidence angle θ_i (the angle between the direction of the incident light and the surface normal) [10].

The above incidence angle can be expressed as the product of two unit vectors, the light direction $\mathbf{L}$ and the surface normal $\mathbf{n}$, $I = \rho\cos(\theta_i) = \rho(\mathbf{L}\cdot\mathbf{n})$.

Let us now consider a Lambertian surface patch with albedo ρ and normal $\mathbf{n}$, illuminated in turn by several fixed and known illumination sources with directions $\mathbf{L}^1$, $\mathbf{L}^2$, ..., $\mathbf{L}^{\tilde{Q}}$. In this case we can express the intensities of the obtained (greyscale) pixels as:

$$I^k = \rho(\mathbf{L}^k\cdot\mathbf{n}), \quad \text{where } \mathrm{k} = 1,2,...,\tilde{\mathrm{Q}}. \tag{1}$$

We stack the pixel intensities to obtain the pixel intensity vector $\mathbf{I} = (I^1, I^2, ..., I^{\tilde{Q}})^T$. Also the illumination vectors are stacked row-wise to form the illumination matrix $[L] = (\mathbf{L}^1, \mathbf{L}^2, ..., \mathbf{L}^{\tilde{Q}})^T$. Equation (1) could then be rewritten in matrix form:

$$\mathbf{I} = \rho[L]\mathbf{n} \tag{2}$$

If there are at least three illumination vectors which are not coplanar, we can calculate ρ and $\mathbf{n}$ using the Least Squares Error technique, which amounts to applying the left inverse of $[L]$:

$$([L]^T[L])^{-1}[L]^T\mathbf{I} = \rho\mathbf{n} \tag{3}$$

Since $\mathbf{n}$ has unit length, we can estimate both the surface normal (as the direction of the obtained vector) and the albedo (as its length). Extra images allow one to recover the surface parameters more robustly.

3 The proposed method for shadow exclusion

If there are no significant image artefacts and no shadows and/or highlights, then all possible Q-tuples of intensities obtained under any given system of illuminats should fall in a 3-dimensional subspace in the Q-dimensional space. This happens because what our imaging process does is to project a 3D vector (vector $\rho\mathbf{n}$) onto a Q-dimensional vector of intensities. Therefore when presented with a new Q-tuple we can decide whether it has any 'defects' or not by checking whether it lies in this 3D subspace. Therefore, we should find a suitable basis, $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{M}_3$ in the Q-space such that 3 basis vectors define the subspace of all permissible Q-tuples, and the rest define the orthogonal subspace of dimensionality $Q-3$. Then to determine how close the given Q-dimensional vector is to the subspace, we should find its projection on the orthogonal subspace.

More formally, the intensity vectors are linear combinations of (Q-dimensional) columns $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{M}_3$ of matrix $[L]$:

$$\mathbf{I} = [L]\mathbf{N} = N_1\mathbf{M}_1 + N_2\mathbf{M}_2 + N_3\mathbf{M}_3 \tag{4}$$

Therefore we need to find vectors $\mathbf{v}_4$,...,$\mathbf{v}_Q$ such that:

$$\mathbf{M}_i^T\mathbf{v}_k = 0, \quad i = 1,2,3, \;\; k = 4,...,Q \tag{5}$$

Let us consider an eigenvector $\mathbf{v}_k$ of matrix $[L][L^T]$:

$$[L][L^T]\mathbf{v}_k = \lambda_k\mathbf{v}_k \tag{6}$$

Let us multiply both sides of the equation from the left with $\mathbf{v}_k^T$:

$$\mathbf{v}_k^T [L][L^T]\mathbf{v}_k = \lambda_k \mathbf{v}_k^T \mathbf{v}_k = \lambda_k \tag{7}$$

Since matrix $[L][L^T]$ has $Q-3$ zero eigenvalues, let us now assume that $\mathbf{v}_k$ has a zero eigenvalue associated with it. This means that the length of vector $[L^T]\mathbf{v}_k$ is zero, which in its turn means that all its three components are zero, and therefore vector $\mathbf{v}_k$ is orthogonal to all three columns of matrix $[L]$. Thus, we can use the eigenvectors $\mathbf{v}_k$ which correspond to zero eigenvalues as the basis orthogonal to the intensity subspace. At the same time the eigenvectors which correspond to non-zero eigenvalues give us an orthonormal basis of the subspace itself. Note that in the case of 4 illuminations the orthogonal subspace is one-dimensional.

Let us now consider a Q-tuple of intensities $\tilde{\mathbf{I}}$ which contains some errors:

$$\tilde{\mathbf{I}} = \mathbf{I} + \mathbf{E} = [L]\mathbf{N} + \mathbf{E} \tag{8}$$

Let us also consider a matrix $[A]: \ (Q-3) \times Q$ which consists of the vectors of the orthogonal basis stacked row-wise. Then let us find the projection of the intensity vector on this orthogonal subspace:

$$[A]\tilde{\mathbf{I}} = [A]\mathbf{I} + [A]\mathbf{E} = [A][L]\mathbf{N} + [A]\mathbf{E} \tag{9}$$

We have already shown that all elements of matrix $[A][L]$ are zero, therefore the first term is zero. What we are left with is the projection of the error vector: $\mathbf{P} \equiv [A]\mathbf{E}$.

Let us now assume that at the observed scene no ambient light is present during image acquisition. This assumption means that pixels in a shadow area have very low grey values (almost 0). On the other hand, highlights resulting from specular reflection will tend to produce very high pixel intensities. Furthermore, when the viewing direction is fixed, a light source can only produce a specular reflection over a relatively limited range of values for the surface normal, at a given specular point on the object surface. Therefore, we may assume that there is at most only one highlight among the Q pixel values corresponding to the same facet, while there is possibility of more than one shadows present among the Q pixel values. Based on the above observations, the proposed algorithm initially ranks the pixel intensities and then temporarily discards the highest one, without knowing yet whether it is a highlight or not. For the remaining $Q-1$ pixel intensities we may assume that all the 'defects' are shadows. The $Q-1$ intensities that remain are ranked in order of magnitude, and then we perform the following algorithm.

1. Consider the K-dimensional intensity space (in the beginning $K = Q-1$). Let us also denote the corresponding intensity vector $\mathbf{I}^K$. Construct matrix $[L^K]: K \times 3$ from the K illuminants.
2. Calculate matrix $[A^K]$ as described above.
3. Find $\mathbf{P}^K$ as product of $[A^K]$ and $\mathbf{I}^K$.
4. Calculate the error measure from $\mathbf{P}^K$ (e.g. its length). If this error is greater than a certain predefined threshold, then the intensity vector contains a defect.
5. Discard the darkest intensity and the corresponding illumination vector. Reform the illumination matrix and the intensity vector as $[L^{K-1}]$ and $\mathbf{I}^{K-1}$.

6. Repeat the process until the K-tuple of intensities falls into an appropriate 3D subspace or until there are only 3 intensities left.

Note that the matrices should be recalculated at every step, because dropping one intensity and illuminant takes away a component in each of the columns in matrix $[L]$, so matrix $[L][L]^T$ will have a completely different form. One can pre-calculate a set of matrices $[A^K]$ (we would need no more than 2^{Q-1} such matrices), and call the appropriate one when needed.

Finally, in order to decide whether the brightest pixel intensity that was discarded initially was really a highlight, we add it to the remaining K shadow-free intensities and then perform again steps 1 to 4 of the above algorithm. If the new error is greater than the predefined threshold the new value that was added is a defect (highlight) and photometric stereo is performed with the $\tilde{Q} = K_{final}$ pixel intensities. Otherwise, if the error is not significant, the brightest pixel is not considered as a highlight and photometric stereo is performed using the $\tilde{Q} = K_{final} + 1$ pixel intensities.

4 Experiments and Results

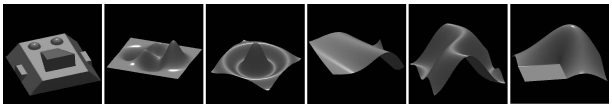

Figure 1: Simulated data used to evaluate the photometric stereo algorithms. All surfaces are of size 128×128.

In case of simulated data, ground truth is known a priori. Therefore, in order to compare the performance of the proposed recursive algorithm with the methods proposed by Coleman and Jain $(C\&J)$ [6], Barsky and Petrou $(B\&P)$[2], Rushmeier et al. (R)[14], Sun et al. (S)[17] and the basic Photometric Stereo technique (PSx where x is the number of lights used), the mean angular error (MAE) of the surface normals is selected as a comparison measure. In particular, the MAE measure suggested by Beauchemin *et al* [1] is used: $\Psi_{AE} = \cos^{-1}(\mathbf{N}_r \cdot \mathbf{N}_e)$ where $\mathbf{N}_r \equiv (x_r, y_r, z_r)^T$ and $\mathbf{N}_e \equiv (x_e, y_e, z_e)^T$ are the real and the estimated surface normals, respectively. To form an estimate for the whole image, the mean value over all pixels is calculated.

The simulated images were produced using the Phong model [13]. This model is mathematically expressed as

$$I = \rho(\mathbf{L} \cdot \mathbf{N}) + k(\mathbf{R} \cdot \mathbf{V})^m \tag{10}$$

where k is the specular reflectance coefficient. $\mathbf{V}$ denotes the viewing direction, $\mathbf{R}$ the perfect reflector vector, and m controls the width of the specular lobe. Therefore, ground truth for shadows could be obtained, by selecting the facets with incident angle greater or equal to $\pi/2$. Regarding the highlights the second term of equation (10) could be used as the ground truth.

For the case of real data there is no ground truth, so we cannot evaluate the performance of the algorithm in the same way. Instead, the reconstructed objects and in some cases the average intensity of the estimated highlights could be used as an indicator of the performance.

Table 1: The mean angular error (MAE) of the surface normals for the simulated surfaces for the methods using 4 (4L), 5 (5L) and 6 (6L) light sources accompanied with the case where the number of lights was chosen to produce the best result, for the methods that this is possible (Best Case). In bold the best result for each surface.

		A	B	C	D	E	F
4L	C & J	1.95737	0.23086	1.45155	**0.82473**	1.40281	0.86270
	PS4	1.94179	0.22989	1.37532	0.84763	1.39952	0.86571
	B & P	1.94177	0.22988	1.37531	0.84760	1.39952	0.86569
	Proposed	**1.93759**	**0.22927**	**1.36486**	0.84731	**1.37819**	**0.86175**
5L	PS5	1.93425	0.23276	1.38181	0.85480	1.40424	0.85522
	R	1.81787	0.23580	1.36705	0.84329	1.34497	0.86239
	Proposed	**1.81787**	**0.22844**	**1.36703**	**0.84328**	**1.34490**	**0.85486**
6L	PS6	1.92466	0.23515	1.38286	0.84399	1.40356	0.85150
	S	1.92348	0.23042	1.37345	0.85025	1.40499	0.84997
	Proposed	**1.90273**	**0.19942**	**1.16508**	**0.84398**	**1.40167**	**0.84825**
Best	C & J	1.95737	0.23086	1.45155	**0.82473**	1.40281	0.86270
Case	PS	1.89896	0.22988	1.35780	0.84398	1.36295	0.85010
	R	1.81787	0.23580	1.36703	0.84328	1.34490	0.86239
	S	1.94177	0.22988	1.37531	0.84760	1.39952	0.86569
	B & P	1.92348	0.23042	1.37345	0.85025	1.40499	0.84997
	Proposed	**1.78636**	**0.19822**	**1.08678**	0.84328	**1.34490**	**0.84064**

Table 2: The mean highlight error (MHE) and the mean shadow error (MShE) of the surface normals for the simulated surfaces. In bold the best result for each surface.

		A	B	C	D	E	F
5L-MHE	R	0.1988	0.1899	0.1864	0.1689	0.1917	0.1901
	Proposed	**0.1063**	**0.0084**	**0.1085**	**0.0881**	**0.1405**	**0.0071**
6L-MHE	S	0.0238	**0.0280**	**0.0024**	0.0373	**0.0207**	**0.0027**
	Proposed	**0.0122**	0.0376	0.0573	**0.0104**	0.0496	0.0068
5L-MShE	R	0.1991	0.1665	0.1462	0.1621	0.1645	**0.1384**
	Proposed	**0.0795**	**0.1048**	**0.0650**	**0.0273**	**0.1216**	0.1703
6L-MShE	S	0.1404	0.1223	0.1877	0.1025	0.2798	0.1877
	Proposed	**0.0110**	**0.0142**	**0.0237**	**0.1013**	**0.0375**	**0.1834**

4.1 Experiments with simulated data

Experiments were performed with simulated data, where six surfaces were artificially created (see figure 1). The surfaces were selected on the basis of containing highlights and shadows, in order to provide a comprehensive evaluation data set. In order to have a more fair comparison, methods with the same number of light sources were grouped together, and additionally the best results of all the methods are presented. In the case of shadows and highlights the comparison was performed only when it was allowable, since some methods are able to estimate either only shadows or only highlights or none of them. The results for the simulated surfaces are shown in Tables 1 and 2. In the case of the mean angular error measure, the proposed method outperforms the other methods resulting in the lowest error. Similar results are obtained for the errors estimating the highlights and shadows. Figure 2 shows the reconstructed surfaces, where it can be observed that the proposed algorithm provides more accurate and smoother results especially at the areas where multiple shadows are present (e.g. simulated surface 'C') indicating that more than six lights may be significantly useful in such cases.

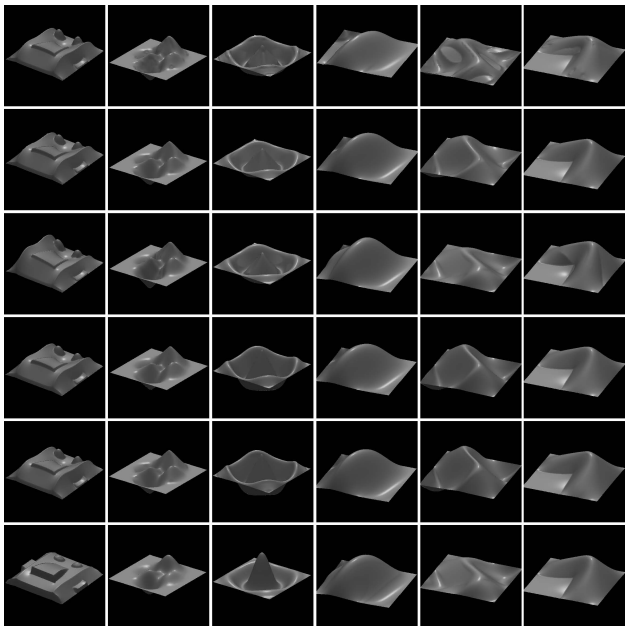

Figure 2: The reconstructed surfaces for all the simulated surfaces using from top to bottom Coleman's and Jain's, Standard PS with 8 lights, Rushmeier's, Sun's, Barsky's and Petrou's and the proposed iterative method.

4.2 Experiments with real data

The proposed algorithm was further applied in the reconstruction of 5 human faces using photometric data captured with maximum eight illumination directions. Figure 3 presents the input images for one of the faces, one for each illumination direction. The lights were

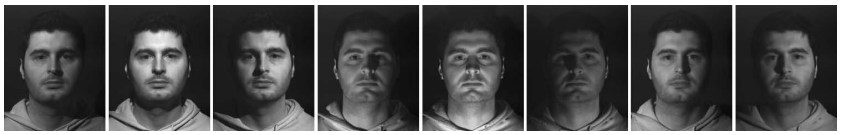

Figure 3: Input images, one for each illumination direction.

Table 3: The mean highlight intensity (MHI) and the mean shadow intensity (MShI) for the real surfaces of all the comparable methods. In bold the best result for each surface.

		Face A	Face B	Face C	Face D	Face E
MHI	R	0.28203	0.32969	0.37047	0.31237	0.35541
	S	0.70501	0.70022	0.72951	**0.71943**	0.76470
	Proposed	**0.71503**	**0.85202**	**0.83626**	0.68986	**0.83668**
MShI	R	0.09697	0.12308	0.12211	0.12139	0.12471
	S	0.24983	0.25706	0.20571	**0.00125**	0.21619
	Proposed	**0.00312**	**0.10675**	**0.08474**	0.09803	**0.00215**

placed on the vertices of a square about $1.5m \times 1.5m$ and the distance between the camera and the observed person was $1.25m$. In such an arrangement the difference of the tilt angles between neighbouring illumination directions is 45^o degrees and the slant angle is equal to 40^o degrees for the corner lights and 31^o degrees for the remaining illumination sources. The person is assumed to be still during the acquisition stage since a high speed camera was used for the acquisition (i.e. 200fps), eliminating the registration problem.

Since ground truth is not available for our real surfaces, the average intensity of the estimated highlights and shadows was used as an indicator of the performance, (see Table 3). The higher values indicate a more accurate approximation of highlights, while lower values for shadows indicate more reliable estimation of shadowed areas. This is because, in general, highlights tend to be bright pixels, while shadowed pixels dark in the absence of ambient light, which is the case for our data. In figure 4 results of the reconstructed faces obtained from the six comparative algorithms are shown. Observing the results it can be inferred that the proposed algorithm outperforms the others especially at the areas shadowed in most of the captured images (e.g. nose, under the jaw) indicating that extra light sources would be useful at these surface patches.

5 Conclusions

In this paper, we propose an iterative algorithm for Q light sources photometric stereo, which is dealing with highlights and shadows allowing one to obtain more reliable estimates of surface parameters. The proposed method iteratively discards the most problematic intensity in a least squares error way until the error becomes smaller than a threshold or the acceptable light sources reduce to three.

Experiments with both artificial and real images were performed in order to evaluate the proposed algorithm. The mean angular error, shadows and highlights ground truth

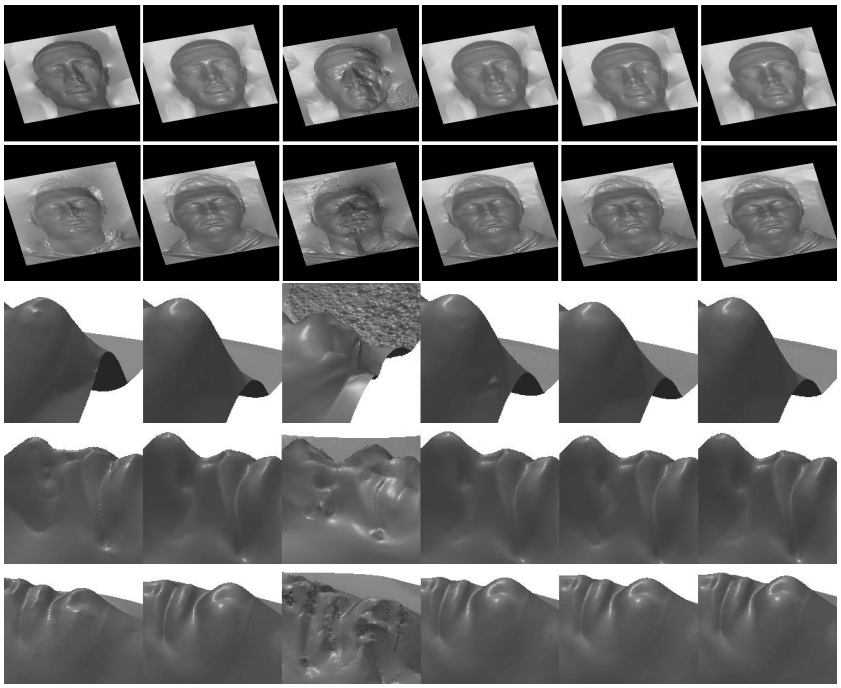

Figure 4: The reconstructed surfaces for two of the real faces using from left to right Coleman's and Jain's, Standard PS with 8 lights, Rushmeier's, Sun's, Barsky's and Petrou's and the proposed iterative method. It could be observed that the inclination below the chin is higher when more lights are used as in the proposed method.

were used in the artificial scenario to compare the methods. From the real images it can observed that the extra light sources could improve the surface normals estimation and consequently the surface reconstruction especially at areas where multiple shadows are present, for example under the chin of a human face.

Acknowledgements: This work was supported by EPSRC grant EP/E028659/1 "Face Recognition using photometric Stereo (PhotoFace)".

References

[1] J.L. Barron, D. Fleet, and S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, 1994.

[2] S. Barsky and M. Petrou. Shadows and highlights detection in 4-source colour photometricstereo. *International Conference on Image Processing*, 3:967–970, October 2001.

[3] S. Barsky and M. Petrou. The 4-source photometric stereo technique for 3-dimensional surfaces in the presence of highlights and shadows. *IEEE Trans. Patt. Anal. Machine Intell.*, 25(10):1239–1252, 2003.

[4] S. Barsky and M. Petrou. Design issues for a colour photometric stereo system. *Journal of Mathematical Imaging and Vision*, 24(1):143–162, January 2006.

[5] M.K. Chandraker, S. Agarwal, and D.J. Kriegman. Shadowcuts: Photometric stereo with shadows. *CVPR*, June 2007.

[6] E.N. Coleman and R. Jain. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *Computer Vision, Graphics and Image Processing*, 18:309–328, 1982.

[7] G.D. Finlayson, M.S. Drew, and C. Lu. Intrinsic images by entropy minimisation. *In Proc. ECCV*, pages 582–595, 2004.

[8] A. Georghiades. Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo. *Ninth IEEE International Conference on Computer Vision (ICCV'03)*, 2:816, 2003.

[9] A. Georghiades. Recovering 3-d shape and reflectance from a small number of photographs. *Proceedings of the 14th Eurographics workshop on Rendering table of contents*, 44:230–240, 2003.

[10] B.K.P. Horn. Understanding image intensities. *Artificial Intelligence*, 8(11):201–231, 1977.

[11] M.D. Levine and J. Bhattacharyya. Removing shadows. *Pattern Recognition Letters*, 26(3):251–265, 2005.

[12] S.K. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Transactions on Robotics and Automation*, 6(4):418–431, 1990.

[13] Bui-Tuong Phong. Illumination for computer generated images. *Comm. ACM*, 18(6):311–317, June 1975.

[14] H. Rushmeier, G. Taubin, and A. Gueziec. Applying shape from lighting variation to bump map capture. *Eurographics Rendering Workshop*, pages 35–44, June 1997.

[15] W. Smith and E.R. Hancock. Estimating cast shadows using sfs and class-based surface completion. *Proceedings of the 18th International Conference on Pattern Recognition*, 4:86–90, 2006.

[16] F. Solomon and K. Ikeuchi. Extracting the shape and roughness of specular lobe objects using four light photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(4):449–454, April 1996.

[17] J. Sun, M. Smith, L. Smith, S. Midha, and J. Bamber. Object surface recovery using a multi-light photometric stereo technique for non-lambertian surfaces subject to shadows and specularities. *Image and Vision Computing*, 25(7):1050–1057, July 2007.

[18] R. Woodham. Photometric stereo: A reflectance map technique for determining surface orientation from image intesit. *SPIE*, 155:136–143, 1978.