Computer Vision and Image Understanding 114 (2010) 887–900

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Photometric stereo with an arbitrary number of illuminants

Vasileios Argyriou, Maria Petrou *, Svetlana Barsky

Communications and Signal Processing Group, Electrical and Electronic Engineering Department, Imperial College London, London SW7 2AZ, UK

ARTICLE INFO

Article history:
Received 28 March 2009
Accepted 15 May 2010
Available online 24 May 2010

Keywords:
Photometric stereo
Phong model
Highlights identification
Shadow identification

ABSTRACT

We present an optimal generalisation of the 4-light photometric stereo technique for an arbitrary number of Q illuminants. We assume that the surface reflectance can be approximated by the Lambertian model plus a specular reflection. 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

Multiple images captured from the same view point under different illumination directions can be used by photometric stereo (PS) methods to derive orientation and reflectance values for each pixel [21]. The recovered surface normals are then integrated, using, for example, a method like the one by Frankot and Chellappa [8], in order to reconstruct the surface. PS has the advantage over conventional stereo that it does not rely on the solution of the correspondence problem and as a consequence, it allows the per pixel surface reconstruction. It is ideal for reconstructing and inspecting surfaces under controlled conditions. PS assumes that neither highlights nor shadows are present in the images, in order to recover gradient information of a surface patch. 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 [15,17].

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, in order 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 are 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 Ikeuchi [19] proposed a 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 [2–4] 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.

It is noted that PS yields the normal field of the observed surface and it is necessary this normal field to be integrated to form the 3D model of the surface [22,8,11]. Integration of the normal field is problematic due to the accumulation of the errors by the process of integration. Many methods have been proposed for the robust integration of a normal field. They may be categorised into two sets. The first approach is to map the obtained noisy non-integrable normal field into an integrable space. Then, a simple path integration is applied to the transformed normal field so that the desired shape is reconstructed. The second approach is to reconstruct the desired shape directly. In this method, some error functional involving surface depth and the normal field is defined and minimised to produce a surface. Either of these approaches could be selected to reconstruct 3D surfaces providing useful

* Corresponding author.
E-mail address: maria.petrou@imperial.ac.uk (M. Petrou).

1077-3142/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.cviu.2010.05.002

information for many computer vision applications, such as 3D face recognition [13].

The motivation of this work is the problem which arises when an arbitrary number of illumination directions are used in capturing images for PS, and we have to select the most appropriate subset of them. Based on the work proposed by Barsky and Petrou in [2,3], a generalisation of the photometric stereo methodology, operating in a least square error manner, is proposed for Q light sources. Both highlights and shadows could be identified and removed resulting in a more accurate estimation of the acquired surface normals.

This paper is structured as follows. In Section 2, we discuss previous work and in Section 3 a short review of photometric stereo is presented. In Section 4, the proposed iterative Photometric Stereo technique for Q light sources and in the presence of highlights and shadows is analysed. Section 5 presents the results of experiments with synthetic and real data. We conclude in Section 6.

2. Previous work on highlight and shadow detection

Coleman and Jain [6] applied the linear photometric stereo algorithm for Lambertian surfaces, to surfaces with non-Lambertian reflectance properties and treated highlights as deviations from the Lambertian law. Their method uses four images of the same surface to detect highlights in the absence of shadows. This is done by comparing the albedos recovered from all four possible triplets of pixels, under the assumption that the specularity regions do not intersect. If the albedos differ significantly, the presence of a highlight is postulated. The three largest triplets must contain highlights and, therefore, the triplet producing the smallest albedo is used for surface recovery.

Solomon and Ikeuchi [19] proposed a variation of the above method taking self-shadows into consideration. This algorithm divides the surface into regions according to the number of sources illuminating each one. Different strategies were suggested for detecting highlights and for the reconstruction of each of the regions. This method has several shortfalls. The most important is that it excludes cast-shadows. Self-shadows restrict the gradient of a facet, whereas the gradient of a cast-shadowed facet has no restrictions at all, and interpreting them as self-shadows leads to an incorrect gradient reconstruction. Furthermore, using a simple threshold to identify shadows is not possible since shadow values in bright areas may be brighter than non-shadows in dark areas.

Chandraker et al. in [5] proposed a reliable algorithm for recovering cast and self-shadows in the presence of albedo variation and complex geometry. This method uses a highly accurate shadow labelling algorithm based on fast graph cuts. Photometric stereo is used to obtain the gradients of each facet and a normal integration technique, that takes into consideration the estimated shadow maps, produces the reconstructed surface. The drawback of this algorithm is that it detects shadows only in the absence of highlights.

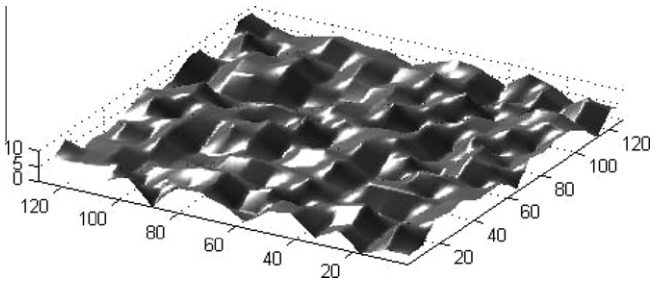

Fig. 1. An example of a simulated surface of size 128 × 128 pixels.

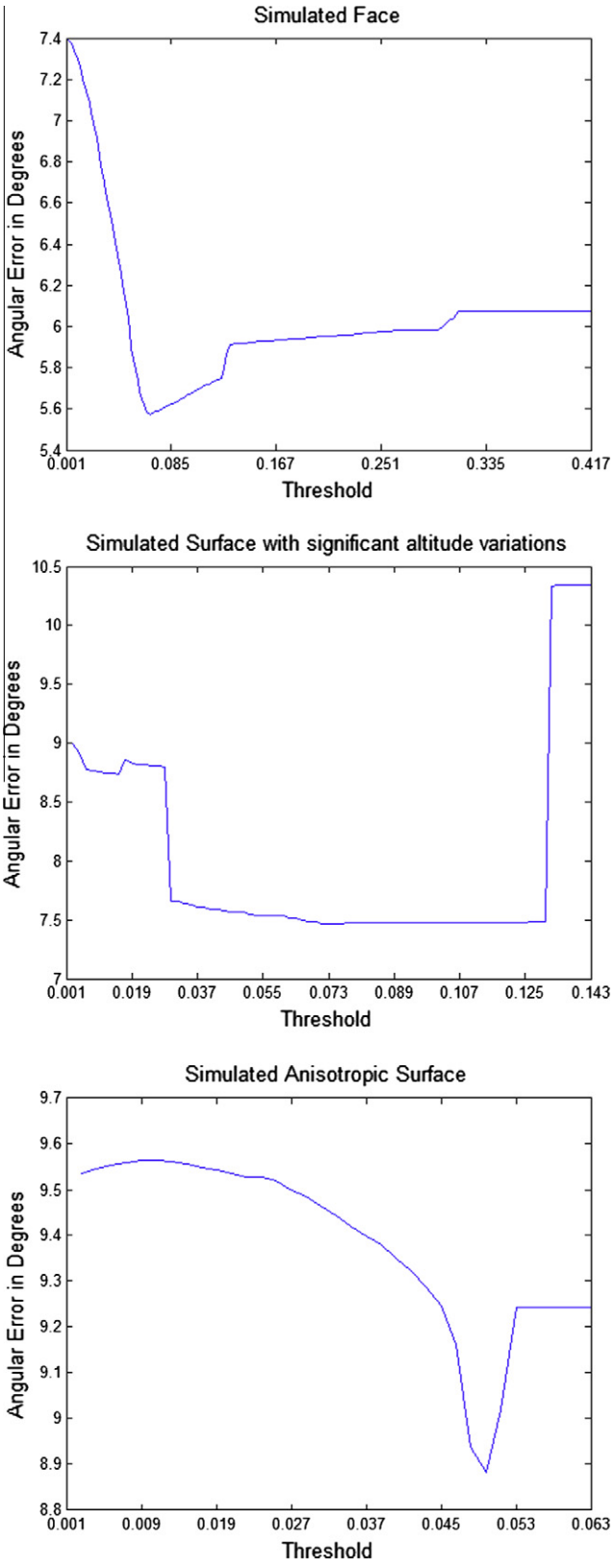

Fig. 2. The angular error as a function of the different threshold values, for a simulated face, a simulated surface with significant altitude variation and an anisotropic simulated surface, from top to bottom, respectively.

Levine and Bhattacharyya [14] used support vector machines to identify shadow boundaries and, based on this information, shadowed regions, without requiring any prior information regarding

Table 1
The mean angular error (MAE) of the surface normals for the simulated surfaces for the methods using four light sources. In bold the best result in each case.

Four lights	Random	Faces	Anisotropic
Photometric stereo-4	1.3776	3.9263	2.6034
Barsky and Petrou	1.9837	**3.0430**	2.2829
Coleman and Jain	3.1733	3.1052	2.6670
Proposed	**0.4232**	3.0671	**2.2829**

Table 2
The mean angular error (MAE) of the surface normals for the simulated surfaces for the methods using five light sources. In bold the best result in each case.

Five lights	Random	Faces	Anisotropic
Photometric stereo-5	1.5676	3.9534	2.9108
Rushmeier	0.1869	3.0103	**2.2014**
Proposed	**0.1683**	**2.9911**	2.2060

Table 3
The mean angular error (MAE) of the surface normals for the simulated surfaces for the methods using six light sources. In bold the best result in each case.

Six lights	Random	Faces	Anisotropic
Photometric stereo-6	1.5794	3.8927	2.8505
Sun	0.6110	3.5086	2.5129
Proposed	**0.1015**	**0.8034**	**0.3439**

the scene. Finlayson et al. [7] proposed a method for recovering intrinsic images that are free from 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 [18]. Rushmeier et al. [17] proposed a five light source photometric stereo system where the highest and lowest values in five components were discarded to avoid highlights and shadows. A six light source Photometric Stereo technique, that employs a slightly more sophisticated decision criterion so as to discard pixels with doubtful values, was suggested by Sun et al. [20]. Georghiades [9,10] proposed a shape recovery method for non-Lambertian surfaces. This work included an approach for excluding shadow and saturated pixels using a more complex model of reflectance (the Torrance and Sparrow model).

In order to separate shadows and highlights, the use of colour information was suggested by Barsky and Petrou in [2–4]. The chromaticity of the brightest pixel was compared with the chromaticity of the darker pixels and, if the difference exceeded a certain threshold, the pixel was 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 proposed 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 reconstructions in such cases. In this paper we 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 $\widetilde{Q}$ lighting directions, where $\widetilde{Q} \leqslant Q$.

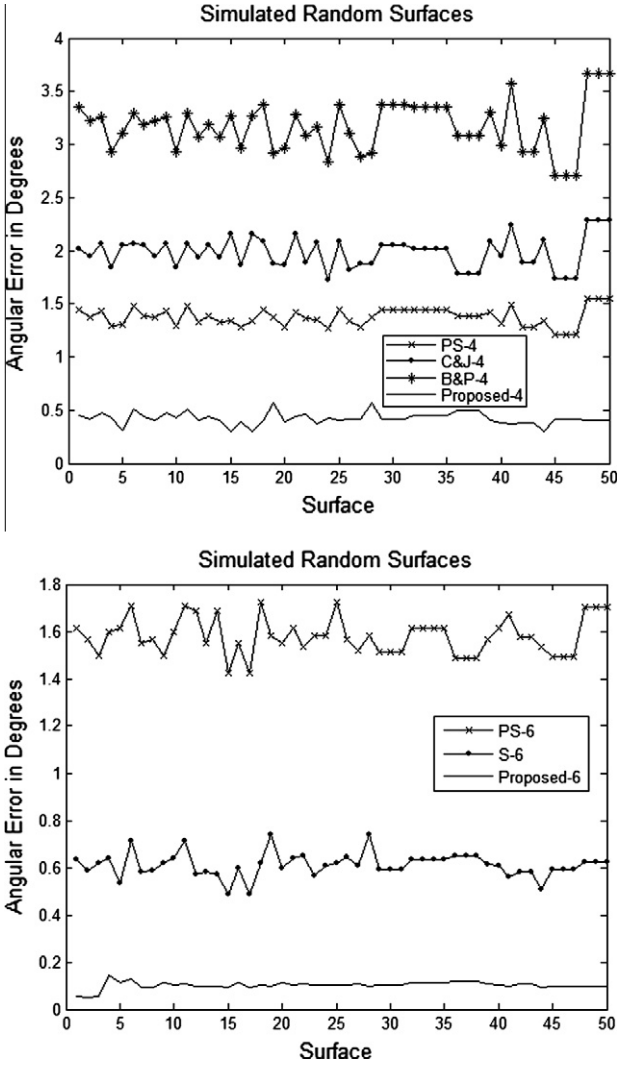

Fig. 3. Angular error in degrees as a function of surface identity. Here (*C&J*) stands for Coleman and Jain (*B&P*) for Barsky and Petrou, (*S*) for Sun et al. and ($PS-x$) for the basic Photometric Stereo technique, where x is the number of lights.

3. 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 the albedo ρ and the cosine of the incidence angle θ_i (the angle between the direction of the incident light and the surface normal) [12]. The above incidence angle can be expressed as the dot 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 $\mathbf{L}^1, \mathbf{L}^2, \ldots, \mathbf{L}^{\widetilde{Q}}$. In this case we can express the intensities of the obtained (grey scale) pixels as:

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

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

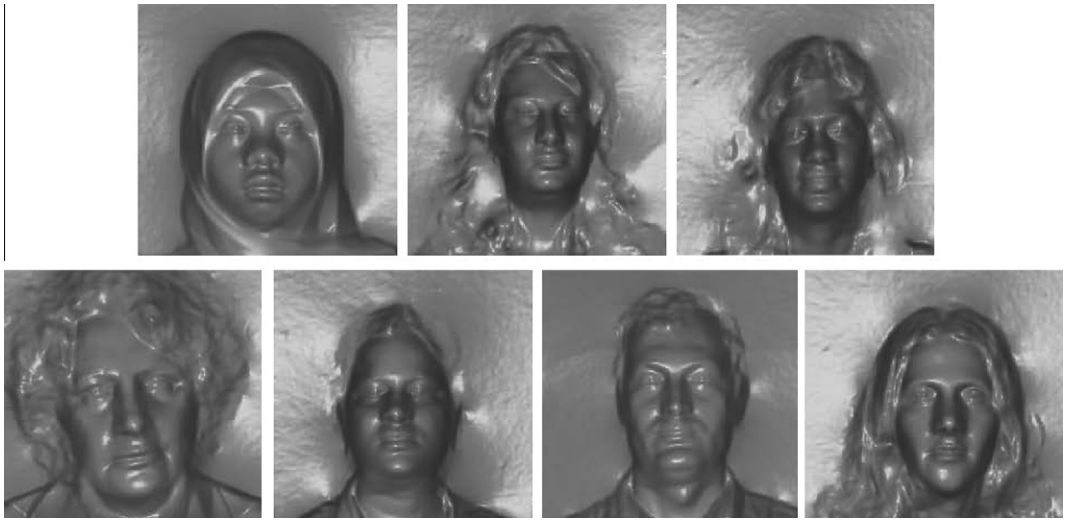

Fig. 4. Examples of simulated faces used in our experiments.

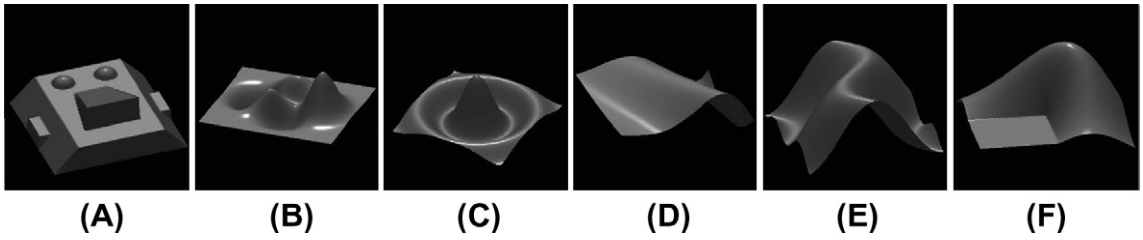

Fig. 5. Simulated data with significant altitude variations used to evaluate the photometric stereo algorithms. All surfaces are of size 128×128.

Table 4
The mean angular error (MAE) of the surface normals for the simulated surfaces, shown in Fig. 5, for the methods using four light sources. In bold the best result for each surface.

Four lights	A	B	C	D	E	F
Coleman and Jain	1.9574	0.2309	1.4515	**0.8247**	1.4028	0.8627
Photometric stereo-4	1.9418	0.2299	1.3753	0.8476	1.3995	0.8657
Barsky and Petrou	1.9418	0.2299	1.3753	0.8476	1.3995	0.8657
Proposed	**1.9376**	**0.2293**	**1.3649**	0.8473	**1.3782**	**0.8617**

Table 5
The mean angular error (MAE) of the surface normals for the simulated surfaces, shown in Fig. 5, for the methods using five light sources.

Five lights	A	B	C	D	E	F
Photometric stereo-5	1.9342	0.2328	1.3818	0.8548	1.4042	0.8552
Rushmeier	1.8179	0.2358	1.3671	0.8433	1.3449	0.8624
Proposed	**1.8179**	**0.2284**	**1.3670**	**0.8433**	**1.3449**	**0.8549**

Table 6
The mean angular error (MAE) of the surface normals for the simulated surfaces, shown in Fig. 5, for the methods using six light sources.

Six lights	A	B	C	D	E	F
Photometric stereo-6	1.9247	0.2352	1.3829	0.8439	1.4036	0.8515
Sun	1.9235	0.2304	1.3734	0.8503	1.4049	0.8499
Proposed	**1.9027**	**0.1994**	**1.1651**	**0.8439**	**1.4017**	**0.8483**

Table 7
The mean angular error (MAE) of the surface normals for the simulated surfaces, shown in Fig. 5, for the case where the number of lights was chosen to produce the best result, for the methods that this is possible (best case).

Best case	A	B	C	D	E	F
Coleman and Jain	1.9574	0.2309	1.4515	**0.8247**	1.4028	0.8627
Photometric stereo	1.8989	0.2299	1.3578	0.8439	1.3629	0.8501
Rushmeier	1.8179	0.2358	1.3670	0.8433	1.3449	0.8624
Sun	1.9418	0.2299	1.3753	0.8476	1.3995	0.8657
Barsky and Petrou	1.9235	0.2304	1.3735	0.8503	1.4049	0.8499
Proposed	**1.7864**	**0.1982**	**1.0868**	0.8433	**1.3449**	**0.8406**

Table 8
The mean highlight error (MHE) of the surface normals for the simulated surfaces, shown in Fig. 5.

		A	B	C	D	E	F
Five lights – MHE	Rushmeier	0.199	0.189	0.186	0.169	0.192	0.190
	Proposed	**0.106**	**0.009**	**0.109**	**0.088**	**0.141**	**0.007**
Six lights – MHE	Sun	0.024	**0.028**	**0.003**	0.037	**0.021**	**0.003**
	Proposed	**0.012**	0.038	0.057	**0.011**	0.049	0.007

Table 9
The mean shadow error (MShE) of the surface normals for the simulated surfaces, shown in Fig. 5.

		A	B	C	D	E	F
Five lights – MShE	Rushmeier	0.199	0.166	0.146	0.162	0.165	**0.138**
	Proposed	**0.079**	**0.105**	**0.065**	**0.027**	**0.122**	0.170
Six lights – MShE	Sun	0.140	0.122	0.188	0.102	0.279	0.188
	Proposed	**0.011**	**0.014**	**0.024**	**0.101**	**0.038**	**0.184**

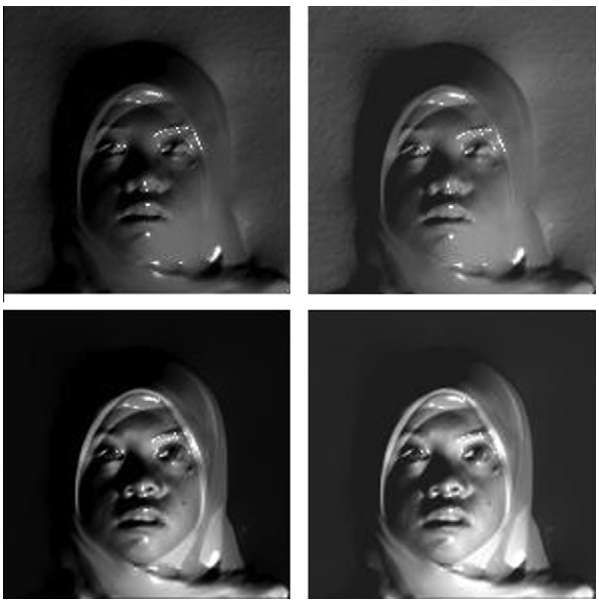

Fig. 6. An example of a simulated face without (left) and with (right) the presence of ambient light.

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

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 $[L]$:

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

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

4. 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 illuminants 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. So, we should find a suitable basis, in the Q-space such that three 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 onto the orthogonal subspace.

More formally, the intensity vectors are linear combinations (Q-dimensional) of the 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 $\mathbf{v}_4,\ldots,\mathbf{v}_Q$ such that:

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

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 four illuminations the orthogonal subspace is one-dimensional.

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

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

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:

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

We have already shown that all elements of matrix $[A][L]$ are zero, therefore the first term in (9) 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 almost 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.

Table 10
The mean angular error (MAE) of the surface normals for the simulated faces with 15% ambient light for 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.

Best case	AF1	AF2	AF3	AF4	AF5	AF6	AF7
Coleman and Jain	5.4525	5.7797	4.9105	4.8823	6.5594	4.9062	4.8944
Rushmeier	5.5489	5.4809	4.9690	5.0075	5.9498	4.7075	4.9716
Sun	5.0306	4.9704	4.8037	4.7728	5.6045	4.5393	4.5648
Barsky and Petrou	5.4945	6.1015	4.9334	4.9182	6.4272	4.9207	4.9433
Proposed	**4.8023**	**4.5334**	**4.6012**	**4.6407**	**5.3423**	**4.3546**	**4.3513**

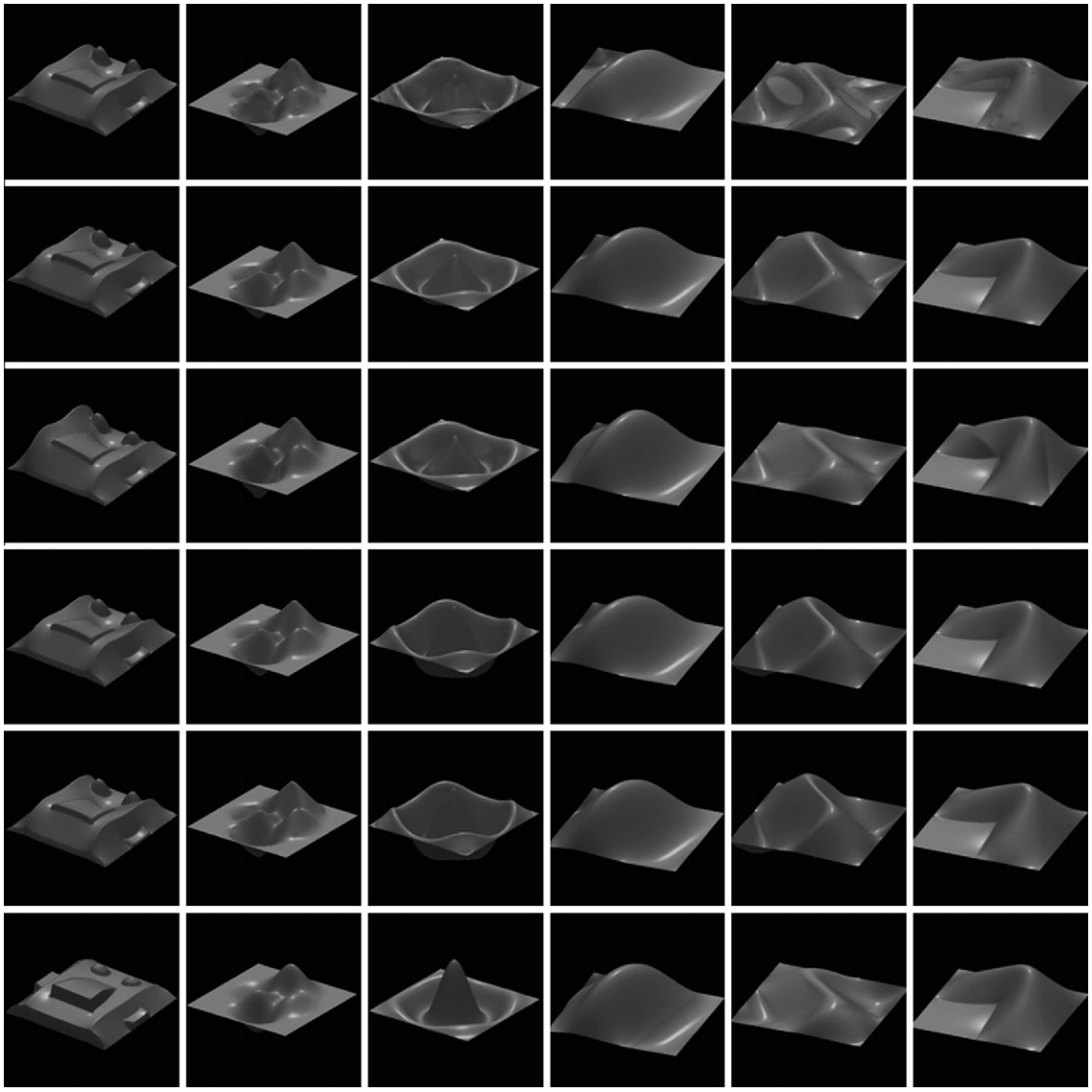

Fig. 7. The reconstructed surfaces, from top to bottom: Coleman's and Jain's, Standard PS with eight lights, Rushmeier's, Sun's, Barsky's and Petrou's, and the proposed iterative method.

1. Consider the K-dimensional intensity space (at 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 an 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 three intensities left.

Note that the matrices should be recalculated at every step, because dropping one intensity and the corresponding illuminant takes away a component in each of the columns of matrix $[L]$, so matrix $[L][L]^T$ will have a completely different form. This does not make the method too slow, because one may pre-calculate a

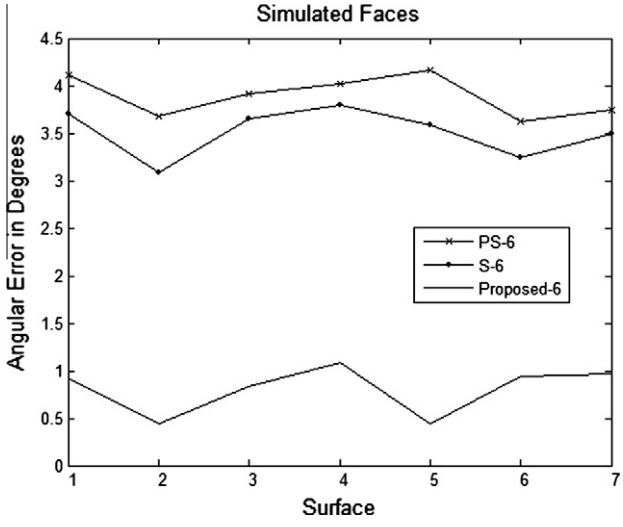

Fig. 8. Angular error in degrees as a function of the surface identity. Here (S) stands for Sun et al. and ($PS - 6$) for the basic Photometric Stereo technique, with six lights.

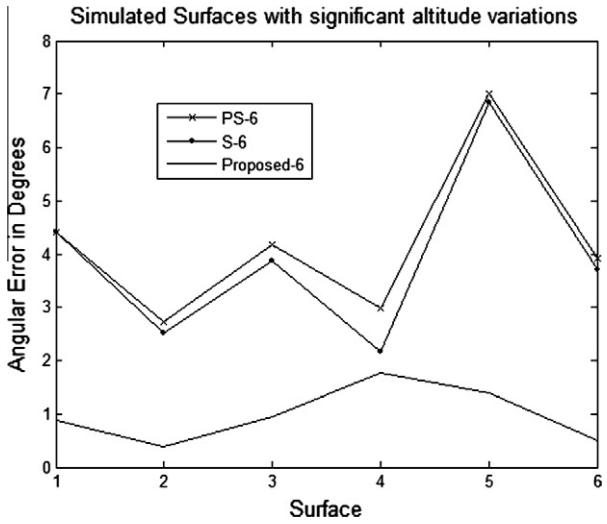

Fig. 9. Angular error in degrees as a function of the surface identity. Here (S) stands for Sun et al. and ($PS-6$) for the basic Photometric Stereo technique, with six lights.

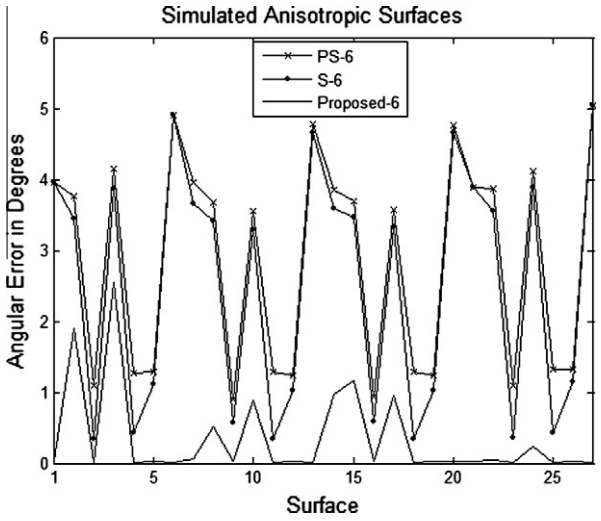

Fig. 11. Angular error in degrees as a function of the surface identity. Here (S) stands for Sun et al. and ($PS-6$) for the basic Photometric Stereo technique, with six lights.

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 1–4 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 $\widetilde{Q} = K_{final}$ 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 $\widetilde{Q} = K_{final} + 1$ pixel intensities.

4.1. Optimal threshold selection

According to the above methodology, a predefined threshold value is required to determine whether a pixel is defected (shadowed or highlighted). Initially this value was selected experimentally without guarantee for an optimal solution. The main problem of estimating the appropriate threshold is that it is content dependent and therefore the variation of the value is significant for different shapes. Observing that objects sharing common features, e.g. faces, require similar threshold values, the threshold may be estimated for each class of object separately. So, we may define the threshold using training data.

Let us assume that for a set of objects with similar distribution of facet normals and of the same class as the surfaces we wish to reconstruct, the normals $\mathbf{N}_r$ of the facets are known. The estimated normals obtained using the proposed algorithm are also known as functions of the threshold value, denoted by $\mathbf{N}_e(t)$, where t is the threshold value. So, the optimal threshold would be the one that minimises

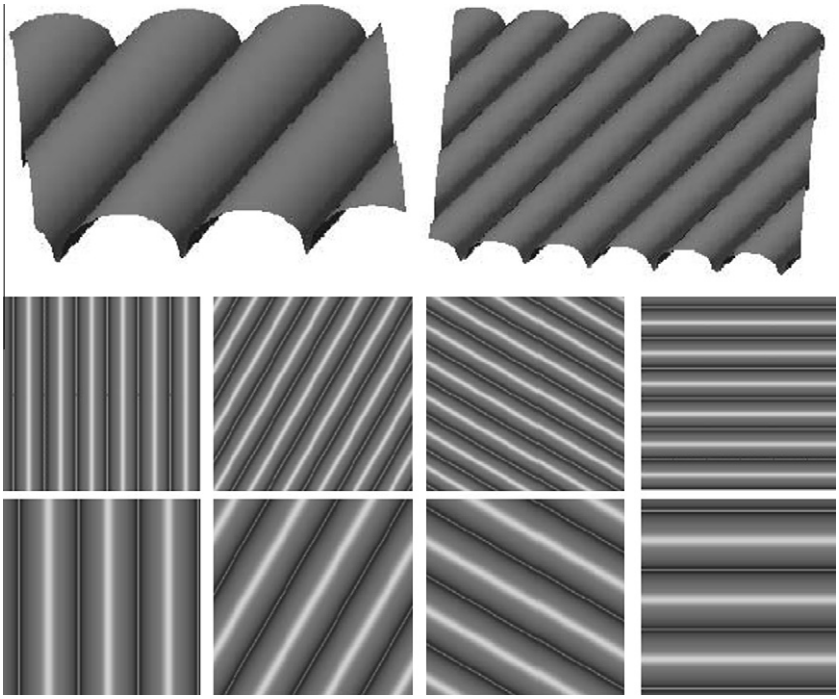

Fig. 10. Examples of anisotropic surfaces used in our experiments.

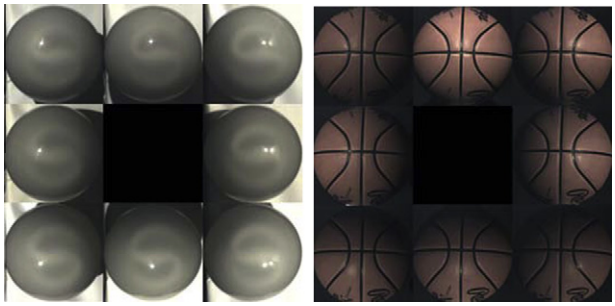

Fig. 12. Spheres of different materials: a light bulb, with a highly reflective surface, on the left, and on the right a basketball, which is much closer to a Lambertian surface.

Table 11
The Hausdorff distance of the spheres' side view with the reconstructed side views by all methodologies tested. In bold, the best result for each surface.

Hausdorff distance	Bulb	Basketball
Rushmeier	39.8209	37.0270
Sun	33.9008	33.8703
Barsky and Petrou	34.8502	34.5392
Coleman and Jain	39.0137	35.6585
Proposed	**33.2224**	**33.7906**

$$f(t) \equiv \frac{1}{S}\sum_{i=1}^{S} |N_{r_i} - N_{e_i}(t)| \tag{10}$$

where S is the total number of facets. Exhaustive search may be used to solve this optimisation problem and obtain the optimal threshold t for each class of object.

5. Experiments and results

In case of simulated data, the 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 [6], Barsky and Petrou [2], Rushmeier et al. [17], Sun et al. [20] and the basic Photometric Stereo technique, the mean angular error (MAE) of the surface normals is selected as a comparison measure. In particular, the MAE measure suggested by Barron 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.

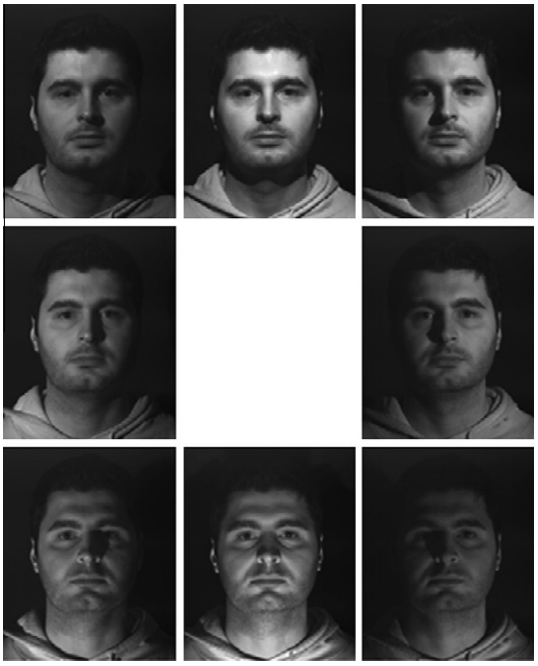

Fig. 15. Input images, one for each illumination direction.

Fig. 13. Real faces used for experiments.

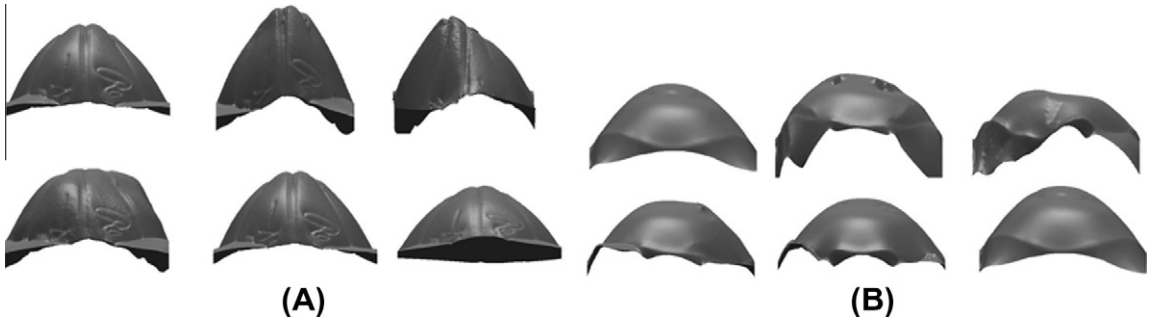

Fig. 14. Side view of the reconstructed spheres for visual performance comparison. (A) Basketball. (B) Light bulb. In lexicographic order, the reconstructions obtained by standard PS, Coleman and Jain, Rushmeier's method, Barsky's and Petrou's, Sun's, and the proposed iterative method.

Table 12
The mean highlight intensity (MHI) and the mean shadow intensity (MShI) for the real surfaces of all comparable methods. In bold the best result for each face.

		Face A	Face B	Face C	Face D	Face E
MHI	Rushmeier	0.2820	0.3297	0.3705	0.3124	0.3554
	Sun	0.7050	0.7002	0.7295	**0.7194**	0.7647
	Proposed	**0.7150**	**0.8520**	**0.8363**	0.6899	**0.8367**
MShI	Rushmeier	0.0969	0.1231	0.1221	0.1214	0.1247
	Sun	0.2498	0.2571	0.2057	**0.0013**	0.2162
	Proposed	**0.0031**	**0.1067**	**0.0847**	0.0980	**0.0022**

The simulated images were produced using the Phong model [16]. This model is mathematically expressed as

$$I = \rho(\mathbf{L}\cdot\mathbf{N}) + k(\mathbf{R}\cdot\mathbf{V})^m \tag{11}$$

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 Eq. (11) 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 indicators of the performance of the algorithm.

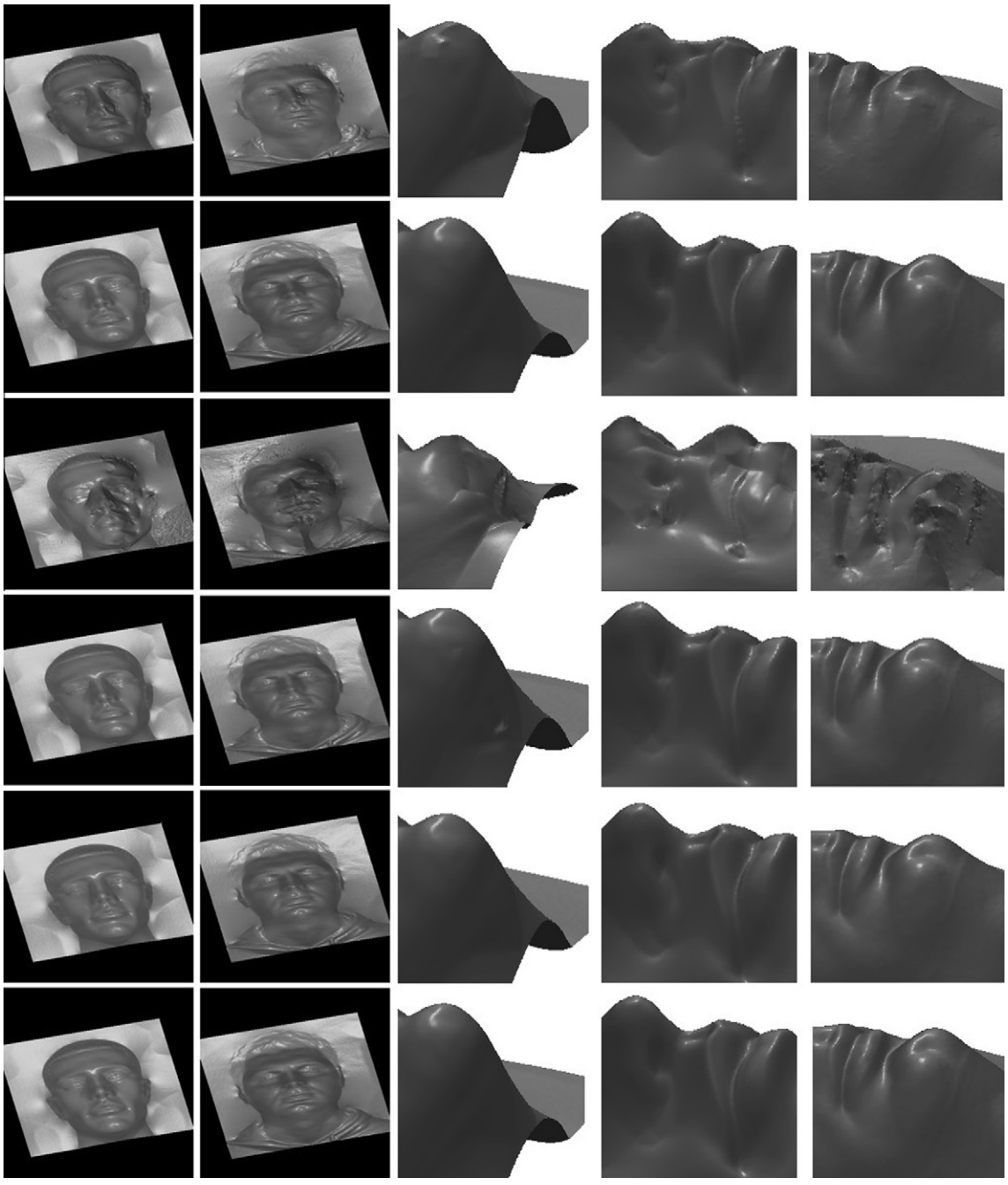

Fig. 16. The reconstructed surfaces for two of the real faces, using, from top to bottom, Coleman's and Jain's, Standard PS with eight lights, Rushmeier's, Sun's, Barsky's and Petrou's, and the proposed iterative method. The rightmost three panels show some details that allow the appreciation of the differences between the different reconstructions. It can be observed that the slope below the chin is higher when more lights are used as in the proposed method.

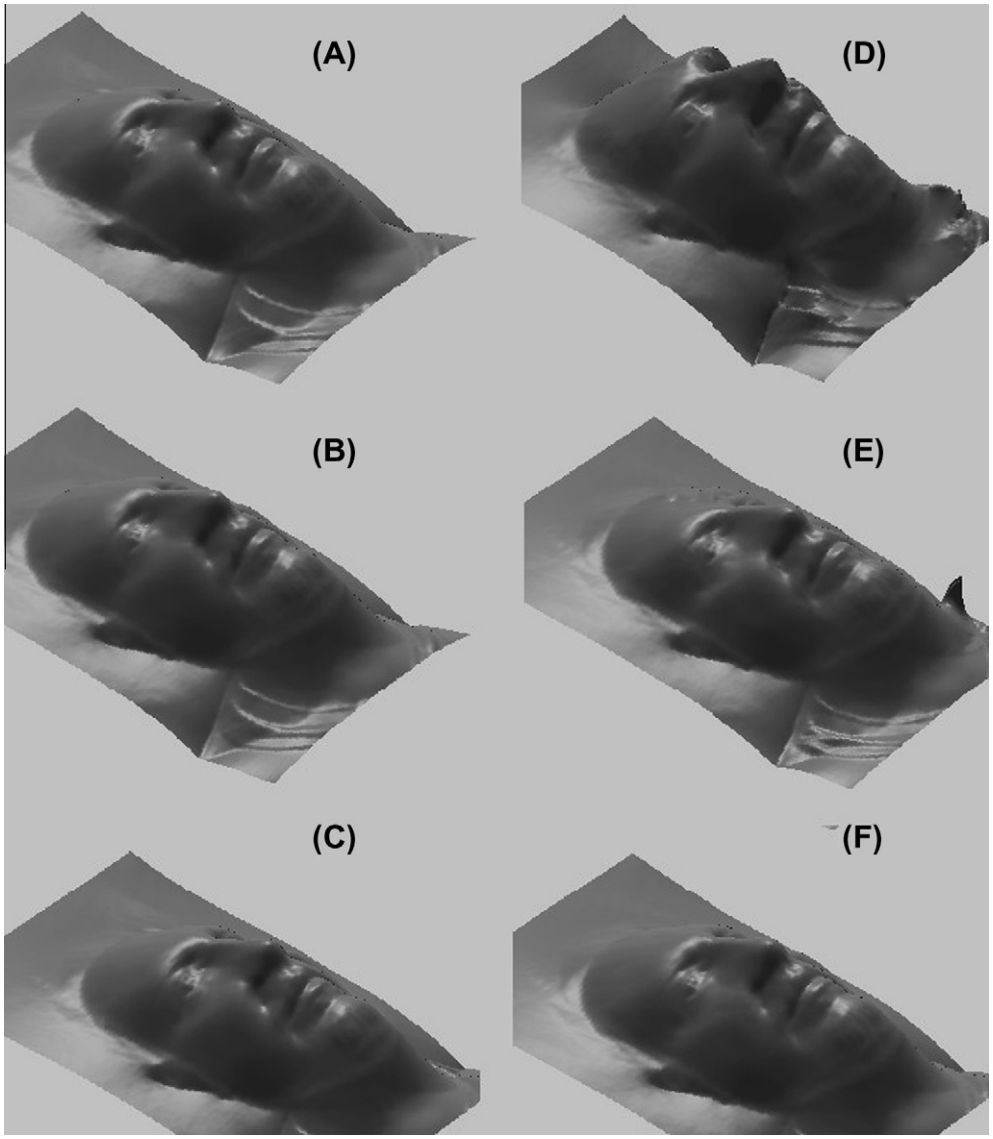

Fig. 17. The reconstructed surfaces for one of the real faces using (A) standard PS, (B) Coleman's and Jain's, (C) Barsky's and Petrou's, (D) Rushmeier's, (E) Sun's, and (F) the proposed iterative method.

In all experiments, after the normal vectors of the surface facets had been obtained, our own implementation of the method proposed in [11] was used to obtain the 3D reconstructed surface.

5.1. Experiments with simulated data

Experiments were performed with four different sets of simulated data. Initially fifty surfaces were artificially created (see Fig. 1) on the basis of containing random peaks with low altitude, in order to avoid for this set of experiments cast-shadows and in order to provide a comprehensive evaluation data set. The optimal threshold was estimated using exhaustive search for each object-class. The results are presented in Fig. 2. The threshold which produced the minimum error in each case was used to produce all results reported in this paper concerning the reconstruction of surfaces of the same class. The surface used to obtain the threshold for each set of surfaces was the first one in the set. In order to have a fair comparison, methods with the same number of light sources were grouped together, and, additionally, the best results of each method are presented. In the case of shadows and highlights the comparison was performed only when this was allowable, since some methods are able to estimate either only shadows or only highlights or none of them. The results for the random simulated surfaces are shown in Tables 1–3. Fig. 3 shows the angular error as a function of the surface identity.

Further experiments were performed using simulated faces (see Fig. 4). The estimated mean angular errors for each method are shown in Tables 1–3. Similar results were obtained using surfaces with significant altitude variations (see Fig. 5). The surfaces were selected on the basis of containing highlights and shadows, in order to provide a comprehensive evaluation data set. The results for the simulated surfaces of Fig. 5 are shown in Tables 4–9. In Tables 8 and 9, the mean highlight error (MHE) and the mean shadow error (MShE) are used to assess performance. For synthetic surfaces, we know the highlighted and shadowed pixels, and we can therefore work out the average error with which these pixels are identified. In the case of the mean angular error measure, the proposed method outperforms the other methods resulting in the lowest error. In addition, experiments were performed with the presence of ambient light for the simulated faces (see Fig. 6 and Table 10) with the proposed method to result the lowest error. Similar results were obtained for the errors estimating the mean

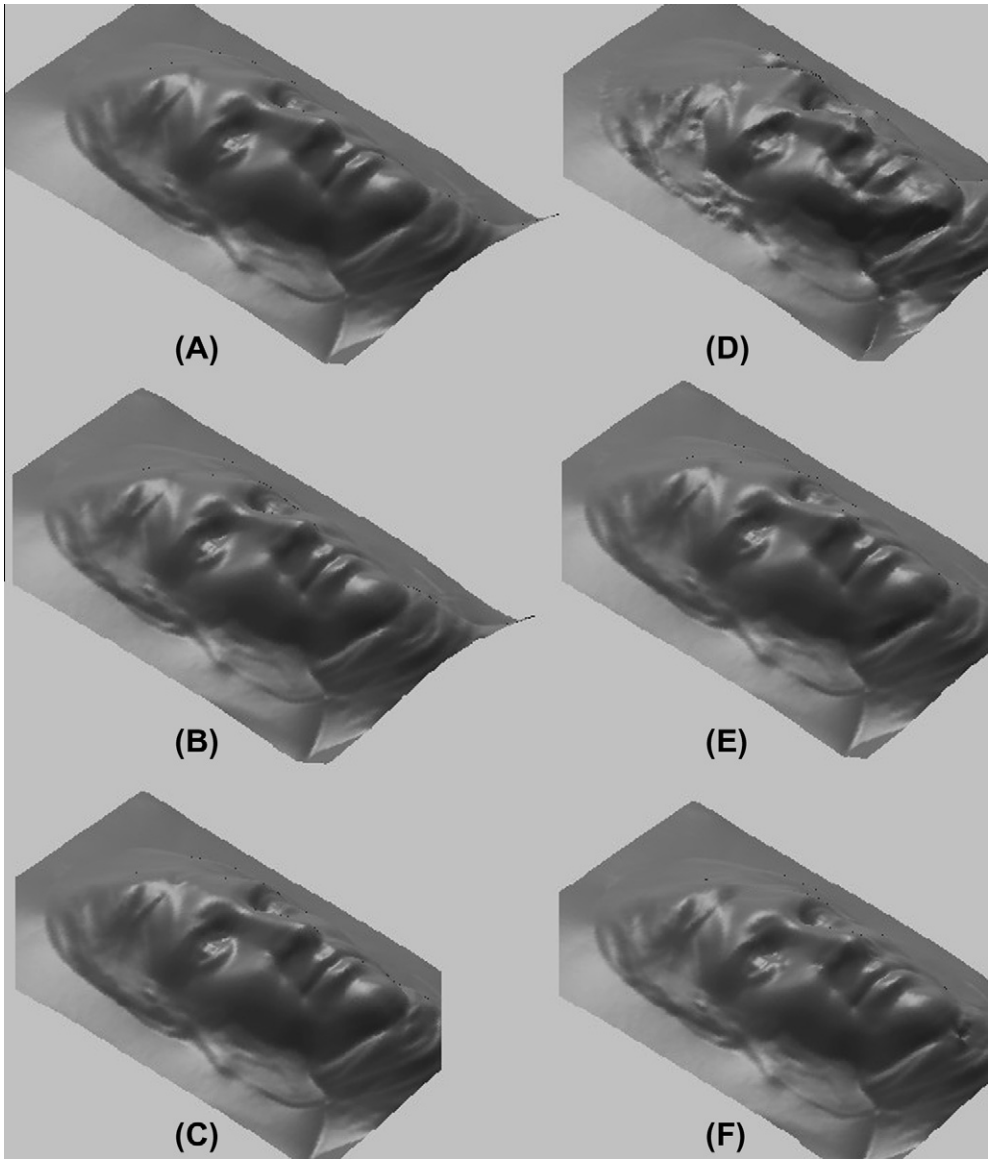

Fig. 18. The reconstructed surfaces for one of the real faces using (A) standard PS, (B) Coleman's and Jain's, (C) Barsky's and Petrou's, (D) Rushmeier's, (E) Sun's, and (F) the proposed iterative method.

Table 13
The Hausdorff distance of the faces' side views from the reconstructed side views for all methodologies tested. In bold, the best result for each face.

Hausdorff distance	Face A	Face B	Face C	Face D	Face E
Rushmeier	81	82	82	95	97
Sun	46	52	57	84	92
Barsky and Petrou	62	64	66	86	91
Coleman and Jain	73	69	71	94	95
Proposed	**37**	**48**	**55**	**82**	**89**

grey value of the highlights and shadows. Fig. 7 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. The average angular error for each face and surface is shown in Figs. 8 and 9.

Finally, a set of rotated anisotropic surfaces was used to evaluate the proposed methodology under these conditions in order to demonstrate further the improvement in the accuracy of the estimated normals (see Fig. 10). The angular error as a function of surface is shown in Fig. 11. The average angular error is shown in Tables 1–3.

5.2. Experiments with real data

The proposed algorithm was further applied to the reconstruction of two spheres of different materials (see Fig. 12) and five human faces (see Fig. 13) using photometric data captured with eight illumination directions. More specifically, a light bulb, with a highly reflective surface, and a basketball, which is much closer to a Lambertian surface, were used to evaluate the proposed methodology. The true shape of each sphere can be easily estimated from the knowledge of its diameter. Using this information as the ground truth, the Hausdorff distance was used to evaluate the accuracy of the reconstructed spheres (see Table 11). Fig. 14 shows the side view of the reconstructed spheres for visual performance evaluation.

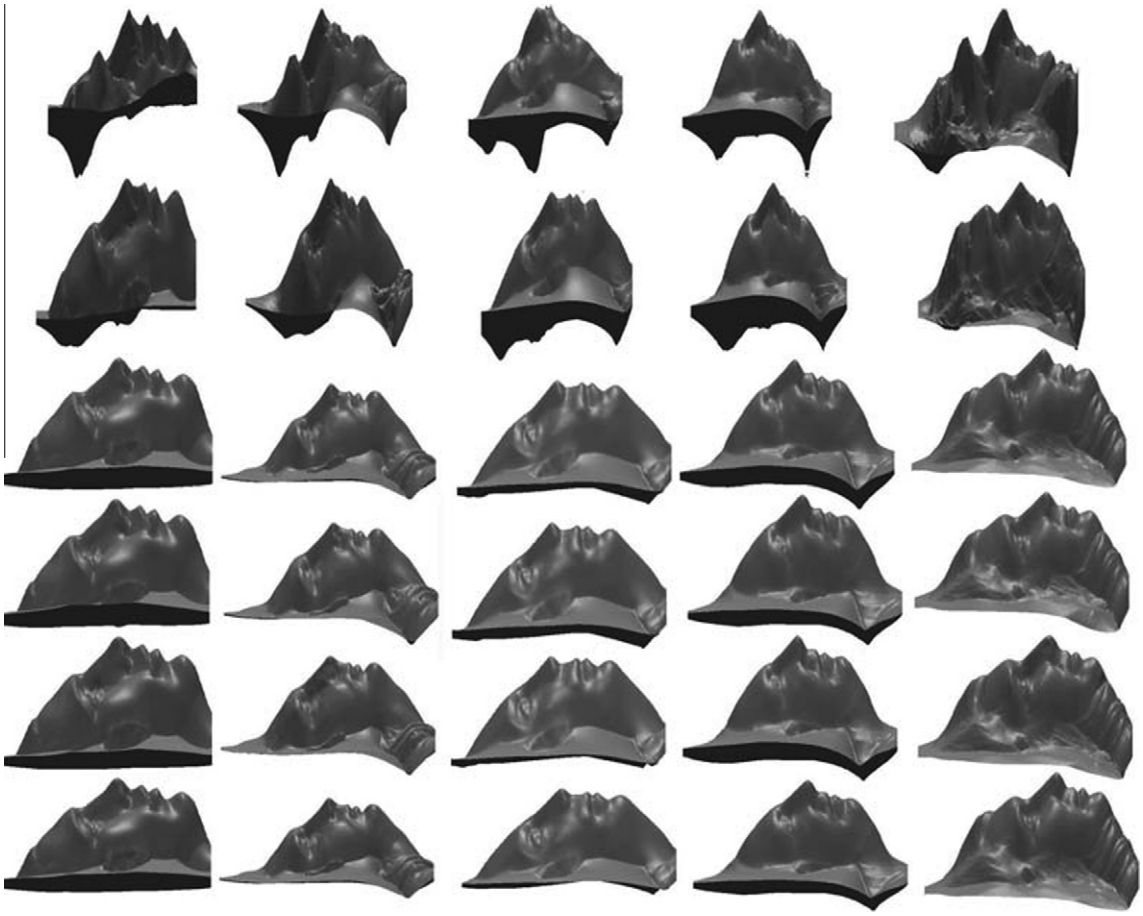

Fig. 19. The reconstructed profiles of the faces using Rushmeier's, Coleman's and Jain's, standard PS, Barsky's and Petrou's, Sun's, and the proposed iterative method from top to bottom, respectively.

Fig. 15 presents the input images for one of the faces, one for each illumination direction. The lights were placed on the vertices and the sides of a square of approximate size 1.5 m $\times$ 1.5 m. The distance between the camera and the observed person was 1.25 m. In such an arrangement the difference of the tilt angles between neighbouring illumination directions is 45° and the slant angle is equal to 40° for the corner lights and 31° 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. 200 frames per second), 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 12). 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 are dark in the absence of ambient light, which is the case for our data. In Figs. 16–18 results of the reconstructed faces obtained from the six compared 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 and region under the jaw) indicating that extra light sources, beyond the conventionally three used, should be useful for these surface patches. Furthermore, the side view was used to evaluate the performance of the proposed iterative approach. The background was extracted manually and the Hausdorff distance was used to compare the reconstructions with the original profiles. Table 13 shows all the results for all the faces and some examples are presented in Figs. 19 and 20.

6. Conclusions

In this paper, we propose an iterative algorithm for Q-light photometric stereo, which is dealing with multiple highlights and shadows, allowing one to obtain more reliable estimates of surface parameters. It should be noted that, on average, and for the face surfaces we used, about 2.4–3.5% of the error is caused by the use of shadowed pixels and about 0.5–1.7% is due to highlighted pixels. The proposed method iteratively discards the most problematic intensity for each pixel, in a least square error sense, until the error becomes smaller than a threshold or the acceptable light sources reduce to three. The threshold value is optimally selected for each class of object, based on the availability of a generic surface model of the reconstructed surfaces. It may appear restrictive to base the selection of the threshold, and as a result the performance of the algorithm, on the availability of a generic model of the reconstructed surfaces. However, one of the most significant applications of PS is in industrial inspection, where the surfaces inspected may belong to the same class, e.g. ceramics of some type, certain type of fruit, etc. The use of PS in a controlled environment also allows the user to reduce the effect of inter-reflections and ambient light that impact, to some extent, on the performance of the PS algorithms.

Experiments with both simulated and real images were performed in order to evaluate the proposed algorithm. The mean angular error of the recovered surface normals, and the average

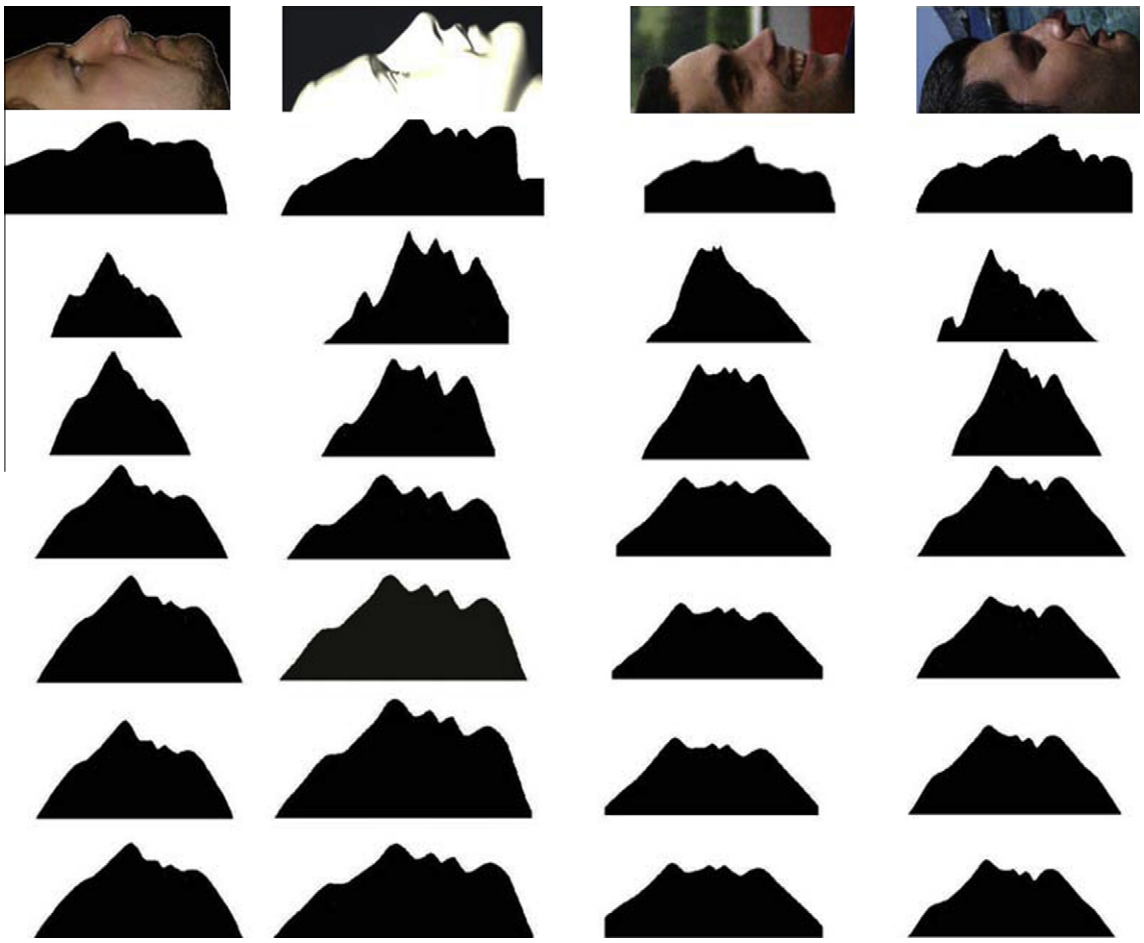

Fig. 20. The reconstructed side views in black and white along side the original profiles. 1st row the real faces, 2nd the manually traced face profiles and then the obtained profiles using Rushmeier's, Coleman's and Jain's, standard PS, Barsky's and Petrou's, Sun's, and the proposed iterative method, from top to bottom, respectively.

grey value of the shadows and the highlights were used to assess the performance of the compared methods in the cases when the ground truth was available. For the cases when the ground truth was not available, only the average grey values of shadows and highlights could be used as performance indicators, apart from visual inspection of the reconstructed surfaces. From the real images it can observed that the extra light sources improve the surface normals' estimation, and consequently the surface reconstruction, especially in areas where multiple shadows are present, for example under the chin of a human face.

Further research is needed to investigate which is the proper way of treating the 'error vector' $\mathbf{P} \equiv [A]\mathbf{E}$ in Eq. (9). We cannot recover the error itself since $\mathbf{E}$ is a Q-dimensional vector, and $\mathbf{P}$ is of dimensionality $Q-3$. The obvious way is to measure the length of $\mathbf{P}$ but the Euclidean metric might not be the best for this purpose.

Acknowledgment

This work was supported by the Project Face Recognition using Photometric Stereo, funded by EPSRC, with Contract Number EP/E028845/1.

References

[1] J.L. Barron, D. Fleet, S. Beauchemin, Performance of optical flow techniques, International Journal of Computer Vision 12 (1) (1994) 43–77.
[2] S. Barsky, M. Petrou, Shadows and highlights detection in 4-source colour photometric stereo, in: International Conference on Image Processing, October 2001, vol. 3, pp. 967–970.
[3] S. Barsky, M. Petrou, The 4-source photometric stereo technique for 3-dimensional surfaces in the presence of highlights and shadows, IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (10) (2003) 1239–1252.
[4] S. Barsky, M. Petrou, Design issues for a colour photometric stereo system, Journal of Mathematical Imaging and Vision 24 (1) (2006) 143–162.
[5] M.K. Chandraker, S. Agarwal, D.J. Kriegman, Shadowcuts: Photometric Stereo with Shadows, CVPR, June 2007.
[6] E.N. Coleman, R. Jain, Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry, Computer Vision, Graphics and Image Processing 18 (1982) 309–328.
[7] G.D. Finlayson, M.S. Drew, C. Lu, Intrinsic images by entropy minimisation, in: Proc. ECCV, 2004, pp. 582–595.
[8] R.T. Frankot, R. Chellappa, A method for enforcing integrability in shape from shading, IEEE Transactions on Pattern Analysis and Machine Intelligence 10 (4) (1988) 439–451.
[9] A. Georghiades, Incorporating the Torrance and Sparrow model of reflectance in uncalibrated photometric stereo, in: Ninth IEEE International Conference on Computer Vision (ICCV'03), 2003, vol. 2, pp. 816–823.
[10] A. Georghiades, Recovering 3-d shape and reflectance from a small number of photographs, in: Proceedings of the 14th Eurographics Workshop on Rendering Table of Contents, 2003, vol. 44, pp. 230–240.
[11] A.S. Georghiades, P.N. Belhumeur, D.J. Kriegman, From few to many: illumination cone models for face recognition under variable lighting and pose, IEEE Transactions on Pattern Analysis and Machine Intelligence 23 (6) (2001) 643–660.
[12] B.K.P. Horn, Understanding image intensities, Artificial Intelligence 8 (11) (1977) 201–231.
[13] I.A. Kakadiaris, G. Passalis, G. Toderici, M.N. Murtuza, Y. Lu, N. Karampatziakis, T. Theoharis, Three-dimensional face recognition in the presence of facial expressions: an annotated deformable model approach, IEEE Transactions on Pattern Analysis and Machine Intelligence 29 (4) (2007) 640–649.
[14] M.D. Levine, J. Bhattacharyya, Removing shadows, Pattern Recognition Letters 26 (3) (2005) 251–265.
[15] S.K. Nayar, K. Ikeuchi, T. Kanade, Determining shape and reflectance of hybrid surfaces by photometric sampling, IEEE Transactions on Robotics and Automation 6 (4) (1990) 418–431.

[16] Bui-Tuong Phong, Illumination for computer generated images, Communications of the ACM 18 (6) (1975) 311–317.
[17] H. Rushmeier, G. Taubin, A. Gueziec. Applying shape from lighting variation to bump map capture, in: Eurographics Rendering Workshop, June 1997, pp. 35–44.
[18] W. Smith, E.R. Hancock, Estimating cast shadows using sfs and class-based surface completion, in: Proceedings of the 18th International Conference on Pattern Recognition, 2006, vol. 4, pp. 86–90.
[19] F. Solomon, 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) (1996) 449–454.
[20] J. Sun, M. Smith, L. Smith, S. Midha, 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) (2007) 1050–1057.
[21] R. Woodham, Photometric stereo: a reflectance map technique for determining surface orientation from image intensity, SPIE 155 (1978) 136–143.
[22] A. Yuille, D. Snow, Shape and albedo from multiple images using integrability, in: Proc. Conf. Computer Vision and Pattern Recognition, 1997.