This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

3D Shape Recovery of Smooth Surfaces: Dropping the Fixed Viewpoint Assumption

Yael Moses *Member, IEEE* and Ilan Shimshoni *Member, IEEE*

***Abstract*— We present a new method for recovering the 3D shape of a featureless smooth surface from three or more calibrated images illuminated by different light sources (three of them are independent). This method is unique in its ability to handle images taken from unconstrained perspective viewpoints and unconstrained illumination directions. The correspondence between such images is hard to compute and no other known method can handle this problem locally from a small number of images. Our method combines geometric and photometric information in order to recover dense correspondence between the images and accurately computes the 3D shape. Only a single pass starting at one point and local computation are used. This is in contrast to methods which use the occluding contours recovered from many images to initialize and constrain an optimization process. The output of our method can be used to initialize such processes. In the special case of fixed viewpoint, the proposed method becomes a new perspective photometric stereo algorithm. Nevertheless, the introduction of the multiview setup, self-occlusions and regions close to the occluding boundaries are better handled, and the method is more robust to noise than photometric stereo. Experimental results are presented for simulated and real images.**

***Index Terms*— 3D shape reconstruction, featureless objects.**

I. Introduction

We present a method for recovering the 3D shape of a smooth featureless surface. Our system accepts as input three or more calibrated images of the surface, taken from different viewpoints (there can be a wide baseline between the camera positions) and illuminated by different known distant point light sources (three of them are independent). The surface is assumed to be Lambertian, and the perspective projection model is used. This is a challenging problem for which classical methods for shape recovery, both geometric or photometric, are inadequate, since correspondence between such images is hard to compute. Geometric methods, such as stereo or structure from motion, are based on the recovery of corresponding points in different images. Determining correspondence for images of the type considered in this paper is hard because the surface is assumed to be featureless and the grey-level values of corresponding points can vary considerably between images when the direction of the light source changes. This is demonstrated in Figure 1. In classical photometric stereo [40], [31], the input images are taken with different lighting directions but from the same viewpoint. The fixed viewpoint assumption provides the correspondence for such methods. In the general setup considered here, the fixed viewpoint assumption does not hold and therefore these methods cannot be directly applied to our images.

Our method accurately computes the surface shape of 3D objects by combining geometric and photometric information. A dense correspondence between the set of input images is recovered, despite the dissimilarity of corresponding features when lighting and viewpoint directions change from image to image. The general idea is to propagate correspondences over the surface and simultaneously recover its 3D shape. Given a set of corresponding points, we can compute the 3D location using geometric stereo triangulation. We can also compute the normal to the point using local photometric information (e.g., classical photometric stereo [40], [31] or more general methods [1]). The recovered 3D point and surface normal define the local shape of the surface, which is then used to propagate the correspondence more accurately than possible by shape-from-shading algorithms [16]. This process is repeated to recover the full 3D shape. A simplified version of this process is illustrated in Figure 2. Noise, self-occlusions, and shadows are overcome by using a few more images than the minimum required for shape recovery from shading. In contrast to whole-object optimization methods (e.g., [19], [34]), our method performs the shape recovery in a single pass, as in [6], [21]. In some applications, the output of our method can also be used to initialize whole-object optimization methods. This replaces, for example, the visual hull created from the occluding contours of a large set of images (e.g.,[20], [41], [39], [15], [2].

Note that traditional geometric methods assume that corresponding points have similar intensities. Differences in intensity values of corresponding points are often regarded by these methods as noise. In our case the intensities of corresponding points are expected to be different since the light source changes from image to image. Thus, these differences, rather than being considered noise, become a source of information for computing the local approximation of the surface shape. Combining geometric and photometric information allows us to draw on both, while overcoming their weaknesses. We overcome the weakness of photometric stereo by dropping the fixed viewpoint assumption while using additional geometric information. In the simple case when a single viewpoint is considered, the algorithm is reduced to a straightforward method for handling perspective photometric stereo. When different viewpoints are considered, it also allows the depth factor ambiguity to be eliminated. In addition, the method can handle self-occlusion because more than a single viewpoint is used. We overcome the weaknesses of geometric stereo through our ability to handle images of featureless, smooth surfaces illuminated by different light sources while using photometric

Y. Moses is with the Dept. of Computer Science, The Interdisciplinary Center, Herzliya 46150, Israel, email:yael@idc.ac.il

I. Shimshoni is with the Dept. of Management Information Systems, University of Haifa, Haifa 31905, Israel, email:ishimshoni@mis.haifa.ac.il

Digital Object Indentifier 10.1109/TPAMI.2008.147

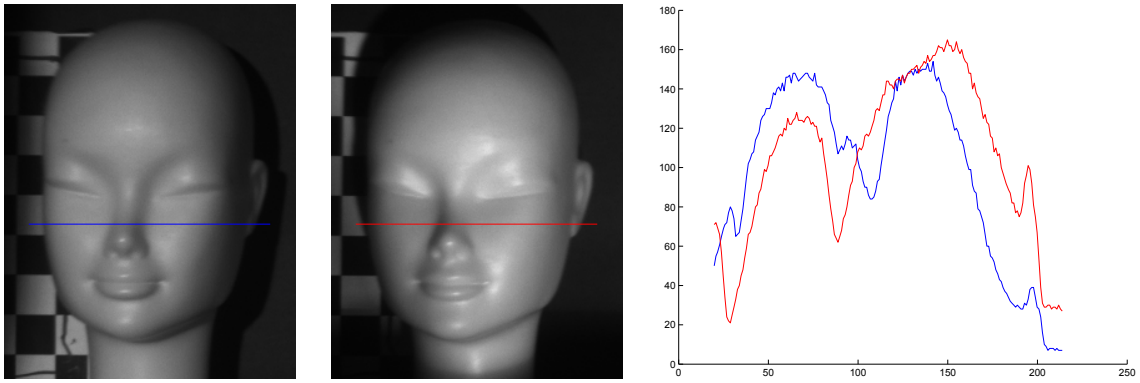

Fig. 1. This figure demonstrates that the intensities of the corresponding points taken under different illumination conditions can differ considerably. The intensity values of corresponding lines in the two images, **(a)** and **(b)**, are shown in **(c)**. If the two images had also been taken from different viewpoints, as in the general setup, an additional non-uniform foreshortening effect would have taken place as well.

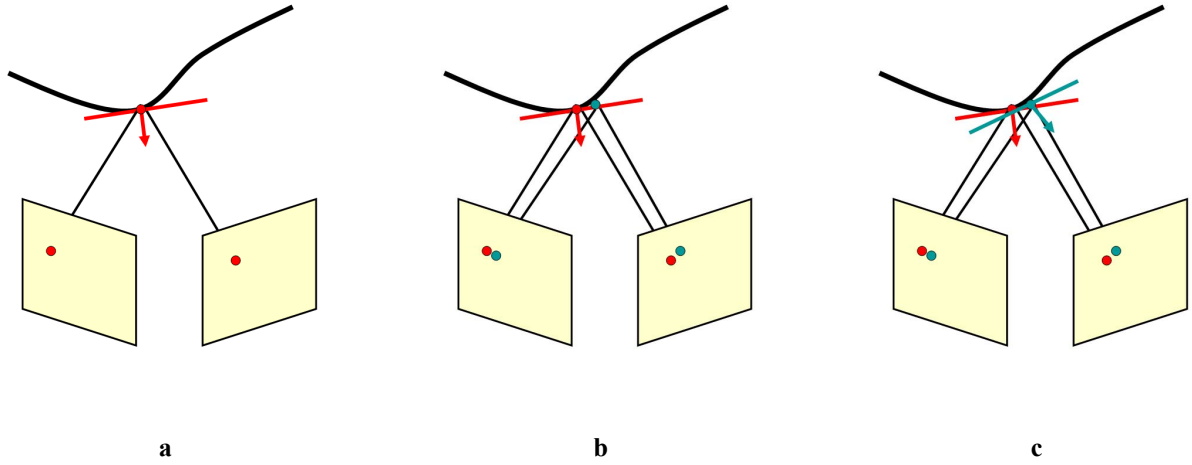

Fig. 2. A simplified illustration of the algorithm. **(a)** Given a pair of corresponding points, the 3D location is computed using triangulation and the normal at that point is computed using photometric stereo. The normal and the 3D point location determine the tangent plane to the surface at that point (the red line). **(b)** A point on this plane is then re-projected to the images and used to compute a new set of corresponding points (the blue points). **(c)** The new correspondence determines a new normal and a new tangent plane (the blue line).

information to obtain support for the reconstructed 3D shape.

Running our algorithm on synthetic and real data reveals that the combined photometric and geometric constraints offer us a powerful tool for handling general images of smooth surfaces. Our one-pass results can be used as a starting point for higher level iterative methods for shape recovery (e.g., [34], [16]). Moreover, since the surface shape is overdetermined when both photometric and geometric information are used, we expect that in future work the camera and light parameters will also be obtainable directly from the images.

The paper continues as follows. We first elaborate on previous related work in Section I-A. In Section II we present the basic derivation of our method and the special case of fixed viewpoint (photometric stereo). Extensions to the basic method are discussed in Section III. We then present experimental results run on simulated and real data in Section IV. Future research directions are discussed in Section V.

A preliminary version of this work was published in [30].

A. Previous work

The direct method for recovering the shape of a smooth Lambertian object from a single image is the shape-from-shading method proposed by Horn [16]. Since the problem is ill-posed, strong assumptions such as smoothness, orthographic projection, and constant albedo must be used to compensate for the lack of information. Numerical methods which solve partial differential equations were developed to enforce the smoothness and the photometric constraints [19], [34]. Please refer to [8], [25], [44] for surveys on this subject. Shape-from-shading algorithms were recently extended to deal with the more accurate perspective projection model [32], [38]. This projection is also used in our algorithm. Shape-from-shading algorithms usually have to be provided with the light source direction and intensity. This requirement has been lifted in [34], [45], which recover the light source information together with the 3D shape. One type of shape-from-shading algorithm starts at a special point at which the normal to the surface is known and propagates the recovered local shape over the whole surface in one pass [6], [21], [23], [22],

[4], [33]. Our method also makes use of one-pass shape propagation. However, the shape information is recovered from three or more images, while the shape-from-shading approach relies only on a single image. As a result, we can relax the assumption of fixed albedo, and obtain more accurate results. This is not surprising: when more images of the surface are used, more information becomes available.

Photometric stereo methods [40], [31], [25] compute the surface normal at each point on the surface without relying on smoothness constraints by using only the local intensity value at the point. This is done by taking several images from the same viewpoint but with different lighting directions. Once the normal has been computed everywhere, the shape is recovered by integrating the normals over the entire image. Thus, the normals are computed more accurately than in shape-from-shading algorithms but the error increases in the final step of the algorithm because no geometric information is available. The quality of the results can be improved by using global integration techniques such as multi-grid [24]. These algorithms have to be supplied with the lighting information about the different images and use the orthographic projection model. Several works have been published that extend the basic method and drop several of its simplifying assumptions [5], [27], [42], [14], [9], [11]. One method for improving the quality of the recovered shape is to supply the algorithm with a small number of anchor points for which the 3D coordinates are known. These points can be obtained using standard stereo methods [17], [18], [12]. Traditional photometric stereo methods assumed the orthographic projection model. A perspective projection photometric stereo algorithm was recently proposed by the authors of [37]. This is an extension of their previous work on shape from shading under perspective projection [38]. [7] proposes an improved method for normal integration for orthographic and perspective photometric stereo. A special case of our method where all images are taken from the same viewpoint provides an alternative, straightforward solution to this problem.

Multiview stereo deals with the problem of 3D reconstruction of smooth Lambertian surfaces taken from multiple viewpoints under fixed illumination. These methods often use a large number of images taken by placing the object on a rotating turntable and rely on the occluding boundaries of each image and photometric information. All this information is used as input for various optimization techniques. In [10] a variational method based on solving PDE's using level sets is presented. The intensity values in the image are used in a smart way only for evaluating the quality of correspondences on the featureless surface. [20] uses a variational optimization method to recover the shape, the lighting, and a piecewise constant albedo. In [41], the visual hull is used as a coarse shape approximation represented as a mesh. The algorithm alternates between improving and refining the mesh and estimating the view independent reflectance model parameters. The reflectance model consists of diffuse and specular components and unknown lighting parameters. Although it seems that these methods could be extendable to deal with varying illumination this extension was not attempted by the authors. In [26] a global continuous optimization scheme for solving this problem is proposed.

Recent studies address the problem of shape recovery of smooth surfaces under non-fixed viewpoint with limited illumination variations. In [36], [29], the shape of a moving object is recovered. The relationship between the changes in lighting and viewing conditions is exploited to yield a modified stereo algorithm. In [43], [28], an iterative scheme is introduced which is able to recover the 3D structure and the camera motion under the same settings. A work that uses an experimental setup similar to ours was presented in [39]. The photometric constraints were only used to verify the 3D structure but not to locally estimate the surface and propagate the correspondences. The advantage of this method is that the surface does not have to be almost entirely smooth. Nonetheless, many more images are needed to determine the 3D shape. An example of a method that fully exploits the photometric constraints is given in [2], where a variational method implemented as a PDE-driven surface evolution interleaved with reflectance estimation is described. Impressive results are obtained by [15] which uses a similar setup. In a preprocessing step the camera positions and the distant point light source directions are estimated from the images. Then, the shape is reconstructed using the silhouettes of the object and the photometric stereo constraints. The algorithm alternates between estimating the 3D surface and computing the normals using photometric stereo. Our method differs from this work in several aspects. We use a relatively small number of images (about 5), we propagate the shape starting from a single point, and do not rely on occluding contours.

The only similar work which strongly integrates geometric and photometric constraints in shape recovery propagation was presented in [35]. There we addressed the special problem of recovering the shape of a smooth, bilaterally symmetric surface from a single image. The geometric and photometric constraints were integrated to compute correspondence between the two halves of the symmetric surface, and hence to compute its 3D shape. In both works the reconstruction starts from a single point and does not rely on the visual hull of the object created from silhouettes of many images to constrain the reconstruction.

II. The basic approach

Consider n calibrated images of a featureless surface taken from different known camera positions under different known lighting conditions. The number n must be sufficient for local normal recovery from photometric information when correspondence is known. The method used to recover the normal is irrelevant for our purposes, as long as it can be locally computed from the intensities of a set of corresponding points in the n images. Here we assume perspective projection and a Lambertian surface. We use classic photometric stereo for recovering the local normal. In this case, $n \geq 3$ images suffice (since we do not assume constant albedo).

In this section we show that a single set of corresponding points is sufficient for propagating the correspondence over the entire image and computing the 3D shape.

A. Notation

The following notations are used in existing geometric and photometric image analysis methods to recover 3D structure.

Geometry: Let M_i, $1 \leq i \leq n$ be the known perspective projection matrices of the n images. Given a 3D surface point $P^{(0)}$, the projection of the point to the n images is given by:

$$p_i^{(0)} \cong M_i P^{(0)}, \quad 1 \leq i \leq n. \tag{1}$$

(Throughout the paper we denote by superscript the point number and by subscript the image number.) The inverse problem is to recover $P^{(0)}$ given its projections to the images $p_1^{(0)}, p_2^{(0)}, \cdots, p_n^{(0)}$ using geometric stereo. In this case, each instance of Eq. 1 can be converted into two linear equations in the coordinates of $P^{(0)}$. Thus, when given two or more projections of a point, its 3D position can be recovered using least squares, or by finding the optimal corrected position of the input points [13].

Photometry: Let l_i, $1 \leq i \leq n$ be the known lighting vectors, where each l_i is pointing to the light source and the magnitude of each l_i is the light source intensity. Denote by $\mathbf{L} = [l_1, \cdots, l_n]^T$ the matrix of all lighting vectors.

Let $P^{(0)}$ be a surface point whose normal and albedo are given by the vector $N^{(0)}$. The direction of $N^{(0)}$ is the normal direction at the point $P^{(0)}$ and its magnitude is the albedo at that point. Denote by $I_i^{(0)}$ the intensities at $p_i^{(0)}$ for $1 \leq i \leq n$. The vector $I^{(0)} = \left[I_1^{(0)}, \cdots, I_n^{(0)}\right]^T$ is the intensity vector of the corresponding points. Under the Lambertian model, where $I_i^{(0)} = \ell_i N^{(0)}$, we obtain:

$$I^{(0)} = L N^{(0)}. \tag{2}$$

Thus, when L and $I^{(0)}$ are given, $N^{(0)}$ can be recovered using a least squares procedure (photometric stereo). That is,

$$N^{(0)} = L^{+} I^{(0)}, \tag{3}$$

where L^+ is the pseudoinverse of L.

B. Combining photometry and geometry

Here we propose our basic method for combining photometric and geometric constraints for computing the 3D shape, starting from a single given correspondence and propagating the computation over the entire image. Extensions of the basic method are presented in the next section.

Given a corresponding set of points, $p_i^{(0)}$, $1 \leq i \leq n$, we can compute the surface point $P^{(0)}$ (by geometric stereo, using Eq. 1) and its normal to the surface $N^{(0)}$ (by photometric stereo, Eq. 3). The task then is to compute a new point on the surface, $P^{(1)}$, based on $P^{(0)}$ and $N^{(0)}$. Consider a small step, δ, on the first image to a neighboring point $p_1^{(1)} = p_1^{(0)} + \delta$. The new surface point that projects to the image point $p_1^{(1)}$ must lie on the ray of points that projects to the point $p_1^{(1)}$. This ray is given by:

$$P(\alpha) = (1-\alpha) O_1 + \alpha P_\delta, \tag{4}$$

where O_1 is the known center of projection of the first camera, $[P_\delta, 1] \cong M_1^{+} p_1^{(1)}$ is a point that projects to $p_1^{(1)}$, and α is a scalar. Any value of α uniquely determines a 3D point in space that projects to $p_1^{(1)}$. That is, $P^{(1)} = P(\alpha)$ for some α. Since we are looking for $P^{(1)}$ on the surface, α can be determined using a local surface approximation that is computed from the surface normals.

Here we consider first order approximation, where the surface is assumed to be locally planar (see Figure 2 for an illustration of the algorithm). The surface is locally approximated by the point $P^{(0)}$ and the tangent plane to that point, which can be computed using the normal $N^{(0)}$. For the point $P^{(1)}$ to lie on the tangent plane at $P^{(0)}$ it must satisfy:

$$N^{(0)}(P^{(1)} - P^{(0)}) = 0. \tag{5}$$

The two constraints on $P^{(1)}$, Eq. 4 and Eq. 5, yield a unique value for α, α_1:

$$\alpha_1 = \frac{(P^{(0)} - O_1)^T N^{(0)}}{(P_\delta - O_1)^T N^{(0)}}. \tag{6}$$

Once $P^{(1)} = P(\alpha_1)$ has been estimated, its projection to all the images $p_i^{(1)}, 1 \leq i \leq n$ is computed, yielding a new set of corresponding points. The surface normal to this point, $N^{(1)}$, can be computed using Eq. 3. Note that any other method for normal recovery based on shading information can be used here instead of the classic photometric stereo method. For highly curved regions, first order approximations might not be sufficient; we consider such cases in Sections III-C and III-D.

To recover the entire surface, the basic scheme starts with a set of corresponding points, and propagates the correspondence as described above. The order of propagation is not expected to affect the recovered surface as long as the data is noise free. To reduce error accumulation due to noise, the propagation order is chosen such that the length of the path between the original pixel and the target pixel is minimal. This can be done by propagating the correspondence in all directions from the given pixel, that is, by a breadth-first search (BFS) traversal of the image. This traversal also circumvents regions for which the propagation cannot be computed (e.g., shadowed pixels).

Our method can be regarded as a traversal of the 3D surface. In our implementation we use a reference image to perform this traversal. The use of a reference image may cause problems when the object is self-occluded. In this case, two or more surface points are projected to a single point on the reference image and only one of these surface points is recovered. Since our method can identify occluding (or self occluding) boundaries, a switch of the reference image may overcome this problem. In addition, in our implementation the size of the step δ is of a single pixel in the reference image. The value of δ can be modified reducing it close to the occluding boundary or at regions with high curvature in order to increase the accuracy. Again, switching to a different reference image, or more generally performing the traversal on the surface itself (as done in [41], [20]), may handle this problem as well. These solutions were not implemented in this paper but can be done in a straightforward manner.

C. Photometric stereo

A special case of our setup is when all images are taken from the same viewpoint. In this case the setup is identical to perspective photometric stereo. In photometric stereo, the normals are estimated from the intensity of corresponding points. In this case the correspondence between image points is trivial. The depth of a point is then computed by some integration method using these normals. The contribution of our method in this case is reduced to the integration process, and to basic error correction (as explained in the next section). In our method, the integration is performed under perspective projection using Eq. 6, as illustrated in Figure 2. Independently, [37] also derived a perspective photometric stereo algorithm, presenting a new method to transform the object coordinate system to the image coordinate system when the perspective projection model is considered.

A comparison of the photometric stereo setup (fixed viewpoint) to the general viewpoint setup reveals several differences. The fixed viewpoint assumption results in a trivial solution to the correspondence problem. Therefore, errors due to incorrect correspondence are eliminated. However, dropping this assumption results not only in a more general setup, which can be used on arbitrary images, but also has several computational advantages. For example, occluded regions in one view can be recovered from a different view. In addition, the overdetermined 3D information from both geometric and photometric data can be used to correct the 3D reconstruction (see Section III-D). Experimental results showing the superiority of the general case are presented in Section IV.

III. Method extensions

The reconstruction method of the basic approach is expected to be sensitive to noise, shadows, and self-occlusions. In particular, since the method is strongly based on propagation, a single error can affect the rest of the reconstruction. In this section we suggest several extensions of the basic method to improve the quality of the reconstruction and reduce the negative impact of noise. It includes using a larger number of input images, computing local error corrections, and considering a higher order approximation of the surface. Finally, we present a method for automatically estimating the initial 3D point used for starting the propagation.

A. Multi-neighbor propagation

In the basic scheme, the correspondence for a given target pixel is estimated using a single neighbor whose 3D location and normal have already been computed. As a result, a single error can affect the rest of the reconstruction. To reduce the errors caused by noise, we use more than a single neighbor to compute the target pixel. A set of candidate locations for the target point is computed from each of its already computed neighbors (in a 8-pixel neighborhood). One possibility is to choose the best location from this set using a score function (see Section III-D). In our scheme, called *multi-neighbor* propagation, the average of the set of locations is chosen as the target point location. As we showed experimentally in Section IV, using multi-neighbor rather than single neighbor propagation improves the reconstruction considerably.

The traversal order is also expected to affect the result. The basic propagation was based on a BFS traversal. The traversal order in this case was dictated by the shortest distance to the initial corresponding point. Since our multi-neighbor propagation uses all of the already computed target point location neighbors, we should choose the traversal order in which, for each target pixel, as many neighbors as possible have already been computed. To choose the traversal order that meets these conditions, the eight neighbors of a given computed target pixel are inserted into the queue. First we insert the four vertical and horizontal neighboring pixels and then the four neighbors on the diagonal. It can be shown that each target pixel will have at least three pre-computed neighbors.

Other choices for the image traversal can improve the reconstruction. In particular, we could set a higher priority for points that are expected to be more reliable, using the score function defined below (Section III-D). In addition, the planarity of a region determines the expected reliability of its reconstruction. Since curved surface regions correspond to image regions with high gradients, higher priority can be assigned to regions with a low intensity gradient. Finally, image regions containing extreme intensity values are expected to be unreliable, since they may contain image highlights. Again, these regions can have lower priority in the traversal queue.

In general, an error term can be computed for each possible candidate for propagation using the scores defined in Section III-D. At each step, the propagation will be performed at the pixel with the lowest score. Then the score of its neighbors will be updated.

B. Using more than three images

A minimum of three intensity values is required for recovering the normal of a given surface point and its albedo using photometric stereo. Therefore, three images were used in the basic scheme. Our method can be easily extended to use a larger set of input images because the same solution to the correspondence problem will also work on a larger number of images. The information available in a larger set of images improves the scheme's robustness and allows occlusions and shadows to be handled without introducing holes in regions for which a set of three images does not include enough information.

When each surface point is visible and not shadowed in all the images in the set, the normal can be computed from more than three intensity values by using Eq. 2. Such a least squares computation improves the system's accuracy. The least squares error,

$$s_{int}^{(0)} = |I^{(0)} - LN^{(0)}|, \tag{7}$$

can also be used to locally evaluate the quality of the estimated normal. Using Eq. 2, we obtain:

$$s_{int}^{(0)} = |I^{(0)} - LL^{+}I^{(0)}|. \tag{8}$$

A surface point is not necessarily visible and illuminated in all images. For non-convex surfaces in particular, occlusion and self-shadowing are expected to occur in some of the images for some of the surface points. When the set of input images is larger than three, a subset of at least three images from which the normal and the albedo can be recovered is sufficient for each surface point. Thus, a larger part of the surface can be recovered. To define the subset of images for a given surface point we have to detect for each of the images shadowed and saturated pixels, as well as occluded pixels. The shadowed points are easy to detect since their intensity value is close to zero. Similarly, we can also avoid using pixels whose intensity values are close to the saturation values of each image.

Occluded points or noisy pixels can be detected using the score function given in Eq. 8. When the score is higher than a given threshold, one of the following cases is implied:

- One or more of the intensity values is noisy.
- The 3D surface point is invisible in one of the images due to occlusion, and therefore its intensity value in that image is the intensity of the 3D point that occludes it.
- The 3D point is incorrect, and therefore its projection to the images results in an arbitrary set of intensities.

In the first two cases, a subset of the intensity values can be used if the score obtained in Eq. 8 is smaller than the threshold. Note that if the size of the set is three, the score will always be zero. Therefore we are looking for the largest set such that the score is still smaller than the threshold. For the last case, a local error correction can be applied, as will be discussed in Section III-D.

The benefits of using more than three images are demonstrated experimentally in Section IV.

C. Second order approximation

For highly curved regions, first order approximations might not be sufficient. We therefore consider a more general case, where the surface curve connecting two points is locally circular. Under this assumption, the average of the normals at $P^{(0)}$ and $P^{(1)}$, $N^{(0)}$ and $N^{(1)}$, must be orthogonal to the vector connecting them:

$$(N^{(1)} + N^{(0)})^T (P^{(1)} - P^{(0)}) = 0.$$

In this case α must satisfy

$$(N(\alpha) + N^{(0)})^T (P(\alpha) - P^{(0)}) = 0, \tag{9}$$

where $N(\alpha) = L^+ I(\alpha)$ and $I(\alpha)$ is the vector of intensities of the projection of $P(\alpha)$ on the n images.

We next show that Eq. 9 can be used to compute the value of α which satisfies it. The value of α can be computed using α_0, the first order approximation computed using Eq. 6, and the image derivatives in the epipolar direction at each of the images.

Let $\alpha = \alpha_0 + \delta_\alpha$. The first order approximations for $I(\alpha)$, $N(\alpha)$, and $P(\alpha)$ are given by:

$$\begin{aligned} I(\alpha) &\approx I(\alpha_0) + \delta_\alpha \frac{\partial I(\alpha_0)}{\partial \alpha} && (10)\\ N(\alpha) &= N(\alpha_0 + \delta_\alpha) \\ &\approx L^+(I(\alpha_0) + \delta_\alpha \frac{\partial I}{\partial \alpha}(\alpha_0)) \\ &= N(\alpha_0) + \delta_\alpha L^+ \frac{\partial I}{\partial \alpha}(\alpha_0) && (11)\\ P(\alpha) &= (1-\alpha)O_1 + \alpha P_\delta \\ &= (1-\alpha_0-\delta_\alpha)O_1 + (\alpha_0+\delta_\alpha)P_\delta \\ &= (1-\alpha_0)O_1 + \alpha_0 P\delta + \delta_\alpha(P_\delta - O_1) \\ &= P(\alpha_0) + \delta_\alpha(P_\delta - O_1). && (12)\end{aligned}$$

Substituting the above terms into Eq. 9 yields:

$$\begin{aligned}(N(\alpha_0) + \delta_\alpha L^+ \frac{\partial I}{\partial \alpha}(\alpha_0) + N^{(0)}) \cdot (P(\alpha_0) + \\ \delta_\alpha(P_\delta - O_1) - P^{(0)}) = 0.\end{aligned}$$

Since all the terms in the above equation except δ_α can be computed in advance, we are left with the following quadratic equation in δ_α:

$$\begin{aligned} L^+ \frac{\partial I}{\partial \alpha}(\alpha_0)(P_\delta - O_1)\delta_\alpha^2 \\ +(L^+ \frac{\partial I}{\partial \alpha}(\alpha_0)(P(\alpha_0) - \\ P^{(0)}) + (P_\delta - O_1)(N(\alpha_0) + N^{(0)}))\delta_\alpha \\ +(N(\alpha_0) + N^{(0)})(P(\alpha_0) - P^{(0)}) = 0. \end{aligned} \tag{13}$$

The second order approximation can be used directly for estimating α. We, however, use it as part of a score function that evaluates the quality of the local shape as described below.

D. Local evaluation and local correction

We base our definition of a score function on the continuity assumptions of the local neighborhood of a target pixel. This score is used for improving the surface reconstruction where a new target point is computed.

The score function is defined as a weighted sum of several subscores. The first subscore is based on the intensity values of the projection of the target point to the set of images, as defined in Eq. 8. We next define subscores which are based on the already computed neighbors of the target pixel. Let $neighb$ be this set of pixels. Let $N^{(i)}$, ρ_i, and P_i be the computed normal, albedo, and 3D location computed for each $p_i \in neighb$. Similarly, let $N^{(t)}$, ρ_t, and P_t be the candidate normal, albedo, and 3D location of the target point. We can now define the local continuity subscores:

Continuity of the normal direction: the angle between the computed target normal and the average normal of its already computed neighborhood,

$$s_{direc} = 1 - \frac{1}{|neighb|} \sum_{p_i \in neighb} |N^{(i)} N^{(t)}|. \tag{14}$$

Continuity of the albedo: the average difference between the computed target albedo and the albedo of its already computed

neighborhood,

$$s_{albedo} = \frac{1}{|neighb|} \sum_{p_i \in neighb} |\rho_i - \rho_t|. \quad (15)$$

Continuity of the shape: the 3D distance between the computed target point and P_{mean}, the average 3D location of its already computed neighborhood,

$$s_{loc} = ||P_{mean} - P_t||. \quad (16)$$

Finally we define a subscore that measures the consistency between the 3D location of the computed points and their normals. This score is based on the analysis in Section III-C.
Consistency of 3D point location and shape normal: The normal to the target point and the 3D location of the target point should be consistent with a second order approximation of the shape. This does not necessarily hold when the target point is computed from the first order approximation. In that case it can be used as a measure for evaluating the target point,

$$s_{shape/norm} = \sum_{p_i \in neighb(t)} \frac{(N^{(p_i)} + N^{(t)})}{||N^{(p_i)} + N^{(t)}||} \frac{(P_i - P_t)}{||P_i - P_t||}. \quad (17)$$

For each subscore, a different threshold is defined, using the subscore's expected values. For example, th_{direc} is the threshold chosen for the continuity of the normal direction. The final score is then defined by:

$$s = \frac{s_{int}}{th_{int}} + \frac{s_{direc}}{th_{direc}} + \frac{s_{loc}}{th_{loc}} + \frac{s_{albedo}}{th_{albedo}} + \frac{s_{shape/norm}}{th_{shape/norm}}. \quad (18)$$

A high error score for a target point implies that either the entire reconstruction so far is incorrect or that the reconstruction of that point is incorrect. Since our method is based on propagation, a single point can affect the rest of the computation. We considered three possibilities for using this score to improve the computation:
Remove unreliable points: target pixels with a high score are removed from the computation. As a result, the full shape is not reconstructed (see Figure 11).
Compute a new target normal: The score of a target point is reduced by choosing a subset of images to compute the new normal.
Correct target point location: The score of a target point is reduced by modifying its 3D location. We search for a 3D point located on the ray whose points project to the target pixel, which minimizes the score. The 3D location of the point is defined by α. We search for the value of α in the range that satisfies $s_{loc} < th_{loc}$. We compute the range of α from the average location of its already computed neighbors P_{mean}. Using Eq. 4, the range of α is the solution of the following quadratic equation in α:

$$((1-\alpha)O_1 + \alpha P_t - P_{mean})^T$$
$$((1-\alpha)O_1 + \alpha P_t - P_{mean}) < th_{loc}^2.$$

In our experiments, the influence of the score function on the computation is demonstrated on synthetic and real images. Note that the score described here is based only on the close neighborhood of a target point. In the next section we describe a score that is based on a larger region. This score is used for automatic detection of the starting point.

E. Automatic detection of the starting point

The starting point of the propagation process can be automatically detected using the integrability constraint. The basic idea is to choose a pixel in one image as an initial point and search for its correct 3D location. The 3D location is determined by a single parameter α, which is the distance along the 3D ray of the image point. We use the integrability constraint to evaluate a given α. This is done in the following manner. We select a closed path in the image starting and ending at the selected pixel. Using the basic scheme, the 3D depth is computed for all pixels on the path starting with the initial point whose depth is determined by α. The integrability constraint stipulates that the depths computed for the first and last pixels in the closed path (which are both the initial point) are the same. Thus, the score of each value of α is the difference in depths between the two points. The algorithm selects the α which minimizes this difference. We select the path to be in a relatively smooth region, where the basic method can be used to recover the 3D shape and the normals quite accurately even though error correction is not performed along the path. For robustness, several different closed paths are used in order to determine the optimal starting point. The initial 3D point for this part of the algorithm is chosen to be a point on a relatively smooth region in one of the images, and the range of α is chosen such that the projection of the 3D point is in the range of all the images. In Section IV-B we show that this method can accurately estimate the correct initial point.

IV. Experimental results

We have implemented several variants of our method and tested them on simulated and real image sets. Simulated image sets enable us to compare the results with the ground truth. Running the algorithm on a real image set is more challenging because we have to supply the algorithm with projection matrices and lighting information, which also have to be recovered from the images. In addition, the algorithm has to deal with image noise and the inaccuracies of the Lambertian model.

A. Simulated images

We generated several sets of simulated images of a bust of Mozart from a 3D range image which was used also in [32]. We chose this object since it has considerable shape variability - high curvature, low curvature, self occlusion etc. The geometric and photometric properties of the generated images match the model used by our algorithm (a perspective projection image of a Lambertian surface with a distant point light source). The size of each image is 280×280. The images of the first set were generated as being shot from different viewpoints and illuminated from different directions. The images and their parameters are presented in the first row of Figure 3. Four variations of this set of images were used for testing various aspects of our method. This includes adding noise (e.g., Figure 4**a**), changing albedo (second row of Figure 3), adding noise and albedo change (e.g., Figure 4**b**), and adding a synthetic shadow (e.g., Figure 4**c**). For testing

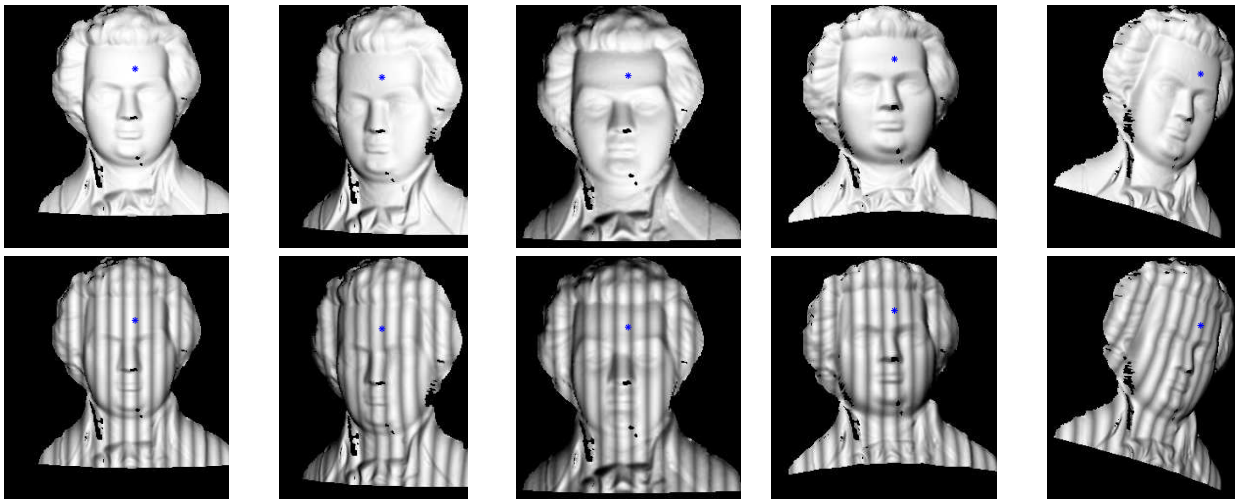

	Image 1	**Image 2**	**Image 3**	**Image 4**	**Image 5**
l :	$(0, 0, -1)$	$(0.45, 0, -0.89)$	$(0.18, -0.44, -0.88)$	$(-0.28, 0.19, -0.94)$	$(-0.57, 0, -0.82)$
θ :	$(0, 0, 0)$	$(0.2, 0.3, 0)$	$(0, 0, 0)$	$(-0.2, 0, 0)$	$(0, -0.5, 0.3)$
T :	$(0, 0, -1000)$	$(300, 0, -1000)$	$(40, -40, -1000)$	$(-140, -400, -1000)$	$(-500, -200, -800)$

Fig. 3. **First row** The five synthetic noise-free images with fixed albedo used in the experiments. The starting corresponding point is marked by a blue star. **Second row** The five synthetic noise-free images with non-fixed albedo used in the experiments. The illumination vector, l, the rotation vector, $\theta = (\theta_x, \theta_y, \theta_z)$, and the translation vector, T, are given for each image.

Fig. 4. The three lines superimposed on (a) indicate the location of the cross-sections presented in Figure 7. (a-b) examples of Gaussian noise added to **Image 1** in the first ($\sigma = 30$) and second ($\sigma = 20$) rows of Fig 3. Similar noise was added to all the images in Figures 3 & 5. (c) the "shadow" patch added to **Image 1**.

Fig. 5. The five noise-free synthetic images used in the photometric stereo experiments. Noise similar to Figure 4a was also added to this set.

our photometric stereo algorithm, we generated two sets of images all being shot from a single viewpoint and illuminated by the point light sources used in the first set. The first set was noise free (see Figure 5) and the second set had additive Gaussian noise with $\sigma = 30$ (see Figure 4**a**).

The basic approach and its extension were implemented in Matlab. In all cases only a first order approximation of the surface was used. The algorithms received the parameters (cameras and illumination) that were used to generate the images, and an initial known 3D point to initialize the propagation. This point is marked by a blue star in each of the images in Figure 3. The reconstructed surfaces for the various algorithms are presented in Figure 6. Cross-sections of the real 3D structure and the recovered ones are then plotted for a more precise comparison in Figures 7 & 9. We next elaborate on the different experiments we performed on the set of synthetic

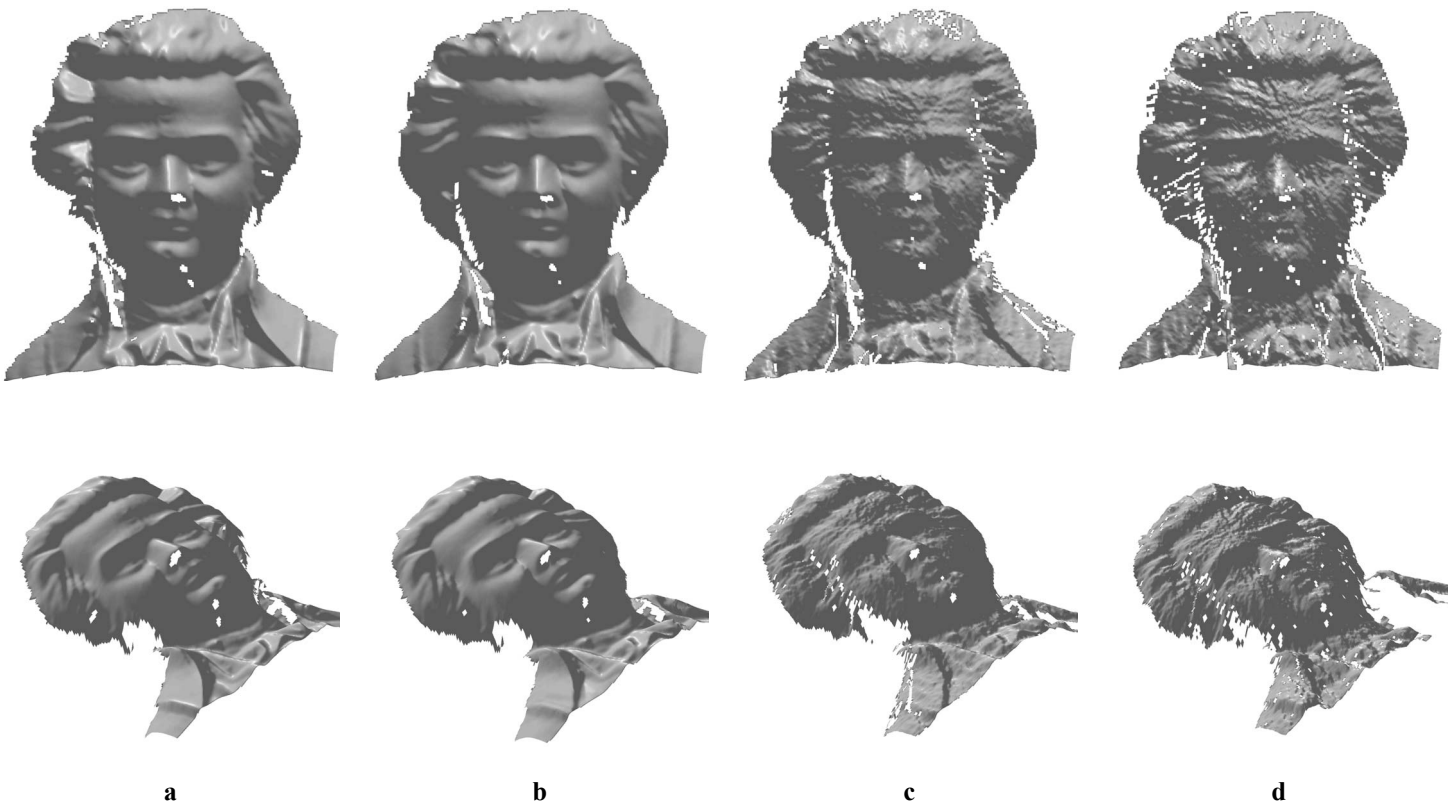

Fig. 6. 3D surfaces obtained by (**a)** the basic scheme run on three noise free images; (**b)** the full scheme run on five noise free images; (**c)** the full scheme run on five noisy images; (**d)** photometric stereo run on five noisy images.

images and their results.

[a] Basic scheme on noise free images: We compare the results of our basic algorithm (without using the score to remove points or correct errors) on three and five images. The propagation begins from the initial point. Cross-sections through the reconstructed surfaces are presented in the first row of Figure 7. The reconstructed surface closely resembles the original data. Note that this is true even for recovered regions far from the starting point (see cross-section C). The reconstruction starts to drift in the hair area of the Mozart head (the left part of cross-section B). This is the result of propagating the 3D shape over unreliable regions (the edge between the face and the hair). When five images instead of three are used, the reconstruction of the hair area is improved, but bears only a distant resemblance to the original data. Thus, error correction is needed. The reconstructed surface obtained by the basic algorithm running on three noise-free images is shown in Figure 6**a**.

[b] Error correction on noise free images: In the second experiment, we apply the error correction on three and five images. The parameters of the score function are: $t_{int} = 8$, $t_{direc} = 0.1$, $t_{loc} = 3$, $t_{albedo} = 0.1$ and $t_{shape/norm} = 0.1$. The global threshold is 6. The second row of Figure 7 shows that error correction improves the results considerably in the problematic hair region even when only three images are used. When five images are used, the results are almost perfect.

[c] Error correction on images with Gaussian noise: In these experiments, we used the first set of images to which Gaussian noise was added (see Figure 4**a** for an example of one of these images). The Gaussian noise was independently added to each pixel, with $\sigma = 30$. The results, are shown in the third row of Figure 7. For the basic algorithm, the results bear little resemblance to the ground truth 3D data. Three different error correction procedures were then compared, using the score functions suggested in Section III-D, and using the same parameters of the score function as in the previous experiments [b]. We first ran the procedure which discards points with overall scores higher than 6, but the reconstruction failed to propagate to regions which are not in the immediate vicinity of the starting point. We therefore do not display this case. We then ran a second procedure. When points with high scores are encountered, it searches for a subset of images which yields a low score value. We use this subset to compute the surface normal. When the score remains high, the point is removed from the computation. In the final run, points for which the first type of correction has failed are corrected by trying to find the optimal depth (value of α). The optimal depth yields the lowest score. Again, points with scores higher than 6 are discarded. The quality of the reconstruction is similar in both cases and is superior to the result of the basic scheme. The advantage of the third type of correction is that the reconstructed part of the face is larger. The reconstructed surface obtained by the full scheme is shown in Figure 6**c**.

We used the same parameters of the score function in all the synthetic experiments. The global threshold is the one which affects the performance of the algorithm. When taken to be

too small, the reconstruction fails to propagate over significant parts of the shape. In the other extreme, when the global threshold is too large, erroneous pixels are reconstructed causing propagation errors as if the error correction mechanism is turned off. The algorithm is not very sensitive to the actual value of the threshold between $6-10$, yielding similar results.

[d] Varying albedo: In this experiment we tested two sets of images which were generated with varying albedo. The albedo used to generate the images is a Sinus function along the x-direction with an amplitude of 0.6 (see second row of Figure 3). The first set consists of noise free images, while the second consists of the same images with added Gaussian noise ($\sigma = 20$). One of the five noisy images is shown in Figure 4**b**. The fourth row of Figure 7 shows the results of our full method on these two sets of images. As can be seen, the reconstruction is not affected when objects with varying albedo are considered. Moreover, the albedo is correctly recovered by our method, as shown in Figure 8**a**.

[e] Shadows: In this experiment, we tested the ability of the algorithm to deal with pixels which do not agree with the photometric model. This occurs, for example, when shadows are cast on the object. We simulate this phenomenon by replacing an arbitrary image patch in one of the images by a constant grey-level (value 70), Figure 4**c**. The algorithm is able to detect and ignore these pixels, while yielding a correct reconstruction using the other images. Figure 8**b** shows the discarded pixels from the first image in the computation. In addition to the patch, other pixels with high levels of noise were also discarded from the computation.

[f] Photometric stereo - fixed viewpoint: The final experiment on synthetic data compares our method run on images shot from a fixed viewpoint (photometric stereo) with the result obtained for images shot from multiple viewpoints. The results are summarized in Figures 6**d** & 9. When noise-free images are considered, the reconstruction obtained for photometric stereo is good for regions close to the initial point but deteriorates in remote regions of the image. This is the result of self-occlusion, which is not dealt with by the normal integration process. When comparing the results for photometric stereo on noisy images ($\sigma = 30$) with the general method, the advantage of using multiple viewpoints is evident (compare to the results plotted in the third row of Figure 7). Using multiple viewpoints provides more reliable information. In particular, the general method can detect and handle self-occluded boundaries and hence propagate correctly the depth over these boundaries. Photometric stereo, in contrast, fails there. In addition, multiple viewpoints handle noise better because the algorithm is able to reason about the depth information of a given point.

B. Real images

We ran our algorithm on five real images of a mannequin taken by a standard CCD camera. The original images are shown in Figure 10. The cameras were internally and externally calibrated using the camera calibration toolbox [3], with the help of a piece of checkerboard paper. The light source direction was estimated using a set of nails and their shadows (see left-hand image of Figure 10). Vectors connecting nail tips to their shadows intersect at the light source position. The 3D tips of the nails were computed by marking their location by hand in a few calibrated images, and then using triangulation. The 3D position of the plane on which the shadows fall was also computed from the projection matrices. Thus the 3D position of the shadows of the tips of the nails can also be computed. The ambient light was determined to be the intensity value in the shadow of the mannequin, and the light source intensity was chosen to be the maximal intensity value on the mannequin after the ambient value was subtracted.

Only the facial part of the images was used in the experiment, as shown on the first row of Figure 12. The images were smoothed with a Gaussian mask, with $\sigma = 2$. The initial corresponding points were chosen by hand and then fine-tuned to minimize the cost function in Eq. 18. An experiment for automatic detection of the initial point is described below.

We present here the results of three versions of our algorithm. The first is the *basic* method, which is based on a single-neighbor propagation with no error correction; The second, more advanced method uses multi-neighbor propagation, and unreliable points, defined by the score function, are removed. Finally, we use the complete scheme with multi-neighbor propagation and full error correction. The parameters used for the score function are $th_{int} = 5$, $th_{direc} = 0.1$, $th_{loc} = 2$, $th_{albedo} = 0.1$ $th_{shape/norm} = 0.1$ and the global threshold is $th_s = 10$.

To evaluate algorithm performance we present in Figure 11 the reconstructed surfaces. In Figure 12 we display the computed correspondence of a set of points in the five images. The correspondence set is computed as the projection of the computed 3D points to each of the images, where corresponding points are marked in the same color. Even in regions which are far from the initial starting point (marked by a black star), the correspondence is good despite the large variations in viewpoint and illumination direction. Finally, several cross-sections of the surfaces recovered by the algorithms are displayed in Figure 13.

Overall, the results demonstrate that the proposed method works very well on real images. The basic scheme already gives a rough 3D shape of the object. Adding multi-neighbor propagation to the basic scheme greatly improves the basic scheme. The score function allows us to stop possible drifting of the 3D shape due to errors. It also causes holes to appear in the reconstructed shape, which stop the reconstruction in unreliable regions. Finally, error correction allows the gaps to be filled in and the reconstructed shape to be extended correctly.

In the final experiment we tested our method for automatic detection of the 3D starting point. The process and the results are shown in Fig 14. An initial 2D point on the forehead in the first image was chosen and marked on Fig 14**a**. The closed cycles were chosen to be rectangles, each marked in a different color on the image. The projection of the search domain is marked on each of the other four images by a yellow line (Fig 14**b-e**). Recall that α (Eq. 4) determines the 3D location

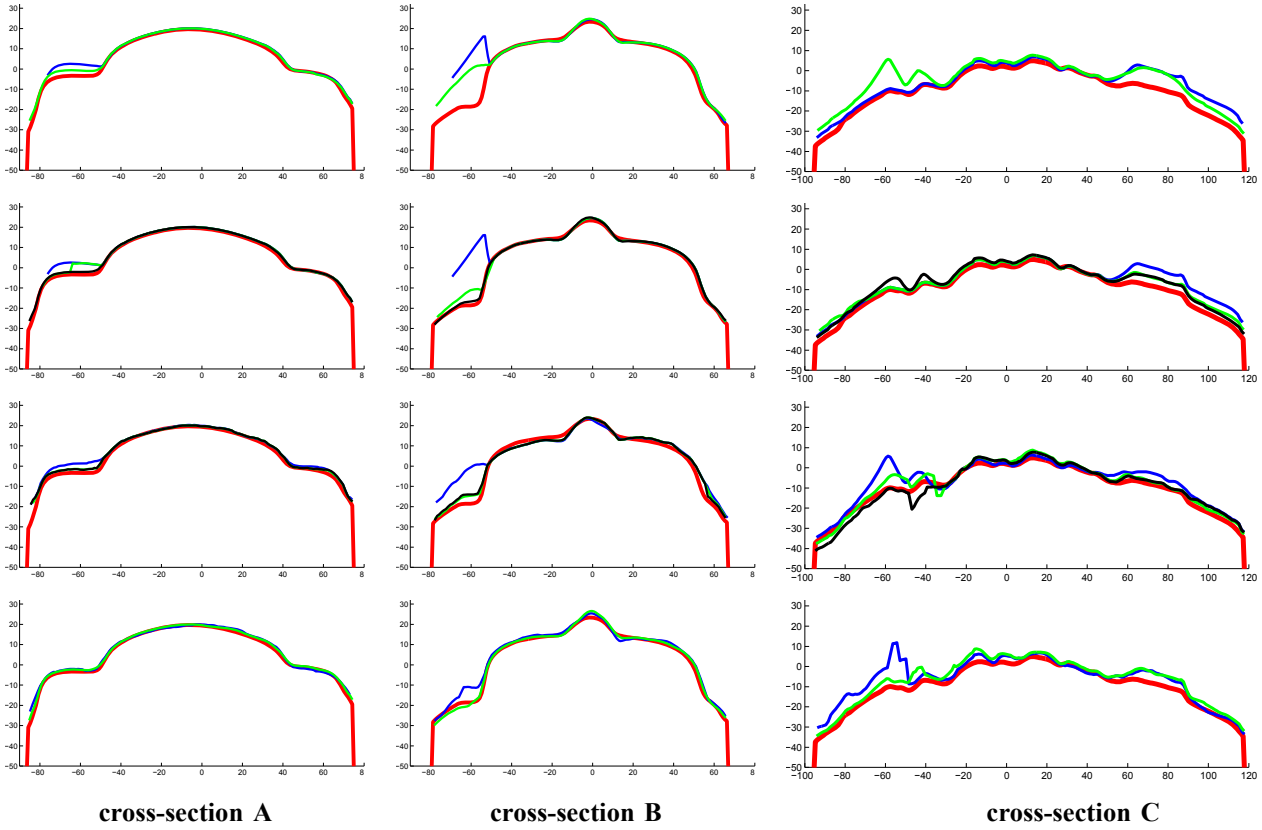

Fig. 7. In this figure we compare different variants of our algorithm running on the synthetic images. Each column is the 3D reconstruction along one of the cross-sections in the 3D shape shown in Figure 4. Each row presents a different experiment, with the red plots representing the ground truth in each. **First row:** the blue and green plots are the results of running the basic scheme running on 3 and 5 noise free images. **Second row:** the blue plot is the result of running the basic algorithm on 3 images and the green plot is the result when error correction was applied. The black plot is the same as the green but when the algorithm was run on 5 images. **Third row:** the blue plot is the result of the basic scheme run on 5 images with noise ($\sigma = 30$). The green and the black plots are the results of applying error correction. The result of changing only the normal at each point is plotted in green, while the black plot consists of correcting the target point location as well when the normal correction is insufficient. **Fourth row:** the blue and the green plots are the results of the full scheme (with error correction) run on the 5 images with varying albedo with and without noise of ($\sigma = 20$), respectively.

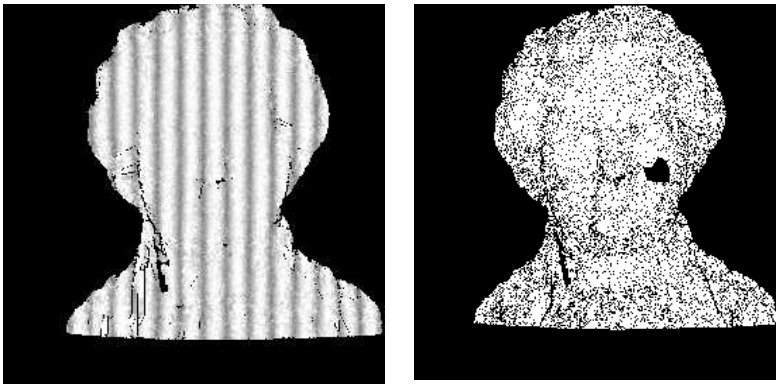

Fig. 8. (a) The recovered albedo in the varying albedo experiment, see image Figure 4(b); (b) the black pixels are those automatically removed from the computation in the shadows experiment, using the image shown in Figure 4(c).

of the starting point. We propagate the 3D location of the points along a closed cycle using the basic approach (the rest of the image was ignored). The 3D distance between the first and the last point along a closed cycle is plotted as a function of α for each of the rectangles (see Fig 14**f**). In general, we search for the minimal value on each of these graphs. As can be seen, three out of the four plots closely agree on the starting point location. The only plot (in pink) that does not agree with the rest is the one that corresponds to a rectangle that passes through relatively high curvature regions. The blue point along the yellow line in each of the images marks the projection of the chosen initial 3D point. Looking at the corresponding points in the four images, we can see that the results are quite accurate.

V. Summary and Conclusions

In this paper we introduced a new single pass shape reconstruction algorithm for smooth featureless surfaces under the

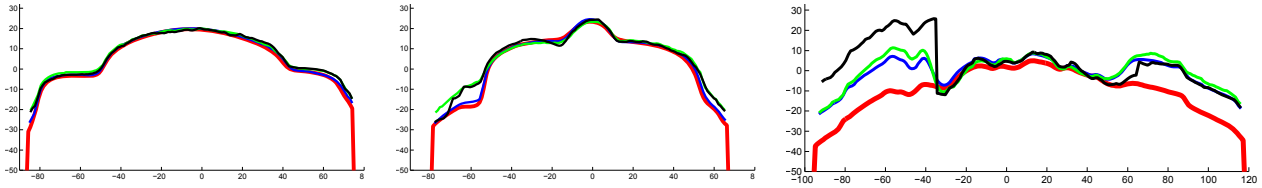

Fig. 9. The photometric stereo experiment: the 3D reconstruction along three cross-sections of the 3D shape as shown in Figure 7. The red plot is the ground truth; the blue plot is the result of running the basic photometric stereo method on 5 noise-free images. The green and black plots are the results of running the photometric stereo method on noisy images ($\sigma = 30$) without and with error correction.

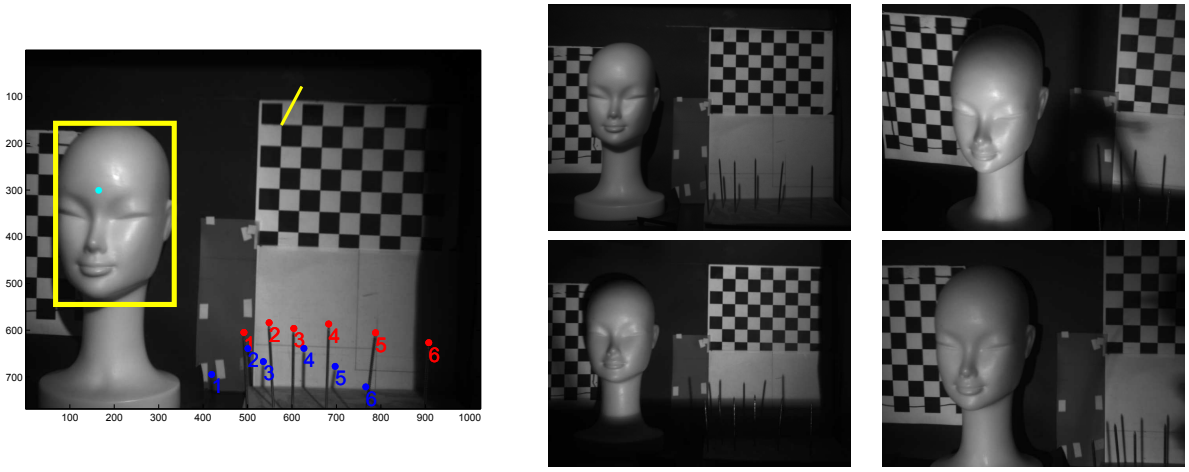

Fig. 10. The real image experimental setup. Calibration paper was used to calibrate the camera location, and a clearly visible subset of the nail shadows were marked by hand to compute the light source direction. The projections of the nail heads are marked in blue and their shadows in red. The yellow line is the projection of the light source position to the image plane. The yellow rectangle is the region on which we run our algorithm.

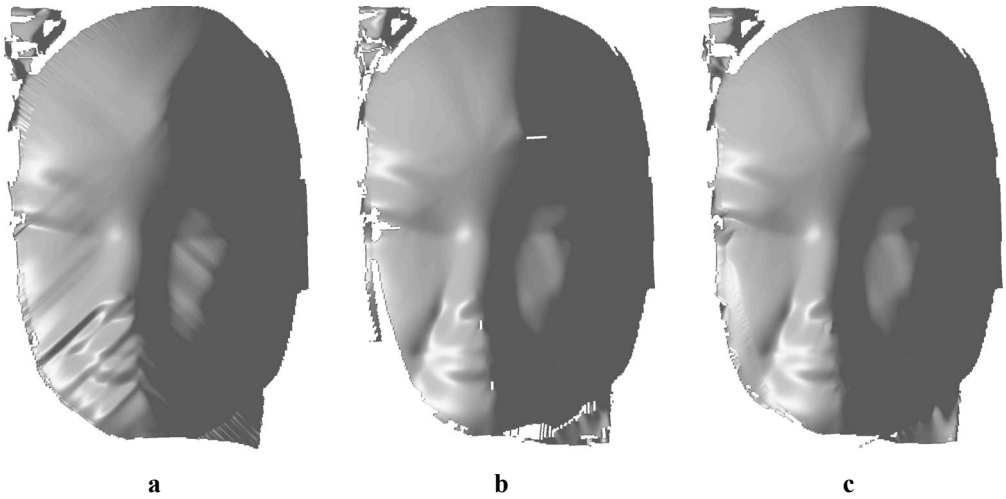

Fig. 11. 3D surfaces obtained by our method. **(a)** Propagation based on a single neighbor without error correction (the naive method); **(b)** and **(c)**, propagation based on multiple neighbors. In **(b)** points that have high scores are removed while in **(c)** points are corrected to have scores below the threshold.

perspective projection model. By enabling independent camera and light source location, an accurate reconstruction algorithm is created. It builds on the strengths of photometric stereo, geometric stereo, and shape-from-shading, where each method alone cannot handle this task adequately.

The algorithm has been tested in realistic settings using an experimental setup that enables us to recover the input parameters to the algorithm from the images. Even though these parameters were estimated from the images, the resulting recovered surfaces were quite accurate. These results can be improved by modifying the reflectance model to deal with specularities and non-distant light sources. This extension is relatively straightforward in our setup because, when the normal to a target point is computed, the 3D location of that point is already known and therefore the lighting direction can be computed locally. A more challenging goal would be to extend the method to deal with other reflectance models. These reflectance models can be easily incorporated within our

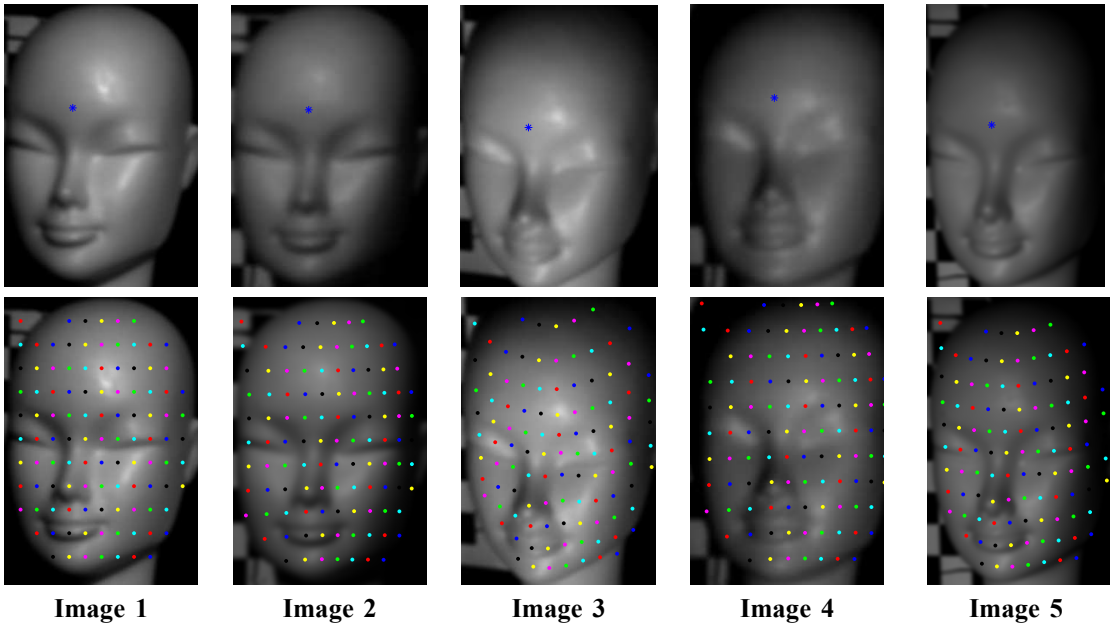

Fig. 12. The first row consists of the five cropped images used in the experiment. Note that each was taken from a different viewpoint and with a different light source direction. The blue star is the corresponding point given to the algorithm. The second row consists of a grid of automatically computed corresponding points on the face images. Corresponding points are marked by the same color on the different images.

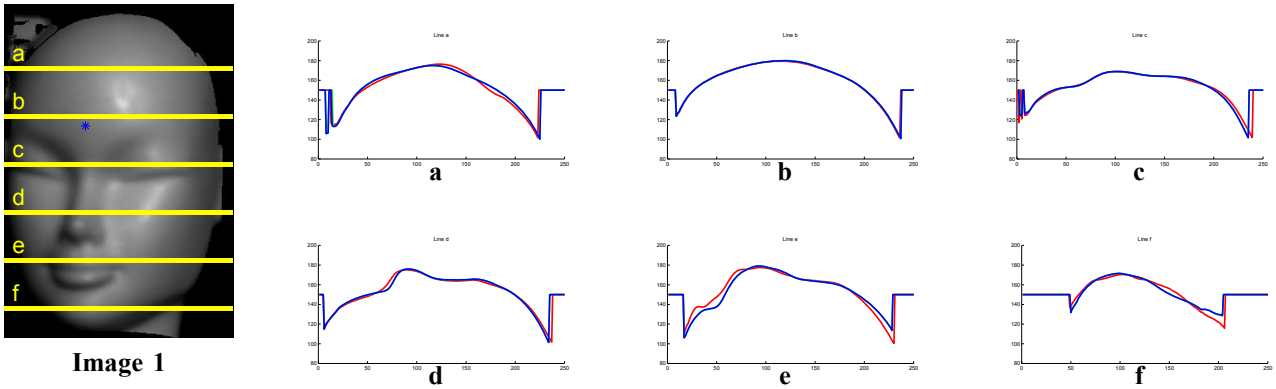

Fig. 13. Cross-sections of the surface shown in **Image 1** are presented. The red is the naive method with a single neighbor propagation and no error correction while the blue is the full method with error correction.

general approach, as long as the normal recovery procedure is local.

Future research will also focus on self-calibration of the system (for camera and lighting parameters). This will allow the reconstruction of smooth featureless surfaces from an arbitrary set of images. In addition, applying whole-object optimization techniques to the output of our method should be explored in order to improve the quality of the recovered surface.

ACKNOWLEDGMENT

This research was supported by the Israel Science Foundation (grant no. 133/0-125). We would like to thank Avi Barliya, Gil Ben-Artzi, and Benjamin Neeman for working on the implementation of our method.

REFERENCES

[1] R. Basri and D. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Trans. Patt. Anal. Mach. Intell.*, 25(2):218–233, 2003.

[2] N. Birkbeck, D. Cobzas, P. Sturm, and M. Jägersand. Variational shape and reflectance estimation under changing light and viewpoints. In *Proc. European Conf. Comp. Vision*, pages I:536–549. Springer, may 2006.

[3] J. Y. Bouguet. Camera calibration toolbox for matlab. http://www.vision.caltech.edu/bouguetj/calib_doc.

[4] A.M. Bruckstein. On shape from shading. *Computer Vision, Graphics, and Image Processing*, 44(2):139–154, 1988.

[5] M.S. Drew. Direct solution of orientation-from-color problem using a modification of pentland's light-source direction estimator. *Comp. Vis. Im. Understanding*, 64(2):286–299, September 1996.

[6] P. Dupuis and J. Oliensis. Direct method for reconstructing shape from shading. In *Proc. IEEE Conf. Comp. Vision Patt. Recog.*, pages 453–458, 1992.

[7] J. D. Durou and F. Courteille. Integration of a normal field without boundary condition. In *Workshop on Photometric Analysis for Computer Vision (ICCV 2007)*, 2007.

[8] J. D. Durou, M. Falcone, and M. Sagona. Numerical methods for shape-from-shading: A new survey with benchmarks. *Comp. Vis. Im. Understanding*, 109(1):22–43, 2008.

[9] J. Fan and L.B. Wolff. Surface curvature and shape reconstruction from unknown multiple illumination and integrability. *Comp. Vis. Im. Understanding*, 65(2):347–359, February 1997.

[10] O. Faugeras and R. Keriven. Complete dense stereovision using level

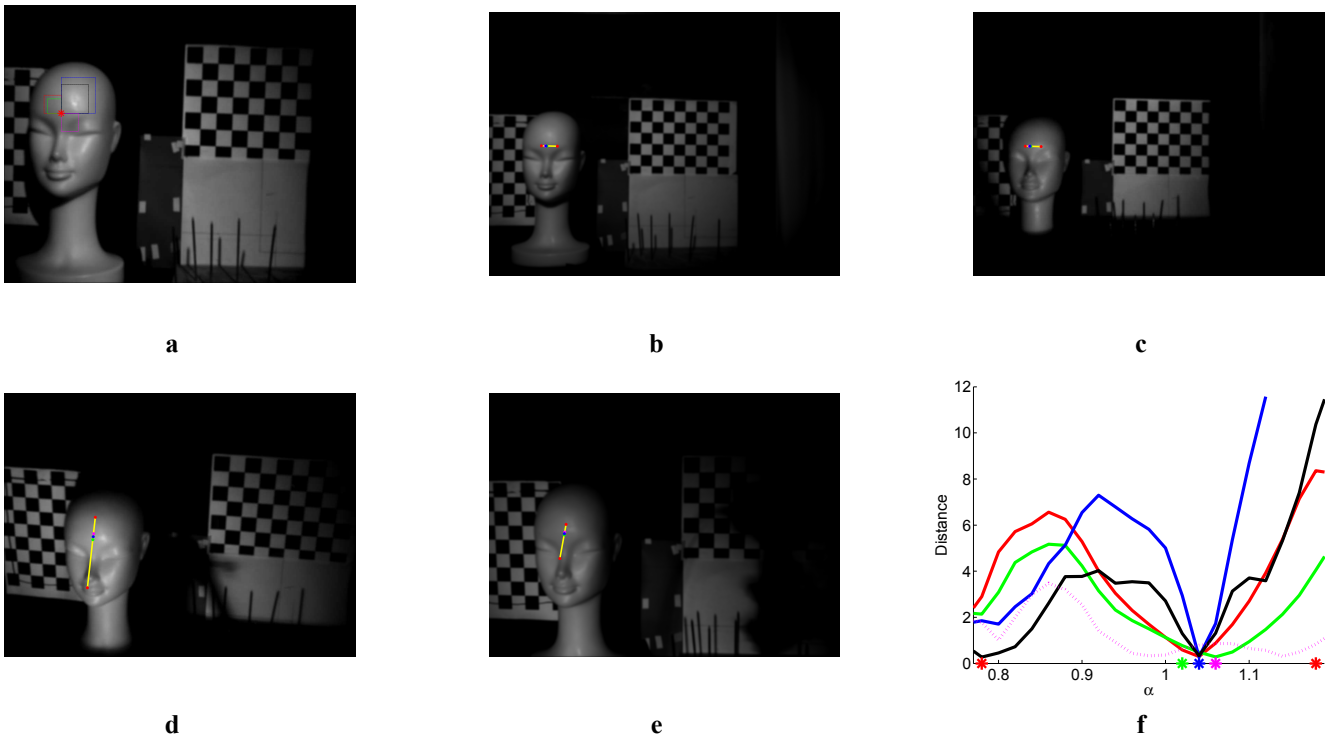

Fig. 14. Detection of the starting point. (**a**) The red star marks the initial point and the five colored rectangles mark the closed cycles. The other four images are shown in **b**-**e**. The yellow line marks the projection of the candidate 3D points to each of the images. The 3D distances between the first and the last point on each of the rectangular paths as a function of α is shown in **f**. The plot colors correspond to the marked rectangles in (**a**). The colored points on the yellow line correspond to the different values of α marked by stars in **f**.

set methods. In *Proc. European Conf. Comp. Vision*, pages 379–393, 1998.

[11] W.T. Freeman and J.B. Tenenbaum. Learning bilinear models for two factor problems in vision. In *Proc. IEEE Conf. Comp. Vision Patt. Recog.*, pages 554–560, 1997.

[12] P. Fua and Y. Leclerc. Object-centered surface reconstruction: Combining multi-image stereo and shading. In *Proc. DARPA Image Understanding Workshop*, pages 1097–1120, 1993.

[13] R.I. Hartley and P. Sturm. Triangulation. *Comp. Vis. Im. Understanding*, 68(2):146–157, November 1997.

[14] H. Hayakawa. Photometric stereo under a light-source with arbitrary motion. *J. Opt. Soc. Am. A*, 11(11):3079–3089, November 1994.

[15] C. Hernández, G. Vogiatzis, and R. Cipolla. Multiview photometric stereo. *IEEE Trans. Patt. Anal. Mach. Intell.*, 30(3):548–554, March 2008.

[16] B. K. P. Horn and M. Brooks. *Shape from Shading*. MIT Press, Cambridge, Mass., 1989.

[17] I. Horowitz and N. Kiryati. Depth from gradient fields and control points: Bias correction in photometric stereo, image and vision computing. *Image and Vision Computing*, 22:681–694, 2004.

[18] D.R. Hougen and N. Ahuja. Estimation of the light source distribution and its use in integrated shape recovery from stereo and shading. In *Proc. Int. Conf. Comp. Vision*, pages 148–155, 1993.

[19] K. Ikeuchi and B.K.P. Horn. Numerical shape from shading and occluding boundaries. *Artificial Intelligence*, 17:141–184, 1981.

[20] H. Jin, D. Cremers, D. Wang, E. Prados, A. Yezzi, and S. Soatto. 3-D Reconstruction of Shaded Objects from Multiple Images Under Unknown Illumination. *Int. J. of Comp. Vision*, 76(3):245–256, 2008.

[21] R. Kimmel and A.M. Bruckstein. Global shape from shading. *Comp. Vis. Im. Understanding*, 62(3):360–369, 1995.

[22] R. Kimmel and A.M. Bruckstein. Tracking level sets by level sets: A method for solving the shape from shading problem. *Comp. Vis. Im. Understanding*, 62(2):47–58, 1995.

[23] R. Kimmel and J. Sethian. An optimal time algorithm for shape from shading. In *Asian Conference on Computer Vision (ACCV)*, 2000.

[24] R. Kimmel and I. Yavneh. An algebraic multigrid approach for image analysis. *SIAM journal on scientific computing*, 24(4):1218–1231, 2003.

[25] R. Klette, R. Kozera, and K. Schluns. Shape from shading and photometric stereo methods. Technical Report CITR-TR-20, University of Auckland, New Zealand,, 1998.

[26] K. Kolev, M. Klodt, T. Brox, S. Esedoglu, and D. Cremers. Continuous Global Optimization in Multiview 3D Reconstruction. In *International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition*. Springer, 2007.

[27] L.L. Kontsevich, A.P. Petrov, and I.S. Vergelskaya. Reconstruction of shape from shading in color images. *J. Opt. Soc. Am.*, 11(3):1047–1052, March 1994.

[28] J Lim, J. Ho, M. Yang, and D. Kriegman. Passive photometric stereo from motion. *Proc. Int. Conf. Comp. Vision*, 2:1635–1642, 2005.

[29] A. Maki, M. Watanabe, and C. Wiles. Geotensity: Computing motion and lighting for 3D surface reconstruction. *Int. J. of Comp. Vision*, 48(2):75–90, 2002.

[30] Y. Moses and I. Shimshoni. 3D shape recovery of smooth surfaces: Dropping the fixed viewpoint assumption. In *Asian Conference on Computer Vision (ACCV)*, pages 429–438, 2006.

[31] R. Onn and A.M. Bruckstein. Integrability disambiguates surface recovery in two-image photometric stereo. *Int. J. of Comp. Vision*, 5(1):105–113, August 1990.

[32] E. Prados and O.D. Faugeras. Perspective shape from shading and viscosity solutions. In *Proc. Int. Conf. Comp. Vision*, pages 826–831, 2003.

[33] E. Prados and S. Soatto. Fast marching method for generic shape from shading. *Proceedings of VLSM*, pages 320–331, 2005.

[34] D. Samaras and D. Metaxas. Incorporating illumination constraints in deformable models for shape from shading and light direction estimation. *IEEE Trans. Patt. Anal. Mach. Intell.*, 25(2):247–264, February 2003.

[35] I. Shimshoni, Y. Moses, and M. Lindenbaum. Shape reconstruction of 3D bilaterally symmetric surfaces. *Int. J. of Comp. Vision*, 39(2):97—110, September 2000.

[36] D. Simakov, D. Frolova, and R. Basri. Dense shape reconstruction of a moving object under arbitrary, unknown lighting. In *Proc. Int. Conf. Comp. Vision*, pages 1202–1207, 2003.

[37] A. Tankus and N. Kiryati. Photometric stereo under perspective projection. In *Proc. Int. Conf. Comp. Vision*, pages 611–616, 2005.

[38] A. Tankus, N.A. Sochen, and Y. Yeshurun. A new perspective [on] shape-from-shading. In *Proc. Int. Conf. Comp. Vision*, pages 862–869, 2003.

[39] M. Weber, A. Blake, and R. Cipolla. Towards a complete dense geometric and photometric reconstruction under varying pose and illumination. In *British Mach. Vis. Conf.*, pages 83–92, 2002.

[40] R. J. Woodham. Photometric stereo: A reflectance map technique for determining surface orientation from image intensity. *In Proc. SPIE*, 155:136–143, 1978.

[41] T. Yu, N. Xu, and N. Ahuja. Shape and View Independent Reflectance Map from Multiple Views. *Int. J. of Comp. Vision*, 73(2):123–138, 2007.

[42] A. L. Yuille, D. Snow, R. Epstein, and P. N. Belhumeur. Determining generative models of objects under varying illumination: Shape and albedo from multiple images using svd and integrability. *Int. J. Comput. Vision*, 35(3):203–222, 1999.

[43] 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 *Proc. Int. Conf. Comp. Vision*, pages 618–625, 2003.

[44] R. Zhang, P.S. Tsai, J.E. Cryer, and M. Shah. Shape from shading: A survey. *IEEE Trans. Patt. Anal. Mach. Intell.*, 21(8):690–706, 1999.

[45] Q. Zheng and R. Chellappa. Estimation of illuminant direction, albedo, and shape from shading. *IEEE Trans. Patt. Anal. Mach. Intell.*, 13(7):680–702, July 1991.