Shape Reconstruction of 3D Bilaterally Symmetric Surfaces

Ilan Shimshoni
Dept. of Computer Science
The Technion
Haifa 32000 Israel

Yael Moses
Dept. of Applied Math
The Weizmann Inst. of Science
Rehovot 76100 Israel

Michael Lindenbaum
Dept. of Computer Science
The Technion
Haifa 32000 Israel

Abstract

The paper presents a new approach for shape recovery based on integrating geometric and photometric information. We consider 3D objects which are symmetric with respect to a plane (e.g., faces) and their reconstruction from a single image. Both the viewpoint and the illumination are not necessarily frontal. In principle, no correspondence between symmetric points is required, but knowledge about a few corresponding pairs accelerates the process.

The basic idea is that an image taken from a general, non frontal viewpoint, under non-frontal illumination can be regarded as a pair of images of half of the object, taken from two different viewing positions and two different lighting directions. Thus, one-image-variants of geometric stereo and photometric stereo can be used. Unlike the separate invocation of these approaches, which require correspondence, we show that integrating the photometric and geometric information yields the unknown lighting and viewing parameters, as well as dense correspondence between pairs of symmetric points. As a result, a dense shape recovery of the object is computed.

Unknown point light source, Lambertian surfaces, unknown constant albedo, and weak perspective projection are assumed. The method has been implemented and tested experimentally on simulated and real data.

1 Introduction

This paper presents a method for shape reconstruction of 3D bilaterally symmetric objects from a single image. A bilaterally symmetric object is an object that is symmetric with respect to a plane (as illustrated in Figure 1). Typical examples of this class of objects are faces, animals, human bodies and many man-made objects. We consider a natural setup where such objects are seen from a non-frontal view and with non-frontal illumination. Note that the resulting images are non-symmetric but, as we will show, are actually more informative than symmetric images.

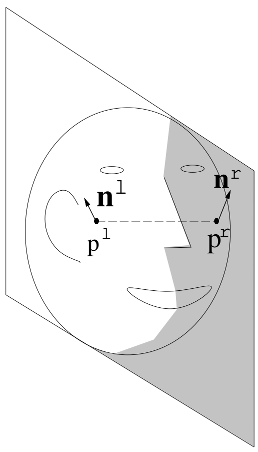


Figure 1: An illustration of a bilaterally symmetric object viewed from an arbitrary viewing direction. For each point $\mathbf{p^l}$ there is a corresponding point $\mathbf{p^r}$ and a pair of corresponding normals at those points $\mathbf{n}^l$ and $\mathbf{n}^r$ respectively.

Methods for general shape reconstruction usually rely on either geometric or photometric constraints. When more than one image is available geometric-stereo [11, 4] or photometric-stereo [24, 19] can be used. When one image is considered, geometric constraints are not sufficient for shape reconstruction and photometric methods such as shape-from-shading [13] are used. These methods require the knowledge of various parameters (e.g., illumination). Class constraints can alleviate the need to be given these parameters and can substantially improve the reconstruction. See for example [9] where a special algorithm to recover the shape of Straight Homogeneous Generalized Cylinders (SHGC's) was presented. Our work can also be classified as a class based 3D shape recovery algorithm, dealing with another, higher dimensional class.

The basic idea is that a single non-frontal image of a bilaterally symmetric object can be regarded as two images of another object taken from two different viewpoints and two different illumination directions. The new object consists of one half of the bilaterally-symmetric object. It follows that special constraints exists for images of the class of 3D bilaterally-symmetric objects. Variants of photometric-stereo [24, 19] and geometric-stereo [11, 4] methods can therefore be applied when the image is taken from a non-frontal viewpoint with a non-frontal illumination.

The straightforward application of the one-image variants of both geometric and photometric stereo methods requires however that the imaging parameters (viewpoint and illumination) as well as the correspondence between symmetric points, denoted *self-correspondence*, be given [8]. Recall that even when the imaging parameters are unknown we can still compute values invariant to them [21, 17] which are sufficient to differentiate between many objects in the class. Finding dense correspondence is generally a non-trivial problem. Correspondences can usually be found only for a sparse set of feature points, which result in only a sparse depth map [11, 4]. Furthermore, as we describe below, the photometric or geometric constraints are insufficient by themselves for computing dense self-correspondence. A special case for which self-correspondence can be found more easily is when the symmetry is in the image plane.

Previous studies on symmetry have focused either on symmetric images [7, 25] or on skewed symmetric images [5, 2, 18, 20]. Note that in both cases, detecting the self-correspondence is a parametric problem, the solution of which depends only on a few parameters (characterizing the symmetry axis and possibly another direction, in the image plane). Under general imaging conditions, neither the image coordinates nor the grey levels associated with 3D symmetric objects, are symmetric. Furthermore, the problem is non-parametric, and finding the self-correspondence is equivalent to recovering the 3D structure. Therefore, it cannot be based on the methods mentioned above.

The essence of our approach for solving the self-correspondence is to specify it as a continuous function, and to solve for its derivatives by combining photometric and geometric constraints. These constraints are local and specified at every pixel. Our approach is to use them first to find the global (pose and illumination) parameters. Once these parameters are known, the local constraints enforce correct reconstruction locally. This stands in sharp contrast to traditional shape from shading solutions, which are neither local nor provide means for recovering the global parameters. Once one pair of corresponding points has been found using the local constraint a combination of the photometric and geometric constraints are used to propagate the self-correspondence function over the entire image, yielding the 3D shape.

Previous attempts to integrate photometric and geometric constraints were based on using each of the methods in different image regions. In some studies the results computed by, say, a stereo algorithm, were transfered to a photometric shape recovery algorithm (e.g., [1, 3, 6, 14]). To the best of our knowledge, our scheme is the first to propose integration in the strong sense. The integration of the photometric and geometric constraints allow us not only to compute the self-correspondence function, but also to recover the shape of the object based on geometric constraints. This is a preferable method over photometric methods for shape reconstruction since it avoids the reconstruction based on the integration of surface normals, which causes error accumulation.

The paper continues as follows: in the next two sections we derive the geometric and photometric constraints on images of a symmetric object. Then in Section 4 we show how to integrate these constraints into one, stronger, constraint on the correspondence function. We examine the easier problem of cylindrical objects and show a specialized solution for them in Section 5. In section 6, we consider the general case and propose a shape reconstruction algorithm that uses the constraints efficiently. All the presented algorithms have been implemented and results of testing them on simulated and real images are presented in the corresponding sections. Finally, a number of issues raised by our algorithm and future research directions are discussed in Section 7.

2 Geometry

In this section we present the geometric setting and constraints associated with bilaterally symmetric objects. In addition the following known results are presented using our notation. In particular we show that when self-correspondence is given, then invariant representation to viewpoint can be computed. If in addition the viewing parameter is given, then the 3D shape of the object can be recovered directly. When self-correspondence is unknown, even if the viewing direction is known, then the 3D shape and the self-correspondence cannot be computed based only on geometrical constraints alone, and the photometric constrains must be added (as shown in the next section). These results are used to build our integrated approach.

2.1 Notations and settings

Let $\mathbf{p^r} = (x, y, z)^T, \mathbf{p^l} = (-x, y, z)^T$ be a pair of symmetric points on a bilaterally symmetric object. We choose our coordinate system such that the *symmetry plane*, the plane relative to which the object is symmetric about, is the $y \times z$ plane. In this case, all lines that connect pairs of symmetric points are parallel to the x-axis. It is usually fairly easy to detect the projections of this axis in the image by, say, finding a pair of corresponding points (e.g., the eyes of a person in the image (e.g. [23])). Therefore, we shall assume that this direction is known.

The object points $\mathbf{p^l}$ and $\mathbf{p^r}$ are projected to the image points $\tilde{\mathbf{p}}^l$ and $\tilde{\mathbf{p}}^r$. We assume here weak perspective projection (i.e., orthographic projection of a rotated and scaled object). The projection of a point $\mathbf{p}$ is $\tilde{\mathbf{p}} = sR\mathbf{p} + \mathbf{t}$ where s is a scale factor, R is a 2×3 matrix of full rank (an orthographic projection of a rotation matrix), and $\mathbf{t}$ is a 2D vector. Under this projection model, all image lines that connect pairs of symmetric points are parallel. Noting the similarity to the stereo setting, which will be discussed below, we refer to these set of image lines as *epipolar lines*. Given the direction of the epipolar lines, which can be computed from a single pair of corresponding points, finding the self-correspondence becomes a one-dimensional search problem.

Without loss of generality, it is sufficient in our case, to set $s = 1$ and $\mathbf{t} = 0$, and to reduce the viewing direction to a single parameter θ, which is the rotation of the object around the y-axis of the object. Thus R is reduced to:

$$R = \begin{pmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \end{pmatrix} \tag{1}$$

To show that a single parameter is indeed sufficient, let us first consider a general rotation matrix. Any 3D rotation can be described by three rotation angles: around the

x, y and z axes (see Figure 2) . When the epipolar direction is known the rotation in the image plane can be directly computed and compensated for. We can therefore assume that the rotation about the z-axis is 0. (Note that this correct even tough the z axis does not coincide with the optical axis.) The inherent ambiguity in the choice of the object coordinate system justifies the choice of zero rotation around the x axis, $s = 1$, and $\mathbf{t} = 0$. This ambiguity results in reconstruction of the object up to a scale factor, translation, and rotation around the x axis.

(a) (b) (c) (d)

Figure 2: A bilaterally symmetric object (face) (a) and its three rotations about the z, x and y axes, respectively.

2.2 Geometric constraints

Under this choice of projection we can now compute the image locations: $\tilde{\mathbf{p}}^{\mathbf{r}} = (\tilde{x}^r, \tilde{y}^r)^T$ and $\tilde{\mathbf{p}}^{\mathbf{l}} = (\tilde{x}^l, \tilde{y}^l)^T$, of a pair of corresponding object points, $\mathbf{p}^{\mathbf{r}} = (x, y, z)^r$ and $\mathbf{p}^{\mathbf{l}} = (-x, y, z)^l$. We obtain that

$$\begin{array}{rcl} \tilde{y}^r & = & y; \\ \tilde{x}^r & = & x\cos(\theta) + z\sin(\theta) \\ \tilde{y}^l & = & y; \\ \tilde{x}^l & = & -x\cos(\theta) + z\sin(\theta). \end{array} \tag{2}$$

These equations can be inverted yielding the following expressions for the 3D coordinates of the object points.

$$\begin{aligned} x &= \tfrac{1}{2\cos(\theta)}(\tilde{x}^r - \tilde{x}^l), \\ y &= \tilde{y}^r, \\ z &= \tfrac{1}{2\sin(\theta)}(\tilde{x}^r + \tilde{x}^l). \end{aligned} \tag{3}$$

We call this shape reconstruction method the *mono-geometric-stereo* because the two sides of the image are interpreted as a stereo pair of one half-object taken from two viewpoints. It is also clear now why we denoted the image lines parallel to the object's transversal axis by epipolar lines.

Finally, we claim that the location of the corresponding image points are not constrained by the fact that the object is symmetric, even if the viewing position θ is known. It is easy to see from Eq. 3 that given a right point in the image, any choice of a left point is consistent with an image of a bilaterally symmetric object. This ambiguity is equivalent to the inherent ambiguity in pairing of points from two stereo images (see, e.g. [4]). Recall however, that even when θ is unknown it is possible to infer invariants to viewing direction directly from Eq. 3 [17, 22].

3 Photometry

In this section we derive the photometric constraint of bilaterally symmetric objects. In a rather artificial way, we ignore for now the evidence available from the location of these image points. This is done in order to isolate the information available from the two sources, and we integrate the two sources of information in the next section. For the case when self correspondence is known (e.g. when the view is frontal) we describe the mono-photometric-stereo algorithm. In addition when the illumination parameters are not given we can use the first derivatives of the image intensity to recover the illumination parameters and the 3D shape. The extension of this result to the classic photometric-stereo with two unknown illuminations is discussed.

3.1 Notations and settings

The analysis is done in the context of a Lambertian surface and a distant light source. Different illumination and shading models lead to other photometric constraints, which may be used as well and are discussed in the last section.

Let $\hat{\mathbf{n}} = (n_x, n_y, n_z)^T$ be the normal at a surface point of the object. Then, using the Lambertian reflectance model, the intensity of light reflected from the point is $I = \rho\, \mathbf{l} \cdot \hat{\mathbf{n}}$, where $\mathbf{l}$ is the light source direction, $|\mathbf{l}|$ is the light source intensity, and ρ is the albedo. We assume that the albedo ρ, is fixed all over the surface. Equivalently, we may assume that $\rho = 1$ and let $|l|$ be the product of the intensity and the albedo. We also assume that no shadows exist. We intend to relax these assumptions in future work.

3.2 Photometric constraints

We denote by $\hat{\mathbf{n}}^{\mathbf{r}}$ and $\hat{\mathbf{n}}^{\mathbf{l}}$ the normals of two symmetric points of the object. Denote the right side of the object surface by the function $(x, y, z(x, y))$. In this case the normals of two self-corresponding points are given by $\hat{\mathbf{n}}^{\mathbf{r}} = \mathbf{n}^r/|\mathbf{n}^r|$ and $\hat{\mathbf{n}}^{\mathbf{l}} = \mathbf{n}^l/|\mathbf{n}^l|$ where $\mathbf{n}^r = (z^r_x, z^r_y, -1)^T$ and $\mathbf{n}^l = (-z^r_x, z^r_y, -1)^T$. The intensity at the two points is

$$\begin{aligned} I^r &= \mathbf{l} \cdot \hat{\mathbf{n}}^{\mathbf{r}}, \\ I^l &= \mathbf{l} \cdot \hat{\mathbf{n}}^{\mathbf{l}} \end{aligned} \tag{4}$$

Taking the sum and the difference of Eq. 4 yields:

$$\begin{array}{rclcl} I_\Delta &=& I^r - I^l &=& 2 l_x n_x \\ I_\Sigma &=& I^r + I^l &=& 2(l_y n_y + l_z n_z) \end{array} \tag{5}$$

Similar to viewpoint invariants, when $\mathbf{l}$ is unknown it is possible to infer invariants to illumination direction directly from Eq. 4 [17].

3.3 Mono-photometric-stereo

An immediate observation is that if the self-correspondence and $\mathbf{l}$ are known, then the intensities reflected from a corresponding pair are sufficient for calculating the normal at these points by solving a quadratic equation up to a twofold ambiguity. First, n_x is calculated from the first part of Eq. 5. Then, the unit size of $\hat{\mathbf{n}}$ is called to provide an additional constraint, yielding a quadratic equation from which the two solutions for the other components of $\hat{\mathbf{n}}$ are recovered. This normal recovery method is the *mono-photometric-stereo* algorithm. It is similar to photometric-stereo where the two sides of the image are interpreted as a stereo pair of one half-object taken with two different illuminations. The classical application for this algorithm is when a frontal view of an object illuminated by a non-frontal light source is considered. Then, the image is geometrically symmetric and finding the self-correspondence is easy.

We next show that by using the derivatives of the intensity, reconstruction is possible even if the illumination $\mathbf{l}$ is unknown. For a frontal view, the image and the object x and y coordinates are identical: $\tilde{x}^r = x$, $\tilde{y}^r = y$, $\tilde{x}^l = -x$, and $\tilde{y}^l = y$. Therefore, the derivatives of the image intensity are equal, up to a sign to the intensity derivatives with respect to the world coordinates x, y. Note that this is no longer true for non frontal views due to the foreshortening effect. In this case the difference in the derivatives between two corresponding points depends on the viewpoint and the 3D shape as we discuss in section 4.2. We denote the derivative $\frac{\partial I}{\partial e}$ by I_e, and the parameters e and $\tilde{e}$ correspond to derivatives with respect to an image coordinate or an object coordinate respectively. The derivatives of the intensity for a frontal image may be written in terms of the normals derivatives:

$$
\begin{aligned}
{I^r}_{\tilde{x}} &= {I^r}_x = \mathbf{l} \cdot \hat{\mathbf{n}}^{\mathbf{r}}_x; \\
{I^l}_{\tilde{x}} &= {I^l}_x = \mathbf{l} \cdot \hat{\mathbf{n}}^{\mathbf{l}}_x \\
{I^r}_{\tilde{y}} &= {I^r}_y = \mathbf{l} \cdot \hat{\mathbf{n}}^{\mathbf{r}}_y; \\
{I^l}_{\tilde{y}} &= {I^l}_y = \mathbf{l} \cdot \hat{\mathbf{n}}^{\mathbf{l}}_y.
\end{aligned}
\tag{6}
$$

Considering smooth integrable surfaces, where $z_{xy} = z_{yx}$, it follows that the normal derivatives, $\hat{\mathbf{n}}^{\mathbf{r}}_x$, $\hat{\mathbf{n}}^{\mathbf{l}}_x$, $\hat{\mathbf{n}}^{\mathbf{r}}_y$, $\hat{\mathbf{n}}^{\mathbf{l}}_y$ are functions of 5 local derivatives of the surface: $z_x, z_y, z_{xx}, z_{xy} = z_{yx}$ and z_{yy}. Substituting into these equations the derivatives of the normals yields:

$$
\begin{aligned}
{I^r}_{\tilde{x}} &= \tfrac{1}{(1+z_x^2+z_y^2)^{3/2}}[\mathbf{l}(1+z_y^2, -z_y z_x, z_x), 0, \mathbf{l}(-z_x z_y, 1+z_x^2, z_y)] \cdot [z_{xx}, z_{yy}, z_{yx}] \\
{I^l}_{\tilde{x}} &= \tfrac{1}{(1+z_x^2+z_y^2)^{3/2}}[\mathbf{l}(-1-z_y^2, -z_y z_x, z_x), 0, \mathbf{l}(z_x z_y, 1+z_x^2, z_y)] \cdot [z_{xx}, z_{yy}, z_{yx}] \\
{I^r}_{\tilde{y}} &= \tfrac{1}{(1+z_x^2+z_y^2)^{3/2}}[0, \mathbf{l}(-z_x z_y, 1+z_y^2, z_y), \mathbf{l}(-1-z_y^2, -z_y z_x, z_x)] \cdot [z_{xx}, z_{yy}, z_{yx}] \\
{I^l}_{\tilde{y}} &= \tfrac{1}{(1+z_x^2+z_y^2)^{3/2}}[0, \mathbf{l}(z_x z_y, 1+z_y^2, z_y), \mathbf{l}(1+z_y^2, -z_y z_x, z_x)] \cdot [z_{xx}, z_{yy}, z_{yx}]
\end{aligned}
\tag{7}
$$

Note that these equations are linear in z_{xx}, z_{xy} and z_{yy}. Therefore, for every symmetric pair, we have 6 equations in 5 local unknowns (2 from Eq. 4 and 4 from Eq. 6). In addition, we have also 3 global unknowns characterizing the illumination and the albedo. Enforcing the intensity constraints for several (n) symmetric pairs, we get $6n$ equations, which, although they are nonlinear, can be solved for any $n \geq 4$. We expect that the relative inaccuracy implied from the use of derivatives is compensated by the many equations that may be written to satisfy the constraints.

This method can be easily adapted for the classical photometric-stereo algorithm, when the illumination directions are unknown (without any points whose intensity is known [10]). In this case, we still have 5 parameters characterizing the local geometry at every point and 6 additional global parameters characterizing the two illuminations. Therefore, the information available from the intensities in $n \geq 6$ points suffice to determine both the illumination and the normals at these points. For non-frontal views, this technique has to be modified, as foreshortening affects the intensity derivatives at corresponding points differently. This more general situation is dealt with in the next section.

4 Integration method

In this section we deal with the general case of a non-frontal viewing direction and non-frontal illumination. In order to recover the shape in this case, the photometric and geometric constraints are integrated and used to recover the self-correspondence. The first derivatives of the image intensity are used to improve the quality of the recovered shape and to recover unknown viewing and illumination parameters when they are not given.

4.1 The self-correspondence function

Our main integration tool is the "self-correspondence" function, which we introduce now. This function, which is initially unknown, can be constrained using the image data. With this function, we may start from a given pair of self-corresponding points, compute dense correspondence in its neighborhood and further propagate the correspondence over the whole image, together with reconstructing the 3D shape.

Given a point $\tilde{\mathbf{p}}^{\mathbf{r}} = (\tilde{x}^r, \tilde{y}^r)^T$ in the image, its corresponding point is given by $\tilde{\mathbf{p}}^{\mathbf{l}} = (\tilde{x}^l, \tilde{y}^l)^T$. Since for two corresponding points we assume that $\tilde{y}^r = \tilde{y}^l$ it is sufficient to define the correspondence function C such that

$$C(\tilde{\mathbf{p}}^{\mathbf{r}}) = C(\tilde{x}^r, \tilde{y}^r) = \tilde{x}^l. \tag{8}$$

The points $\tilde{\mathbf{p}}^{\mathbf{r}}$ and $\tilde{\mathbf{p}}^{\mathbf{l}}$ are functions of x, y, z and θ. Taking the derivatives of the function C with respect to the object coordinates x and y we obtain:

$$\begin{array}{rcl} C_x & = & C_{\tilde{x}}\tilde{x}_x + C_{\tilde{y}}\tilde{y}_x \\ C_y & = & C_{\tilde{x}}\tilde{x}_y + C_{\tilde{y}}\tilde{y}_y \end{array} \tag{9}$$

The derivatives of the image coordinates $\tilde{x}^r$, $\tilde{x}^l, \tilde{y}^r$ and $\tilde{y}^l$ with respect to the object coordinates x and y can be computed from Eq. 2:

$$\begin{array}{rcl} \tilde{x}^r_x & = & \cos(\theta) + z_x \sin(\theta); \\ \tilde{x}^r_y & = & z_y \sin(\theta) \\ \tilde{x}^l_x & = & -\cos(\theta) + z_x \sin(\theta); \\ \tilde{x}^l_y & = & z_y \sin(\theta) \\ \tilde{y}^r_x & = & \tilde{y}^l_x = 0; \\ \tilde{y}^r_y & = & \tilde{y}^l_y = 1 \end{array} \tag{10}$$

These derivatives can be used to compute from Eq. 9 the derivatives of the correspondence function with respect to the image coordinates as a function of the surface normals (z_x, z_y):

$$\begin{array}{rcccccl} C_{\tilde{x}^r} & = & C_x/\tilde{x}^r_x & = & \frac{\tilde{x}^l_x}{\tilde{x}^r_x} & = & \frac{-\cos(\theta)+z_x\sin(\theta)}{\cos(\theta)+z_x\sin(\theta)} \\ C_{\tilde{y}^r} & = & C_y - \tilde{x}^r_y C_{\tilde{x}^r} & = & \tilde{x}^r_y - \tilde{x}^r_y\tilde{x}^r_x/\tilde{x}^r_x & = & \frac{z_y\sin(2\theta)}{\cos(\theta)-z_x\sin(\theta)}. \end{array} \tag{11}$$

When given a corresponding pair, the illumination and viewing direction, the normal at the point can be computed using Eq. 5. Then, using Eq. 11 the derivatives of the correspondence function at the image can also be computed. These derivatives are then used to estimate the corresponding position of a point close to $\tilde{\mathbf{p}}^{\mathbf{r}}$ using the first order

Taylor expansion of the correspondence function. This procedure can be repeated again and again propagating the correspondence function to large regions of the image (as illustrated in Figure 3).

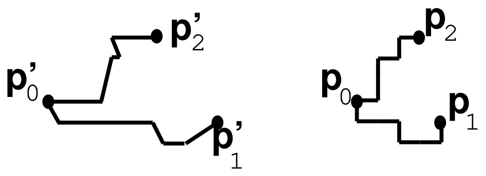


Figure 3: An illustration of the correspondence propagation process. Note the difference between the propagation of the correspondence on the two sides. Corresponding pairs are always on the same horizontal line but on a horizontal move d_x on the right hand side the x coordinate of the corresponding point changes by $d_x C_{\tilde{x}}$ and for a vertical move of d_y on the right hand side the x coordinate of the corresponding point changes by $d_y C_{\tilde{y}}$. and the y coordinate changes by d_y.

4.2 Derivatives of the intensity

In Section 3.3 it was shown that given the derivatives of the intensity it is possible to overcome unknown illumination. In this section we show how to compute the image derivatives of the image intensity I at two corresponding points for non-frontal views. In this case it is necessary to deal with the different foreshortening effect at the two points. Let $I^r(\tilde{\mathbf{p}}^{\mathbf{r}})$ and $I^l(\tilde{\mathbf{p}}^{\mathbf{l}})$ be the intensity at two corresponding points. In section 3.3 we have shown that the partial derivatives of I^r and I^l with respect to x and y can be expressed as functions of $\mathbf{l}$ and the local shape derivatives $z_x, z_y, z_{xx}, z_{xy}, z_{yy}$. However in the image we can measure only derivatives with respect to image coordinates. Similar to the previous section we compute the partial derivatives of I with respect to x and y and by using the geometric constraints we can relate the image and surface intensity derivatives as follows:

$$
\begin{aligned}
I^r{}_{\tilde{x}^r} &= \mathbf{l}\cdot\hat{\mathbf{n}}^{\mathbf{r}}_x/(cos(\theta)+z_x sin(\theta)) \\
I^l{}_{\tilde{x}^l} &= \mathbf{l}\cdot\hat{\mathbf{n}}^{\mathbf{l}}_x/(-cos(\theta)+z_x sin(\theta)) \\
I^r{}_{\tilde{y}^r} &= \mathbf{l}\cdot\hat{\mathbf{n}}^{\mathbf{r}}_y - z_y sin(\theta) I^r{}_{\tilde{x}^r} \\
I^l{}_{\tilde{y}^l} &= \mathbf{l}\cdot\hat{\mathbf{n}}^{\mathbf{l}}_y - z_y sin(\theta) I^l{}_{\tilde{x}^l}.
\end{aligned}
\tag{12}
$$

Thus we now have in addition to the two equations in z_x and z_y (Eq. 4), four additional equations which are linear in the three local second derivatives of the surface z_{xx}, z_{xy} and z_{yy}, which can be used to recover and verify the values of unknown global and local parameters. We denote this set of equations by the *integration constraint.* When the global parameters are known a self-corresponding point can be found by eliminating from the equations all the local partial derivatives of z leaving us with a single integration constraint in the image intensity values and their first derivatives at the two points. These values can be searched in the image. Conversely when given four corresponding pairs we can use these equations to solve for the unknown global parameters. When even this information is not available we use the integration constraint to solve a global optimization problem using numerical methods [15]. This method is demonstrated in our experiments.

5 Cylindrical Objects

Before turning to the reconstruction of general bilaterally symmetric objects we will first consider the simpler case of cylindrical objects. Such objects are characterized by a surface $z = f(x)$. For this class, neither the object nor the intensity changes in the y direction, $\partial f/\partial y = \partial I/\partial y = 0$, and the $y-$component of the surface normal is zero. This case is much more complicated than say, the mono-photometric-stereo, and requires all the tools that we developed for the general case. Still, the setting is simpler which makes its description more intuitive.

Knowing the albedo and illumination, ρ and $\mathbf{l}$, makes this task simple, as the basic Lambertian constraints $I = \rho\mathbf{l} \cdot \mathbf{n}$ together with the unity size of the normal $\mathbf{n}$ suffice to determine this normal (up to twofold ambiguity). The symmetry is thus not even needed for calculating the normals, and is used only for inferring the self-corresponding and the final reconstruction. Here, we consider a slightly more difficult task, of carrying out the reconstruction task, when illumination/pose parameters are unknown.

5.1 Additional constraints

For the simplified objects the grey level constraints are also simplified to

$$\begin{array}{rcl} I_\Delta & = & I^r - I^l = 2\rho l_x n_x \\ I_\Sigma & = & I^r + I^l = 2\rho l_z n_z. \end{array} \tag{13}$$

Inserting Eq. (13) into the unity size constraint, $n_x^2 + n_z^2 = 1$, gives

$$1 = n_x^2 + n_z^2 = \frac{1}{\rho^2 l_x^2 l_y^2}(l_z^2(I^l - I^r)^2 + l_x^2(I^l - I^r)^2) \tag{14}$$

Separating the space independent fixed illumination and albedo from the space dependent intensities, this constraint may be written in a matrix form:

$$\begin{array}{ccc} \begin{pmatrix} (I_1^l - I_1^r)^2 & (I_1^l + I_1^r)^2 & -1 \\ (I_2^l - I_2^r)^2 & (I_2^l + I_2^r)^2 & -1 \\ . & . & -1 \\ . & . & -1 \\ (I_n^l - I_n^r)^2 & (I_n^l + I_n^r)^2 & -1 \end{pmatrix} & \begin{pmatrix} l_z^2 \\ l_x^2 \\ \rho^2 l_x^2 l_z^2 \end{pmatrix} & = \quad 0 \\ \\ A & V & = \quad 0 \end{array} \tag{15}$$

These constraints yield two important conclusions: First they imply that knowing ρ and $\mathbf{l}$ allows us to calculate I^r for every corresponding grey level value I^l. That is, we can predict the whole right part of the image intensities from the left part. Given the actual intensities on the right side (from the image) we may warp the predicted intensity profile so that it matches the given one, and thus find the correspondence function and its derivative $C_{\tilde{x}}$ everywhere. A second implication is that given two pairs of corresponding grey levels, allows us to solve for the unknown ρ and $\mathbf{l}$ (up to the l_y which does not affect the image). In principle, with the self-correspondence function solved for, we may proceed and reconstruct both the 3D shape and the pose parameter θ. In order to accelerate the process it is better to find θ first.

5.2 Algorithm outline

Basically, the algorithm consists of the following four stages:

1. **Choose** V **-** Choose a candidate for the illumination and albedo (the vector V).

2. **Find some corresponding pairs -** Compute the intensities at the right hand side of the image using Eq. 15 and the hypothesized value of V.

3. **Verify** V **-** Use some pairs of points from the left side and their hypothesized corresponding points in the right side, to solve for θ (as described below). A consistent solution is obtained only when V is correct. If no consistent solution exists, return to step 1.

4. **Recover shape -** Starting from a correct match, found in the previous stage propagate the self-correspondence by warping the grey levels. This readily yields the 3D shape.

Following is the algorithm in more detail.

1. **Choose V.** The naive approach to implement the iterative process of stages 1-3, would be to perform a two parameter search for two matched intensity pairs, which, by Eq. (15), determines V. Alternatively, we may search directly for $\rho l_x, \rho l_z$. Such a search requires a function which vanishes for the correct V, which is described below as part of stage 3. In general a two dimensional minimization process is not a difficult task, however, here we may reduce the search cost by exploiting the extrema of the image intensities.

Shortcuts for choosing V - First we use the common heuristic that the maximal intensity corresponds to a normal which maximally aligns with the illumination. That is, $n_{max} = (l_x, 0, l_z)$. Therefore, the illumination vector components must satisfy the constraint $I_{max} = \rho(l_x^2 + l_z^2)$, implying that the search may be one dimensional over, say, $\alpha = \arctan(l_z/l_x)$. This is a task easily carried out by standard numerical optimization methods. Even this one dimensional search can be avoided by using intensity maxima which do not correspond to normal identical to the illumination vector. These points correspond to inflection points of the cylindrical surface [16]. Symmetry implies that such inflections and hence intensity extrema exist on both sides, and allow for easy detection of a corresponding pair. These pairs can be used to recover α using Eq. 15. Combining these two shortcuts determines a small discrete set of solutions for V. If such special points are not available, then the search described above is performed.

2. **Find some corresponding pairs.** Choose some pairs of close points on the left. Using V from stage 1 and Eq. 15, calculate the intensity of the corresponding points on the right, and use it to find several candidates for these corresponding points. Note that we find here sets of 4 points: two (close) points on the left and their suggested corresponding points on the right (which are also close).

3. **Verify** V. For every set of 4 points, computed above, it is possible to locally compute the viewing direction, θ. The verification of the illumination V and the chosen self-corresponding sets is done by comparing the computed θ for different sets of 4 points in different places of the image. Clearly, this viewing direction must be fixed all over the image.

Consider as set of 4 points consisting 2 left points (p_1^l, p_2^l) and their corresponding right points (p_1^r, p_2^r). To compute θ, we first compute the average normal at the left points using the intensity constraint (Eq. 13). Note that this normal, and the viewing direction θ determine the self-correspondence function locally. That is, they determine its derivative $C_{\tilde{x}^r}$. Therefore, we may calculate θ from the normal $\mathbf{n}$, and $C_{\tilde{x}^r}$. The latter may be approximated by finite differences as

$$C_{\tilde{x}^r} = \frac{p_2^l - p_1^l}{p_2^r - p_1^r}. \tag{16}$$

Then, θ is recovered using the first part of Eq. (11):

$$\theta = \arctan\left(\frac{n_z(C_{\tilde{x}^r} + 1)}{n_x(C_{\tilde{x}^r} - 1)}\right). \tag{17}$$

Note that θ is computed using only local information. A function evaluating the consistency of all correspondences and normals with a unique θ value (e.g. the variance of the set of θ values solved for different 4 point sets) is then used to verify the hypothesized V and the self-correspondences.

4. **Correspondence and shape reconstruction.** Starting from a corresponding pair of points (such as those found above) the algorithm calculates the surface normal and using θ, solves for $C_{\tilde{x}^r}$. The 3D position of the points may be recovered using Eq. 3. Now the algorithm starts a propagation step: a new point, close to the left point is chosen, and the rough position of it corresponding right point is determined using $C_{\tilde{x}^r}$. Note that

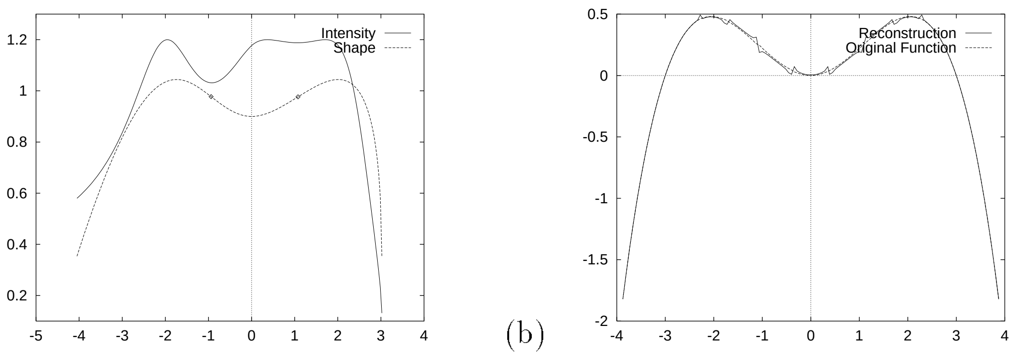
the intensity of the corresponding right point is known from Eq. 13. Thus, a search for a point associated with this intensity is done in the vicinity of this estimated position. A unique point is found in almost all cases. The 3D position of the new corresponding pairs is calculated and the propagation continues, yielding the 3D shape. An important feature of this reconstruction approach is that although there can be mismatches due to noise in this technique it always returns to the correct match because of the local nature of the algorithm. Thus, error is not accumulated.

5.3 Results

We now demonstrate this algorithm on synthetic data. Figure 4(a) shows the object and its intensity profile. Note that the asymmetric profile is asymmetric and that five extrema can be seen: Three maxima, which are extrema of the first type (normal close to the lighting direction) and two minima which are due to inflection points (and are marked on the object). The algorithm uses the extrema of the first type to estimate $\rho(l_x^2 + l_z^2)$ and uses pairs of extrema of the second type (those associated with non-maximal intensity) suspected to be corresponding pairs, to recover the components l_x and l_z. Because the correct correspondence between these extrema is unknown, there might be several candidates for these illumination vector components.

Using the verification, as described in stage 3, our algorithm easily detected the correct lighting parameters from the small number of candidates and an estimate for θ. The values of these parameters are fine-tuned using the Levenberg-Marquardt optimization procedure which finds the set of parameters which minimize the θ variance criterion.

Once the global parameters have been recovered the shape recovery procedure follows (stage 4). As predicted, error does not accumulate and we get a nearly perfect

(a) 

(b)

Figure 4: (a) A cylindrical object and its intensity profile. Note that there are intensity extrema at inflection points of the object. (b) The recovered shape of the object and the correct shape.

reconstruction. Figure 4(b) shows the recovered shape of the object and the ground truth.

The algorithm for the class of cylindrical objects demonstrated the following three components which appear also in the general algorithm but in a more intuitive way:

- A method for recovering the global lighting and pose parameters.
- A method to propagate the self-correspondence function and the resulting 3D shape over the whole image.
- A method to verify that the self-correspondence is correct.

6 General algorithm and results

In this section we present an algorithm based on the proposed method. The results of running the algorithm on simulated and real data are presented. In future work we intend to test our method on more images. In particular, since we intend to drop the Lambertian assumption, it will allow us to handle more natural objects.

6.1 Results on simulated data

We first show the results of testing our algorithm on the simulated image shown in Figure 5(b) whose frontal view, frontal illumination image is shown in 5(a). Although we could have recovered the global pose and lighting parameters using a minimization process we use shortcuts to obtain initial estimates for them and use the minimization process only to fine-tune the initial estimates.

We estimate ρ as the maximal intensity in the image. We use the boundary points of the object in the image to estimate $\mathbf{l}$. There are two types of boundary points: occluding boundary points where $\mathbf{n} \cdot \mathbf{v} = 0$, where $\mathbf{v}$ is the viewing direction and shading boundary points where $\mathbf{n} \cdot \mathbf{l} = 0$. They can be distinguished by the fact that the intensity close to the shading boundary points reduces gradually to zero. Where as at occluding contour the intensity can drop abruptly to zero. We are only interested in occluding boundary points (Figure 5(c)). At these points the normal is known and lies in the image plane. We extrapolate the intensity at the boundary point from neighboring points and obtain at each point a linear equation in the components of lighting direction l_x and l_y. Using a least-squares procedure we recover l_x and l_y and obtain l_z from the unity constraint.

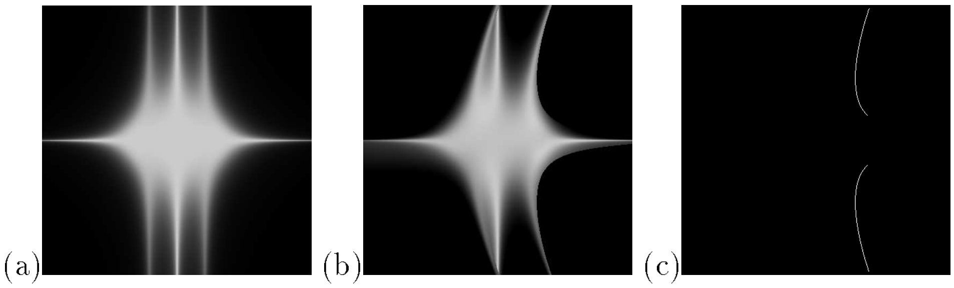


Figure 5: (a) A frontal view and frontal illumination image of a symmetric object. (b) An image of the object taken with $\theta = 0.2$rad. (c) The occluding contour of the object.

We recover θ by performing a two parameter numerical search for corresponding points for several points chosen arbitrarily. For each point the θ and corresponding point which minimize the integration constraint are found. All the values of θ found are then tested on a different set of points and a sequence of points recovered using the propagation method. For each point the corresponding point which minimizes the integration constraint is found and the values of the integration constraint measure for all the points are summed up yielding a measure for the quality of the candidate θ. This measure is used to verify the values of the global parameters.

Figure 6(a) shows the values of the measure for various values of θ on several propagated sequences. Figure 6(b) shows the values of the measure for the 9 θ's found in the first stage of the algorithm. Note that most of the θ's are close to the correct value of 0.2rad and have low scores. The minimal value was achieved for $\theta = 0.204$rad. It is important to note that this value of θ was achieved with the approximate value of $\mathbf{l}$ found in the previous step.

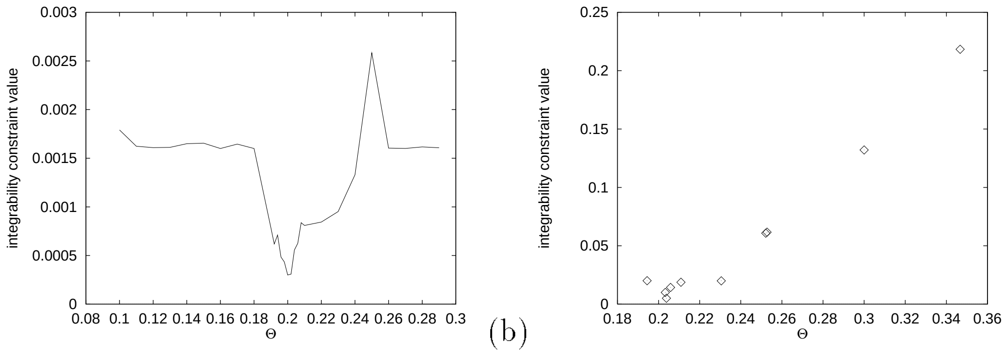
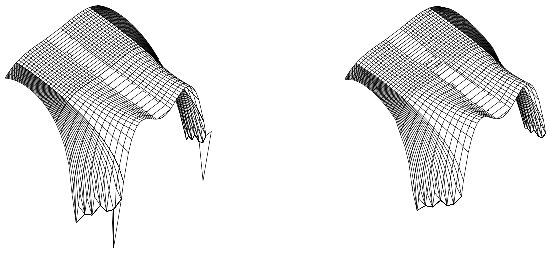
(a) 

Figure 6: (a) The results of searching for θ using 10 initial matches where from each point a sequence of twenty pairs of corresponding points are propagated. The y coordinate is the value of the aggregated integration constraint. If for a certain value of θ the value of the integration constraint exceeded 0.016 the computation terminated with a failure. (b) The candidate θ's suggested by 9 arbitrary points. These are verified on 10 other sequences of length 10. Note that even in the initial stage many results were close to $\theta = 0.2$rad and the best result was achieved for $\theta = 0.204$rad.

Once the global parameters are estimated a global optimization procedure is run to fine-tune their values minimizing the value of the integration constraint.

In the final step the 3D shape is recovered using a combination of the propagation procedure and the integration constraint to correct correspondence errors accumulated during the procedure. Once the correspondences are found the 3D structure is recovered using Eq. 3 without needing to perform integration on the recovered normals [12]. Figure 7(a) shows the ground-truth of the 3D structure and Figure 7(b) shows the recovered shape.



(a) (b)

Figure 7: (a) The ground truth of the 3D structure of the object studied. (b) The results of the shape recovery algorithm.

6.2 The mono-photometric-stereo experiment

Next we tested the mono-photometric-stereo algorithm on a frontal image of a face with non-frontal and ambient illumination (Figure 8(a)). The algorithm was given a rough estimate of the lighting information and the location of the symmetry line in the image. The mono-photometric-stereo algorithm was used to recover the normals. Points for which there was no solution were left blank. This was mainly due to changes in albedo and due to shadows. As this is a frontal view Eq. 3 could not be used as in the previous example. Instead a simple normal integration procedure was employed to recover the shape. The results are displayed by overlaying a contour map over the original image (Figure 8(b)). The results are only qualitatively correct and could be improved when a more exact reflectance model specially developed for human skin is used.

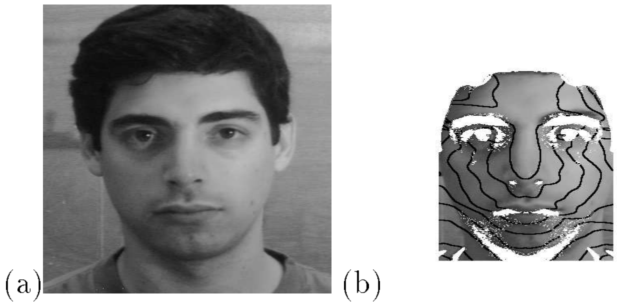


Figure 8: (a) A frontal image of a face with non-frontal illumination. (b) An image which overlays a contour map over the original image. White points in the image represent points for which no normals could be recovered. This is usually due to albedo changes.

6.3 Results for the mannequin image

In another experiment we took a picture of a mannequin painted with matte white paint (Figure 9(a)) to produce a close approximation to a Lambertian surface. This demonstrates our general algorithm for a non-frontal image with non-frontal illumination. We estimated the albedo as the maximum intensity and estimated the lighting direction from the occluding contour. We chose two pairs of corresponding points by taking the corners of the eyes (which can be recovered automatically) and moved vertically to points on the forehead and the cheeks. We gave the algorithm an estimate of θ as 0.2rad. We applied the propagation procedure to these two pairs of points recovering the normals and the 3D coordinates in object coordinates at each point. The propagation procedure was run on the left and right sides of the image independently. The results are displayed by overlaying a contour map of the z coordinate of the object coordinates, over the original image (Figure 9(b)).

(a) (b)

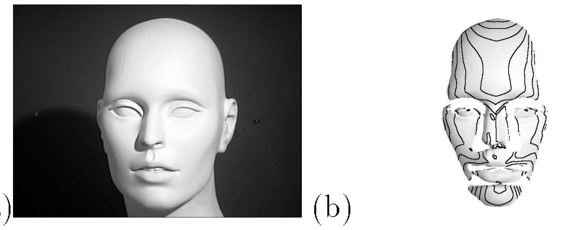


Figure 9: (a) A non-frontal image of a face with non-frontal illumination. (b) An image which overlays a contour map over the original image. White points in the image represent points for which no normals could be recovered. This is usually due to shadows.

7 Summary and discussion

In this paper we have presented a shape recovery method for bilaterally symmetric objects. The major contributions of the paper are: a strong, integrated, local constraint which combines photometric and geometric information; the correspondence extraction method; and the implied correspondence-less 3D shape reconstruction algorithm. The local constraints are used to recover the shape locally, improve the quality of the extracted correspondence and most importantly to recover unknown lighting and viewing parameters.

The derivation has been performed on Lambertian surfaces with a fixed unknown albedo. Extending this method to other viewpoint independent reflection models and dealing with specular points is straightforward [11]. However, dealing with more complex reflectance models and with shadows will require additional research.

Acknowledgments

Ilan Shimshoni was supported in part by the Koret and Goldschmidt Foundations. We would also like to thank Dalit Ramon for donating the mannequin used in the experiments. Yael Moses would like to thank Mike Brady for useful discussion in early stages of this study.

References

[1] A. Blake, A. Zisserman, and G. Knowles. Surface description from stereo and shading. *Image Vision Computation*, 3(4):183–191, 1985.

[2] T. Cham and R. Cipolla. Symmetry detection through local skewed symmetries. *Image and Vision Computing Journal*, 1995.

[3] J.E. Cryer, P.S. Tsai, and M. Shah. Integration of shape from shading and stereo. *PR*, 28(7):1033–1043, July 1995.

[4] O.D. Faugeras. *Three-Dimensional Computer Vision*. MIT Press, Boston MA, 1993.

[5] S.A. Freiberg. Finding axes of skewed symmetry. *Computer Vision Graphics and Image Processing*, 34:138–155, 1986.

[6] 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.

[7] Y. Gofman and N. Kiryati. Detecting symmetry in grey level images: The global optimization approach. In *ICPR*, page A94.2, 1996.

[8] G. G. Gordon. Shape from symmetry. In *Proc. of SPIE Intellent Robots And Computer Vision VIII*, volume 1192, November 1989.

[9] A.D. Gross and T.E. Boult. Recovery of shgcs from a single intensity view. *IEEE Trans. Patt. Anal. Mach. Intell.*, 18(2):161–180, February 1996.

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

[11] B.K.P. Horn. *Computer Vision.* MIT Press, Cambridge, Mass., 1986.

[12] B.K.P Horn and M.J. Brooks. The variational approach to shape from shading. *Comp. Vis. Graph. Im. Proc.*, 33:174–203, 1986.

[13] B.K.P. Horn and M.J. Brooks. *Shape from Shading.* MIT Press, 1989.

[14] 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.

[15] D. Kahaner, C. Moler, and S. Nash. *Numerical Methods and Software.* Prentice Hall, Englewood Cliffs, NJ, 1989.

[16] J.J. Koenderink and A.J. Van Doorn. Photometric invariants related to solid shape. *Optica Acta*, 27(7):981–996, 1980.

[17] Y. Moses and S. Ullman. Limitation of non-model-based recognition schemes. In G. Sandini, editor, *Proc. European Conf. Comp. Vision*, pages 820–828. Springer-Verlag, 1992.

[18] D.P. Mukherjee, A. Zisserman, and J.M. Brady. Shape from symmetry - detecting and exploiting symmetry in affine images. Technical Report Report No. OUEL 1988/93, University of Oxford, Department of Engineering Science, 1993.

[19] R. Onn and A.M. Bruckstein. Integrability disambiguates surface recovery in two-image photometric stereo. *IJCV*, 5(1):105–113, August 1990.

[20] J. Ponce. On characterizing ribbons and finding skewed symmetries. *Comp. Vis. Graph. Im. Proc.*, 52:328–340, 1990.

[21] C. A. Rothwell, A. Zisserman, D. A. Forsyth, and J. L. Mundy. Canonical frames for planar object recognition. In G. Sandini, editor, *Proc. European Conf. Comp. Vision*, pages 757–772. Springer-Verlag, 1992.

[22] C.A. Rothwell. Object recognition through invariant indexing. In *Oxford University Press*, 1995.

[23] B. Takacs and H. Wechsler. Detection of faces and facial landmarks using iconic filter banks. *PR*, 30(10):1623–1636, October 1997.

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

[25] H. Zabrodsky. *Computational aspects of pattern characterization - continous symmetry.* PhD thesis, Hebrew University, Jerusalem, 1993.