Reconstruction of shapes of 3D symmetric objects by using planarity and compactness constraints

Yunfeng Li & Zygmunt Pizlo
Purdue University

ABSTRACT

We present a new algorithm for reconstructing 3D shapes. The algorithm takes one 2D image of a 3D shape and reconstructs the 3D shape by applying *a priori* constraints: symmetry, planarity and compactness. The shape is reconstructed without using information about the surfaces, such as shading, texture, binocular disparity or motion. Performance of the algorithm is illustrated on symmetric polyhedra, but the algorithm can be applied to a very wide range of shapes. Psychophysical plausibility of the algorithm is discussed.

Keywords: Shape reconstruction, 3D symmetric shape, planarity, compactness, orthographic projection

INTRODUCTION

The problem of veridical perception of a 3D scene based on one or more 2D images is computationally difficult. Despite this difficulty, the human visual system solves this problem quickly and accurately. There have been a number of approaches to explain and model this ability. So far, all these efforts failed. The best known is the approach by David Marr [1]. Marr assumed that the first step in visual processing involves reconstructing visible surfaces (so called 2.5D sketch). Once the 3D surfaces are reconstructed, the 3D shape is produced by expressing the surfaces in the object's centered representation and filling-in the back, invisible part of the shape by using information stored in the memory. Biederman [2] emphasized the role of recognition, as opposed to reconstruction of shapes. Recognition involved a small set of simple 3D elementary parts, called geons. The third approach, represented by Poggio & Edelman [3], assumed that perceptual representation of 3D scenes involves a set of 2D images of the scenes and recognition involves evaluating similarities between the current, and the remembered images.

In this paper we present a new theory, in which a 3D shape percept is produced by applying simplicity constraints (priors) to a single 2D view of the shape. The following constraints are used: symmetry of a shape, planarity of the shape's contours and compactness of the shape (maximal volume given its surface area). Compactness has never been used before in algorithms for reconstructing 3D shapes. The new theory is illustrated on randomly generated symmetric polyhedra.

ALGORITHM

A single 2D image of a 3D shape is not sufficient for a unique reconstruction of the shape. The problem is ill-posed because the family of possible 3D interpretations is infinitely large [4]. Despite this inherent ambiguity, a human observer perceives a single 3D shape when presented with its 2D image, and the percept is usually veridical (the percept agrees with the shape "out there"). It is obvious that the human visual system "regularizes" the inverse problem, by imposing constraints on the family of possible 3D interpretations [5]. Several constraints were used in previous models of 3D shape reconstruction: symmetry of the shape, planarity of contours and minimum variance of angles [6]. The main limitation of these constraints was that they could not be applied to a wide range of shapes. For example, when the shape is not a polyhedron, minimum variance of angles cannot be used.

Constraints, such as symmetry, can be used in two ways in regularization models: implicitly or explicitly. Implicit constraint is equivalent to an assumption. For example, in Ullman's structure from motion algorithm [7], object's rigidity was an assumption, which allowed Ullman to reconstruct the 3D structure and its motion. Rigidity can also be used as an explicit constraint in a regularization algorithm [8]. In such a case, the reconstructed shape does not have to

Vision Geometry XV, edited by Longin Jan Latecki, David M. Mount, Angela Y. Wu,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6499, 64990B, © 2007 SPIE-IS&T · 0277-786X/07/$15

be perfectly rigid. Instead, the shape is as rigid as possible, and at the same time, as consistent with the data as possible. The compromise between fitting the image data and satisfying the constraint is controlled by a regularization parameter[4].

Symmetry was used by Vetter and Poggio [9] as an implicit constraint in their algorithm for 3D shape reconstruction. They considered a single orthographic image of a 3D wire (transparent) shape. When the shape has two planes of symmetry, the reconstruction is unique (up to depth reversal). But when the shape has only one plane of symmetry, the reconstruction is not unique. Again, when a human observer is presented with a single orthographic image of a symmetric shape, he or she perceives a single shape (see Figure 1). The question arises about the constraints that the human visual system uses. According to our theory, the human visual system chooses, from the infinitely many symmetric 3D interpretations, the one that has maximal volume for a given surface area. Maximizing volume, while keeping the surface area constant is equivalent to maximizing 3D compactness of the shape. 2D compactness was used by Brady & Yuille [10] to reconstruct the slant of surfaces. To our knowledge, 3D compactness has never been used before in 3D shape reconstruction.

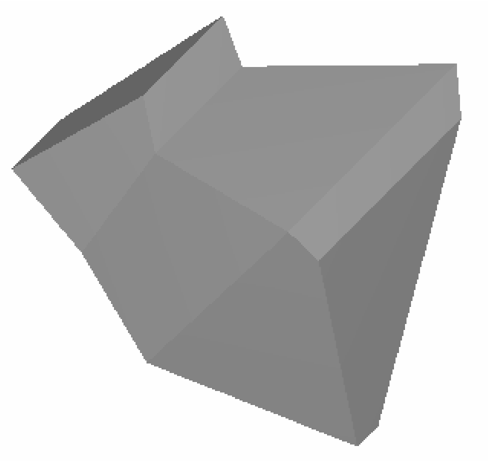

Figure 1: A single orthographic image of a symmetric shape leads to a unique percept.

We begin with describing how Vetter & Poggio used symmetry to restrict the family of 3D interpretations. Given an orthographic image of a transparent mirror-symmetric 3D shape, and assuming known correspondence of symmetric points, Vetter & Poggio showed how to compute a virtual view $p_{virtual}$ of the shape, given its real view p_{real}:

$$p_{virtual} = Dp_{real} \tag{1}$$

$$D = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Under this transformation, for any symmetric pair of points $p_{real} = [X_L \quad Y_L \quad X_R \quad Y_R]^T$ in the real (given) view, their corresponding pair of points in the virtual view is $p_{virtual} = [-X_R \quad Y_R \quad -X_L \quad Y_L]^T$. The virtual view is another orthographic view of the same 3D shape. Specifically, given an orthographic image of a 3D shape, which has one plane of symmetry, a second orthographic image can be computed directly from the given image, without the knowledge of the 3D shape, itself. Degenerate cases involve images, which themselves are mirror symmetric. One way to explain the construction of a virtual view is to observe that when a shape is mirror symmetric, a 3D reflection of the shape relative to a plane, can be undone by a 3D rigid motion. An orthographic image of a 3D reflection of a shape with respect to the plane X=0, is equivalent to a 2D reflection of an orthographic image with respect to the line x=0. This 2D reflection is represented by substituting each x coordinate in a 2D image by –x. It follows that a virtual image is a valid image of the 3D symmetric shape obtained from a different viewing direction. Figure 2 shows an example of a real and virtual view of a symmetric wire (transparent) shape.

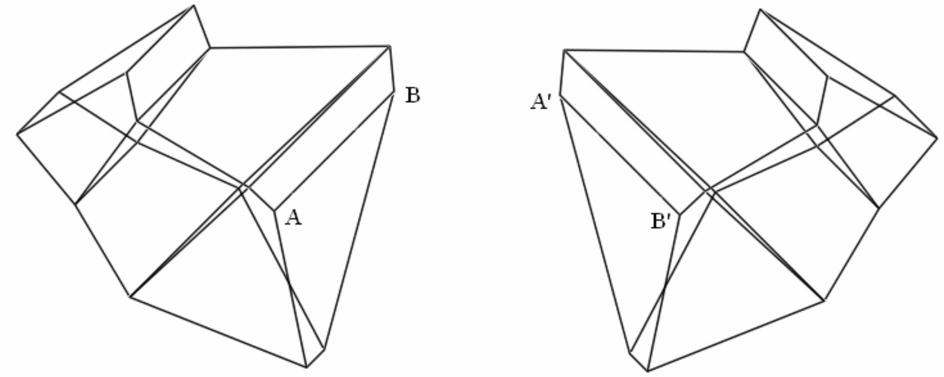

Figure 2. A real (left) and virtual (right) images of a 3D symmetric shape. A, B are images of a symmetric pair of points in the 3D shape. A′ and B′ are the corresponding points in the virtual image. Note that when the virtual image was produced, A′ was obtained (computed) from B. But in the 3D representation, A′ is produced after a 3D rigid rotation of A.

Two orthographic views are not sufficient to reconstruct a 3D shape [7, 11, 12]. Specifically, two orthographic views determine a family of 3D shapes, and the family is characterized by one parameter. It follows that a single orthographic view of a mirror-symmetric shape is consistent with a one-parameter family of 3D symmetric shapes. Next, we show how this family is determined.

In orthographic projection, 3D translation does not affect the shape or size of the 2D image. Specifically, translations along the direction orthogonal to the image plane have no effect on the image, and translations parallel to the image plane result in translations of the image. It follows that the 3D translation of the shape can be eliminated by translating the real view or virtual view, or both, so that the corresponding points in the two images coincide. Let this common point be the origin O of the image plane. Without restricting generality we can assume that the corresponding 3D point coincides with O, as well. After these translations, the real view can be considered an orthographic projection of the 3D shape at its original orientation and a virtual view is the one produced after some rotation (R) of the 3D shape around the origin O. Let the 3D coordinates of a vertex at its original orientation be $\overrightarrow{X_i} = \begin{bmatrix} X_i & Y_i & Z_i \end{bmatrix}^T$ and its corresponding vertex after rotation (R) be $\overrightarrow{X_i}' = \begin{bmatrix} X_i' & Y_i' & Z_i' \end{bmatrix}^T$. The following relation is satisfied:

$$\overrightarrow{X_i}' = R\overrightarrow{X_i}$$

This can be written as follows:

$$\begin{bmatrix} X_i' \\ Y_i' \\ Z_i' \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} \tag{2}$$

Consider the first two elements of the column vector $\overrightarrow{X_i}'$:

$$\begin{bmatrix} X_i' \\ Y_i' \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \end{bmatrix} + \begin{bmatrix} r_{13} \\ r_{23} \end{bmatrix} Z_i \qquad (3)$$

In equation (3), the points $\begin{bmatrix} X_i & Y_i \end{bmatrix}^T$ and $\begin{bmatrix} X_i' & Y_i' \end{bmatrix}^T$ in real and virtual views are known. Huang and Lee [11] derived the following relation between $\begin{bmatrix} X_i & Y_i \end{bmatrix}^T$, $\begin{bmatrix} X_i' & Y_i' \end{bmatrix}^T$ and R :

$$r_{23} X_i' - r_{13} Y_i' + r_{32} X_i - r_{31} Y_i = 0 \qquad (4)$$

Let's put the four elements of the rotation matrix R in a vector $\begin{bmatrix} r_{23} & r_{13} & r_{32} & r_{31} \end{bmatrix}$. Using equation (4), three pairs of corresponding points between real and virtual views are sufficient to compute the direction of this vector. The length of this vector can be derived from the constraint that the rotation matrix is orthonormal:

$$r_{13}^{\;2} + r_{23}^{\;2} = r_{31}^{\;2} + r_{32}^{\;2} = 1 - r_{33}^{\;2} \qquad (5)$$

Thus, if r_{33} is given, $\begin{bmatrix} r_{23} & r_{13} & r_{32} & r_{31} \end{bmatrix}$can be computed from the two images. The remaining elements of the rotation matrix can be determined based on the orthonormality of R (see Appendix). It follows that two orthographic images (real and virtual) determine R up to r_{33} which remains unknown. This unknown characterizes the family of 3D symmetric shapes consistent with the given 2D orthographic image of a given symmetric shape. Usually for eachr_{33}, two different rotation matrices are produced. This is because if $\begin{bmatrix} r_{23} & r_{13} & r_{32} & r_{31} \end{bmatrix}$ is the solution of equations (4) and (5), $\begin{bmatrix} -r_{23} & -r_{13} & -r_{32} & -r_{31} \end{bmatrix}$ is also a solution. Consequently, two 3D shapes are reconstructed, related to one another by depth reversal.

As shown just above, the family of 3D reconstructed shapes can be determined from four corresponding points in the real and virtual images. One point is the origin(O). It follows that three other points have to be selected to compute the rotation matrix(R). Note that these four points cannot be coplanar in the 3D shape. In application, three visible pairs of symmetric points (i.e., 6 points) in the real view are selected and their corresponding points in the virtual view are computed using Vetter and Poggio's method (Equation (1)). From these three pairs of points, four points are chosen (two of them are chosen from one symmetric pair and the other two are chosen from the other two pairs) to compute$R$. To guarantee that the corresponding 3D vertices of these four points are not coplanar, which is equivalent to the fact that these three pairs of points are not coplanar in 3D, we only need to verify that the midpoints (v_1, v_2, v_3) of these three pairs in the orthographic view are not collinear:♣

$$\left\| (v_1 - v_2) \times (v_1 - v_3) \right\| \neq 0$$

After the four points are chosen, the rotation matrix R is computed, as described above. Recall that R depends on the unknown value of r_{33}. Then the following steps are performed:

1. All visible symmetric pairs of vertices are reconstructed using equation (3);
2. If there are pairs of symmetric points, whose vertices are both invisible, reconstruction fails. The reason is that if both $\begin{bmatrix} X_i & Y_i \end{bmatrix}^T$ and $\begin{bmatrix} X_i' & Y_i' \end{bmatrix}^T$ are unknown, Z_i cannot be computed;

♣ In some cases, the corresponding 3D vertices of those three symmetric pairs are not coplanar, but their midpoints in the view are collinear. This can happen when the 3D shape is symmetric with respect to the YZ plane. In such a case, all midpoints in the orthographic image are on the y axis. In this case, their real view and virtual view are dependent and the reconstruction cannot be performed. Therefore, verifying that the image midpoints are not collinear implies that 3D points are not coplanar and the image is not degenerate.

3. For those pairs of symmetric points for which one point is visible and the other is occluded, planarity constraint is applied. Symmetry in conjunction with planarity of contours of faces is sufficient to compute the coordinates of the occluded vertex. In order to use a planarity constraint, at least three points from a given face (S) have to be reconstructed first. Let the image of the visible vertex on this face be $P = \lfloor P_x \quad P_y \rfloor$. The visible vertex (V), whose image is P, is reconstructed as an intersection of the face S and the line $\lfloor P_x \quad P_y \quad 0 \rfloor + \lambda [0 \quad 0 \quad 1]$. The hidden counterpart of V is reconstructed by reflecting (V) with respect to the symmetry plane of the shape. The symmetry plane is determined by the midpoints of three reconstructed pairs. Figure 3 shows a real and a virtual view of an opaque shape which can be reconstructed completely, i.e., not only the front (visible), but also the back (invisible) parts can be reconstructed. For transparent (wire) shapes, steps 2 and 3 can be omitted because all vertices are visible in the image. For such shapes, two 3D shapes are produced (they are depth reversal of one another). For an opaque shape, on the other hand, occlusion eliminates one of these two interpretations. So, paradoxically, opaque shapes, which provide less information in the image, lead to less ambiguity.

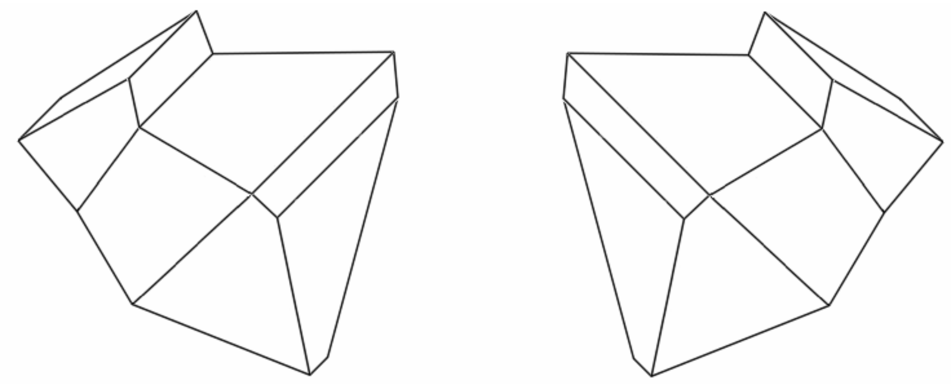

Figure 3. A real (left) and a virtual (right) view of a 3D symmetric opaque shape.

Up to this point, we described how the one-parameter family of 3D shapes is determined. This family is characterized by the value of r_{33}. For each value of r_{33}, one, or at most two, shapes are reconstructed. All 3D shapes from this family project to the same 2D image (the real view). From this infinitely large family of 3D shapes, we choose a shape with maximum compactness (V^2 / S^3, V and S are the volume and surface area of the reconstructed shape, respectively). Shape whose compactness is maximal is the shape whose volume is maximal for a given surface area. For the class of shapes we used, there was always a unique local maximum of compactness. Interestingly, the 3D shape whose compactness is maximal is usually very close to the original shape that produced the 2D image from which the reconstruction was performed. Figure 4 shows an example. Several images of the original and reconstructed 3D shape are shown and it is clear that there two 3D shapes are almost identical. One reason for why maximum compactness constraint works so well is that compact 3D shapes always produce compact 2D images, whereas non-compact 3D shapes are unlikely to produce compact 2D images.

The next section presents preliminary psychophysical results that provide support for the new algorithm.

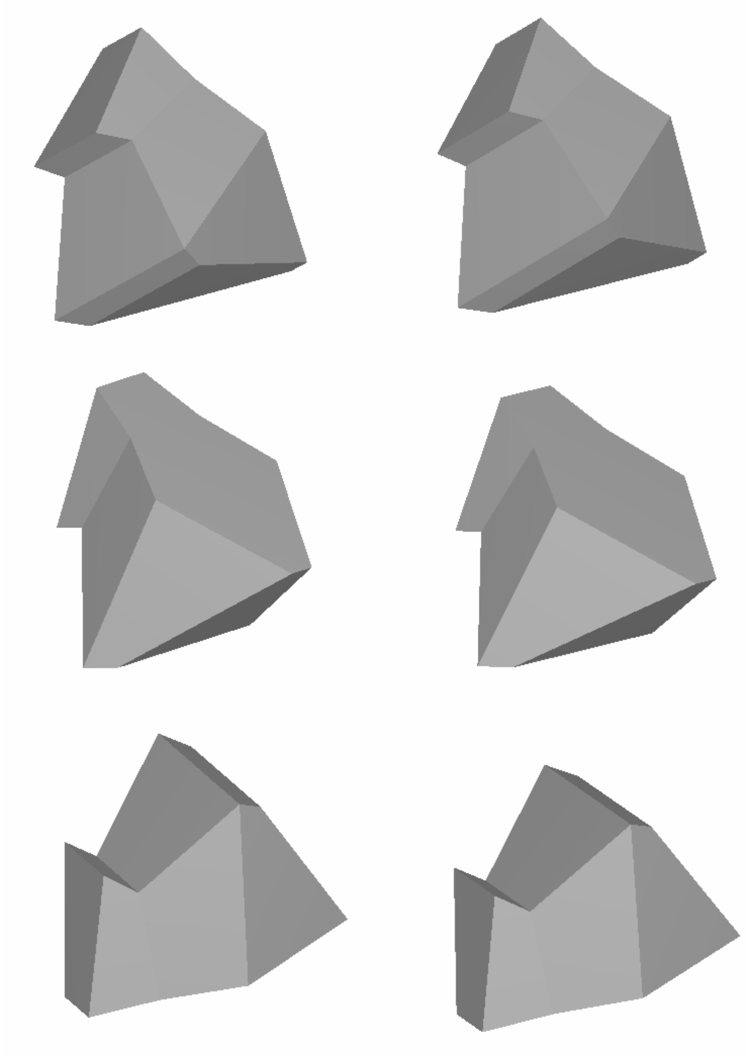

Figure 4. Three images of the original (left) and reconstructed (right) 3D symmetric shape. The reconstruction is close to perfect.

PSYCHOPHYSICAL TEST OF THE NEW ALGORITHM

To test whether the perceived 3D shape is close to the one with maximal compactness, reconstruction experiment was designed. In this experiment, one 2D image of a 3D symmetric polyhedron (like that in Figure 1) was shown on the top of a computer monitor. When the subject looks at a 2D image of a symmetric polyhedron, he or she perceives a 3D symmetric polyhedron. The question is, which shape, from the infinitely many 3D symmetric shapes that are consistent with this image, the subject actually perceives. In order to determine this, a 3D rotating shape was shown on the bottom. The 3D shape was taken from the one-parameter family of shapes consistent with the given 2D image. That is, the 3D shape was symmetric, but its shape could be adjusted by changing the value of r_{33} (see description of the algorithm). The subject was asked to adjust the position of a slide bar representing the value of r_{33} until the rotating 3D shape was same as the perceived shape when looking at the 2D image on the top of the monitor. The compactness Cp of the perceived shape, of the original (real) shape Cr as well as the maximal compactness Cm were computed and stored for each trial. Then, compactness of the perceived shape and of the original shape were normalized to the maximal compactness:

$$RCP = (Cm - Cp) / Cm$$

$$RCR = (Cm - Cr) / Cm$$

One of the authors (YL) was tested in 120 trials, each trial with a randomly generated polyhedron. For each polyhedron, a random 3D orientation was chosen and a 2D image computed. The results are shown in Figures 5 and 6. Figure 5 shows a joint frequency histogram of the relative compactness RCP and RCR. It can be seen that about 1/3 of the original shapes had compactness very close to maximal (RCR less than 0.01) and they were perceived as such, by the subject (RCP less than 0.01). Only a handful of shapes had perceived compactness very different from the maximal compactness (RCP >0.03). Figure 6 shows cumulative relative frequency distribution for RCP and RCR. It can be seen that the curve for perceived shapes increases somewhat faster than that for real shapes. For example, 70% of perceived shapes had relative compactness less than 0.01, while only 60% of original shapes had relative compactness within this range. This means that the human percept is biased towards shapes with maximum compactness.

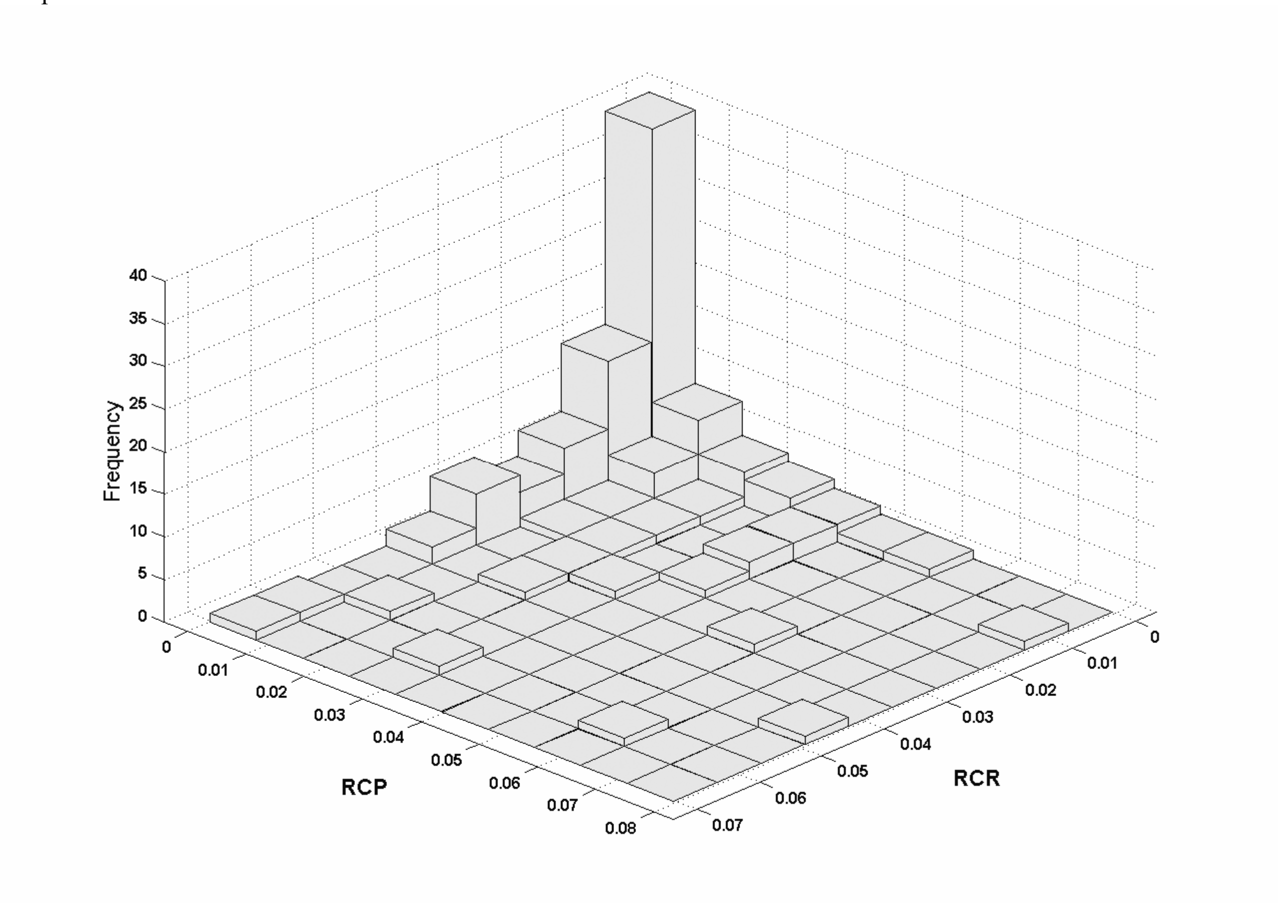

Figure 5: Frequency histogram of relative compactness for real shapes (RCR) and perceived shapes (RCP)

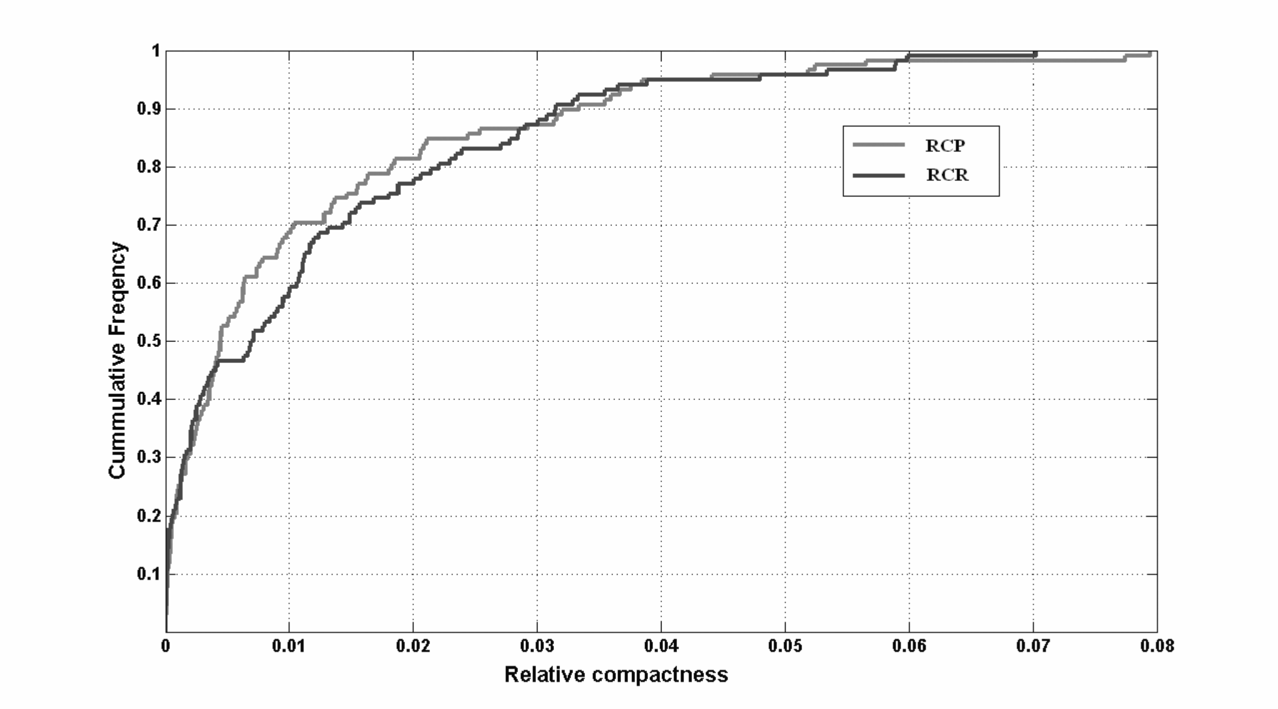

Figure 6: Cumulative relative frequency distribution of relative compactness for real and perceived shapes

SUMMARY AND DISCUSSION

We described a new algorithm for reconstructing 3D shapes. The main idea of the algorithm is based on the regularization method of solving inverse problems [4, 5]. There are two main differences between the conventional regularization methods and the one described here. First, our algorithm uses spatially global, rather than spatially local constraints. Spatially local constraints, such as smoothness of surfaces are attractive because they are quite general. However, they are not strong enough, and, as a result, they cannot produce a unique 3D shape from a single 2D image. This can be accomplished only by spatially global constraints, such as symmetry. The second difference is the use of 3D compactness in our algorithm. This constraint has not been used before in 3D shape reconstruction. Note that symmetry and, especially compactness are very general. They can be applied to polyhedral, as well as smoothly curved shapes. Considering the fact that maximum compactness leads to reconstructions that are similar to the percept of a human observer, it seems reasonable to expect that our algorithm will provide computational basis for a model of human 3D shape perception. Note that our model is different from prior models of human shape perception. We do not use 3D surfaces, as Marr did, nor do we use geons. Our theory is an extension of the theories of Gestalt Psychologists[13], of Hochberg & McAlister [14] and Perkins [15].

The current implementation of our algorithm uses orthographic images. If a single perspective image of a symmetric shape is given, a unique reconstruction is possible, in principle, even without constraints. However, the reconstruction is computationally unstable in the presence of noise, for the same reason that binocular reconstruction is unstable [16]. Therefore, compactness constraint has to be used with perspective images, as well. Next, we want to point out that our algorithm assumes that the image has already been segmented and the vertices and contours of faces have been established. Finally, we assume that it is known which points in the image represent pairs of symmetric points. Our future work will address the issue of symmetry detection, as well as robust image segmentation.

ACKNOWLEDGEMENT

Yunfeng Li was supported by a fellowship from Purdue Research Foundation.

REFERENCES

1. D. Marr, *Vision*, Freeman, San Francisco, 1982.
2. I. Biederman, “Recognition-by-components: A theory of human image understanding”, *Psychological Review,* 94, 115-147 (1987).
3. T. Poggio, & S. Edelman, “A network that learns to recognize three-dimensional objects”, *Nature,* 343, 263-266 (1990).
4. T. Poggio, V. Torre, C. Koch, “Computational vision and regularization theory”, *Nature,* 317, 314-319 (1985).
5. Z. Pizlo, “Perception viewed as an inverse problem”, *Vision Research,* 41(24), 3145-3161 (2001).
6. M.W. Chan, A.K. Stevenson, Y. Li, & Z. Pizlo, “Binocular shape constancy from novel views: the role of a priori constraints”. *Perception & Psychophysics (in press).*
7. S. Ullman, *The interpretation of visual motion*, MIT press, 1979.
8. S. Ullman, “Maximizing rigidity: the incremental recovery of 3D structure from rigid and nonrigid motion”. *Perception,* 13, 255-274 (1984)
9. T. Vetter, T. Poggio, “Symmetric 3D shapes are an easy case for 2D object recognition”. In C.W. Tyler (Eds.), *Human symmetry perception and its computational analysis*, 349-359. Lawrence Erlbaum, New Jersey, 2002.
10. M. Brady, & A. Yuille, “Inferring 3D orientation from 2D contour (an extremum principle)”. In: Richards, W. (Ed.), *Natural computation*, 99-106, Cambridge, MA: MIT Press, 1983
11. T.S. Huang, C.H. Lee, “Motion and structure from orthographic projections”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, 536-40 (1989).
12. J.J. Koenderink, J.A. van Doorn, “Affine structure from motion”, *J. Opt. Soc. Am. A,* 8, 377-385 (1991).
13. K. Koffka, *Principles of Gestalt Psychology*, Harcourt Brace, New York, 1935.
14. J. Hochberg, E. McAlister, “A quantitative approach to figural “goodness””, *J. Exp. Psychol.,* 46, 361-364 (1953).
15. D.N. Perkins, “How good a bet is good form?”, *Perception,* 5, 393-406, 1976.
16. M.W. Chan, Z. Pizlo, D.M. Chelberg, “Binocular shape reconstruction: psychological plausibility of the 8 point algorithm”. *Computer vision & image understanding,* 74, 121-137 (1999).

APPENDIX A: DERIVATION OF $[r_{11}, r_{12}, r_{21}, r_{22}]$

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \text{ is orthonormal}$$

$$\begin{aligned}
& r_{11}r_{22} - r_{12}r_{21} = r_{33} && 1' \\
& r_{13}r_{21} - r_{23}r_{11} = r_{32} && 2' \\
\Rightarrow\quad & r_{23}r_{12} - r_{13}r_{22} = r_{31} && 3' \\
& r_{32}r_{21} - r_{31}r_{22} = r_{13} && 4' \\
& r_{31}r_{12} - r_{32}r_{11} = r_{23} && 5'
\end{aligned}$$

These are five equations with four unknowns. The last four equations are dependent[♠], so only three of them can be used to calculate $r_{11,}r_{12}, r_{21}, r_{22}$

Suppose the first four of the five equations are used.

From Equation 3’ $r_{12} = \frac{r_{31} + r_{13}r_{22}}{r_{23}}$ 6’

From Equation 4’ $r_{21} = \frac{r_{13} + r_{31}r_{22}}{r_{32}}$ 7’

From Equation 2’ $r_{11} = \frac{-r_{32} + r_{13}r_{21}}{r_{23}} = \frac{-r_{32} + r_{13}\frac{r_{13} + r_{31}r_{22}}{r_{32}}}{r_{23}} = \frac{-r_{32}^2 + r_{13}^2 + r_{13}r_{31}r_{22}}{r_{23}r_{32}}$ 8’

Combine equations 1’, 6’, 7’ and 8’

$$r_{11}r_{22} - r_{12}r_{21} = r_{33}$$

$$\Leftrightarrow \frac{-r_{32}^2 + r_{13}^2 + r_{13}r_{31}r_{22}}{r_{23}r_{32}} r_{22} - \frac{r_{31} + r_{13}r_{22}}{r_{23}} \frac{r_{13} + r_{31}r_{22}}{r_{32}} = r_{33}$$

$$\Leftrightarrow \frac{(-r_{32}^2 + r_{13}^2)r_{22} + r_{13}r_{31}r_{22}^2}{r_{23}r_{32}} - \frac{r_{31}r_{13} + (r_{13}^2 + r_{31}^2)r_{22} + r_{13}r_{31}r_{22}^2}{r_{23}r_{32}} = r_{33}$$

$$\Leftrightarrow \frac{-r_{31}r_{13} + (-r_{32}^2 - r_{31}^2)r_{22}}{r_{23}r_{32}} = r_{33}$$

$$\Leftrightarrow \frac{-r_{31}r_{13} - (1 - r_{33}^2)r_{22}}{r_{23}r_{32}} = r_{33}$$

$$\Rightarrow r_{22} = -\frac{r_{33}r_{23}r_{32} + r_{31}r_{13}}{1 - r_{33}^2}$$

[♠] The four equations can be written as $\begin{bmatrix} -r_{23} & 0 & r_{13} & 0 \\ 0 & r_{23} & 0 & -r_{13} \\ 0 & 0 & r_{32} & -r_{31} \\ -r_{32} & r_{31} & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \end{bmatrix} = \begin{bmatrix} r_{32} \\ r_{31} \\ r_{13} \\ r_{23} \end{bmatrix}$

$$\det \begin{bmatrix} -r_{23} & 0 & r_{13} & 0 \\ 0 & r_{23} & 0 & -r_{13} \\ 0 & 0 & r_{32} & -r_{31} \\ -r_{32} & r_{31} & 0 & 0 \end{bmatrix} = -r_{23}r_{13}r_{32}r_{31} + r_{13}r_{23}r_{31}r_{32} = 0$$