

Representing and Comparing Shapes Using Shape Polynomials

Gabriel Taubin[†], Ruud M. Bolle[‡], and David B. Cooper[†]

[†]Laboratory for Engineering Man/Machine Systems
Brown University, Division of Engineering,
Providence, RI 02912

[‡]Exploratory Computer Vision Group
IBM T. J. Watson Research Center,
P.O. Box 704, Yorktown Heights, NY 10598

Abstract

We address the problem of multiresolution 2D and 3D shape representation. Shape is defined as a probability measure with compact support. Both object representations, typically sets of curves and/or surface patches, and observations, sets of scattered data, can be represented in this way. Global properties of shapes are defined as expectations (statistical averages) of certain functions. In particular, the moments of the shapes are global properties. For any shape S , and every integer $d > 0$, we associate a *shape polynomial* of degree $2d$, whose coefficients are functions of the moments of S . These polynomials are related to the shape S in an affine invariant way. They yield small values near S and large values far away and their level sets approximate S . With the shape polynomials we define two distances between shapes. An asymmetric distance measures how well one shape fits as a *subset* of another one; a symmetric version indicates how *equal* two shapes are. The evaluation of these distance measures is determined via a sequence of computationally very fast matrix operations. The distance measures are used for recognition and positioning of objects in occluded environments.

1 Introduction

Much computer vision research has been focussed on how to represent two-dimensional and three-dimensional objects. Since the central problems in model-based computer vision are object modeling, sensor data (images, depth maps, etc.) generation modeling, and model inferring from sensor data, the representation of 2D and 3D shapes is a basic problem. A useful theory of shape will also have applications in other areas. Such applications include computer graphics, design automation, manufacturing automation, terrain mapping, vehicle guidance, surveillance, and in-

telligent robots.

One can distinguish boundary and region/volumetric representations. For the 3D case, the distinction between boundary and volumetric representations is somewhat artificial when dealing with visual data. What can be measured as input for a visual system are points on the object surface, obtained either by passive or active means [3]. If a shape-from- x technique [1] is used, the input may be surface normals. Information about the surface of the scene, and not about volume, is the input. Hence, the accessibility criterion of Marr and Nishihara [16] suggests a boundary representation for 3D objects. Also, Brady's criterion [8], that shape should be locally reconstructible, suggests a boundary representation for 2D shapes. In this paper, we will concentrate on 2D and 3D boundary representations; the theory can be extended to any dimension.

1.1 Some existing boundary representations

Let us review some 2D/3D shape representations. A 2D object can be represented by a 1D curve that bounds a "well-behaved" region [3]. Examples are the chain code [13], or implicit functions describing the 2D boundary. In the 3D case, implicit functions, for example, quadric surfaces popularized by Faugeras et al. [12], Bolle and Cooper [5, 6], and others [9, 14, 10], can be used to describe the surface of an object. Here, a complex surface is represented by patches of primitive surfaces. Superquadrics [4, 2, 7, 17], an extension of quadric surfaces, fall in the same category. An example of a superquadric is an ellipsoid that can smoothly deform into various exotic shapes. More recently, Taubin [18] has been using high-degree polynomials which permit the representation of complex collections of 2D curves and 3D surfaces with a single analytic function.

1.2 A new shape representation

A representation of an object is typically a vector of features or a specification of the boundary of the object. Usually, a boundary representation is given by a finite set of curve segments and/or surface patches. Since 3D sensors provide scattered sets of points in three-space, which can be sparse for passive stereo (using isolated features) or tactile data, or dense for active stereo, shape reconstruction has to be achieved from sets of points, and the reconstruction has to be compared to the representation of the object. In this paper, we present a new view on boundary-based representations of shape and observations of such shapes. This view allows for a unified treatment of both data sets corresponding to observations of objects and representations of objects. In Section 2, we define *shape* as a probability measure with compact support, i.e., a distribution of *mass*. We define *properties* of shapes as expectations, or averages, of functions in Section 3. In particular, all the moments of a shape can be expressed in this way. In Section 4, we construct the *shape polynomials* from the higher-order moments of a shape. These polynomials yield low values near the original shape and large values elsewhere. Section 5 introduces distances between shapes by computing the expectation of a shape polynomial, corresponding to one shape, with respect to another shape. In Section 6, we discuss some applications and we describe a small part of an experimental 2D object recognition system.

2 Definition of shape

Let n be the dimension of the space. Although the theory is independent of the dimension, in this paper, $n = 2$ or $n = 3$. Hence, a point $\mathbf{x} = (x, y)^t$ or $\mathbf{x} = (x, y, z)^t$ refers to a point in two-space and three-space, respectively.

2.1 Motivation

Let $\mathcal{D} = \{\mathbf{x}_1, \dots, \mathbf{x}_q\}$ be a set of scattered data points in n -space. The mean value $\mu = \frac{1}{q} \sum_{i=1}^q \mathbf{x}_i$ and covariance matrix $\Sigma = \frac{1}{q} \sum_{i=1}^q (\mathbf{x}_i - \mu)(\mathbf{x}_i - \mu)^t$, provide information about the position, orientation, and spread of \mathcal{D} . Particularly, the polynomial

$$(\mathbf{x} - \mu)^t \Sigma^{-1} (\mathbf{x} - \mu) + 1 \quad (1)$$

summarizes all this information. Now, by a *sublevel set*, we will mean the set of points where this polynomial is less than or equal to a constant. These sublevel sets are a family of nested hyperellipsoidal regions centered at μ and with principal axes in the direction of the eigenvectors of Σ . In almost all the cases, the sublevel sets do not provide a good description of the shape of the original data set \mathcal{D} , because the polynomial (1) is a function

of only low-order moments of \mathcal{D} . We will extend (1) by including higher order moments of the data, thereby capturing the information required to better approximate the data set. We will call these polynomials *shape polynomials*.

2.2 Shape as a probability measure

In Mechanics one usually deals with *point masses*, and with *curve, surface, and volume mass densities*. In Electrostatics, the objects of study are *point charges* and *curve, surface, and volume charge densities*. A unified foundation for both cases is provided by Measure and Probability Theory.

Before introducing our definition of shape, let us recall some definitions. A probability measure P is *concentrated* in a set S , or, alternatively, the set S *supports* P if $P(S) = 1$. The *support* of P is the smallest closed set which supports P . The probability measure P has *compact support* if its support is bounded. And finally, if the support of P does not contain any open set in the usual Euclidean norm topology, then P is a *singular* probability measure.

Although our typical data sets can be represented by singular probability measures because the data lies on points, curves and surfaces, we define a *shape* as a probability measure with compact support. With this definition, a single point is a *shape*, the well-known Dirac δ . A finite set of q points, as the data set \mathcal{D} from above, can be represented in a similar fashion, as a sum of δ 's with uniform weight $1/q$.

In Mechanics and Electrostatics, we can measure how much mass or charge lies in a certain region of space by computing the curvilinear or surface integral of the mass or charge density, along the part of the curve or surface present inside the given region. In our case, the amount of *shape* present in a specific region of space is just the fraction of length or area of the curve or surface included in the given region. The shape of a set of curve segments or surface patches is the weighted linear combination of the corresponding parts.

3 Properties of shapes

A property of a single point *shape* is just the value that a certain function attains at the given point. For a finite set of points, a property is the average value of the same function over all the points in the set. In general, we define a *property* of a shape as the expected value of a certain function.

For example, for a finite set of points $\mathcal{D} = \{\mathbf{x}_1, \dots, \mathbf{x}_q\}$, the mean of \mathcal{D} is the expectation of the vector function \mathbf{x} , i.e., $\mu = \frac{1}{q} \sum_{i=1}^q \mathbf{x}_i$. The covariance of \mathcal{D} is the expectation of the matrix function $(\mathbf{x} - \mu)(\mathbf{x} - \mu)^t$, i.e., $\Sigma = \frac{1}{q} \sum_{i=1}^q (\mathbf{x}_i - \mu)(\mathbf{x}_i - \mu)^t$.

3.1 General properties

For any measurable function $\phi(\mathbf{x})$, the property it defines for a single point shape \mathbf{x}_0 is $E_{\mathbf{x}_0}[\phi] = \int \phi(\mathbf{x}) dP_{\mathbf{x}_0} = \phi(\mathbf{x}_0)$. For a finite set of data points \mathcal{D} , the expressions of the previous paragraph can be generalized to the function ϕ . And, for example, the property defined by the function ϕ for a curve segment L yields

$$E_L[\phi] = \int \phi(\mathbf{x}) dP_L = \frac{1}{|L|} \int_L \phi(\mathbf{x}) dl(\mathbf{x}),$$

where $|L|$ is the length of the curve segment L , $|L| = \int_L dl(\mathbf{x})$. Similarly, for a surface patch A , properties are surface integrals. This can be extended to a set of curve segments or surface patches. Hence, it is possible to represent both local and global properties of shape.

3.2 Transformations of shapes and equivalence

In general, a surface or object can be sensed from any position. The points on a surface that are sensed from two sensing positions are related by an affine transformation. We wish to study how our shapes and their properties change with respect to these transformations.

Given a shape S and some affine transformation $\mathbf{T}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, a new shape $\mathbf{T}S$ is induced by the transformation. That is, $\mathbf{T}S$ is the probability measure defined by the equation

$$\mathbf{T}S(B) = S(\mathbf{T}^{-1}[B]),$$

where B is a measurable set. Equivalently, since a probability measure is uniquely defined by the expectations of all possible measurable functions on the space, $\mathbf{T}S$ is also defined by the following equation

$$E_{\mathbf{T}S}[\phi] = E_S[\phi \circ \mathbf{T}],$$

where ϕ is a measurable function. Then, two shapes S and Q are *equivalent* if there exists some affine transformation \mathbf{T} such that $Q = \mathbf{T}S$. Since the family of affine transformations is a group, this is a true relation of equivalence on the shapes.

4 Shape polynomials

In this section, a systematic generalization of the polynomial (1) is introduced, which will enable us to compare shapes.

4.1 Higher order shape polynomials

Let S be a shape. We rewrite (1) in the following way

$$(\mathbf{x} - \mu)^t \Sigma^{-1} (\mathbf{x} - \mu) + 1 = \mathbf{X}^t \begin{pmatrix} 1 & \mu^t \\ \mu & \Sigma + \mu\mu^t \end{pmatrix}^{-1} \mathbf{X} \quad (2)$$



Figure 1: Ellipse

where $\mathbf{X}^t = (1, \mathbf{x}^t)$. Also note that the matrix

$$\begin{pmatrix} 1 & \mu^t \\ \mu & \Sigma + \mu\mu^t \end{pmatrix} = E_S \left[\begin{pmatrix} 1 & \mathbf{x}^t \\ \mathbf{x} & \mathbf{x}\mathbf{x}^t \end{pmatrix} \right] = E_S[\mathbf{X}\mathbf{X}^t].$$

The polynomial of (2) can be generalized to higher order. Given a positive integer d , let \mathbf{X} be the vector of all the monomials of degree $\leq d$. For example, with $d = 2$ and $n = 3$, we have

$$\mathbf{X} = (1, x, y, z, x^2, xy, xz, y^2, yz, z^2)^t.$$

In general, \mathbf{X} is a column vector of $m = \binom{n+d}{n}$ elements. Now, let \mathbf{M}_S be the nonnegative and symmetric matrix $\mathbf{M}_S = E_S[\mathbf{X}\mathbf{X}^t]$. Unless S is perfectly represented by a polynomial in \mathbf{x} of degree d , this matrix is nonsingular. Hence, with certain minor restrictions we can assume that this is the case for the classes of shapes involved [20].

The *shape polynomial* of order d associated with the shape S is

$$\Phi_{S,d}(\mathbf{x}) = \frac{1}{m} \mathbf{X}^t [\mathbf{M}_S]^{-1} \mathbf{X},$$

where m is the dimension of the vector \mathbf{X} . We will write $\Phi_S(\mathbf{x})$ when the order d is clearly understood. Note that $\Phi_{\mathcal{D},1}(\mathbf{x})$ is the polynomial (2). Since the matrix \mathbf{M}_S is positive definite, so is $[\mathbf{M}_S]^{-1}$, and Φ_S is a positive polynomial of degree $2d$, in fact $\Phi_S(\mathbf{x}) \geq 1/m$ for all \mathbf{x} .

To motivate the term *shape polynomials*, we show some examples. These are the graphs of shape polynomials represented as gray-level images, for different values of d . Darker regions correspond to low values and brighter regions to higher values. These are 2D examples, because they are easier to visualize, but the same conclusions hold in the 3D case. Figure 1 shows a set of noisy points located around an ellipse. Figure 2 shows a more complex shape, the boundary of an artificially generated wrench. Figure 3 shows another complex shape but produced by thresholding a real image. Note that as the degree, d , of the shape polynomial increases, $\Phi_{S,d}(\mathbf{x})$ approximates the original data set better. However, even for $d = 2$, the polynomial captures the essence of the data set well.

4.2 Properties of the shape polynomials

We state some properties of the shape polynomials. A detailed analysis, along with the proofs, will be found

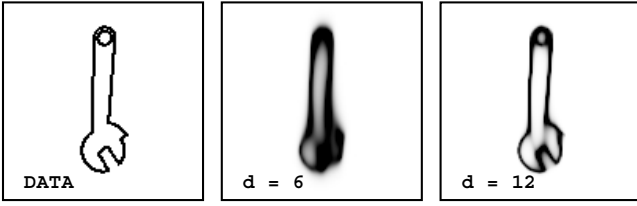


Figure 2: Wrench



Figure 3: Plier

in [20]. The most important property is the invariance. For any shape S and affine transformation T , the expression

$$\Phi_{TS}(\mathbf{x}) = \Phi_S(T^{-1}(\mathbf{x})),$$

is a polynomial identity. Hence, the value of the shape polynomial at a points \mathbf{x} depends only on the *relative* position of the point with respect to the original shape.

Let S be a shape and Φ_S its corresponding shape polynomial. Then there exists a positive number λ such that for every \mathbf{x} , $\Phi_S(\mathbf{x}) \geq \lambda \|\mathbf{x}\|^{2d}$. Hence, $\Phi_S(\mathbf{x})$ grows to ∞ as $\|\mathbf{x}\|$ grows to ∞ . An implication is that at least all of the sublevel sets are *bounded*. This is a very weak way to say that they approximate in some sense the shape of the original data set.

In the same way as we have done here, we can define *trigonometric* shape polynomials, replacing the vector of monomials \mathbf{X} by a vector of complex exponentials. For a large family of shapes we can prove that the sequence of trigonometric shape polynomials converges to the original shape [20]. The present case of regular shape polynomials is very closely related to the trigonometric case.

5 Comparing shapes

In this section, we will use the shape polynomials $\Phi_S(\mathbf{x})$ to define distance measures between shapes.

5.1 The Δ -distance between shapes

Let S and Q be two arbitrary shapes, we define the Δ -distance from S to Q as

$$\Delta_S(Q) = |E_Q[\Phi_S] - 1| = \left| \int \Phi_S(\mathbf{x}) dP_Q - 1 \right|.$$

This distance measure is not a real distance in the sense of metric spaces. For example, we have $\Delta_S(Q) \neq$

$\Delta_Q(S)$. The asymmetry is, in fact, required since we want to measure the distance between an object S and a subobject Q . A symmetric version is useful for other applications

$$\Delta(S, Q) = \frac{E_Q[\Phi_S] + E_S[\Phi_Q]}{2} - 1.$$

5.2 Properties of the Δ -distances

This section is a consequence of the properties of the shape polynomials. The Δ -distance from a shape S to itself is zero, i.e.,

$$\Delta_S(S) = 0.$$

The Δ -distance, $\Delta_S(Q)$ is invariant with respect to simultaneous affine transformations of shapes S and Q . If $T(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ is an affine transformation, then

$$\Delta_{TS}(TQ) = \Delta_S(Q).$$

Finally, if we move shape Q to ∞ keeping S fixed, the Δ -distance also goes to ∞ .

Similar statements hold for the symmetric version, along with the property

$$\Delta(S, Q) \geq 0.$$

5.3 Computation of the Δ -distance

We give a fast algorithm for the evaluation of $E_Q[\Phi_S]$, which is the essential part of both distance measures. We are interested in the two cases, the distance between the shapes S and Q , $\Delta_S(Q)$, and the distance between S and TQ ,

$$\Delta_S(TQ) = \Delta_{T^{-1}S}(Q).$$

The latter matches two shapes by transforming one with an affine transformation T . It can be shown [20] that

$$\begin{aligned} E_Q[\Phi_S] &= \int \Phi_S(\mathbf{x}) dP_Q \\ &= \frac{1}{m} \text{trace} \{ \mathbf{M}_Q \mathbf{M}_S^{-1} \} = \frac{1}{m} \|\mathbf{L}_S^{-1} \mathbf{L}_Q\|_2^2, \end{aligned}$$

where $\|A\|_2^2$ is the sum of the squares of the entries of the matrix A , and \mathbf{L}_S and \mathbf{L}_Q are the Cholesky decompositions of \mathbf{M}_S and \mathbf{M}_Q , respectively, with \mathbf{L}_S a lower triangular matrix with positive diagonal elements, and similarly for \mathbf{M}_Q and \mathbf{L}_Q . The right-hand side of the above expression is the sum of the squares of the entries of the matrix. The use of the Cholesky decomposition improves the numerical accuracy of the computation.

Given an affine transformation T , there exists a unique matrix T^* , which is a function of T only, such that $\mathbf{X}(T(\mathbf{x})) \equiv T^* \mathbf{X}$ is a polynomial identity. The entries of the matrix T^* are polynomials in the parameters of T for which analytic expressions are known [18]. To compute $\Delta_S(TQ)$, we have

$$E_{TQ}[\Phi_S] = \|\mathbf{L}_S^{-1} T^* \mathbf{L}_Q\|_2^2.$$

6 Applications

We start this section using the symmetric Δ -distance to estimate the position and orientation of one known and unoccluded object. Then we extend the procedure to the classification of one unoccluded object among a finite set of known models. Finally, we describe how to use both distance measures for recognizing objects in a cluttered environment. For the first two problems, the model of an object is a matrix of moments. For the latter, the object model is a collection of such matrices, corresponding to significant parts of the object.

6.1 One unoccluded object and one model

Let the data be represented by the shape S , the model by the shape Q , and let \mathbf{T} be the linear transformation that matches the model with the data. For a fixed order d , the *best match* is produced by the minimizer $\hat{\mathbf{T}}$ of the expression $\Delta(S, \mathbf{T}Q)$. This is a multimodal polynomial function of the parameters \mathbf{T} . To compute $\hat{\mathbf{T}}$ we need a good initial approximation. We begin by using second degree polynomials, $d = 1$, for the approximation. Then the matrix M_S is just the usual scatter matrix of the data in the plane or 3-space. For $d = 1$ and \mathbf{T} a rigid body transformation, determination of the estimate $\hat{\mathbf{T}}$ has an analytic solution – two solutions in 2-space and four in 3-space, in most of the cases. One can first rotate the model until the eigenvectors of its scatter matrix align with the corresponding ones of the data, thus estimating the orientation of the object being measured, and then translate the model until its mean coincides with the mean of the data, thus estimating the location of the object. Since there is more than one rotation possible, the ambiguity is then resolved by computing the symmetric distance for a higher degree approximation, i.e., some $d > 1$. Then, this initial transformation can be improved by a gradient descent technique. However, the initial transformation is a very good approximation to the global minimum.

6.2 One unoccluded object and several models

Here, we have to determine which model, Q_1, \dots, Q_r , is the one which *best* matches the observed data, S . For each model Q_i , we compute an initial transformation \mathbf{T}_i , as described in the previous section. Then we classify S as belonging to the class of Q_j , if

$$\Delta(S, \mathbf{T}_j Q_j) = \min \{ \Delta(S, \mathbf{T}_i Q_i) : 1 \leq i \leq r \}.$$

Only for the model that minimizes this expression, the Δ -distance for some $d > 1$ is iteratively minimized. Although the computation of the Δ -distances is not very expensive, the number of evaluations can be reduced. If the Δ -distance for $d = 1$ from the data to an object is

large, there is no good match. For $d = 1$ the Δ -distance between two shapes is an analytic expression, a sum of quotients of corresponding eigenvalues [20]. Low value of order one Δ -distance corresponds to almost equal eigenvalues of the data and model scatter matrices. Therefore, the first step of the classification can be implemented with a hash table [15] using a properly quantized function of the 2D or 3D eigenvalue space. That is, the set of eigenvalues is used as the hash index and allows for quickly pruning the search space.

6.3 Modeling in terms of important parts

Our goal is to use the above procedures to develop techniques for recognizing objects in such hostile environments as, for example, a bin of parts or in a cluttered environment. Recognition is based on matching small regions of observed data with parts of known models. The small regions will be assumed to constitute *unoccluded* observations of the corresponding parts of the models. If an object is modeled as a collection of geometrically related subobjects, the problem is reduced to finding the best collection of pairings between model subobjects and regions of observed data that satisfies the same set of global constraints.

We need a procedure to chose *important* parts of a model or data set. This procedure is described below. To speed up the recognition steps, the models are preprocessed. For each model, a *standard position* is computed, using the mean vector and eigenvectors of its scatter matrix. Every *important* part of a model is transformed to its own standard position relative to the model coordinate system. This transformation is stored, along with the Cholesky decomposition of the matrix of moments and the inverse of this matrix. This significantly reduces the number of operations in the recognition phase. All of the important parts are organized in the hash table described in the previous subsection. That is, the information about all the important parts of all the models is stored in a hash table indexed by eigenvalues.

The important parts were chosen as the data within a circular (rotation invariant) window of fixed radius. The center of each window was chosen at random and then adjusted, iteratively, until it was close to the mean of the data inside the window. This procedure tends to find windows containing high-curvature points and, more generally, regions with much structure, i.e., many data points. Other criteria for locating special interest regions are under study. Figure 4 shows all the windows computed for the data set and the two models used in the experiments.

Both the observation and model data of Figure 4 were acquired from images of tools taken against a white background. These images were thresholded to extract object boundaries. These boundaries, including deformations because of specular reflections and shad-

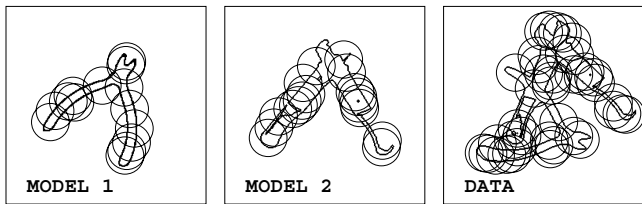


Figure 4: Important Parts

ows were used as data.

6.4 Occluded objects and several models

In the recognition phase, the important parts of the data are extracted following the same procedure as for the models (Figure 4). A standard position is computed for each part, relative to a global coordinate system associated with the data. Hence, windows of data are chosen by a criterion that should result in windows corresponding to those chosen in the model. The data in a window is translated and rotated to put the mean and scatter matrix eigenvalues in standard position. Initial local matches are made of the eigenvalues for a window with those for model subparts; the eigenvalues of the data within the windows are used as an index in the hash table that contains the models.

The next stage in the recognition process is the combinatorial correspondence problem. Suppose that at a certain point in the recognition process, several important parts of the data are assigned to certain parts of the same model, and these assignments are spatially consistent. The union of these data points constitutes a subobject of the corresponding model and the asymmetric Δ -distance is used to measure how well this data fits as a subset of the model. Simultaneously, other groups of important data parts will be assigned to other models or the same model, in a different position. The sum of these asymmetric Δ -distances, from groups of consistent data assignments to models, is a global consistency measure to be minimized.

One of the possibilities is a sequential algorithm based on the classification procedure described above. The best match is computed for every important part of the data; then consistent pairs are grouped together following a verification procedure. Another possibility is to use some kind of stochastic minimization technique, for example, simulated annealing for optimizing the global performance function. Yet another possibility is explored in the following. This possibility can be seen as a way to compute an initial configuration for stochastic relaxation.

Global hypotheses are generated by a weighted voting scheme based on purely local matching. Every possible pair of data and model regions, determined by entries in the hash table, generated an hypothesis for a model along with a transformation. The votes for

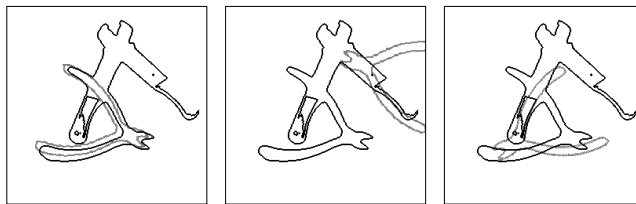


Figure 5: Hypothesis generated for model 1

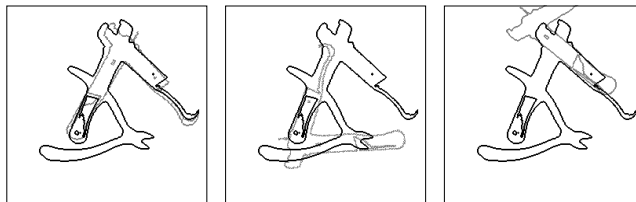


Figure 6: Hypothesis generated for model 2

model and transformation pairs were accumulated in a parameter space corresponding to the group of transformations. The weight of each vote was computed as $1/(1 + K\Delta)$, with Δ the symmetric distance between model and data part and K a positive constant. Finally, peaks were extracted as global hypotheses about objects at certain locations.

This experiment was designed to understand and show the usefulness of the Δ -distance. Figures 5 and 6 show the hypotheses generated for the models and data set shown in figure 4. These hypotheses were generated by matching procedures based on the eigenvectors and eigenvalues of the local scatter matrices. That is, the initial transformations estimated from the distance measure of order $d = 1$ were used. The hypotheses have to be verified, using the global consistency measure defined by the asymmetric Δ -distance. The location estimates of objects can be improved by a gradient descent technique applied to the global asymmetric Δ -distance. These techniques have been applied in closely related problems [18].

7 Discussion

By defining shape as a probability measure with compact support, object representations (curves, surfaces) and observations of such objects (scattered data points) are treated in a unified fashion. The concept of shape polynomials captures the properties of these fundamentally different shape types at multiple resolutions. Using the shape polynomials, two meaningful and computationally efficient measures of shape similarity are defined, a symmetric and an asymmetric version. It is shown that these tools can be used to recognize, classify, and estimate the positions of unoccluded objects, which is a part of the general problem of recognizing occluded objects in a complicated scene. Finally, it is

demonstrated that the asymmetric version of the distance measure can be used to implement a global consistency measure for the correspondence problem.

Some of the concepts introduced in this paper are generalizations of well-established concepts. The Δ -distances provide meaningful measures to fit data points or surface specifications of one object to that of other objects. Special cases of this are well-known, for example, methods to match points to planar patches. Though representing models and data by one or more polynomial surface patches [18] is an attractive approach there are two difficulties that are eliminated with the approach in this paper. The first is that the zeros of a polynomial generally extend out to infinity. Hence, data points far from and not associated with an object can still be very close to its polynomial (e.g., a plane extends to infinity). The other problem is that polynomial surface fitting requires on-line matrix eigenvalue finding. Shape polynomials in this paper take low values only in the vicinity of the model, and the computation to compute the distance measure is less than for the polynomial surface fitting.

The representation of one object as a collection of subobjects has been used before, however, the subobjects have been simpler, such as quadric patches. For 2D edge images, our representation allows us to use the information provided, not only by the silhouette, but by the internal boundary structure too. This representation is a generalization of the concept of interest points, such as intersections of lines and high curvature points. Our interest points are subregions with a relatively complex structure, i.e., regions that can only be represented with higher d .

Finally, we are in the process of implementing a general vision system. The results are very promising, and we believe that the methods for shape representation and comparison are extremely consequential.

8 Acknowledgement

This work was partially supported by an IBM Manufacturing Research Fellowship and National Science Foundation Grant IRI-8715774.

References

- [1] J. Aloimonos, "Visual shape computation," *Proc. of the IEEE*, Vol. 76, no. 8, Aug. 1988, pp. 899-916.
- [2] R Bajcsy & F. Solina, "Three dimensional object representation revisited," *Proc. First Int'l Conf. on Comp. Vision*, London, June 1987, pp 231-240
- [3] D.H. Ballard & C.M. Brown, *Computer Vision*, New Jersey, Prentice-Hall, Inc., 1982
- [4] A.H. Barr, "Superquadrics and angle-preserving transformations," *IEEE Comp. Graphics and Applications*, Jan. 1981, pp. 11-23.
- [5] R.M. Bolle & D.B. Cooper, "Bayesian recognition of local 3-D shape by approximating image intensity functions with quadric polynomials," *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 6, July 1984, pp. 418-429.
- [6] R.M. Bolle & D.B. Cooper, "On optimally combining pieces of information, with application to estimating 3-D complex-object position from range data," *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-8, Sept. 1986, pp. 619-638.
- [7] T.E. Boult & A.D. Gross, "Recovery of superquadrics from depth information," *Proc. AAAI Workshop on Spatial-Reasoning and Multisensor Integration*, Oct. 1987.
- [8] M. Brady, "Criteria for representing shape," *Human and Machine Vision*, A. Rosenfeld and J. Beck (eds.), New York, Academic Press, 1983.
- [9] B. Cernuschi-Frias, "Orientation and location parameter estimation of quadric surfaces in 3-D from a sequence of images," *Brown University Ph. Dissertation*, May 1984; also *Brown Univ TR LEMS-5*, Feb. 1984.
- [10] R.D. Rimey & F.S. Cohen, "Maximum Likelihood Segmentation of Range Data," *IEEE Trans. on Robotics & Automation*, April 1988.
- [11] D.B. Cooper, Y.P. Hung & G. Taubin, "A new model-based stereo approach for 3D surface reconstruction using contours on the surface pattern," *Proc. Second Int'l Conf. on Comp. Vision*, December 1988.
- [12] O.D. Faugeras & M. Hebert, "A 3-D recognition and positioning algorithm using geometric matching between primitive surfaces," *Proc. 8th Int. Joint Conf. on Artificial Intell.*, Aug. 1983, pp. 996-1002.
- [13] H. Freeman, "Computer processing of line drawing images," *Computer Surveys*, Vol.6, No. 1, March 1974, pp. 57-98
- [14] E.L. Hall, J.B.K. Tio, C.A. McPherson, & F.A. Sadjadi, "Measuring curved surfaces for robot vision," *IEEE Computer*, Dec. 1982, pp. 42-54.
- [15] Y. Lamdan, J.T. Schwartz & H.J. Wolfson, "Object Recognition by Affine invariant Matching," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, June 1988
- [16] D. Marr & H.K. Nishihara, "Representation and recognition of the spatial organization of three dimensional structure," *Proc. R. Soc. London B*, Vol. 200, 1978, pp. 269-294.
- [17] A.P. Pentland, "Recognition by parts," *Proc. First Int'l Conf. on Comp. Vision*, London, June 1987, pp. 612-620.
- [18] G. Taubin, "Algebraic Nonplanar Curve and Surface Estimation in 3-Space with Application to Position Estimation," *Brown Univ. TR LEMS-43*; also *IBM Technical Report RC 13873*, February 1988.
- [19] G. Taubin, R.M. Bolle and D.B. Cooper, "Representing and Comparing Shapes Using Shape Polynomials," *IBM Technical Report RC 14292*, December 1988.
- [20] G. Taubin, "About Shape Descriptors and Shape Matching," *Brown Univ. TR LEMS* (in preparation).