MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 1060 July 1988

The Alignment of Objects with Smooth Surfaces

Shimon Ullman* and Ronen Basri†

Abstract: This paper examines the recognition of rigid objects bounded by smooth surfaces, using an alignment approach. The projected image of such an object changes during rotation in a manner that is generally difficult to predict. An approach to this problem is suggested, using the 3-D surface curvature at the points along the silhouette. The curvature information requires a single number for each point along the object's silhouette, the magnitude of the curvature vector at the point. We have implemented this method, and tested it on images of complex 3-D objects. Models of the viewed objects were acquired using three images of each object. The implemented scheme was found to give accurate predictions of the objects' appearances for large transformations. Using this method, a small number of (viewer-centered) models can be used to predict the new appearance of an object from any given viewpoint.

©Massachusetts Institute of Technology (1988)

This report describes research partially done at the Massachusetts Institute of Technology within the Artificial Intelligence Laboratory. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-85-K-0124.

*Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139

†Dept. of Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel

1 Introduction

Visual object recognition requires the identification of objects observed from different viewpoints. In recent years several attempts have been made to approach this problem using an alignment approach [5, 7, 9, 12, 17, 18]. In this paper we shall consider the recognition of rigid objects bounded by smooth surfaces, using an alignment approach.

Alignment is a two stage process. Given a model object and an image object, in the first stage a transformation is sought, that would bring the model object to a position and orientation in space that corresponds to the projected image. Second, the appearance of the model following the transformation is predicted. The result is compared with the actual image, and the degree of match is used to decide whether the image is in fact an instance of the model.

The first stage of the alignment process, namely the search for an aligning transformation, will not be discussed here. The transformation may be determined by a small set of corresponding features, identified in both the model and the image. For example, three non-colinear points on the image, and their corresponding points on the model determine uniquely the transformation [8, 9, 12, 18]. Two points and a line or three lines may also serve for this purpose.

1.1 The Prediction Problem

In this paper we address ourselves to the second stage of the alignment process. We present an approach for solving the following problem: Let $\mathcal{M} = \{M_1, M_2, ...M_n\}$ be a set of object models. Let $\mathcal{T}$ be a set of transfomations, that include rotations in 3-D space, translations and scale changes, followed by an orthographic projection. This projection model assumes that the object is not too close to the camera. Given a model $M \in \mathcal{M}$ and an aligning transformation $T \in \mathcal{T}$, predict the appearance in the image of M following the application of T.

The above definition of the set $\mathcal{T}$ of allowed transformations enables the prediction of the appearance of rigid objects from any given viewpoint. The general prediction problem extends the set $\mathcal{T}$ with other types of transformations, including for example articulated motion and distortion. This extension is beyond the scope of this paper.

In order to utilize edge maps in the image, we make the following definitions: Given an object O and a viewpoint v, the *rim* is the set of all the points on the object's surface, whose normal is perpendicular to the visual axis [11]. This set is also called the *contour generator* [13]. A *silhouette* is an image generated by the orthographic projection of the

rim. In the analysis below we assume that every point along the silhouette is generated by a single rim point, and that the object lies at the neighbourhood of a rim point only on one side of the tangential line of sight.

The prediction of the appearance of smooth objects is not at all straightforward. An edge map of an object usually contains the silhouette, which is generated by its rim. A rim that is generated by a sharp edge is stable on the object as long as the edge is visible. In contrast, a rim that is generated by a smooth surface changes continuously with the viewpoint.

The problem of predicting the new appearance of a smooth object following a rotation is illustrated in Figure 1. The figure shows a bird's eye view of two rotating objects, a cube (a & b) and an ellipsoid (c & d). For both objects points $\mathbf{p}, \mathbf{q}$ lie on the object's rim, and therefore their projections lie in the image on its silhouette. When the cube rotates from position (a) to (b), $\mathbf{p}, \mathbf{q}$ remain on the rim. Their new 3-D position is easily determined, therefore the new silhouette can be predicted in a straightforward mannar. In contrast, when the ellipsoid rotates from position (c) to (d), the new 3-D position of $\mathbf{p}, \mathbf{q}$ is no longer relevant since these points no longer lie on the object's rim. The silhouette is now generated by a new set of points $\mathbf{p}', \mathbf{q}'$ in (d). Figures 1 e & f show the ellipsoid in a Defense view before and after the rotation, compared to its appearance if the rim, as a space curve, had been rotated by the same amount. The conclusion is that the prediction problem for smooth objects is in general significantly more complicated than that of objects with sharp edges.

1.2 Previous Approaches

Two approaches have been suggested in the past to solve this prediction problem. The first approach approximates the object's surface by a set of 3-D wires [1]. The transformation T is applied to each wire. The extreme wires are taken to be the object silhouette. The second approach describes an object as a composition of either surface or volumetric primitives, that have simple geometrical structures [2, 3, 6, 7, 14, 16]. The transformation T is applied to each primitive. Since the primitives have simple geometrical structures, their silhouette can be easily predicted. The extreme points of the collection of the primitives silhouettes are taken to be the object's silhouette.

The decomposition approach works well for simple objects, but usually not for complex ones. The wire approach is often costly from a computational standpoint due to the large number of wires required, and the need to perform "hidden line elimination". Finally, these approaches usually enable the prediction of the bounding contours only.

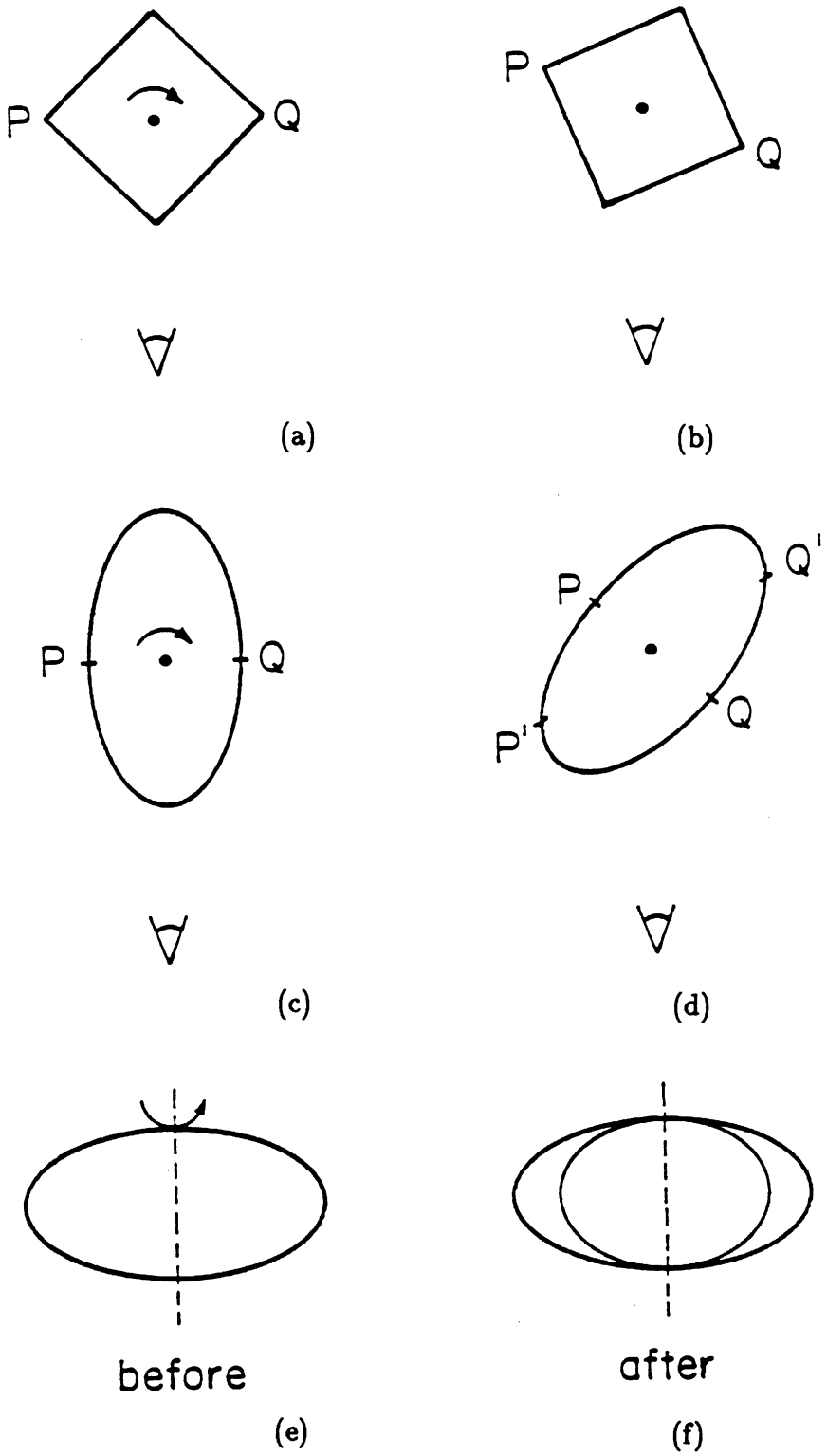


Figure 1: Changes in the rim during rotation. (a) A bird's eye view of a cube. (b) The cube after rotation. In both (a) and (b) points **p**, **q** lie on the rim. (c) A bird's eye view of an ellipsoid. (d) The ellipsoid after rotation. The rim points **p**, **q** in (c) are replaced by $\mathbf{p}'$, $\mathbf{q}'$ in (d). (e) An ellipsoid in a frontal view. (f) The ellipsoid rotated (outer), superimposed on the appearance of the rim, as a space curve after rotation by the same amount (inner).

Internal contours and surface markings, that may have a significant role in shape based recognition, are often not treated.

This paper presents an alternative approach for the prediction problem. In this approach an object is represented by its silhouette, as seen from a particular viewpoint. Using the 3-D surface curvature of each point along the silhouette, it is possible to make an accurate estimation of the silhouette after large transformations. A few models of this kind are sufficient for predicting the object's appearance from any given viewpoint.

2 The Curvature Method

The method is based on representing surface curvature of points along the silhouette. The basic idea is shown in Figure 2. Let X and Y be the main axes of the image plane, and Z-axis be the line of sight. Consider an object O rotating by a rotation R around the vertical axis Y. Let $\mathbf{p}$ be a point on its rim. The figure shows a section of the object through $\mathbf{p}$, that is perpendicular to Y. Let r_x be the curvature radius of $\mathbf{p}$ in this section, and let $\mathbf{r_x}$ be a vector of length r_x parallel to the X-axis. When the object rotates by R, point $\mathbf{p}$ ceases to be a rim point, and it is replaced by a new point $\mathbf{p}'$ approximated by:

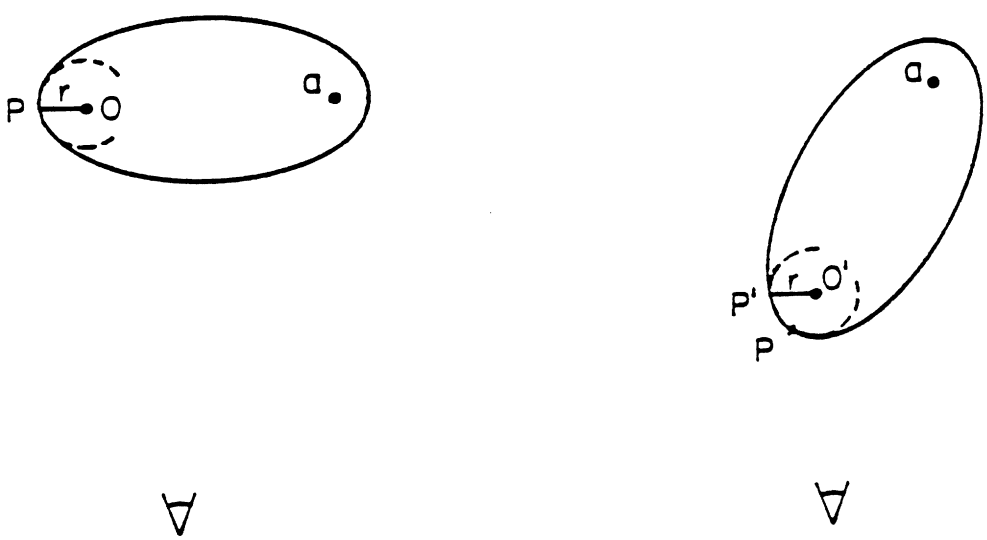
$$\mathbf{p}' \approx R(\mathbf{p} - \mathbf{r_x}) + \mathbf{r_x} \tag{1}$$

The meaning of eq. (1) is the following. The point $\mathbf{o} = \mathbf{p} - \mathbf{r_x}$ is the center of the circle of curvature of $\mathbf{p}$. To predict the new rim point we first apply R to $\mathbf{o}$. Let $\mathbf{o}' = R(\mathbf{p} - \mathbf{r_x})$. The new rim point is then $\mathbf{p}' = \mathbf{o}' + \mathbf{r_x}$. This approximation holds as long as the circle of curvature provides a good approximation to the section at $\mathbf{p}$.

It is worth noting that "sharp" boundaries, such as the cube edges in Fig. 1, or markings on the surface itself, do not require a special treatment. They are included in eq. (1) as the special case $r = 0$.

So far we have considered rotations around the vertical Y axis. We shall next consider the effect on the silhouette of a rotation around an arbitrary axis in space. Any 3-D rotation can be decomposed into two successive rotations: a first rotation around some axis V in the image plane, followed by a second rotation around the Z-axis. The effect of rotating the object around the line of sight Z is, of course, easy to predict. The problem, therefore, is to predict the appearance of the object following a rotation about an axis V lying in the image plane.

In general, the vector of curvature radius $\mathbf{r_x}$ used in eq. (1) would depend on the rotation axis. Let r_x, r_y be the radii of curvature at $\mathbf{p}$ for rotations about the Y and X



(a) (b)

Figure 2: The curvature method. (a) A horizontal section of an ellipsoid. **p** is a point on the rim, r is the radius of curvature at **p**, **o** is the center of the curvature circle, and **a** is the intersection of the Y-axis with this section. (b) The section rotated. **p'** is the new rim point, and it is approximated by eq. (1).

axes respectively. Proposition 1 below states that the radius of curvature for a rotation around any axis can be determined from r_x, r_y alone.

Proposition 1: Let $\mathbf{p}$ be a rim point, and let V_α be an axis lying within the image plane and forming an angle α with the X-axis. The curvature radius at $\mathbf{p}$ for rotations around V_α is given by:

$$r_\alpha = r_y \cos\alpha - r_x \sin\alpha \tag{2}$$

(A proof is given in the appendix).

From this proposition and eq. (1) we can predict the position of $\mathbf{p}'$, the new rim point, for a rotation around an arbitrary axis within the image plane, and consequently any 3-D axis as well, using the two parameters r_x, r_y at $\mathbf{p}$. Proposition 2 below shows that, in fact, a single parameter suffices.

Proposition 2: Let $\mathbf{r} = (r_x, r_y)$ be the curvature vector at $\mathbf{p}$, and let $\mathbf{t}$ be the tangent vector to the silhouette at $\mathbf{p}$. Then $\mathbf{r} \cdot \mathbf{t} = 0$ that is, $\mathbf{r}$ is perpendicular to $\mathbf{t}$ (A proof is given in the appendix).

It follows from the two propositions above that a single number is sufficient to determine the radius of curvature for a rotation around any axis in the image plane. This number is the magnitude of the curvature vector, $\|\mathbf{r}\|$. All other parameters can be derived from it as follows. Let θ be the angle between the tangent vector to the silhouette $\mathbf{t}$ and the X-axis, then:

$$\begin{aligned} r_x &= \|\mathbf{r}\| \sin\theta \\ r_y &= \|\mathbf{r}\| \cos\theta \\ r_\alpha &= \|\mathbf{r}\| \cos(\theta + \alpha) \end{aligned} \tag{3}$$

The scheme is therefore the following. An object model M is a 2-D (orthographic) projection of its visible contours (including its sharp and smooth boundaries, as well as internal markings), as observed from a particular viewing direction. To represent the entire object, a number of views would be required [10]. As shown in the examples below, this number is usually small. Each point along the silhouette has associated with it, along with its spatial coordinates, the magnitude $\|\mathbf{r}\|$ of its curvature vector in eq. (3). Given a transformation T, translation, scaling and rotation around the line of sight are applied to M in a straightforward manner. The effect of rotation around an arbitrary axis in the image plane is computed as follows: first, for each point on the model, the radius of curvature r with respect to the rotation axis is determined using eq. (3). Once r is known, the new position of the point in the image is determined

using eq. (1), where instead of $\mathbf{r_x}$ a vector of size r perpendicular to the rotation axis is plugged in.

In this approach an object is represented using a number of viewer-centered descriptions, rather than a single object-centered representation. Each description covers a range of possible viewing angles, and to represent the entire object a number of descriptions are required. This number depends on the object shape and on the complexity of its aspect graph [10]. As shown in the examples below, this number is expected to be small. The cars in these examples required 4 models to cover all common views, which included all vertical rotations and elevation of $\pm 30°$. because of the orthographic projection approximation, if the object is to be recognized from both large distances as well as close-up views, additional models will be required. The computations required in this scheme during the prediction stage are simple, for example, no hidden-line elimination is necessary.

It is of interest to note that the use of multiple viewer-centered descriptions appears to be consistent with a number of recent psychophysical and physiological investigations. For example, Perret *et al* [15] have investigated the response properties of face-sensitive cells in area STS of the macaque's visual cortex. They have found that cells typically respond to a wide range of 3-D orientations, but not to all viewing directions. A face-selective cell that responds to e.g. a face-on view will typically not respond to a profile view, but will respond to a wide range of intermediate orientations. Perret *et al* conclude that "High level viewer-centered descriptions are an important stage in the analysis of faces" [p. 314]. The use of viewer-centered rather than object-centered descriptions [14], each covering a wide range of orientations, is consistent with the representation scheme described above.

3 Model Construction

We have implemented the alignment scheme described above, and tested it on images of 3-D objects. To apply the scheme, models of the viewed objects must be acquired. For our purpose, an object model must contain the spatial coordinates and the radii of curvature of the object's visible contours. The required 3-D information can be obtained during a learning period using various 3-D cues, such as stereo information and shading.

To estimate the radii of curvature we have used three edge pictures of each object. The radii of curvature are estimated using the following procedure. Consider three silhouette pictures (denoted by A, B and C) that are taken from three different viewpoints along a circle in space perpendicular the Y-axis (Figure 3). Suppose α is the

rotation angle between pictures A and B, and β is the rotation angle between pictures A and C. Let $\mathbf{p_1}$, $\mathbf{p_2}$ and $\mathbf{p_3}$ be three corresponding points in A, B and C respectively. Since the camera was rotated around the Y-axis between successive pictures, all three points share the same y coordinate, so that $\mathbf{p_1} = (x_1, y, z_1)$, $\mathbf{p_2} = (x_2, y, z_2)$ and $\mathbf{p_3} = (x_3, y, z_3)$. According to eq. (1):

$$x_2 = (x_1 - r_x)\cos\alpha + z_1 \sin\alpha + r_x \tag{4}$$

$$x_3 = (x_1 - r_x)\cos\beta + z_1 \sin\beta + r_x \tag{5}$$

These are two linear equations of two unknown parameters z_1 and r_x. Hence:

$$z_1 = \frac{x_1(\cos\alpha - \cos\beta) - x_2(1 - \cos\beta) + x_3(1 - \cos\alpha)}{(1 - \cos\alpha)\sin\beta - \sin\alpha(1 - \cos\beta)} \tag{6}$$

$$r_x = \frac{x_1 \sin(\alpha - \beta) + x_2 \sin\beta - x_3 \sin\alpha}{(1 - \cos\alpha)\sin\beta - \sin\alpha(1 - \cos\beta)} \tag{7}$$

In the range $-\frac{\pi}{2} < \alpha, \beta < \frac{\pi}{2}, \alpha \neq \beta$, the denominator does not vanish.

Using a similar technique we can calculate also r_y. For most of the points an independent estimation of r_y is in principle redundant (see eq. (3)), but it can be used to improve the estimate of the curvature. In this manner five pictures can be used to create a model, three for a rotation around X-axis, and three for a rotation around Y-axis, with the central picture common to both sets. The final model consists of an edge map of the central picture, together with the depth coordinates and the estimated magnitude of the curvature vector (r_x, r_y) at each point.

Note that identifying corresponding points in the pictures is straightforward in this procedure. When the rotation is about the Y-axis, the corresponding points must lie on the same horizontal line. Each contour point therefore usually has a small number of candidate corresponding points to be considered. Details of this matching procedure will not be discussed here.

By setting $\beta = -\alpha$ the equations above can be simplified as follows:

$$z_1 = \frac{x_2 - x_3}{2\sin\alpha} \tag{8}$$

$$r_x = \frac{x_2 + x_3 - 2x_1\cos\alpha}{2(1 - \cos\alpha)} \tag{9}$$

If the angle α is not known, but assuming that α is small we can define new quantities $\hat{z}$, $\hat{r}_x$ as follows:

$$\hat{z} = z\alpha \approx \frac{x_2 - x_3}{2} \tag{10}$$

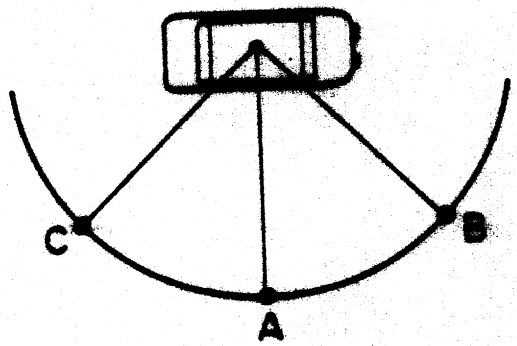


Figure 3: Building a model using three-edge pictures, a bird's eye view. Points A, B and C are the three common locations along a circle in space perpendicular to the Y-axis.

$$\hat{r}_x = (r_x - x_1)\alpha^2 \approx x_2 + x_3 - 2x_1 \tag{11}$$

This approximation uses $\sin\alpha \rightarrow \alpha$ and $\cos\alpha \rightarrow 1 - \frac{\alpha^2}{2}$. In this case the aligning transformation should provide, instead of a rotation angle θ, the ratio $\frac{\theta}{\alpha}$:

$$x' \approx x + z\theta + (r_x - x)\frac{\theta^2}{2} = x + \hat{z}(\frac{\theta}{\alpha}) + \frac{\hat{r}_x}{2}(\frac{\theta}{\alpha})^2 \tag{12}$$

The ratio $\frac{\theta}{\alpha}$ can be determined during the alignment process if we take these approximations into account. Suppose for instance the alignment transformation is determined by a three points correspondence. In this case a set of six equations describing rotation in 3-D space, translation and scale must be solved [9, 18]. If the three points lie on the object's contour, we can substitute two of the six equations, those describing rotations around the X and Y axes, by eq. (12). Consequently we obtain a new set of six equations with six unknown parameters to solve. This set will usually have a small number of solutions, but the details will not be considered here further. The range of rotations covered by a single model would depend on the object's shape and on the similarity to other models. The results shown in section 4 also hold for this approximation.

4 Implementation

A prototype system for object recognition using alignment, that predicts the appearance of objects using the curvature method was implemented on a Symbolics 3670 machine. Pictures comprising of 512x512 pixels were taken, using a vidicon Cohu camera. Edge maps of the pictures were created using the Canny edge-detector [4]. The system first constructs object models comprising of depth values and curvature radii as described in Section 3. Models can be constructed in this system using either three images using rotations around the Y-axis, or five images using rotations around both the Y and X axes. The internal model can then be used to predict the appearance of the model following 3-D rotation, translation and scaling, using the curvature method described in Section 2.

Two models of similar cars, a VW and a Saab, were created (Figure 4). For each model three pictures were taken, with α and β (the angles between successive pictures, see Section 3) being $\pm 30°$ around the Y-axis. For each car, the procedure resulted in a single model, comprising of the edge map of the central image, together with the approximated depth and curvature along the edges. It was found that a single model of this type yields accurate predictions to the appearance of the object within the entire $60°$ of rotation around the Y-axis.

Figure 5 shows four pictures, two of each car, rotated by $\pm 15°$. Such rotations already create large deformations of the images (Figure 6). Figure 7 shows the results of aligning the models with the images. An approximation to the transformation (rotation, translation and scale) can be supplied by different alignment routines, e.g., using three corresponding points [18]. It can be seen that, by using the alignment procedure, a single model gives accurate fits to the object seen from different viewing positions. Figure 8 shows the result of matching the two cars with the incorrect models. The discrepancy between the image and the aligned model is significantly higher than in Figure 7. A simple distance metric between the image contours and the aligned model was therefore sufficient to select the correct model. It is worth noting that accurate predictions were obtained despite the fact that (1) the objects have complex 3-D shapes, and (2) we have used crude approximations to the radii of curvature using three pictures.

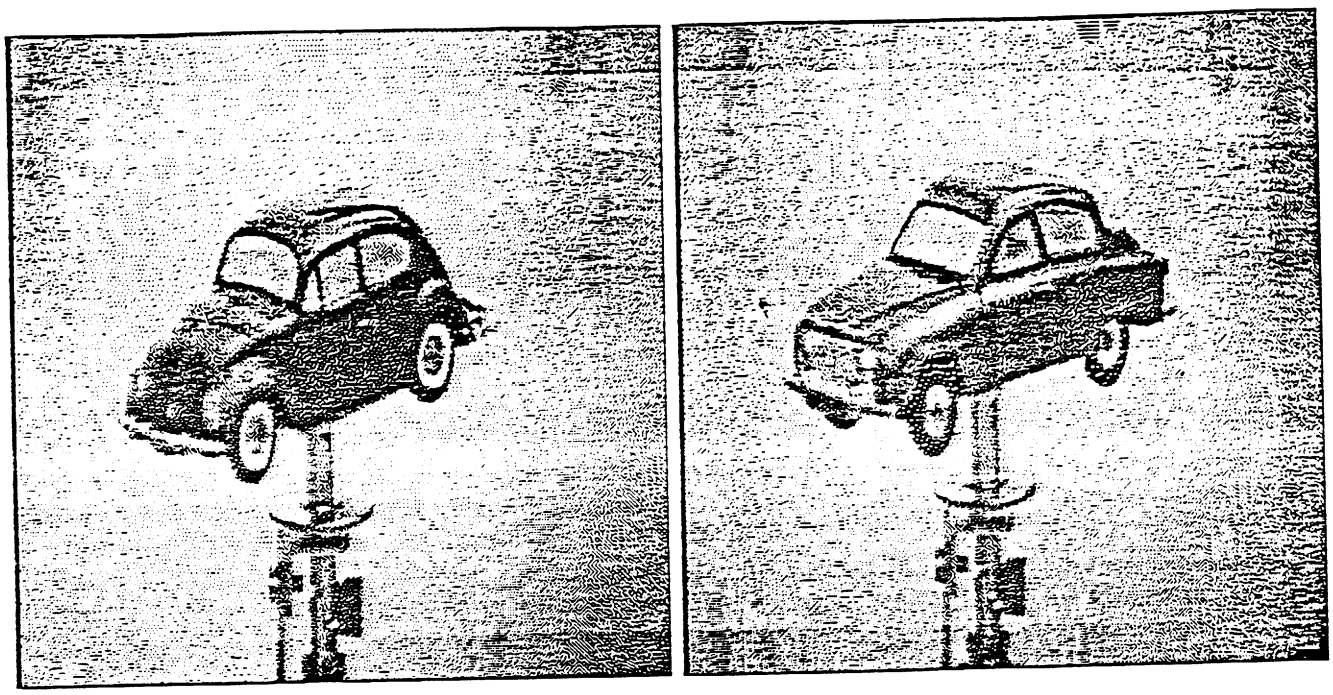
5 Summary

In this paper we have proposed a method for predicting the new appearance of an object with smooth surface following a similarity transformation (3-D rotation, translation and scaling). The method uses the 3-D surface curvature along the object contours. We have shown that a single parameter, the magnitude of the curvature vectors at these points, is sufficient to recover their curvature radii for a rotation around any given axis. Three pictures are in principle sufficient for approximating the radii of curvature for most contour points, and five can be used to estimate the components r_x, r_y independently.

The implemented scheme was found to give accurate results for large transformations. In the scheme we have proposed, each object is represented by a number of models, each covering a range of potential viewpoints. The results suggest that only a small number of such models are required to predict the new appearance of an object from any viewpoint.

Appendix

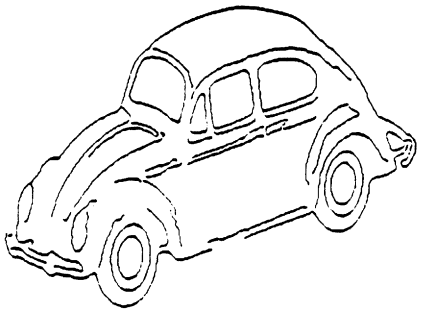
Consider a surface defined by the implicit function $F(x, y, z) = 0$. Assuming an orthographic projection, where Z is the visual axis, the rim is defined by the set of points on the surface where $F_z(x, y, z) = 0$. Let $\mathbf{p_0} = (x_0, y_0, z_0)$ be a rim point, that is $F(\mathbf{p_0}) = F_z(\mathbf{p_0}) = 0$. We consider the cases where F has second derivatives at $\mathbf{p_0}$, the surface normal there is not zero (that is either $F_x(\mathbf{p_0}) \neq 0$ or $F_y(\mathbf{p_0}) \neq 0$). In case one of them is zero we redefine the axes such that both $F_x(\mathbf{p_0})$ and $F_y(\mathbf{p_0})$ will be nonzero.



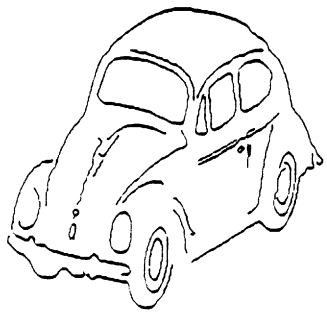
(a) (b)

(c) (d)

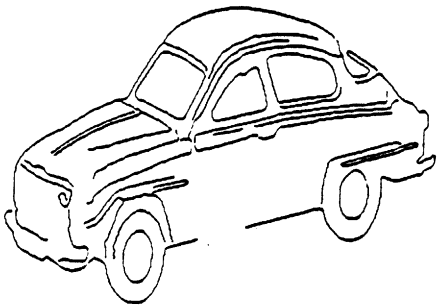
Figure 4: The model objects. (a) A picture of the model VW car. (b) A picture of the model Saab car. (c) An edge map of the VW car. (d) An edge map of the Saab car.



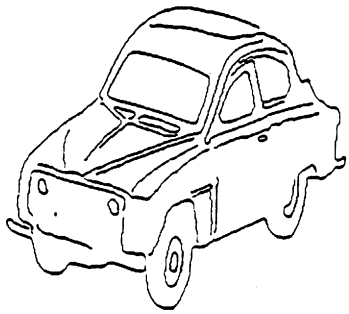
(a)



(b)



(c)



(d)

Figure 5: The iamge objects. (a) A VW car rotated by $-15°$ with respect to the model. (b) A VW car rotated by $+15°$ with respect to the model. (c) A Saab car rotated by $-15°$ with respect to the model. (d) A Saab car rotated by $+15°$ with respect to the model.

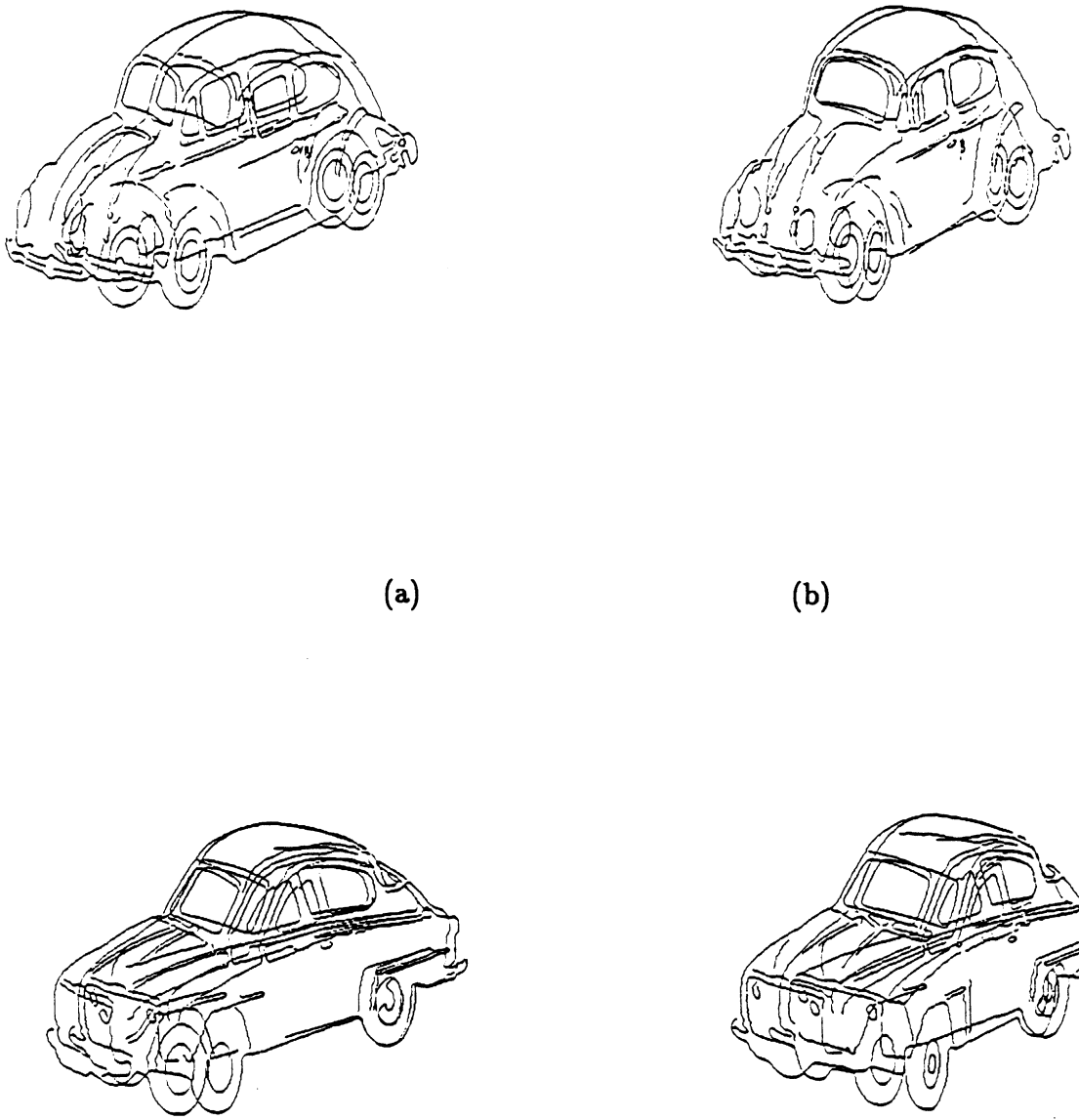


Figure 6: Deformation of the images with respect to the models. (a) A deformation of the VW car following a rotation of $-15°$. (b) A deformation of the VW car following a rotation of $+15°$. (c) A deformation of the Saab car following a rotation of $-15°$. (d) A deformation of the Saab car following a rotation of $+15°$.

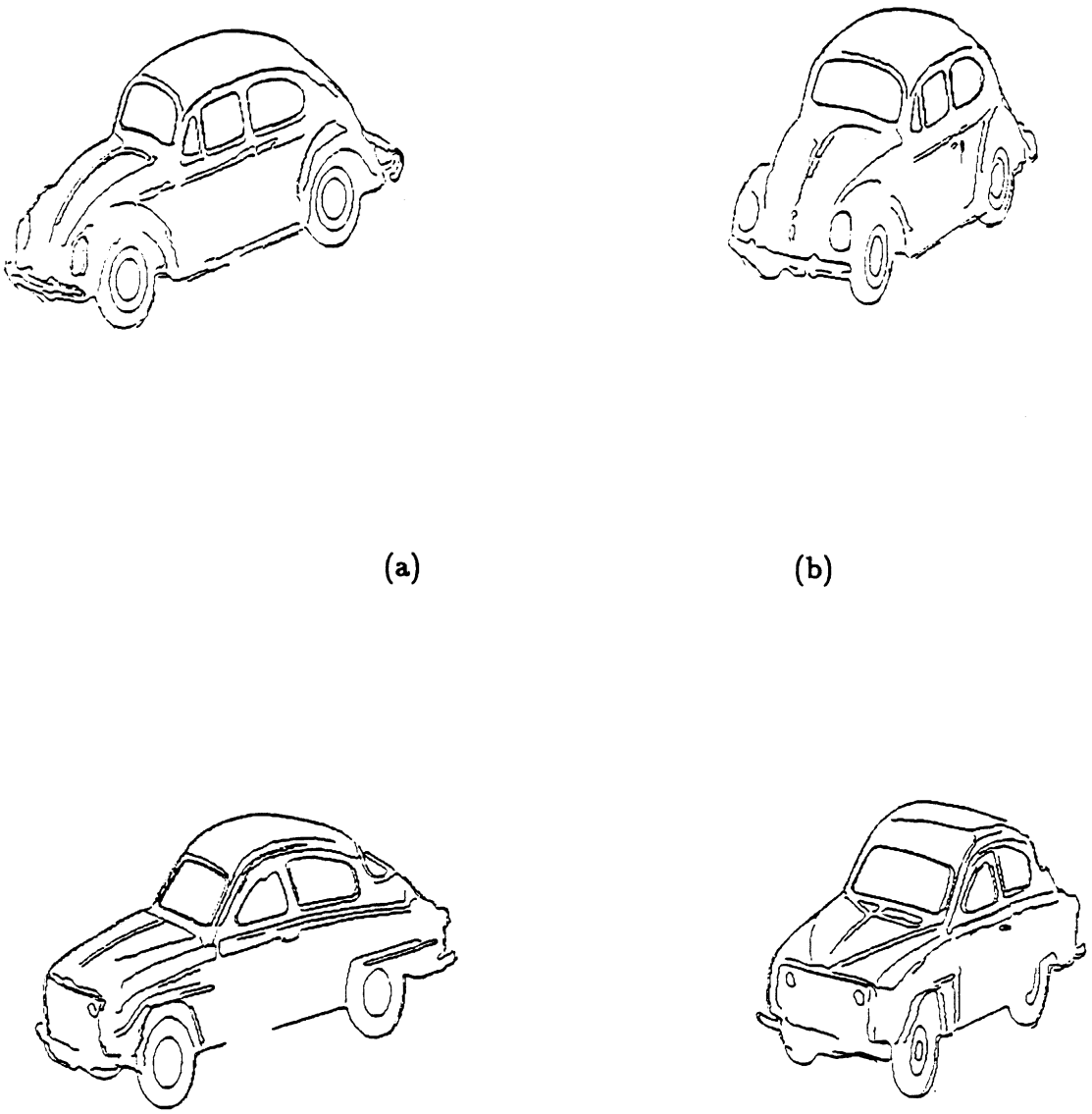
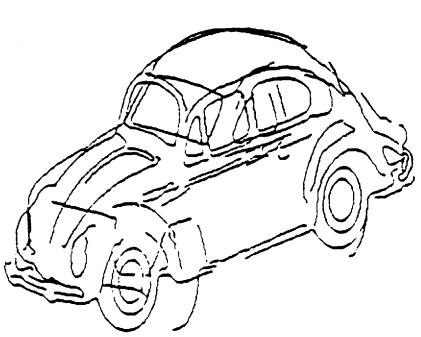
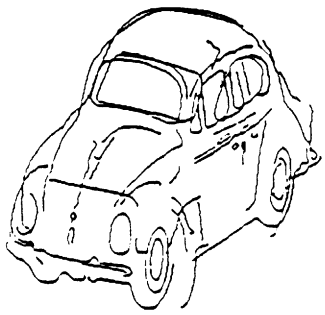


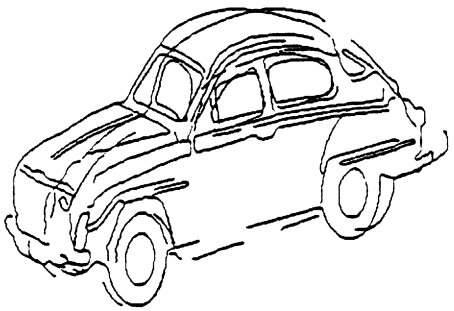
Figure 7: Correct alignment of the models with the images. (a) Alignment of the VW model with the first VW image. (b) Alignment of the same VW model with the second VW image. (c) Alignment of the Saab model with the first Saab image. (d) Alignment of the same Saab model with the second Saab image.



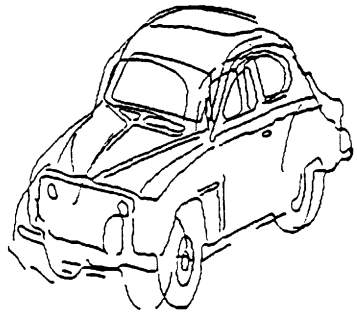
(a)



(b)



(c)



(d)

Figure 8: Matching the images with incorrect models. (a) Matching the first VW image to the Saab model. (b) Matching the second VW image to the same Saab model. (c) Matching the first Saab image to the VW model. (d) Matching the second Saab image to the same VW model.

Moreover we assume that $F_{zz}(\mathbf{p_0}) \neq 0$. By this we ignore points where a continuous line yields a single point in the silhouette, and points with infinite radius of curvature. These points may change their place unexpectedly during rotation.

Lemma 1: Let $F(x,y,z) = 0$ be a surface description, and let $\mathbf{p_0} = (x_0, y_0, z_0)$ be a rim point, i.e. $F(\mathbf{p_0}) = F_z(\mathbf{p_0}) = 0$. The curvature radii of $\mathbf{p_0}$ with respect to the Y and X axes are given by:

$$r_x = -\frac{F_x}{F_{zz}}$$

$$r_y = -\frac{F_y}{F_{zz}}$$

Proof: Consider the space curve defined by the implicit function $F(x, y_0, z) = 0$. According to the implicit function theorem, since $F_x(\mathbf{p_0}) \neq 0$ and $F_{zz}(\mathbf{p_0}) \neq 0$, $x(z)$ is a well defined function in a neighborhood of $\mathbf{p_0}$, and:

$$\frac{d}{dz}F_z = F_{zz} + F_{zx}(\frac{dx}{dz}) = \frac{1}{F_x}(F_{zz}F_x - F_{zx}F_z)$$

$$\frac{d}{dz}F_x = F_{xz} + F_{xx}(\frac{dx}{dz}) = \frac{1}{F_x}(F_{xz}F_x - F_{xx}F_z)$$

And since:

$$\frac{d^2x}{dz^2} = \frac{d}{dz}(\frac{dx}{dz}) = \frac{d}{dz}(-\frac{F_z}{F_x}) = \frac{-(\frac{d}{dz}F_z)F_x + (\frac{d}{dz}F_x)F_z}{F_x^2}$$

We obtain:

$$\frac{d^2x}{dz^2} = \frac{-F_{zz}F_x^2 + F_{zz}F_xF_z + F_{zz}F_xF_z - F_{xx}F_z^2}{F_x^3}$$

$F_z(p_0) = 0$, therefore:

$$\frac{dx}{dz}(z_0) = 0$$

$$\frac{d^2x}{dz^2}(z_0) = -\frac{F_{zz}}{F_x}$$

For a curve $x(z)$, the radius of curvature at z_0 is given by:

$$r(t) = \frac{1}{k(t)} = \frac{(1 + \frac{dx}{dz}(z_0))^{3/2}}{\frac{d^2x}{dz^2}(z_0)}$$

Substituting the appropriate terms we obtain:

$$r_x = -\frac{F_x}{F_{zz}}$$

And in a similar way -

$$r_y = -\frac{F_y}{F_{zz}}$$

Proposition 1: Let $F(x,y,z) = 0$ be a surface description, and let $\mathbf{p_0}$ be a rim point, i.e. $F(\mathbf{p_0}) = F_z(\mathbf{p_0}) = 0$. Let V_α be an axis lying in the image plane and forming an angle α with the positive X-axis. The radius of curvature at $\mathbf{p_0}$ with respect to V_α is given by:

$$r_\alpha = r_y \cos\alpha - r_x \sin\alpha$$

Proof: Let $G(x',y',z) = 0$ be the surface $F(x,y,z) = 0$ rotated around Z-axis by the angle $-\alpha$, i.e.

$$G(x',y',z) = F(x'\cos\alpha - y'\sin\alpha, x'\sin\alpha + y'\cos\alpha, z)$$

After such a rotation V_α coincides with X-axis, therefore:

$$r^F_\alpha = r^G_{y'}$$

Where r^F, r^G are radii of curvature for the surfaces F, G respectively. According to lemma 1:

$$r^G_{y'} = -\frac{G_{y'}}{G_{zz}}$$

Since:

$$G_{y'} = -F_x \sin\alpha + F_y \cos\alpha$$

$$G_z = F_z = 0$$

$$G_{zz} = F_{zz}$$

We obtain:

$$r_\alpha = \frac{-F_y \cos\alpha + F_x \sin\alpha}{F_{zz}} = r_y \cos\alpha - r_x \sin\alpha$$

Proposition 2: Let $F(x,y,z) = 0$ be a surface description, and let $\mathbf{p_0}$ be a rim point, i.e. $F(\mathbf{p_0}) = F_z(\mathbf{p_0}) = 0$. Let $\mathbf{r} = (r_x, r_y)$ be the curvature vector at $\mathbf{p_0}$, and let $\mathbf{t}$ be the tangent vector to the silhouette at $\mathbf{p_0}$. Then $\mathbf{r} \cdot \mathbf{t} = 0$ that is, $\mathbf{r} \perp \mathbf{t}$.

Proof: The point $\mathbf{p_0}$ satisfies the two constraints $F(\mathbf{p_0}) = 0$ and $F_z(\mathbf{p_0}) = 0$. According to the implicit function theorem, since $F_y(\mathbf{p_0}) \neq 0$, $F_{zz}(\mathbf{p_0}) \neq 0$, $y(x)$, $z(x)$ are well defined functions in a neighborhood of $\mathbf{p_0}$. The tangent vector $\mathbf{t}$ to $y(x)$ is in the

direction $(1, \frac{dy}{dx})$ in the XY plane, and since $\frac{dy}{dx} = -\frac{F_x}{F_y}$, $\mathbf{t}$ is the direction $(-F_y, F_x)$. According to Lemma 1, the vector of curvature radii is given by:

$$\mathbf{r} = (r_x, r_y) = (-\frac{F_x}{F_{zz}}, -\frac{F_y}{F_{zz}})$$

Therefore:

$$\mathbf{r} \cdot \mathbf{t} = \frac{F_x F_y}{F_{zz}} - \frac{F_x F_y}{F_{zz}} = 0$$

References

[1] Baker, H., 1977. Three-dimensional modeling. *Proc. 5th Int. Joint Conf. on Artificial Intelligence (Cambridge, Mass., Aug 22-25)*, IJCAI, pp 649-655.

[2] Brady, M., Ponce., J., Yuille, A. & Asada, H., 1985. Describing surfaces. *A.I. Memo 882, The Artificial Intelligence, 17*, 285-349.

[3] Brown, C.M., 1981. Some mathematical and representational aspects of solid modeling. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, Vol. PAMI-3, No. 4, July.

[4] Canny, J., 1983. Finding edges and lines in images. *MIT Artificial Intelligence Laboratory Report,* AI-TR-720.

[5] Chien, C.H. & Aggarwal, J.K., 1987. Shape recognition from single silhouette. *Proc. of ICCV Conf. (London)* 481-490.

[6] Dane, C. & Bajcsy, R., 1982. An object-centered three-dimensional model builder. *Proc. 6th Int. Conf. Pat. Recog., Munich, West Germany, Oct 19-22*, 348-350.

[7] Faugeras, O.D. & Hebert, M., 1986. The representation, recognition and location of 3-D objects. *Int. J. Robotics Research* 5(3), 27-52.

[8] Fishler, M.A. & Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography. *Communications of the ACM* 24(6), 381-395.

[9] Huttenlocher, D.P. & Ullman, S., 1987. Object recognition using alignment. *Proc. of ICCV Conf. (London)* 102-111.

[10] Koenderink, J.J., & Van Doorn, A.J., 1979. The internal representation of solid shape with respect to vision. *Biol. Cybernetics 32*, 211-216.

[11] Koenderink, J.J., & Van Doorn, A.J., 1981. The shape of smooth objects and the way contours end. *Perception 11*, 129-137.

CS-TR Scanning Project
Document Control Form

Date : 4/27/95

Report # AIM-1060

Each of the following should be identified by a checkmark:
Originating Department:

☒ Artificial Intellegence Laboratory (AI)
☐ Laboratory for Computer Science (LCS)

Document Type:

☐ Technical Report (TR) ☒ Technical Memo (TM)
☐ Other:_____

Document Information

Number of pages: 20 (26-IMAGES)
Not to include DOD forms, printer intstructions, etc... original pages only.

Originals are:
☒ Single-sided or
☐ Double-sided

Intended to be printed as :
☐ Single-sided or
☒ Double-sided

Print type:
☐ Typewriter ☐ Offset Press ☒ Laser Print
☐ InkJet Printer ☐ Unknown ☐ Other:_____

Check each if included with document:

☒ DOD Form (2) ☐ Funding Agent Form ☐ Cover Page
☐ Spine ☐ Printers Notes ☐ Photo negatives
☐ Other: _____

Page Data:

Blank Pages (by page number):_____

Photographs/Tonal Material (by page number):_____

Other (note description/page number):

Description :	Page Number:
IMAGE MAP (1) UN#'ED TITLE PAGE	
(2-20) PAGES #'ED 1-19	
(21-23) SCONCONTROL, DOD (2)	
(24-26) TRGTS (3)	

Scanning Agent Signoff:

Date Received: 4/27/95 Date Scanned: 5/2/95 Date Returned: 5/4/95

Scanning Agent Signature: Michael W. Cook

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AIM 1060	2. GOVT ACCESSION NO.: AD-A214892
3. RECIPIENT'S CATALOG NUMBER:	
4. TITLE (and Subtitle): The Alignment of Objects with Smooth Sufaces	5. TYPE OF REPORT & PERIOD COVERED: Memorandum
	6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Shimon Ullman and Ronen Basri	8. CONTRACT OR GRANT NUMBER(s): N00014-85-K-0124
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Artificial Intelligence Laboratory, 545 Technology Square, Cambridge, MA 02139	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:
11. CONTROLLING OFFICE NAME AND ADDRESS: Advanced Research Projects Agency, 1400 Wilson Blvd., Arlington, VA 22209	12. REPORT DATE: July 1988 13. NUMBER OF PAGES: 20
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office): Office of Naval Research, Information Systems, Arlington, VA 22217	15. SECURITY CLASS. (of this report): UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Distribution is unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

object recognition
recognition
visual recognition
alignment

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Abstract: This paper examines the recognition of rigid objects bounded by smooth surfaces, using an alignment approach. The projected image of such an object changes during rotation in a manner that is generally difficult to predict. An approach to this problem is suggested, using the 3-D surface curvature at the points along the silhouette. The curvature information requires a single number for each point along the object's silhouette, the magnitude of the curvature vector at the point. We have

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block 20 cont.

implemented this method, and tested it on images of complex 3-D objects. Models of the viewed objects were acquired using three images of each object. The implemented scheme was found to give accurate predictions of the objects' appearances for large transformations. Using this method, a small number of (viewer-centered) models can be used to predict the new appearance of an object from any given viewpoint.

Scanning Agent Identification Target

Scanning of this document was supported in part by the **Corporation for National Research Initiatives,** using funds from the **Advanced Research Projects Agency** of the **United states Government** under Grant: **MDA972-92-J1029.**

The scanning agent for this project was the **Document Services** department of the **M.I.T Libraries.** Technical support for this project was also provided by the **M.I.T. Laboratory for Computer Sciences.**

Scanned

Date: 5/2/1995

M.I.T. Libraries
Document Services