
Curve and Surface Duals and the Recognition of Curved 3D Objects from their Silhouette

Amit Sethi (asethi@uiuc.edu)
David Renaudie* (david.renaudie@ensimag.imag.fr)
David Kriegman (kriegman@uiuc.edu)
Jean Ponce (ponce@cs.uiuc.edu)
Department of Computer Science and Beckman Institute
University Of Illinois
Urbana, IL 61801, USA

Abstract. This article addresses the problem of recognizing a solid bounded by a smooth surface in a single image. The proposed approach is based on a new representation for two- and three-dimensional shapes, called their signature, that exploits the close relationship between the dual of a surface and the dual of its silhouette in weak-perspective images. Objects are modeled by rotating them in front of a camera without any knowledge of their motion. The signatures of their silhouettes are concatenated into a single object signature. At recognition time, the signature of the contours extracted from a test photograph is matched to all modeled objects’ signatures. This approach has been implemented, and recognition examples are presented.

Keywords: Three-dimensional object recognition, invariants, duals, pedal curves.

1. Introduction

Most approaches to model-based object recognition are based on establishing correspondences between viewpoint-independent image features and geometric features of object models (Huttenlocher and Ullman, 1987; Lowe, 1987). For objects with smooth surfaces, few surface markings and little texture, the most reliable image feature is the object’s silhouette, i.e., the projection into the image of the curve, called the occluding contour, where the cone formed by the optical rays grazes the surface. The dependence of the occluding contour on viewpoint makes the construction of appropriate feature correspondences difficult. Appearance-based methods do not rely on such correspondences, and they are suitable for recognizing objects bounded by smooth surfaces, but they generally require a dense sampling of the pose/illumination space to be effective (Murase and Nayar, 1995). Methods for relating image features to 3D geometric models of curved surfaces have been developed for surfaces of revolution (Kriegman and Ponce, 1990b; Glachet et al., 1991), generalized cylinders (Ponce and Chelberg, 1987; Richetin et al., 1991; Liu et al., 1993; Zeroug and Medioni, 1995), algebraic surfaces (Kriegman and Ponce, 1990a; Ponce et al., 1992), and triangular

* D. Renaudie is with the Ecole Nationale Supérieure d’Informatique et Mathématiques Appliquées de Grenoble, France. This work was done while he was visiting the Department of Computer Science and the Beckman Institute at the University of Illinois at Urbana-Champaign.
splines (Sullivan and Ponce, 1998). These approaches require separate processes for the construction of 3D models either from image data or using CAD tools, and for the extraction/segmentation of the image contours associated with each object.

An alternative is to replace a parametric description of the object surface by an empirical representation of contour features constructed from sampled image data. We proposed in (Joshi et al., 1997; Vijayakumar et al., 1998) two variants of this approach where contour bitangents and inflections are recorded in an image sequence and serve as the basis for object recognition: in (Joshi et al., 1997), the trajectory of the camera is assumed to be known, and it is used to explicitly reconstruct the surface curves giving rise to bitangents and inflections during modeling. In turn, these curves are used to predict the appearance of image features observed at recognition time. In (Vijayakumar et al., 1998) on the other hand, the contour tangents parallel to each bitangent and inflection serve as image features, and the distances between successive parallels form coordinates for the corresponding feature space used for classification. The features recorded during a modeling session trace a curve in the feature space that is independent of the camera trajectory. At recognition time, the bitangents, inflections, and the corresponding parallel tangents present in the test image are matched to the closest model curve in feature space. Here we propose to generalize this method by replacing a sparse set of silhouette features with a much denser set offering greater discriminatory power. Our work builds on geometric insights about the occluding contour and silhouettes of smooth surfaces (Koenderink, 1984; Giblin and Weiss, 1995) and their use in determining geometric structure from sequences of images (Arbogast and Mohr, 1991; Cipolla and Blake, 1992; Vaillant and Faugeras, 1992; Boyer and Berger, 1996; Cipolla et al., 1995). See (Cipolla and Giblin, 2000) for an overview of this line of research.

The basic processing steps for each image include detecting the silhouette curve $\Gamma$, computing its pedal curve (a representation of its dual) $\Gamma'$, and constructing its signature $\Sigma''$, a family of curves embedded in $\mathbb{R}^d$, where $d \geq 2$ depends on the geometric complexity of the observed object. The signature only depends on the projection direction and is unaffected by changes in the other viewing parameters. When a solid with a smooth surface $\Sigma$ is observed by a moving camera, the signatures of the successive silhouettes sweep a family $\Sigma'$ of two-dimensional surface patches in $\mathbb{R}^d$, also called the signature of $\Sigma$. This surface is independent of the viewing conditions, and can thus be constructed without any knowledge of the camera motion, each object in the database being represented by a different signature. At recognition time, the signature of the silhouette found in the test image is matched with the signatures of all modeled surfaces, and the closest model is recognized. The hierarchy of curve and surface representations used in this paper is illustrated below. Its components are introduced in the next sections.

\[
\begin{align*}
\text{Surface } \Sigma \text{ in } \mathbb{R}^3 & \rightarrow \text{ Pedal Surface } \Sigma' \text{ in } \mathbb{R}^3 & \rightarrow \text{ Signature } \Sigma'' \text{ in } \mathbb{R}^d \\
\downarrow \text{ projects onto} & \uparrow \text{ is a planar section of} & \uparrow \text{ lies on} \\
\text{Silhouette } \Gamma \text{ in } \mathbb{R}^2 & \rightarrow \text{ Pedal Curve } \Gamma' \text{ in } \mathbb{R}^2 & \rightarrow \text{ Signature } \Gamma'' \text{ in } \mathbb{R}^d
\end{align*}
\]

A preliminary version of this paper appeared in (Renaudie et al., 2000).
2. Duals and Pedal Curves and Surfaces

Let us consider a smooth closed curve $\Gamma$ and some point $O$ in the plane. We will call this point the origin of the plane, but refrain from introducing a full-fledged (and at this point fully unnecessary) coordinate system. We define the dual $\mathcal{D}$ of $\Gamma$ as the set of its tangent lines. The dual also forms a curve in the two-dimensional space formed by all lines in the plane. Since this space is a bit difficult to visualize (it is topologically equivalent to a cylinder), as is the dual embedded in it, we will use the pedal curve $\Gamma'$ associated with $\Gamma$ as a convenient device for depicting $\mathcal{D}$ in this paper. Figure 2 illustrates the construction of that curve: we associate with each point $P$ on $\Gamma$ the orthogonal projection $P'$ of $O$ onto the tangent line $T$ in $P$; the pedal curve is the curve $\Gamma'$ traced by $P'$ as $P$ varies along $\Gamma$. If $\vec{N}$ denotes the unit normal to $\Gamma$ in $P$, the corresponding point $P'$ can also be defined by $\overrightarrow{OP'} = (\overrightarrow{OP} \cdot \vec{N})\vec{N}$.

![Figure 1. Construction of the pedal curve.](image)

Like the dual, the pedal curve depends on the choice of origin. The embedding of the dual in the pedal curve is not injective: indeed, any tangent line passing through the origin maps onto the origin. Although there is no such tangent for the curve and origin shown in Figure 2, tangents passing through the origin are guaranteed to exist when this point lies outside the curve, and may exist even when this is not the case (see Figure 5 for an example). To simplify the discussion, we will pretend in most of this paper that the pedal curve does not pass through the origin, and identify the dual and the pedal curve. We will come back to the general case during the presentation of our implementation.

The pedal curve $\Gamma'$ associated with $\Gamma$ enjoys the following properties:

(A) The inflections of $\Gamma$ (the points $B$ and $C$ in Figure 2) map onto cusps of $\Gamma'$ ($B'$ and $C'$ in this case);

(B) The bitangents lines to $\Gamma$ (like the line $L$ that passes through the points $A$ and $D$ in Figure 2) map onto double points of $\Gamma'$ (the point $A' = D'$ in this case);

(C) The points of $\Gamma$ whose tangents are parallel to each other map onto the intersections of the pedal curve with a line through the origin whose direction is orthogonal.
to the common tangent direction (consider for example the two points $G$ and $H$ and their images $G'$ and $H'$ in Figure 2). Conversely, the intersection points of $\Gamma'$ with a line passing through the origin are the images of points with parallel tangents on $\Gamma$.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{The intersections between various lines passing through the origin and the pedal curve are the images of point sets with parallel tangents. The number of intersections is normally even (e.g., two intersections for the line joining $G'$ and $H'$) but is odd at cusps and double points, corresponding to inflections and bitangents of the original curve. Lines intersecting $\Gamma'$ in four points are not shown here to avoid clutter.}
\end{figure}

Property (C) will in fact be the basis for the approach to object modeling and recognition presented in this paper. In general, a line passing through the origin will intersect the pedal curve in an even number of points: for example, the line passing through $G'$ and $H'$ in Figure 2 only intersects $\Gamma'$ at these two points; lines passing through the counterclockwise angular sector defined by $\overline{OA'}$ and $\overline{OB'}$, on the other hand, will intersect $\Gamma'$ four times. For a rotating line passing through $O$, the number of intersection points can only change at cusps and double points, where the number of intersections is (exceptionally) odd: for example the line passing through $O$ and $B'$ also intersects $\Gamma'$ in $I'$ and $F'$, and the line passing through $O$ and $A' = D'$ also intersects $\Gamma'$ in $E'$ and $J'$.

As noted earlier, the pedal curve associated with a planar curve depends on the choice of origin. However, Properties (A) to (C) are independent of this choice, and they will be used in Section 4.1 to map the pedal curve onto another curve which is invariant under rigid transformations of the plane.

The definition of dual and pedal curves generalizes naturally to three dimensions: consider a smooth surface $\Sigma$ in $\mathbb{R}^3$ and pick some origin $O$; the dual of $\Sigma$ is defined as the set of its tangent planes, and it forms a two-dimensional surface in the three-dimensional space of all planes. To represent the dual in a convenient manner, we associate with every point $P$ on $\Sigma$ the orthogonal projection $P'$ of the origin $O$ onto its tangent plane. The surface swept by $P'$ as $P$ varies over $\Sigma$ is the pedal surface $\Sigma'$ associated with this surface. As in the two-dimensional case, this embedding of the dual, like the dual itself, depends on the origin. And like pedal curves, the pedal surface

\footnote{In principle, it may also change at points where the rotating line is tangent to the pedal curve, but these points are easily shown to correspond to cusps of the second kind of the curve $\Gamma$, which do not exist in the case of silhouettes of generic smooth surfaces (Koenderink, 1984).}
may contain singularities, including swallowtails and cuspidal edges corresponding to parabolic lines of $\Sigma$ (Bruce and Giblin, 1992).

3. Occluding Contours and their Projection

The brightness discontinuities in the image of an untextured solid bounded by a smooth surface form a curve, called the image contour, silhouette or outline. Under perspective projection, this curve is the intersection of the image plane with a viewing cone whose apex coincides with the center of projection and whose generators graze the object along a second (generically nonplanar) curve, called the occluding contour or rim, and the tangent plane at an occluding contour point projects onto the tangent line at the corresponding silhouette point (Figure 3.a). Under orthographic projection, the center of projection moves to infinity, the viewing cone becomes a cylinder whose generators are parallel to the (fixed) viewing direction, and the normal to the image contour is the same as the surface normal at the corresponding occluding contour point (Figure 3.b).

![Figure 3. Occluding boundaries under (a) perspective and (b) orthographic projection.](image)

Orthographic projection can be thought of as an approximation of perspective when the scene is far from the camera observing it. When the scene lies at a finite distance from its observer but its depth range is small compared to this distance, we obtain the weak-perspective (or scaled-orthography) approximation to perspective projection. In this model, all scene points are assumed to lie in the same fronto-parallel plane. The projection geometry is the same as in the orthographic case, and the silhouette and occluding contour enjoy the same properties, but the distance between any pair of image points is a constant multiple (or magnification) of the distance obtained under orthographic projection. We will assume either orthographic or weak-perspective projection in the rest of this paper.
Most real objects are opaque of course, but we will assume in most of the rest of this paper that all objects are translucent. This is just to simplify the upcoming discussion, since in this case a necessary and sufficient condition for a point to project onto the silhouette is that the viewing direction belongs to its tangent plane. The approach proposed in this paper is not limited in any sense to transparent objects, and we will come back to the case of opaque objects when we discuss our implementation.

4. The Signatures of Curves and Surfaces

We begin by clarifying the relation between pedal surfaces and the pedal curves of image silhouettes. Since these curves and surfaces depend on a choice of origin, we then introduce a novel representation of these shapes, called their signatures, that can be used to relate the geometry of a surface and its projections but is independent of any choice of origin.

Let us consider a regular surface Σ, an origin O, and the corresponding pedal surface Σ'. Given some viewing direction ť, we denote by Π the plane perpendicular to ť passing through O and define Γ as the silhouette of Σ. Let o denote the image of O, and Γ' denote the pedal curve of Γ defined using o as the origin of the plane Π. We have the following result.

**Lemma 1.** Under orthographic (resp. weak-perspective) projection, the pedal curve Γ' can be mapped onto the intersection of the pedal surface Σ' with the plane Π via a rigid transformation (resp. a rigid transformation followed by a scaling).

We only give an informal sketch of the proof. Let us first note that the image of the occluding contour under the mapping that associates with every point on Σ the corresponding pedal point on Σ' is the curve σ' formed by the intersection of Σ' and Π. As mentioned earlier, the normal at a point on the image contour is the same as the surface normal at the corresponding occluding contour point under orthographic projection. It is easy to see that the distance between the point O and the tangent plane at an occluding contour point is the same as the distance between the image of O and the tangent line to the corresponding silhouette point. It follows that Γ' maps onto σ' via the translation mapping o onto O, and the lemma follows for orthographic projection. The weak-perspective case is similar but involves the scaling inherent in this projection model.

Informally, Lemma 1 can be restated as saying that the dual of the image contour can be identified with a planar slice of the dual of the surface. This leads to the following critical observation: consider a moving camera following some trajectory and assume that the viewing direction ť is known at all times; according to the lemma, the pedal surface can be recovered from the successive pedal curves as long as the projection of the origin O can be determined in the corresponding images. If the set of viewing directions covers (say) half a great circle of the unit sphere, every point on the surface will lie on the occluding contour for some viewing direction (barring self occlusion), and the entire pedal surface will be revealed. The object may then be recognized from any view (including ones never seen before) by matching the corresponding pedal curve to a planar section of the pedal surface. Note that, for
some camera motions, parts of the pedal surface may be missed (or in fact be covered twice), and the object will only be recognizable from a subset of all possible views in this case.

4.1. The Signatures of Curves and Surfaces and Their Properties

The approach to object recognition sketched in the previous section requires that the projection of the origin be identified during both modeling and recognition, and that the viewing direction be known during modeling. We do not know of any geometric property of arbitrary smooth surfaces that would allow this. Instead, we now define a new representation for curves and surfaces that is independent of the choice of the origin (and in fact of rigid image transformations) and whose construction does not require knowing the camera motion during modeling.

Let us start by defining the partition of the pedal curve $\Gamma'$ (Figure 4.1): we consider an oriented line $\Delta$ with orientation $\theta$ passing through the origin associated with $\Gamma'$. As noted earlier, the number of intersections of $\Delta$ and $\Gamma'$ only changes at cusps and double points of this curve. Let $\theta_i (i = 1, \ldots, p)$ denote the corresponding orientations of $\Delta$. The partition of $\Gamma'$ is defined as the set of pairs $(\theta_i, n_i) (i = 1, \ldots, p)$, where the number of intersections of $\Gamma'$ and $\Delta$ is equal to $n_i$ for any $\theta$ in $(\theta_i, \theta_{i+1})$ (here index addition is performed modulo $p$ so $p + 1 \equiv 1$).

![Figure 4](image.png)

*Figure 4.* The partition of a pedal curve: a rotating line $\Delta$ passing through $O$ intersects the pedal curve $\Gamma'$ in two points when $\theta$ is in the (open) range ($\theta_5, \theta_4$) or ($\theta_6, \theta_1$), and it intersects $\Gamma'$ in four points when $\theta$ is in one of the (open) intervals $(\theta_1, \theta_2), (\theta_2, \theta_3), (\theta_4, \theta_5)$ and $(\theta_5, \theta_6)$.

We now define the signature $\Gamma''$ of the curve $\Gamma$. We consider again an oriented line $\Delta$ passing through the origin with orientation $\theta$ in the $(\theta_i, \theta_{i+1})$ range, and denote by $AB$ the signed distance between two points $A$ and $B$ on $\Delta$. Let us denote by $P_k$ ($k = 1, \ldots, n_i$) the intersections of $\Delta$ with $\Gamma'$, sorted in increasing $\overrightarrow{OP_k}$ order, and define $d_k = P_kP_{k+1} (k = 1, \ldots, n_i - 1)$. We define $\Gamma''$ as the curve traced in $\mathbb{R}^{n_i - 1}$ by the points $(d_1, \ldots, d_{n_i-1})^T$ as $\theta$ varies over $(\theta_i, \theta_{i+1})$ and define the signature $\Gamma''$ of the curve $\Gamma$ as the tuple $(\Gamma''_1, \ldots, \Gamma''_p)$.

Note that the scalars $d_k$ associated with $\Delta$ are simply the (signed) distances between the tangent lines of $\Gamma$ that are parallel to each other and are perpendicular to $\Delta$. In particular, we obtain immediately the following result:
LEMMA 2. The signature $\Gamma''$ of a curve $\Gamma$ is invariant under rigid transformations of $\Gamma$ and the choice of the origin $O$ used to define its pedal curve $\Gamma'$. 

Lemma 2 states the fundamental property of signatures, and it is the key to their usefulness in recognition tasks: unlike pedal curves, that completely capture the shape of the dual of a curve and therefore the shape of the curve itself, but unfortunately depend on the choice of the origin, signatures omit some of the shape information (namely the orientation of the parallel tangents and their position relative to a fixed point), but gain a complete independence on any choice of origin.

It is also possible to define the signature $\Sigma''$ of a surface as a tuple $(\Gamma_j^\prime, \ldots, \Gamma_q^\prime)$ of two-dimensional surface patches $\Sigma_j''$ embedded in $\mathbb{R}^{n_j-1}$ for $j = 1, \ldots, q$, each patch being swept by the signed distances between parallel tangent planes as its surface normal varies. Like the signature of a pedal curve, the signature of a pedal surface is independent of the choice of the origin $O$ and is invariant under rigid transformations of the original surface. Another fundamental result follows immediately from Lemmas 1 and 2, namely:

LEMMA 3. As a surface $\Sigma$ is observed by a moving camera, the surfaces patches swept by the signatures of its silhouettes form a subset of the signature $\Sigma''$ of $\Sigma$.

In particular, the signature of a surface can be constructed from a sequence of images without any knowledge of the camera motion. This will be the basis for the approach to object recognition presented in the next section.

Note that the family of patches swept by the silhouette signatures is actually equal to $\Sigma''$ for sufficiently rich camera trajectories, e.g., when the viewing directions along the camera path cover a half great circle of the unit sphere. A better understanding of the situation can be gained by considering the close relation between the Gaussian image of a surface and its dual. In particular, for a given viewing direction $\hat{v}$, the Gaussian image of the occluding contour of a surface is the great circle formed by the intersection of the surface's Gaussian image with the plane orthogonal to $\hat{v}$ and passing through the origin. For a moving camera, the great circles associated with the successive viewing directions cover a subset of the Gauss sphere. If the camera trajectory is sufficiently rich to guarantee full coverage of the entire Gauss sphere, every point on the surface will have been observed (up to occlusion) for some viewing direction, and the successive silhouette signatures completely sweep out the entire signature of the surface.

It should also be noted that the dimension of the space in which a patch $\Sigma_j''$ of the signature is embedded depends on the number of intersections of a line $\Delta$ with the pedal surface $\Sigma'$. If the number of intersections is 2, then $\Sigma_j''$ is embedded in $\mathbb{R}^1$ (i.e., it is simply an interval of $\mathbb{R}$) and is unlikely to offer much discriminatory power for recognition. In practice, we only retain those components of the signature surface for which $\Delta$ intersects $\Sigma'$ at least four times so the patch is embedded in $\mathbb{R}^n$ where $n \geq 3$. Also note that our discussion has assumed that the silhouette is a regular curve. Generically, the image contour of a smooth surface may in fact be singular and contain cusps and crossings (Koenderink and Van Doorn, 1976). For a moving camera, the trajectory of the viewing direction may cross a visual event boundary for which other singularities are observed (i.e., tangent crossings, triple points, cusp
crossings, swallowtails, lips, and beaks) (Kergosien, 1981; Koenderink and Van Doorn, 1976). These have been studied extensively, particularly within the context of aspect graph construction. Since these singularities are not observed in our implementation, we leave a fuller characterization of their corresponding pedal curves and signatures for future research.

Under weak perspective, the image magnification is an unknown additional parameter that may vary with each image (i.e., it may change over the camera trajectory used to model an object when the distance from the camera to the object varies). We can eliminate the dependency of the signature on magnification by normalizing the distances $d_k$ by the largest one (which is by construction $d_{n-1}$). This yields the *quotient signature* $\hat{\Gamma}' = (\hat{\Gamma}_1', \ldots, \hat{\Gamma}_p')$, where $\hat{\Gamma}_i'$ is the curve formed in $\mathbb{R}^{n-2}$ by the tuples $(\hat{d}_1 = d_1/d_{n-1}, \ldots, \hat{d}_{n-2} = d_{n-2}/d_{n-1})$ when $(d_1, \ldots, d_{n-1})$ varies over the corresponding component $\Gamma_i'$ of the signature $\Gamma'$. The quotient signature is invariant under affine transformations of the curve $\Gamma$. Over a sequence of images, the quotient signatures of successive silhouettes sweeps out the *quotient signature* of the corresponding surface $\Sigma$.

5. **Object Modeling and Recognition: An Implementation and Results**

As suggested in the previous section, the signatures of surfaces and their silhouettes can be used as the basis for object modeling from image sequences and object recognition in a single image. We discuss below an implemented approach to these two problems. The experimental results presented in this section are not intended as a definitive characterization of the capabilities and limitations of signatures as a representation for recognition: they merely demonstrate that signatures are indeed easy to compute from real images and can support the recognition of objects with complex shapes. Our results also demonstrate that curved 3D objects can be modeled from 2D images with unknown camera motion and recognized from novel viewpoints.

5.1. **Object Modeling**

The images used in our modeling experiments have good contrast with light objects and a dark background (or vice versa), allowing us to use thresholding to detect the silhouette $\Gamma$ of an object as a linked list of discrete edge points (note that we miss most internal edges by using thresholding instead of a full-fledged edge detector). The normal vector at each point of $\Gamma$ is then computed using linear least squares, and the pedal curve $\Gamma'$ is finally computed in a straightforward manner (Figure 5.a,b). The origin in the pedal curve computation is (arbitrarily) taken to be the center of mass of the edge points.

As shown by Figure 5.b, the noise in the detected edge point position is amplified by the pedal curve construction process, and it is thus necessary (in general) to smooth the silhouette $\Gamma$ before constructing $\Gamma'$. We use recursive Gaussian smoothing as suggested in (Mackworth and Mokhtarian, 1988) in our implementation, and results are shown in Figure 5.c,d.

The signature $\Gamma''$ associated with the curve $\Gamma$ is also computed in a straightforward manner: the range of orientations between 0 and $2\pi$ is sampled uniformly, and the
lines passing through the origin at the corresponding orientations are intersected with the pedal curve. The intersections are sorted along each line, and the corresponding distances $d_i$ are then computed. The quotient signature curve is computed if necessary by dividing all distances by the last one. Figure 6 shows an example signature projected to $\mathbb{R}^3$ for the silhouette in Figure 5, and the corresponding quotient signature.

Signature and quotient surfaces are constructed by stitching together the signature and quotient curves found in successive images. Figure 7 shows ten images of an 87-image sequence (4° sampling) taken as a telephone handset rotates about a fixed axis, and Figure 8 shows the signature $\Sigma^q$ of its surface.

5.2. Object Recognition

We have constructed a simple recognition system. Figure 9 shows images of six objects, modeled using the technique proposed in the previous section: a camel, a toy stuffed animal, a dolphin, a telephone, a pig, and a duck. Recall that objects are modeled by rotating them about a fixed axis over 180°. Five test images of each object were also acquired from novel viewing directions, and Figure 10 shows some examples.

The principle of the recognition method is straightforward. Each modeled object is represented by its signature. The signatures of the silhouettes found in a test image are
computed and matched to the stored surface signatures. In practice, however, some care must be given to the construction of indexing schemes adapted to the signature representation: recall that a silhouette signature $\Gamma''$ is actually a collection of curves $\Gamma''_i$ embedded in $\mathbb{R}^{n_i-1}$ ($i = 1, \ldots, p$), while a surface signature is a collection of patches $\Sigma''_j$ embedded in $\mathbb{R}^{n_j-1}$ ($j = 1, \ldots, q$). In particular, the computation of the distance from a point on $\Gamma''$ to the signature of the corresponding surface must be performed at the appropriate dimension. In addition, while the discussion so far has assumed translucent objects, our test objects are actually opaque, and some parts of the surface may be self-occluded during modeling and recognition. Clutter and occlusion by other objects may also change the dimension of the signature embedding during recognition. This has prompted us to implement the following matching strategy: For each surface signature patch $\Sigma''_j \subset \mathbb{R}^{n_j-1}$, we project points on $\Sigma''_j$ to $\mathbb{R}^d$ ($d = d_{\text{min}}, \ldots, n_i - 1$) by taking all $\binom{n_i - 1}{d}$ combinations of the $d$ coordinates of the point in $\mathbb{R}^{n_i-1}$. Each original patch of the signature is thus represented by $\sum_{d=d_{\text{min}}}^{n_i-1} \binom{n_i - 1}{d}$ new patches. Since there can also be self-occlusion in the test image, for each point on $\Gamma''_j \subset \mathbb{R}^{n_j-1}$, we form all

Figure 6. a. The signature curve projected to $\mathbb{R}^3$ computed from the smoothed silhouette of the telephone shown in Figure 5; b. The corresponding quotient signature curve.

Figure 7. Images of a telephone handset used to construct its signature. The images were acquired by rotating the phone by 180 degrees about the vertical axis.
Figure 8. a. A single patch of the surface signature projected to $\mathbb{R}^3$ associated with the telephone shown in Figure 5; one of the curve components shown in Figure 6 was used to construct this patch; b. The full surface signature of the phone projected to $\mathbb{R}^3$.

Figure 9. Six objects used in the recognition experiments. The objects are resting on a platform and were rotated about an axis which was approximately parallel to the vertical axis of the image plane.

$\binom{m_j - 1}{d}$ combinations of $d$ coordinates for all $d = d_{\text{min}}, \ldots, m_j - 1$. For each dimension $d < \min(n_i, m_j) - 1$, we then perform indexing with each possible combination to find the closest match. As a last step, each point on the test silhouette signature $\Gamma''$ casts votes for those models in the database for which the distance to the model surface signature is less than some experimentally determined threshold. The votes are tallied over all points on $\Gamma''$, and an object is identified as the model with the most votes.

The above recognition procedure can be applied under orthographic viewing conditions and when the distance of the camera to all model and test images is the same for all objects. Since it is generally unreasonable to expect that the distance to all the test images to be the same as the distance to the model images, we face the following
alternative: (1) use the quotient signature surface, with a price that the dimension of the embedding space is one lower than the embedding space of the signature surface — i.e., some discriminatory information may be lost; or (2) repeatedly scale the test curve over some range, and for each scale perform the pure orthographic projection recognition procedure described above. In our implementation, we have chosen the latter method which is valid when the camera-object distance remains constant while gathering model images of each object. Test images can then be acquired at different distances.

Two minor improvements to the proposed scheme have been used: along the oriented line $\Delta$, the intersections with a pedal curve alternate between background-to-object or object-to-background transitions, and so the points on $\Gamma''$ and $\Sigma''$ can be labeled accordingly. During indexing, we enforce the fact that these labels should agree. Second, the convex hull of the surface $\Sigma$ projects onto the convex hull of the image curve $\Gamma$. For each point on the signature $\Gamma''$, the largest coordinate corresponds to a diameter of $\Gamma$ (distance between two parallel tangent lines of the hull of $\Gamma$). If there is no inter-object occlusion, then the entire convex hull will be visible. When performing the combinatorial comparison described above, it is therefore known that the largest coordinate of each point on $\Gamma''$ in the test image will also be present in the model image.

We have tested the method using 30 test images of six objects. As in the modeling case, these images have high contrast and no clutter or partial occlusion, and thresholding can be used to find the silhouette of the observed object. The test images were taken for viewing directions not used for modeling. One test image was taken from nearly overhead (i.e. a viewing direction orthogonal to the great circle of directions used during modeling). Objects were correctly identified in 27 images, or an overall recognition rate of 90%. Table I presents a confusion matrix that tabulates the correct
and incorrect matches (e.g., of the five images of a pig, the pig was recognized correctly in four images but misidentified as the Dolphin in one image).

Table I. The confusion matrix for our recognition experiments.

<table>
<thead>
<tr>
<th>Test Object</th>
<th>Number of Test Images</th>
<th>Camel</th>
<th>Duck</th>
<th>Pig</th>
<th>Toy</th>
<th>Phone</th>
<th>Dolphin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camel</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Duck</td>
<td>5</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pig</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Toy</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Phone</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Dolphin</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>

Since the approach is not based on global image features, the method is expected to perform well when there is either partial occlusion or background clutter. Given the use of a trivial segmentation scheme in our implementation, occlusion and clutter have the same effect as introducing additional, unmodeled parallel tangents while obscuring relevant parallel tangents that could have been used for recognition. In this case, the indexing and voting scheme will have to consider many more possible groupings of parallel tangent lines. We have conducted preliminary experiments with clutter (Figure 12). In this case thresholding returned a single closed contour that encompasses both the camel and also the background objects, yet the recognition scheme was able to classify a sufficient fraction of the pedal curve as coming from the camel.
6. Conclusion

We have introduced in this paper a new representation for two- and three-dimensional shapes, called their signature, that exploits the close relationship between the dual of a surface and the dual of its silhouette in weak-perspective images. Unlike duals and pedal curves and surfaces, the signatures of curves and surfaces do not depend upon the choice of an origin and they are invariant under rigid transformations. They have been used as the basis for a simple approach to object recognition from a single image. Unlike most methods for recognizing smooth curved 3D objects, this technique does not assume that objects come from a limited class such as surfaces of revolution, generalized cylinders or algebraic surfaces, and it does not require the construction of an explicit 3D model. We have presented preliminary recognition experiments, and believe that the results obtained are promising. Constructing and evaluating a true recognition system capable of recognizing complex shapes in cluttered scenes is obviously the next step in our research.

The basis for the presented recognition method is that the set of points on an object’s surface with parallel tangent planes project under orthographic projection to image curve points with parallel tangent lines. Between a test image and each model image, the signature curves and signature surfaces are essentially being used to identify candidate stereo frontier points (Giblin and Weiss, 1995). This correspondence also provides a constraint on the relative camera pose between each pair of images in an image sequence. Hence, we expect that it can be used in a process to recover both the camera motion and the 3-D structure from the silhouettes detected in a sequence of images.

This paper generalizes our earlier results (Vijayakumar et al., 1998) which defined invariants from silhouette tangent lines that were parallel to the tangent lines at inflections or bitangents. It turns out that these features are the singularities (cusps and crossings) of the dual of the silhouette. In (Vijayakumar et al., 1998), this led us to represent an object by a set of “invariant curves” while in this paper, an object is represented by a set of “invariant surfaces,” namely the signature surfaces. Consequently, the presented representation retains much more information about the object’s shape and should provide greater discriminatory power than the curves used in (Vijayakumar et al., 1998).
ACKNOWLEDGMENTS.

This work was supported in part by the Beckman Institute and the UIUC-CNRS collaboration in Computer Science, D. Kriegman and A. Sethi were supported in part by NSF ITR IIS 00-85980. J. Ponce was supported in part by the National Science Foundation under grant IRI-990709.

References


