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

Amit Sethi (asethi@uiuc.edu)
David Renaudie* (david.renaudie@ensimag.imag.fr)
David Kriegman* (kriegman@cs.ucsd.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 or constraints on their motion. The signatures of their silhouettes are concatenated into a single object signature. To recognize an object from novel viewpoint other than those used during modeling, the signature of the contours extracted from a test photograph is matched to the signatures of 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

* D. Renaudie is with the Ecole Nationale Supérieure d'Informatique et Mathématiques Appliquées de Grenoble, France, and David Kriegman is with the Department of Computer Science at the University of California at San Diego. This work was done while they were with the Department of Computer Science and the Beckman Institute at the University of Illinois at Urbana-Champaign.

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 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 successive distances between these parallel lines are used as the basis 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 the feature space. Here we propose to replace the sparse set of silhouette features used in that method 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, 1997; 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 Γ, computing its *pedal curve* (a representation of its dual) Γ', and constructing its

signature Γ'', 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 Σ is observed by a moving camera, the signatures of the successive silhouettes sweep a family Σ'' of two-dimensional surface patches in $\mathbb{R}^d$, also called the signature of Σ. 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 is then 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.

Surface Σ in $\mathbb{E}^3$	$\longrightarrow$	Pedal Surface Σ' in $\mathbb{E}^3$	$\longrightarrow$	Signature Σ'' in $\mathbb{R}^d$
$\downarrow$ projects onto		$\uparrow$ is a planar section of		$\uparrow$ lies on
Silhouette Γ in $\mathbb{E}^2$	$\longrightarrow$	Pedal Curve Γ' in $\mathbb{E}^2$	$\longrightarrow$	Signature Γ'' in $\mathbb{R}^d$

A preliminary version of this paper appeared in (Renaudie et al., 2000).

2. Duals and Pedal Curves and Surfaces

Let us consider a smooth C^2 closed curve Γ in $\mathbb{E}^2$. We define the *dual* $\mathcal{D}$ of Γ as the set of its tangent lines. The dual also forms a closed curve in the projective plane formed by all lines of $\mathbb{E}^2$. It will prove convenient to represent the dual by yet another planar curve, the *pedal curve* introduced by Maclaurin in 1718 (Bruce and Giblin, 1992; Maclaurin, 1718; Lockwood, 1967). Like the original curve Γ, the pedal curve Γ' "lives" in $\mathbb{E}^2$, but unlike the dual, its definition depends on the choice of some *origin* O in the plane. Figure 1 illustrates its construction: We associate with each point P on Γ the orthogonal projection P' of O onto the tangent line T in P; the pedal curve is the curve Γ' traced by P' as P varies along Γ. If $\vec{N}$ denotes the unit normal to Γ in P, the corresponding point P' on the pedal curve can also be defined by

$$\overrightarrow{OP'} = (\overrightarrow{OP} \cdot \vec{N})\vec{N}. \tag{1}$$

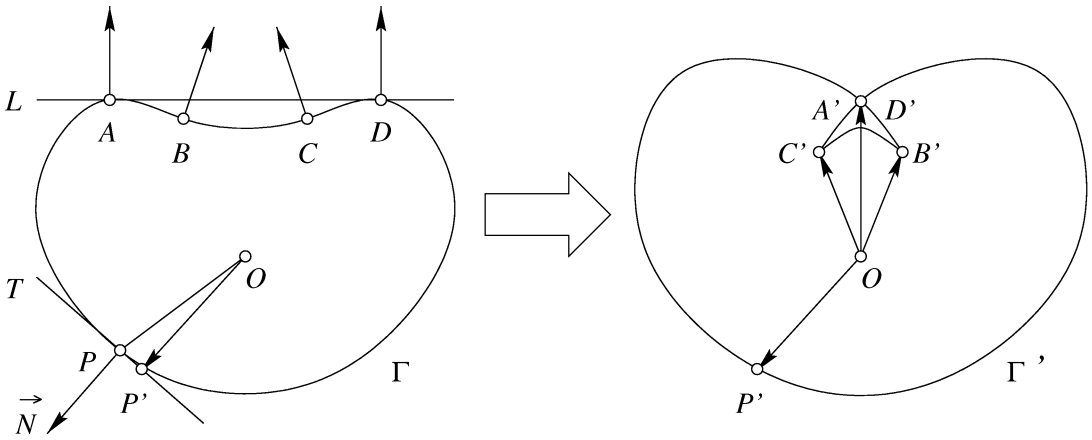

Figure 1. Construction of the pedal curve.

The pedal curve can be thought of as the image of the dual in $\mathbb{E}^2$, that associates with each tangent line T to Γ the orthogonal projection P' of O onto this line. The mapping from the dual curve to pedal curve is *not* injective in general: Indeed, any tangent passing through the origin O maps onto this point. Although there is no such tangent for the curve and origin shown in Figure 1, tangents passing through the origin are guaranteed to exist when the origin lies outside the curve, and may exist even when this is not the case (see Figure 6 for an example). To simplify the discussion, we will assume in most of this paper that the pedal curve does not pass through the origin, and identify it with the dual. We will come back to the general case during the presentation of our implementation.

As shown for example in (Bruce and Giblin, 1992)[pp. 166], the pedal curve Γ' associated with Γ has the following properties:

(A) It is smooth at all points whose preimages on Γ are not inflections.

(B) The inflections of Γ (the points B and C in Figure 1) map onto cusps of Γ' (B' and C' in this case).

(C) The lines bitangent to Γ (like the line L that passes through the points A and D in Figure 1) map onto double points of Γ' (the point $A' = D'$ in this case).

(D) The points of Γ 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 Γ' with a line passing through the origin are the images of points with parallel tangents on Γ.

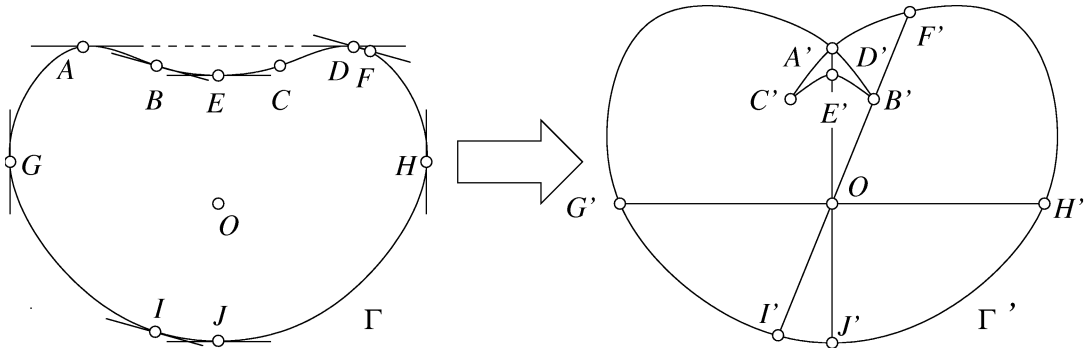

Figure 2. The intersections between various lines passing through the origin and the pedal curve are the images of point sets with parallel tangents. The number of intersections is normally even (e.g., two intersections for the line joining G' and H') but is odd at cusps and double points, corresponding to inflections and bitangents of the original curve. Lines intersecting Γ' in four points are not shown here to avoid clutter.

Property (D) 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 Γ' at these two points; lines passing through the counterclockwise angular sector defined by $\overrightarrow{OB}'$ and $\overrightarrow{OA}'$, on the other hand, will intersect Γ' four times. For a rotating line passing through O, the number of intersection points can only change at cusps and double points,[1] where the number of intersections is (exceptionally) odd: For example the line passing through O and B' also intersects Γ' in I' and F', and the line passing through O and $A' = D'$ also intersects Γ' 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 (D) 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.

[1] 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 Γ, which do not exist in the case of silhouettes of generic smooth surfaces (Koenderink, 1984).

The definitions of the dual and pedal curves generalize naturally to three dimensions (Bruce and Giblin, 1992): Consider a smooth C^2 surface Σ in $\mathbb{E}^3$; the dual of Σ 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 choose an origin O in $\mathbb{E}^3$ and associate with every point P on Σ the orthogonal projection P' of O onto its tangent plane. The surface swept by P' as P varies over Σ is the *pedal surface* Σ' associated with this surface. As in the two-dimensional case, the pedal surface is an image of the dual in $\mathbb{E}^3$ that depends on the choice of origin. And like pedal curves, the pedal surface may contain singularities, including swallowtails and cuspidal edges corresponding to parabolic lines of Σ (Bruce and Giblin, 1992).

3. Occluding Contours and their Projections

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).[2] Under orthographic projection, the center of projection is at 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).

Under orthographic projection, the imaging process is simply modeled as an orthogonal projection onto the image plane. This is a reasonable approximation of perspective projection for distant objects lying at a roughly constant distance from the cameras observing them. The *weak-perspective* (or *scaled-orthography*) projection

[2] In general, the silhouette is a curve whose only singularities are cusps and crossings (T-junctions). In this paper, we will mostly ignore these singularities so the discussion from Section 2 applies to the dual and pedal curves of silhouettes. It should be noted that the results obtained in that section can be extended to cusps and crossings since they have well-defined normal(s). The image of a cusp is simply an inflection of the pedal curve while a crossing maps to two points on the pedal curve corresponding to the two normals at the crossing.

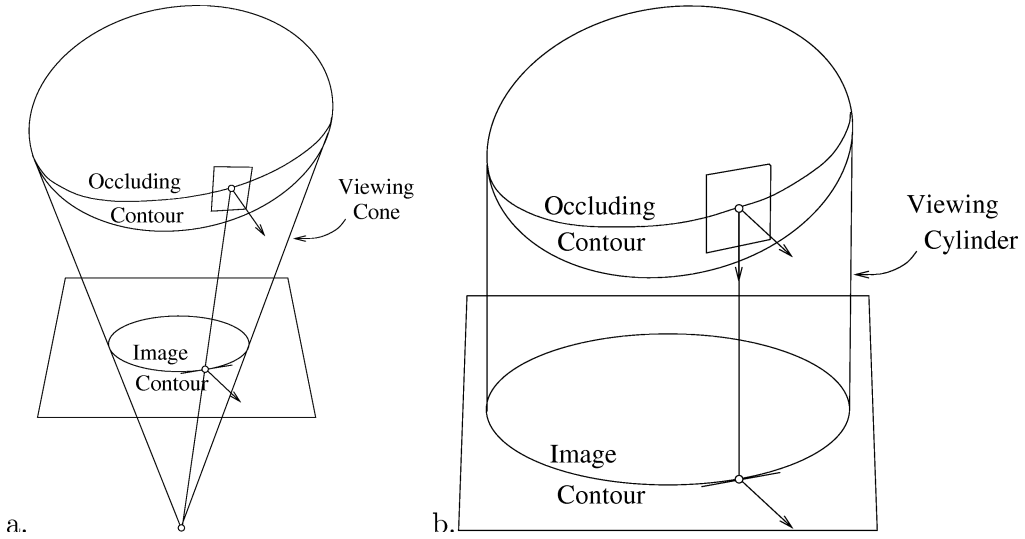

Figure 3. Occluding boundaries under (a) perspective and (b) orthographic projection.

model generalizes the orthographic one to allow for variations in the depth of an object relative to the camera observing it. The projection geometry is the same as in the orthographic case, and the silhouette and occluding contour have 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 for these objects, 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 $\vec{v}$, we denote by $\Pi_{\vec{v}}$ the plane perpendicular to $\vec{v}$ passing through O and define Γ as the silhouette of Σ formed under orthographic projection onto some image plane parallel to $\Pi_{\vec{v}}$. Let o denote the image of O, and Γ' denote the pedal curve of Γ defined using o as the origin of the plane $\Pi_{\vec{v}}$. 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 $\Pi_{\vec{v}}$ via a translation (resp. a translation followed by a scaling).*

Proof. Consider a point Q in $\Sigma' \cap \Pi_{\vec{v}}$. Since Q belongs to Σ', there exists (at least) one point P on the surface Σ such that $\overrightarrow{OQ} = [\overrightarrow{OP} \cdot \overrightarrow{N}(P)]\overrightarrow{N}(P)$. Since Q also belongs to $\Pi_{\vec{v}}$, $\overrightarrow{OQ}$ is orthogonal to $\vec{v}$ and, assuming as usual that the pedal surface does not pass through the origin, $\overrightarrow{N}(P)$ must also be orthogonal to $\vec{v}$. In other words, the point P belongs to the occluding contour associated with the projection direction $\vec{v}$.

Now, let o and p denote respectively the orthographic projections of the points O and P onto some image plane parallel to $\Pi_{\vec{v}}$. We pick o as the origin for that plane and denote respectively by Γ and Γ' the silhouette of Σ and the corresponding pedal curve. The surface Σ at P and the silhouette Γ at p have the same normal, and it follows that p maps onto the point q of Γ' defined by

$$\overrightarrow{oq} = [\overrightarrow{op} \cdot \overrightarrow{N}(P)]\overrightarrow{N}(P) = [(\overrightarrow{oO} + \overrightarrow{OP} + \overrightarrow{Pp}) \cdot \overrightarrow{N}(P)] = [\overrightarrow{OP} \cdot \overrightarrow{N}(P)] = \overrightarrow{OQ},$$

since the projection vectors $\overrightarrow{Oo}$ and $\overrightarrow{Pp}$ are by definition parallel to $\vec{v}$. It follows that $\overrightarrow{Qq} = \overrightarrow{Oo}$, and that the two curves $\Sigma' \cap \Pi_{\vec{v}}$ and Γ' are separated by the translation $\overrightarrow{Oo}$. The weak-perspective case is similar but involves the scaling inherent in this projection model. ∎

Lemma 1 identifies planar slices of the pedal surface with the pedal curves of the image contour in a coordinate-free manner. Exploiting this lemma in recognition tasks requires (1) *identifying the projection o of the point O in every image*, and (2) handling

the fact that the corresponding measurements necessarily depend on the choice of a world coordinate system and appropriate image coordinate frames. The latter problem is not difficult when the former one is solved: Indeed, it follows immediately from Lemma 1 that the measured (and therefore coordinate-dependent) descriptions of the pedal curve and the corresponding planar slice of the pedal surface are related by a rigid transformation in the orthographic case, and by a similarity transformation in the weak-perspective one. We concentrate on the orthographic projection setting here (the weak-perspective case easily follows), assuming that both the camera motion and the position of o in the image plane are known. In this case, the pedal curve of the silhouette is easily constructed, and the viewing direction $\vec{v}$ is used to determine the sectioning plane $\Pi_{\vec{v}}$; the position of o is then used to determine the translation component of the rigid transformation within $\Pi_{\vec{v}}$, while the camera motion is used to determine the rotation angle. 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 passing through the origin of the pedal surface. Note that, for some camera motions, parts of the pedal surface may be missed or in fact be covered multiple times, and the object may 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 to be determined from a silhouette. Instead, we now define a new representation for curves and surfaces that is independent of the choice of the origin (and in fact of arbitrary rigid transformations) and whose construction does not require knowing the camera motion during modeling.

We first use the pedal curve Γ' as a device for constructing the *partition* of the associated curve Γ (Figure 4): Pick some arbitrary direction in the plane as the horizontal

direction with orientation $\theta = 0$, and consider the family of oriented lines Δ_θ with orientation θ passing through the origin O associated with Γ'. As noted earlier, the number of intersections of Δ_θ and Γ' only changes at cusps and double points of this curve. Let θ_i $(i = 1, \ldots, p)$ denote the corresponding orientations of Δ_θ. The *partition* of Γ is defined as the set of triplets $(\theta_i, \theta_{i+1}, n_i)$ $(i = 1, \ldots, p)$, where the number of intersections of Γ' and Δ_θ is equal to n_i for any $\theta \in (\theta_i, \theta_{i+1})$, and index addition is performed modulo p so $p + 1 \equiv 1$.

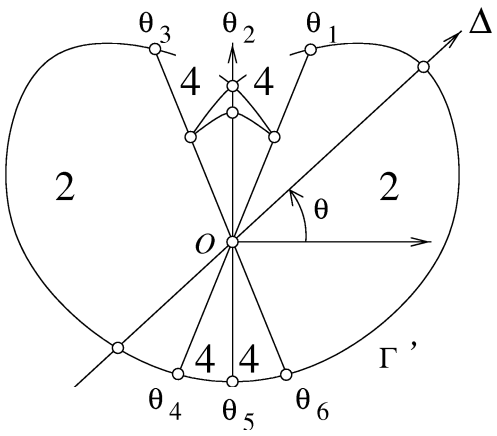

Figure 4. The partition of a pedal curve: A rotating line Δ_θ passing through O intersects the pedal curve Γ' in two points when θ is in the open range (θ_3, θ_4) or (θ_6, θ_1), and it intersects Γ' in four points when θ is in one of the open intervals (θ_1, θ_2), (θ_2, θ_3), (θ_4, θ_5) and (θ_5, θ_6). Hence, the partition for this pedal curve is $\{(\theta_1, \theta_2, 4), (\theta_2, \theta_3, 4), (\theta_3, \theta_4, 2), (\theta_4, \theta_5, 4), (\theta_5, \theta_6, 4), (\theta_6, \theta_1, 2)\}$.

We now define the *signature* Γ'' of the curve Γ. We consider again an oriented line Δ_θ passing through the origin with orientation $\theta \in (\theta_i, \theta_{i+1})$, and denote by $\overline{AB}$ the *signed* distance between two points A and B on Δ_θ. The sign is determined by the orientation of the line Δ_θ. Let us denote by P_k $(k = 1, \ldots, n_i)$ the intersections of Δ_θ with Γ', sorted in increasing $\overline{OP}_k$ order, and define $d_k = \overline{P_k P}_{k+1}$ $(k = 1, \ldots, n_i - 1)$. We define Γ''_i as the curve traced in $\mathbb{R}^{n_i - 1}$ by the points $(d_1, \ldots, d_{n_i - 1})^T$ as θ varies over (θ_i, θ_{i+1}) and define the *signature* Γ'' of the curve Γ as the *unordered set* $\{\Gamma''_1, \ldots, \Gamma''_p\}$ (Figure 5).

Note that the scalars d_k associated with some orientation θ are simply the (signed) distances between the tangent lines to Γ that are parallel to each other and perpendicular to Δ_θ. We then have:

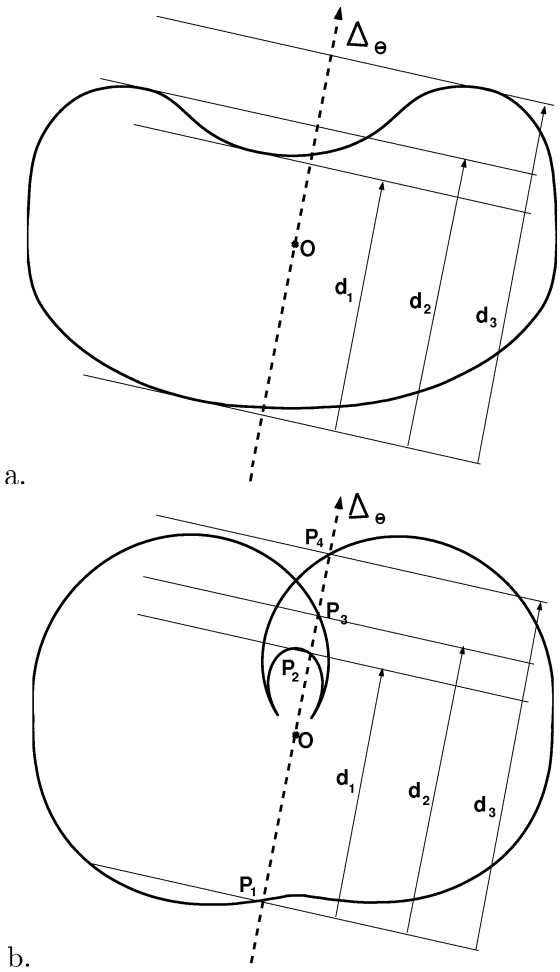

Figure 5. The scalars d_1, d_2, and d_3 defining a point on the signature are defined using distances between (a) parallel tangents on the original closed curve, or equivalently (b) intersections P_1, ... , P_4 of a rotating line Δ_θ passing through the origin (dashed) with the associated pedal curve.

LEMMA 2. *The signature* Γ'' *of a curve* Γ *is independent of the choice of the origin used to define its pedal curve* Γ'*, and it is invariant under rigid transformations of the plane.*

Proof. As noted earlier, the scalars d_k associated with some orientation θ are the signed distances between the tangent lines to Γ that are parallel to each other with orientation $\theta \pm \pi/2$. In particular, they are independent of the choice of the origin used to define Γ'. Unlike the partition, the signature does not depend on the choice of horizontal direction either, since changing this direction amounts to applying a circular permutation to the partition and the set of curves Γ''_i, but does not affect the signature

as an *unordered* set of curve segments. Finally, since rigid transformations of the plane preserve the parallelism of lines and the distance between parallel lines, the signature is also invariant under rigid transformations. ▮

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 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. In this context, the pedal curve is simply a convenient bookkeeping device for constructing the signature.

It is also possible to define the signature Σ'' of a surface as a set $\{\Sigma''_1, \ldots, \Sigma''_q\}$ of two-dimensional surface patches Σ''_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 their surface normal varies. Like the signature of a curve, the signature of a surface is independent of the choice of the origin O and is invariant under rigid transformations. Another fundamental result follows from Lemmas 1 and 2, namely:

PROPOSITION 1. *Under orthographic projection, the curve segments forming the signature of the silhouette of a surface Σ lie on the surface patches forming the signature of Σ.*

Proof. Under orthographic projection, two silhouette points with parallel tangent lines are the projection of surface points with parallel tangent planes defined by the tangent directions and the projection direction. The distance between these planes is the same as the distance between the lines, and it follows that the silhouette's signature point $(d_1, \ldots, d_{n-1})$ associated with some orientation in the image plane belongs to the signature of the surface. The proposition follows immediately. ▮

In particular, a subset of the signature of a surface can be constructed from a set of training images without any knowledge of the corresponding camera configurations. As shown in the next section, this subset can be thought of as a model of the corresponding surface, and object recognition can be formulated as the problem of deciding what surface model contains (in practice, lies close to) the signature of some test silhouette.

The model of a surface will consist of the whole signature Σ'' for sufficiently rich set of input pictures, e.g., when the viewing directions associated with a moving camera 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 $\vec{v}$, 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 $\vec{v}$ 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 will completely sweep out the entire signature of the surface.

It should also be noted that the dimension of the space in which a patch Σ''_i of the signature is embedded depends on the number of intersections of a line Δ with the pedal surface Σ'. If the number of intersections is 2, then Σ''_i 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 Δ intersects Σ' at least four times in which case these patches are embedded in $\mathbb{R}^n$ where $n \geq 3$. Also note that our discussion has assumed that the silhouette is a regular curve. In general, 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 detected in our implementation, we leave a more complete 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_i-1}). This yields the *quotient signature* $\hat{\Gamma}'' = (\hat{\Gamma}''_1, \ldots, \hat{\Gamma}''_p)$, where $\hat{\Gamma}''_i$ is the curve formed in $\mathbb{R}^{n_i-2}$ by the points $(\hat{d}_1, \ldots, \hat{d}_{n_i-2})^T = (\frac{d_1}{d_{n_i-1}}, \ldots, \frac{d_{n_i-2}}{d_{n_i-1}})^T$, as the point $(d_1, \ldots, d_{n_i-1})^T$ varies over the corresponding component Γ''_i of the signature Γ' as described in Section 4.1. The quotient signature is easily shown to be invariant under affine transformations of the curve Γ. Over a sequence of images, the quotient signatures of successive silhouettes sweeps out the *quotient signature* of the corresponding surface Σ.

5. Object Modeling and Recognition: 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. It is very important to note that 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 motions and recognized from novel viewpoints.

5.1. Object Modeling

The Canny edge detector is used to obtain object boundaries as linked, closed curves. To prevent the program from getting confused by the internal edges while extracting the silhouette, some of the internal edges are removed by hand. The normal vector at each point of Γ is then computed using linear least squares, and the pedal curve Γ' is finally computed in a straightforward manner (Figure 6.a,b). The origin O in the pedal curve computation is (arbitrarily) taken to be the center of mass of the edge points.

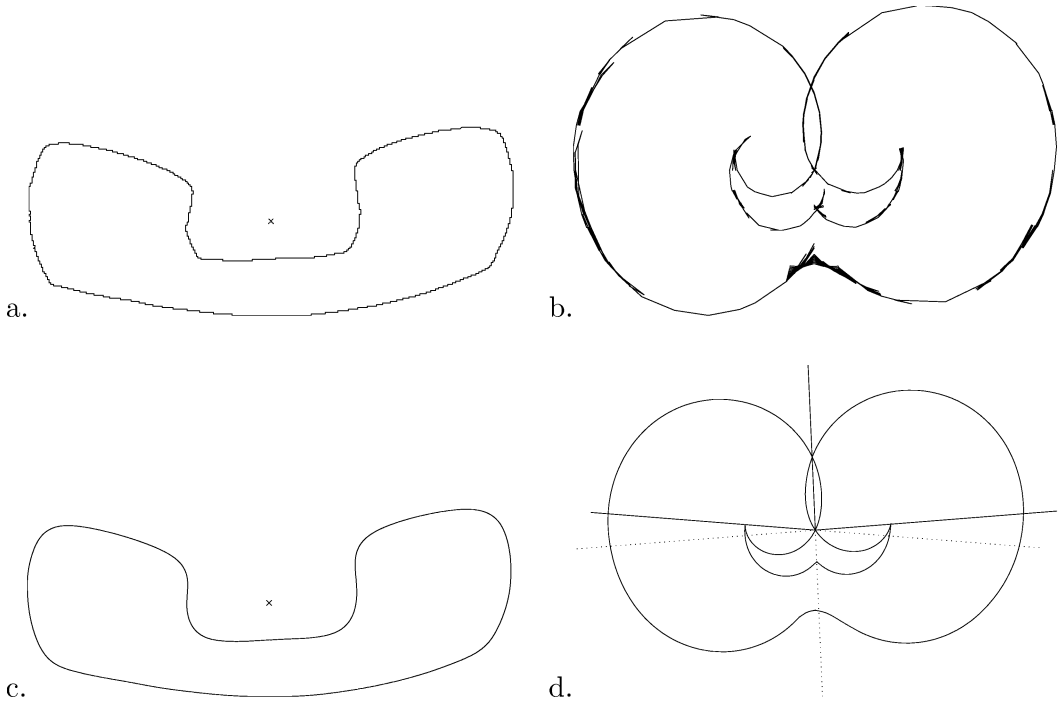

Figure 6. Pedal curve construction: (a) the raw silhouette of a telephone handset and (b) its pedal curve; (c) smoothed silhouette and (d) the corresponding pedal curve and the directions giving the partition.

As shown in Figure 6.b, the pedal curve construction process amplifies the noise in the detected edge point position. Consequently, we smooth the silhouette Γ using active contours before constructing Γ'. The active contour is initialized with the points obtained from the Canny edge detector, and the results are shown in Figure 6.c,d.

The signature Γ'' associated with the curve Γ is also computed in a straightforward manner: The range of orientations between 0 and π is sampled uniformly to give a set of oriented lines $\{\Delta_i\}$ passing through the origin. Every line Δ_i is intersected with the pedal curve, at points P_k. The intersections P_k $(k = 1, \ldots, n_i)$ are now sorted based on their signed distance x_k from the origin (where the sign is given by the orientation of Δ_i). The smallest distance x_1 is subtracted from all the signed distances of the intersections of a line from the origin to give a point $(d_1, \ldots, d_{n_i-1})^T = (x_2 - x_1, \ldots, x_{n_i} - x_1)^T$, which forms one point of the signature Γ''. The quotient signature is computed by dividing all coordinates of each point on the signature by the last one,

i.e., $(\hat{d}_1, \ldots, \hat{d}_{n_i-2})^T = (\frac{d_1}{d_{n_i-1}}, \ldots, \frac{d_{n_i-2}}{d_{n_i-1}})^T$. Since the components of $(x_1, \ldots, x_{n_i})^T$ are sorted in increasing order, the components of $(d_1, \ldots, d_{n_i-1})^T$, and $(\hat{d}_1, \ldots, \hat{d}_{n_i-2})^T$ are also sorted. Figure 7 shows an example signature projected to $\mathbb{R}^3$ for the silhouette in Figure 6, and the corresponding quotient signature.

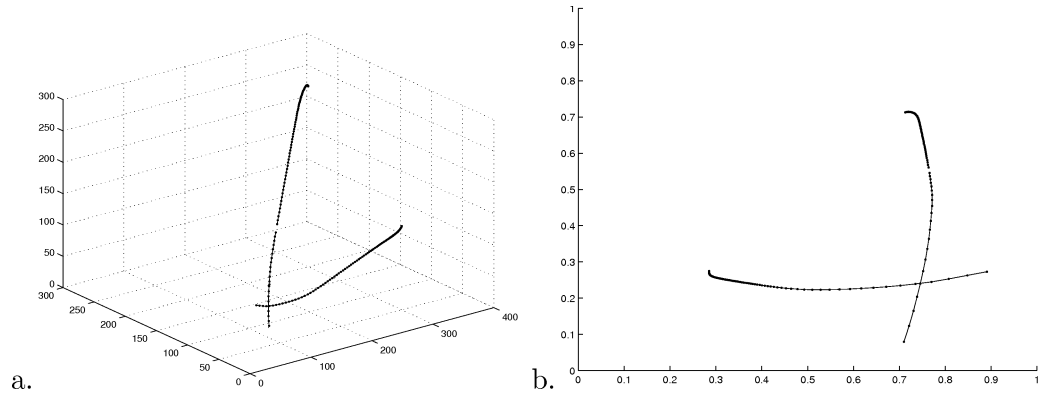

Figure 7. a. The signature curve projected to $\mathbb{R}^3$ computed from the smoothed silhouette of the telephone handset shown in Figure 6; b. The corresponding quotient signature curve.

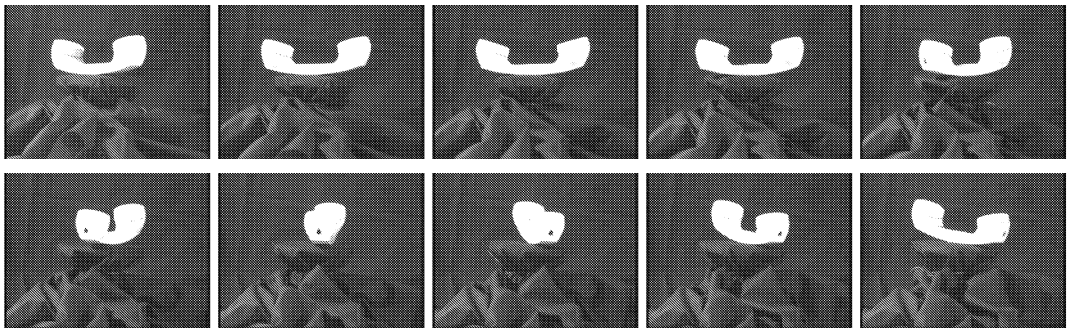

Figure 8. Images of a telephone handset used to construct the signature surface. The images were acquired by rotating the phone by 180 degrees about the vertical axis.

Note that the orientation of the sample lines in the $[0, \pi]$ interval intersecting the pedal curve and defining the signature was chosen arbitrarily. Reversing this orientation gives a second valid point on the signature for each line. As shown in the next section, rather than explicitly storing these extra points, we take them into account in the matching phase of our approach.

Signature and quotient surfaces can be constructed by concatenating together the signature and quotient curves found in successive images. Since the silhouettes of a solid observed from opposite directions are the same under orthographic projection, we only sample an 180° interval of viewing directions. Figure 8 shows ten images of a 40-image sequence (4.5° sampling) taken as a telephone handset (from here on, *phone*) rotates about a fixed axis. Obviously, the algorithm for constructing the signature knows nothing about the trajectory. Figure 9 shows the signature Σ'' of its surface.

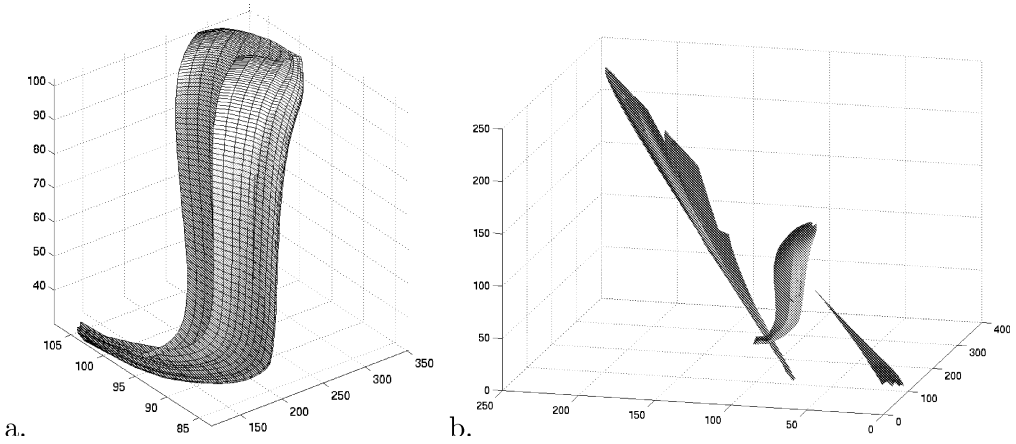

Figure 9. a. A single patch of the surface signature projected to $\mathbb{R}^3$ associated with the phone shown in Figure 6; one of the curve components shown in Figure 7 was used to construct this patch; b. The full surface signature of the phone projected to $\mathbb{R}^3$.

5.2. Object Recognition

We have constructed a simple recognition system. Figure 10 shows images of six objects, modeled using the technique proposed in the previous section: a camel, a dolphin, a duck, the phone, a pig, a stuffed toy (from here on, *toy*). Recall that objects are modeled by rotating them about a fixed axis over 180°. Ten test images of each object were also acquired from novel viewing directions, and Figure 11 shows some examples.

The principle of the recognition method is straightforward. Each modeled object is represented by its signature. The signature of the silhouette extracted from a test image is computed and matched to the stored surface signatures. In practice, some care must be given to the construction of indexing schemes adapted to the signature

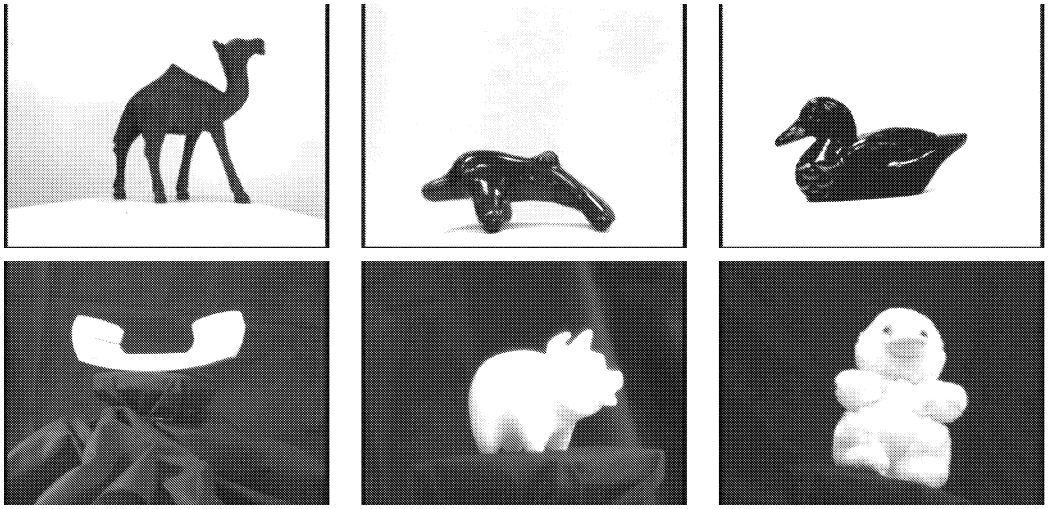

Figure 10. Six objects used in the recognition experiments. The objects are resting on a platform and were rotated about an axis which was approximately parallel to the vertical axis of the image plane.

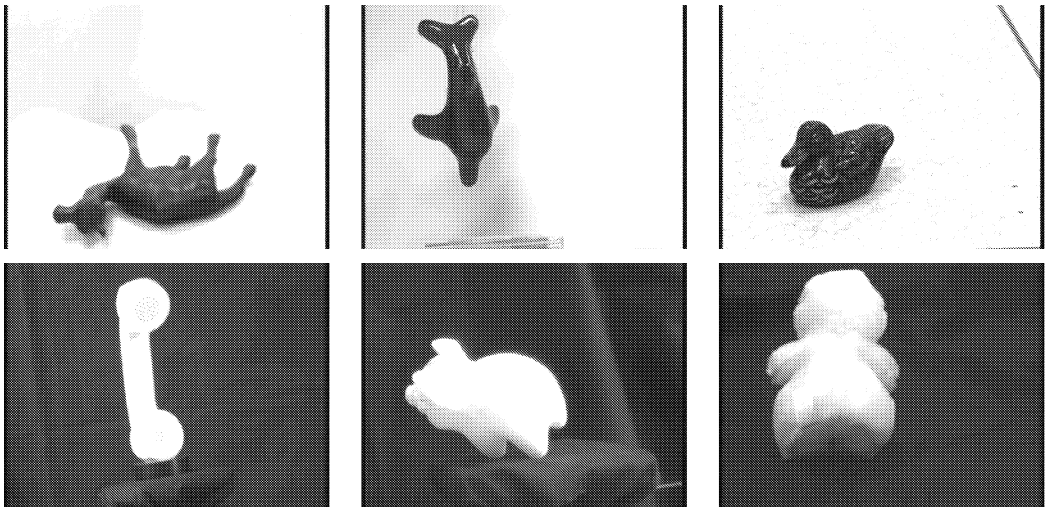

Figure 11. Sample test images for each of the objects. Note that the viewing direction is different than any of those used during modeling.

representation: Recall that a silhouette signature Γ'' is actually a collection of curves Γ''_i embedded in $\mathbb{R}^{n_i-1}$ $(i =, 1, \ldots, p)$, while a surface signature is a collection of patches Σ''_j embedded in $\mathbb{R}^{n_j-1}$ $(j = 1, \ldots, q)$. If there were no occlusion due to opacity or other objects, we would only need to compare each point p in Γ''_i to the components of Σ''_j for which $n_i = n_j$. But because of occlusion and clutter during modeling and recognition, only a subset of the features (coordinates) on Γ''_i will match those on Σ''_j.

The above reasons have prompted us to implement the following modeling and matching strategy: All images—whether acquired for modeling or as a test image—undergo the same feature extraction process. The silhouette from the image is detected, and its pedal curve is computed. The 180° range of orientations of the lines passing through the pedal curve's origin is sampled at 3° intervals, and a direction is chosen arbitrarily for each line. The signed distances of the intersections of each line with the pedal are sorted, offset and normalized as described in Section 5.1 to yield a point of the quotient signature. Thus, we obtain 60 sample points of the quotient signature curve from an image. In the experiments, an object is modeled from 40 images taken by rotating the object in 4.5° increments about a fixed axis. Thus, the signature surface of an object given by a collection of 40×60 sample points. Similarly, the test image is represented by 60 sample points of the quotient signature curve extracted from a silhouette. For each sample point on a detected quotient signature curve in a test image, the sample point on the closest quotient signature surface is determined according to the distance criterion described below, and a vote is cast for the corresponding model.

To determine the distance between two signature points $X = (x_1, \ldots, x_i, \ldots, x_n)^T$ and $Y = (y_1, \ldots, y_j, \ldots, y_m)^T$ where m may not equal n, we follow the robust matching techniques presented in (Torr and Zisserman, 2000; Forsyth and Ponce, 2002) for example. We assume that the distance between two correctly matching coordinates x_i and x_j is normally distributed with variance σ, and that the distribution of distances for all other (incorrect) matches is uniform. The maximum-likelihood match between X and Y is determined by taking the log likelihood and assuming independence. In turn, the distance between x_i and y_j is computed as the Lorentzian of their differences $d_{ij} = x_i - y_j$, or

$$l_\sigma(d_{ij}) = 1 - \frac{d_{ij}^2}{d_{ij}^2 + \sigma^2} = \frac{\sigma^2}{d_{ij}^2 + \sigma^2}. \tag{2}$$

Note that a perfect match gives a Lorentzian of 1, whereas a large mismatch gives a Lorentzian approaching 0. For all i, j, we can define an m by n matrix whose entries are $d_{i,j}$, and the best match between X and Y is taken as the path (non-decreasing function $j(i)$) that maximizes the sum of the Lorentzians. This optimal path can be found efficiently using dynamic programming. For voting, the sum of the Lorentzians is normalized by dividing by $\max(m, n)$.

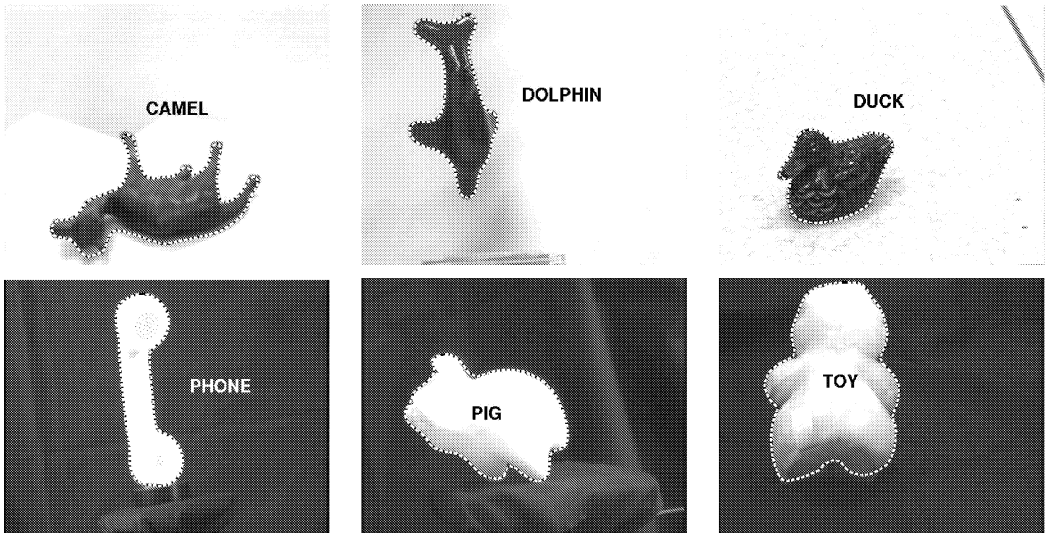

Figure 12. Sample recognition results.

Note that this simple scheme allows for the occlusion of some of the internal parallel tangents but requires that both extremal tangents be visible. This has proven sufficient in the experiments presented in the next section, where few of the extremal tangents are ever occluded.

When defining the signature, the orientation of the line intersecting the pedal curve was chosen arbitrarily. As noted earlier, for each sample point $X = (x_1, \ldots, x_n)^T$ on the signature or quotient signature, a second valid point $X' = (1 - x_n, \ldots, 1 - x_1)^T$ can be associated with the same line with opposite direction. Thus, the match score is taken as the greater of the scores of matching X with Y and X' with Y.

Hence, for each sample point on the quotient signature curve computed in a test image, a vote is cast for the model having the greatest match score. The votes are tallied over all sample points, and the model with the highest number of votes is declared to be the recognized object.

We have tested the method using 60 test images: ten images of six objects each. As in the modeling case, these images have high contrast and no clutter or partial occlusion, and Canny edge detection can be used to find the silhouette of the observed object. The test images are taken from viewing directions not used for modeling. One test image is taken from nearly overhead, with a viewing direction nearly orthogonal

to the great circle of directions used during modeling. Objects were correctly identified in 56 images, or an overall recognition rate of above 93%. It was also observed that in all the cases where the object was misclassified, it received the second highest number of votes. Table I presents a confusion matrix that tabulates the correct and incorrect matches (e.g., the toy was recognized correctly in eight of its ten images, but misidentified as the pig in two images).

Table I. The confusion matrix for our recognition experiments.

Test Object	Number of Test Images	Recognized model					
		Camel	Dolphin	Duck	Phone	Pig	Toy
Camel	10	10	0	0	0	0	0
Dolphin	10	0	10	0	0	0	0
Duck	10	0	0	9	0	0	1
Phone	10	0	0	0	10	0	0
Pig	10	0	0	0	0	9	1
Toy	10	0	0	0	0	2	8

Since our approach is based on local (or rather semi-local) image features, it is expected to exhibit some robustness to background clutter and partial occlusion. Occlusion and clutter have the same effect of introducing additional, unmodeled parallel tangents while obscuring relevant parallel tangents that could have been used for recognition. In this case, the proposed matching scheme will have to consider many more possible groupings of parallel tangent lines. We have conducted preliminary experiments with clutter (Figure 13). In this case, we always picked the longest contour segment returned by the Canny edge detector as the input to our algorithm. In our experiments, this segment encompasses both the object of interest and some of the background objects, yet our recognition scheme was able to classify a sufficient fraction of the image contour as coming from the correct object model. Similar results with occlusion are shown in Figure 14. It is clear, however, that truly accounting for occlusion and clutter in a systematic manner will require handling the combinatorics of matching in a more satisfactory manner.

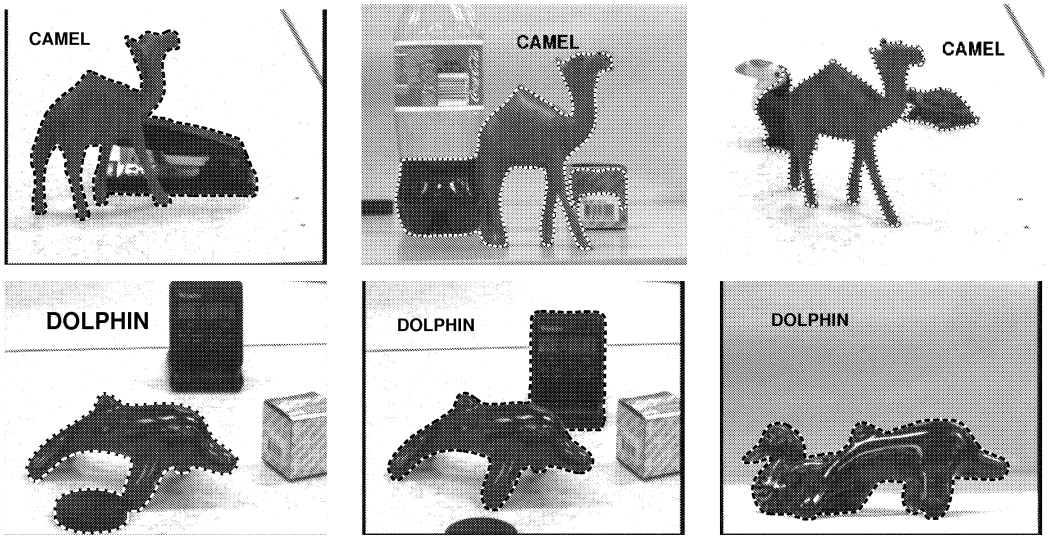

Figure 13. Three examples each of the camel and the dolphin being successfully recognized with a cluttered background. Note that the input to our matching algorithm in this case is a single closed contour encompassing the camel or the dolphin and background objects.

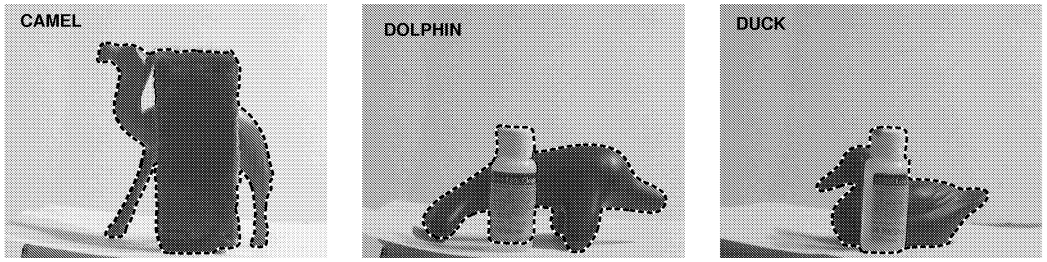

Figure 14. One example each of the camel, the dolphin, and the duck being successfully recognized with partial occlusion. Note that the input to our matching algorithm in this case is a single closed contour encompassing the object to be recognized and the occluding object.

6. Discussion and Conclusions

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 orthographic and weak-perspective images. Unlike 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 believe that our preliminary recognition experiments demonstrate the promise of the proposed approach. We would like to stress, however, that our somewhat modest set of test images is not intended to provide reliable statistics for the performance of a full-blown, operational recognition system. It is also clear that our approach is limited to shapes complex enough to have signature surfaces embedded in three- (or higher-) dimensional spaces. The aim of our experiments is not so much to characterize the recognition performance of the proposed approach as to demonstrate that (1) signature curves and surfaces can readily be extracted from image data, and (2) they form a representation powerful enough to recognize an object from viewing directions differing from those used during modeling. Ultimately, recognition accuracy will be a function of the geometric complexity of the objects, the similarity of different objects' shape, how frequently distinct objects will appear to be similar and over what range of viewpoints, and the accuracy of extracted silhouetes and tangent lines. Constructing and evaluating a true recognition system capable of recognizing complex shapes in cluttered scenes is obviously a 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. The signature curves and surfaces are essentially used to identify candidate stereo frontier points (Giblin and Weiss, 1995) between a test image and each model image. 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.

The choice of the viewpoints, and consequently the order of the sample points of the signature surface is immaterial to the modeling scheme. Indeed, although the input pictures used in our modeling experiments were always part of a continuous image sequence, we have not exploited this information in recognition tasks. However, it is

worth mentioning that we expect neighboring sample points from a signature curve obtained from a test image to match with neighboring points on the signature surface of the correct model. Therefore, the ordering of the sample points of the signature curves and surfaces could be used to add greater discriminatory power to the recognition system when input pictures are part of an image sequence. We plan to experiment with this idea in the future.

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 grants IRI-990709 and IIS-0308087. Many thanks to Martial Hebert for revealing how the Lorentzian could be used for matching.

References

Arbogast, E. and R. Mohr: 1991, '3D structure inference from image sequences'. *Journal of Pattern Recognition and Artificial Intelligence* **5**(5).

Boyer, E. and M. Berger: 1997, '3D Surface Reconstruction Using Occluding Contours'. *Int. J. Computer Vision* **22**(3), 219–233.

Bruce, J. and P. Giblin: 1992, *Curves and Singularities*. Cambridge University Press.

Cipolla, R., K. Aström, and P. Giblin: 1995, 'Motion from the Frontier of Curved Surfaces'. In: *Int. Conf. on Computer Vision*. pp. 269–275.

Cipolla, R. and A. Blake: 1992, 'Surface Shape from the Deformation of the Apparent Contour'. *Int. J. Computer Vision* **9**(2), 83–112.

Cipolla, R. and P. Giblin: 2000, *Visual Motion of Curves and Surfaces*. Cambridge: Cambridge University Press.

Forsyth, D. and J. Ponce: 2002, *Computer Vision: A Modern Approach*. Prentice Hall.

Giblin, P. and R. Weiss: 1995, 'Epipolar Curves on Surfaces'. *Image and Vision Computing* **13**(1), 33–44.

Glachet, R., M. Dhome, and J. Lapersté: 1991, 'Finding the perspective projection of an axis of revolution'. *Pattern Recognition Letters* **12**, 693–700.

Huttenlocher, D. and S. Ullman: 1987, 'Object recognition using alignment'. In: *Int. Conf. on Computer Vision*. London, U.K., pp. 102–111.

Joshi, T., B. Vijayakumar, and D. Kriegman: 1997, 'HOT curves for modeling and recognition of smooth curved 3D objects'. *Image and Vision Computing* **15**(7), 479–498.

Kergosien, Y. L.: 1981, 'La famille des projections orthogonales d'une surface et ses singularités'. *C.R. Acad. Sc. Paris* **292**, 929–932.

Koenderink, J. J.: 1984, 'What does the occluding contour tell us about solid shape?'. *Perception* **13**, 321–330.

Koenderink, J. J. and A. J. Van Doorn: 1976, 'The Singularities of the Visual Mapping'. *Biological Cybernetics* **24**, 51–59.

Kriegman, D. and J. Ponce: 1990a, 'On Recognizing and Positioning Curved 3D Objects from Image Contours'. *IEEE Trans. Pattern Anal. Mach. Intelligence* **12**(12), 1127–1137.

Kriegman, D. J. and J. Ponce: 1990b, 'Computing exact aspect graphs of curved objects: Solids of revolution'. *Int. J. Computer Vision* **5**(2), 119–135.

Liu, J., J. Mundy, D. Forsyth, A. Zisserman, and C. Rothwell: 1993, 'Efficient recognition of rotationally symmetric surfaces and straight homogeneous generalized cylinders'. In: *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.* New York City, NY, pp. 123–128.

Lockwood, E. H.: 1967, 'Pedal Curves'. In: *A Book of Curves*. Cambridge University Press, Cambridge, England, pp. 152–155.

Lowe, D. G.: 1987, 'The Viewpoint Consistency Constraint'. *Int. J. Computer Vision* **1**(1), 57–72.

Maclaurin, C.: 1718, 'Tractatus de Curvarum Constructione & Mensura'. *Philosophical Transactions* **30**(356), 803–812.

Murase, H. and S. Nayar: 1995, 'Visual Learning and Recognition of 3-D Objects from Appearance'. *Int. J. Computer Vision* **14**(1), 5–24.

Ponce, J. and D. Chelberg: 1987, 'Finding the limbs and cusps of generalized cylinders'. *Int. J. Computer Vision* **1**(3).

Ponce, J., A. Hoogs, and D. Kriegman: 1992, 'On Using CAD Models to Compute the Pose of Curved 3D Objects'. *CVGIP: Image Understanding* **55**(2), 184–197.

Renaudie, D., D. Kriegman, and J. Ponce: 2000, 'Duals, Invariants, and the Recognition of Smooth Objects from their Occluding Contour'. In: *Proc. European Conf. on Computer Vision*. pp. Vol. I, 784–798.

Richetin, M., M. Dhome, J. Lapresté, and G. Rives: 1991, 'Inverse perspective transform from zero-curvature curve points: Application to the localization of some generalized cylinders from a single view'. *IEEE Trans. Pattern Anal. Mach. Intelligence* **13**(2), 185–191.

Sullivan, S. and J. Ponce: 1998, 'Automatic Model Construction, Pose Estimation, and Object Recognition from Photographs using Triangular Splines'. *IEEE Trans. Pattern Anal. Mach. Intelligence* **20**(10), 1091–1096.

Torr, P. and A. Zisserman: 2000, 'MLESAC: A New Robust Estimator with Application to Estimating Image Geometry'. *CVIU* **78**(1), 138–156.

Vaillant, R. and O. Faugeras: 1992, 'Using extremal boundaries for 3D object modeling'. *IEEE Trans. Pattern Anal. Mach. Intelligence* **14**(2), 157–173.

Vijayakumar, B., D. Kriegman, and J. Ponce: 1998, 'Invariant-Based Recognition of Complex Curved 3-D Objects from Image Contours'. *Computer Vision and Image Understanding* pp. 287–303.

Zeroug, M. and G. Medioni: 1995, 'The Challenge of Generic Object Recognition'. In: M. Hebert, J. Ponce, T. Boult, and A. Gross (eds.): *Object Representation for Computer Vision*. Springer-Verlag, pp. 271–232.