CAR-TR-632
CS-TR-2942

N0004-91-J-1222
DACA76-92-C-0009
August 1992

GEOMETRIC INVARIANTS AND OBJECT RECOGNITION

Isaac Weiss

Center for Automation Research
University of Maryland
College Park, MD 20742-3275

Abstract

In this paper we discuss the role of the general invariance concept in object recognition, and review the classical and recent literature on projective invariance. Invariants help solve major problems of object recognition. For instance, different images of the same object often differ from each other, because of the different viewpoint from which they were taken. To match the two images, common methods thus need to find the correct viewpoint, a difficult problem that can involve search in a large parameter space of all possible points of view and/or finding point correspondences. Geometric invariants are shape descriptors, computed from the geometry of the shape, that remain unchanged under geometric transformations such as changing the viewpoint. Thus they can be matched without search. Deformations of objects are another important class of changes for which invariance is useful.

Keywords: object recognition, matching, invariance, projective, perspective.

The author is grateful for the support of the Office of Naval Research under Grant N0004-91-J-1222, the Defense Advanced Research Projects Agency (ARPA Order No. 8459) and the U.S. Army Topographic Engineering Center under Contract DACA76-92-C-0009.

1. Introduction

Object recognition is a major goal of computer vision, but many obstacles remain on the road towards effective recognition systems. In this paper we discuss ways of overcoming many of the difficulties by using invariants of shapes.

A typical problem is that an object can be seen from different points of view, resulting in different images which we would like to recognize as portraying the same object. In a typical recognition task we have one image stored in a database, and we need to compare it with an image of an unknown object observed from an unknown point of view. This difficult task can be greatly facilitated by using suitable invariants. These are shape descriptors computed from the image which are independent of the viewpoint, i.e. they are the same regardless of which point of view the image was taken from.

It can be argued that object recognition is the search for invariants. Given an image of an object, we want to extract one unique invariant: a name or a similar ultimate descriptor. Given another image of the same object, differing from the first by, e.g., viewpoint, we want to extract the same unique descriptor. To do that, we have to eliminate in some way the effect of the transformations that gave rise to the differences between the images.

There are several methods of eliminating transformations between images. The simplest is by performing every possible transformation of one image and see if any of its transformed versions matches the other image. For instance, in template matching [Ballard and Brown 1982], it is assumed that a template and an image differ only by translation, and the template is moved pixel by pixel over the image until a match is found. However, when more complicated transformations are involved, such as rotation, projection, etc. this search space becomes overwhelmingly large.

To reduce the search space, "invariant features" can be used [Lowe 1985]. These are features in the image that stay invariant under some transformation and can be matched directly between the two images. For example, an edge remains an edge, so edges can be used for matching. The problem here is that the kinds of features usually used do not have much distinctiveness. Any edge in one image can match any edge in the other. This leads to the correspondence problem, which can easily lead to a combinatorial explosion. Invariant constraints [Grimson 1987] can also be used but they still leave a large space to search in.

Other methods aimed at viewpoint invariance have their own drawbacks. Fourier descriptors are not fully invariant and suffer from occlusion problems. Hashing methods such as the Hough transform break down when a large number of parameters is involved.

The correspondence problem can be solved by using more distinctive invariant descriptors, i.e. descriptors that are invariant only to the transformation we are interested in and not to others. For instance, a shape descriptor of a fish should be distinct from a descriptor of a frog, i.e. it should not be invariant to a transformation that maps the shape of the fish into that of the frog. Edges, of course, are invariant to this since they can appear in both shapes; they are "too" invariant, i.e. they are invariant to too wide a set of transformations. Thus, we must try to find features that are invariant only to the transformations that we want to eliminate and to no others, so they are distinctive enough to match without ambiguity.

Change in the point of view is only one kind of geometric transformation that images can undergo. For instance, we would like to identify an object as a "fish" even if the particular example of a fish we are looking at is somewhat thinner or fatter than some standard fish. In this case we need invariants to *deformations*, i.e. quantities that will not change under a not-too-great deformation of the object. It is again important not to seek invariance to transformations that are too general, because then the descriptors will blur the distinction between different objects.

A fundamental question immediately arises: what transformations do we want to eliminate? When do we decide that two images come from the same object, even though they are different? Viewpoint change is one example; other transformations will probably depend on the types of objects in question.

Another consideration in choosing the kind of invariance we need is that the larger the set of transformations, the harder it is to extract meaningful distinctive quantities that are invariant to it. (For example: distance, a Euclidean invariant, is not preserved under projection, a larger group.) Yet the need for invariants is much more acute, because the larger set of transformations has more unknown parameters and requires a search in a much bigger space. This consideration thus leads to the same conclusion as the distinctiveness argument: we have to find optimal invariants, i.e. ones that will stay invariant under the set of transformations that we want to eliminate, but not under a larger set.

A paradigm for object recognition can thus include the following:

1) Identify the transformations that an image can undergo and still describe the same object, i.e. the transformations that we want to eliminate for particular classes of objects.

2) Find descriptors that are invariant to these transformations but not to others.

3) Use these descriptors for indexing of shapes and matching.

In the next section we discuss point (1) above, and in the rest of the paper we carry out points (2), (3) for projective and related transformations. For other transformations these points have yet to be investigated.

2. Which invariants?

Here we only deal with purely geometric invariants, i.e. ones that can be calculated from the shape alone. Other surface properties such as shading, reflectance, color, etc. can also be considered as invariants, subject to the same considerations as above, but are not treated here.

The most obvious invariants useful in vision are the Euclidean ones. A simple example is the length of a rod, which is invariant under rotation. In a simple world consisting of rods that lie in a plane, and with images that can only rotate, one can identify a particular rod by measuring its length on the image and comparing it to a database of rod lengths. The rod's orientation is irrelevant and can be ignored. As another example, when a 2-D curve is rotated or translated in the plane, its curvature at each point does not change. Thus curvature is an invariant of the Euclidean transformations. It is common to plot the curvature of such a curve as a function of its arc-length (which is invariant up to a starting point) to obtain a 2-D Euclidean invariant representation of the curve. Curvatures of surfaces have also been used when they can be measured, e.g. from range data.

The formation of images in general involves a larger set of transformations (containing the Euclidean group). A projective transformation, for example, more general than Euclidean, and involves non-parallel projection onto a plane. The number of free parameters in this case is eight, so finding the correct point of view can involve a search in an 8-dimensional space! Clearly projective invariants, namely quantities that are unchanged under this transformation, are of crucial importance.

When enlarging the transformation set, the problem arises that the invariants of the smaller set do not remain invariant. The length of a rod is no longer invariant under projective transformation. Similarly, an oblique view of a circular disc yields an ellipse, and obviously neither arc-length nor curvature is preserved under such projections.

To find invariants of larger sets, one has to extract more information from the image. While finding length requires two points, a similar projective invariant needs four, so we need to extract more data from the image to obtain reliable results. This is more than offset by the enormous saving of eliminating the search. However, it does lead us to conclude that we should not enlarge the transformation group beyond what is absolutely necessary. The distinctiveness argument mentioned before leads to the same conclusion.

Projective transformations (projectivities) are the smallest group that includes all possible viewpoint-related changes in images, and therefore we concentrate on them. Apart from the invariants issue, using projective geometry can unify and simplify the treatment of perspective and orthographic projections, which are often treated separately.

The most readily useful projectivities are the ones operating on a 2-D plane. One view is sufficient to reconstruct a planar shape (except for the projectivity). Therefore invariants by themselves are sufficient as means for indexing and recognizing planar shapes. They are also applicable to 3-D objects, since many objects contain planar shapes, such as facets, symmetry planes, etc., which are generally projected onto the image as planes. In addition, small areas of a 3-D surface can be approximated as planar. Thus, 2-D projective transformations and their invariants can be used for recognition of many 3-D objects.

Smaller subsets of the projective transformations are often quite useful. When the object is distant from the camera, one can assume that the projection rays are nearly parallel, which defines the affine transformations. If we can find one feature point that can be regarded as unchanged in the projection, we have a perspective transformation. Euclidean motions are a common subset of both the affine and perspective transformations.

In 3-D, one rarely needs to consider a full projection. A surface in 3-D can be translated, rotated or perhaps scaled, but not projected. However, 3-D projective invariants of curves and surfaces do exist and they are summarized in [Weiss 1988]. The Euclidean and affine 3-D invariants have the same role of indexing of 3-D shapes as the projective invariants have in 2-D.

The case of projecting a 3-D object into a 2-D image is of a different nature. In this case, true invariants cannot be found because the depth information is missing and cannot be recovered by purely geometrical methods. Additional, "model-based" knowledge is needed to reconstruct the missing information, and this is beyond the capacity of invariants alone. However, invariants can be useful here too. We will see some projection examples. Deformation invariants can also be useful here. When trying to identify a pair of stereo images as belonging to the same object, we can regard small parts of the object as nearly planar, with the deviation from planarity giving rise to a small deformation in the image. Thus a combination of projective and deformation invariants can be of use in problems of reconstructing shapes from stereo, motion or other geometric information.

As mentioned before, invariants of deformations are valuable in their own right. The same problem immediately arises: what kind of deformation? Obviously too general invariance will defeat the goal of distinctiveness. One possibility is to restrict ourselves to small, or quasi-linear deformations. This is a very promising topic for current investigation.

3. History of Geometric Invariants

From a purely abstract point of view, it can be argued [Klein 1926] that geometry is in essence the study of invariants. In Klein's view, abstract ("synthetic") objects such as "points" or "lines", which do not necessarily have a real world interpretation, are invariant objects, and geometry deals with abstract operations on these objects. This is the Klein "Erlangen program" of 1872.

Here we are interested in invariants that are more analytic. The first one was discovered by Lagrange [1773] who showed that the the discriminant of a quadratic polynomial is invariant under translation along the x axis. (However, it is claimed that this invariant was discovered in India much earlier [Bhaskaracharya 1150].) Geometrically, the vanishing of the discriminant indicates that the two roots of the polynomial coincide, a translationally invariant property.

In the last half of the 19th century, there developed an extensive study of invariants. Two main tracks evolved: algebraic and differential invariants. The algebraic track is concerned with invariants of algebraic forms, namely homogeneous polynomials. The field was advanced in England by Salmon, Elliot, Cayley, Sylvester, Grace and Young. A systematic "symbolic" method was developed in Germany by Aronhold, Clebsch and Gordan. Of central interest was the question of whether a complete system of fundamental invariants exists for a given set of algebraic forms, from which any other invariant can be derived. The question was finally answered in the affirmative by [Hilbert 1890, 1893] in a famous set of theorems that ended the search for polynomial invariants, and has become the foundation of algebraic geometry.

On the other track, progress was made in finding invariants of general parametrized curves and surfaces (rather than algebraic forms). These differential invariants are local to points on a shape and can be used for arbitrary shapes. They were studied by Halphen [1880], Wilczynski [1906, 1907, 1908] and Fubini [1927]. Lane [1942] describes some of this work.

A more modern, abstract approach was taken by [Weyl 1939], Cartan, Mumford [1965] and Nagata [1963], who developed theories of invariants of general Lie group transformations. The mathematical field is still active [Abhyankar 1990].

In computer vision, only very restricted kinds of invariants were used until recently. The curvature, a Euclidean invariant, is common. An algebraic projective invariant, the cross ratio of four points on a line, was used by several authors: Duda and Hart [1973]; Chang *et al.* [1987]. Tensor invariants for camera calibration were studied by Kanatani [1986].

Projective invariants for curves and surfaces were first introduced in vision by Weiss [1988]. This paper reviewed some of the classical literature on algebraic and differential invariants, which had previously been ignored, and pointed out their importance to object recognition. Since then many researchers have developed various aspects of the subject, and some of this work is summarized here.

4. Overview of Geometric Invariants

Basic Definitions

We describe here some general characteristics of invariants of a general transformation. The geometric shape itself is a fixed entity in space, but its analytic representation necessitates choosing some coordinates and parameters, and it is their transformation which raises the invariance issue. There are two main ways of representing shapes: the implicit and the explicit representations. In the implicit approach, the shape is represented as a relation between coordinates x_i

$$f(a_k, x_i) = 0$$

with a_k being coefficients characterizing the shape, such as line or conic coefficients—namely, they are mainly global descriptors. This is mostly associated with algebraic, global invariants. In the explicit approach the coordinates of the shape points are functions of some local curve parameter t (or surface parameters t_l)

$$x_i = x_i(t)$$

The shape descriptors here are the derivatives $\partial^n x_i / \partial t^n$, so this is mostly associated with differential, local invariants. There are also mixed, or hybrid approaches. An invariant is a function derived from either the global or local descriptors whose value does not change under a transformation of the coordinates x_i and parameters t_l, or changes in a limited way defined below.

We define a *relative invariant I of weight w* an a function of the shape descriptors that transforms as

$$\tilde{I} = J^{-w} I \tag{1}$$

with the tilde indicating a quantity in the new system. J is the Jacobian of the appropriate transformation. There are in general different weights for different transformations: the coordinate transformation T, and the parameter change $d\tilde{t}/dt$. A similar change can result from multiplication of x_i homogeneously by some factor λ. This is of importance in projective homogeneous coordinates. In this case the invariant can change as

$$\tilde{I} = \lambda^d I \tag{2}$$

with d being the *degree* of the invariant.

An invariant of weights and degrees zero is *absolute*.

The Jacobians and λ can very from one point to another, i.e. they depend on x_i, t, but they do not depend on the descriptors of the shape itself, i.e. a_i or $\partial x_i^n / \partial t^n$.

General Properties

Among the general properties that one is interested in are the questions of uniqueness and completeness of sets of invariants. For differential invariants a fundamental theorem can be stated [Guggenheimer, p. 144]:

Theorem 1. *All differential invariants of a (transitive) transformation group in the plane are functions of the two lowest order invariants and their derivatives.*

This is part of the *completeness* property, which states that the original curve can be reconstructed from the two independent invariants that exist at each point, except for the relevant transformation.

This leads to the possibility of creating invariant "signatures" of curves. For example, in the Euclidean case, all invariants can be derived from the curvature and the arc-length at each point, $\kappa(s)$. Thus, we can use the signature, or the plot of the curvature vs. the arc-length, to identify the curve up to a Euclidean transformation. Similar plots can be drawn in the affine case (Section 9). In the projective case one does not have a natural arc-length but there are still two independent invariants I_1, I_2 at each curve point. Thus one can plot I_1 against I_2 in an invariant plane obtaining an invariant signature curve.

The method is illustrated in Figs. 1–3. Fig. 1 shows a shape to be recognized. Fig. 2 is a projection of this shape. At each point of the shape of Fig. 1 we have calculated two invariants, I_1, I_2, and plotted an invariant curve with coordinates I_1, I_2 (Fig. 3). Here the invariants are the affine arc-length and curvature. Repeating the process for the projected curve in Fig. 2, we obtain another invariant curve which is superimposed on the first one in Fig. 3. Since the match is close we are able to conclude that Figs. 1 and 2 differ only by a projection. No search is needed!

Similar completeness properties were proved for algebraic invariants of homogeneous polynomials, or algebraic forms. These shapes include points, lines, conics and higher order shapes. Hilbert's fundamental theorem (Section 9) in its various versions ensures the existence of a finite base of invariants from which all other invariants can be derived, for any finite set of algebraic forms.

An interesting general question is how much information has to be obtained from the given shape in order to calculate invariants. To find invariants, the parameters of the transformation between the object and the image, which are unknown, have to be eliminated. For example, for a rod under Euclidean transformations, the angle at which a rod lies (the rotation parameter) and its position (the translation parameters) are of no interest and we only want the rod's length. Thus, from the coordinates of the two end points (four measured quantities) we have to eliminate three by calculating the length, the only invariant. In general, we need to extract more quantities from the image than the number of transformation parameters, so that we can later eliminate the transformation parameters mathematically and be left with invariants. For the general planar projectivity, for example, the number of measured quantities has to exceed eight, the number of projection parameters. In differential methods, the information extracted from the image is in the form of derivatives, while in the algebraic methods it consists of global shape parameters. The curve's arbitrary parametrization, however, complicates matters since it has to be eliminated also. We will return to the parameter issue and propose a differential method that does not depend on a parameter.

This counting argument is only a guide, since it changes when a shape has an internal symmetry or degeneracy. For example, if instead of a rod we had a ring, than the rotation angle does not play a role. The ring has three parameters, the center and the radius, of which only the two center coordinates have to be eliminated, while the radius is a Euclidean invariant. In the planar projective case, four collinear points (eight parameters) have one invariant, the cross ratio. This is because the configuration is symmetric with respect to one of the projectivity's parameters, namely the tilt whose axis is the line. We will later encounter other such degeneracies.

A rigorous treatment of the amount of information needed can be obtained from the Lie prolongation method, in which the ordinary space in "prolonged" to include all the information needed for invariants [Guggenheimer 1963, Olver 1986, Weyl 1939]. For Euclidean distance, for instance, this space made up of pairs of points rather than the original space of single points. For differential invariants the space is prolonged to include derivatives. In principle one can derive the invariants from studying transformations in this space. In practice, the differential theory leads to systems of partial differential equations so other, more specific methods are easier to use.

For vision purposes, one can compare the usefulness of algebraic versus differential invariants.

As we saw, the amount of information needed to be extracted from the image depends on the generality of the transformation, not on the method of computing the invariants. If we need high derivatives in the differential method, we also need a rather large number of coefficients for fitting an algebraic curve. Thus, the decision as to which method to use will be based on other considerations, such as suitability for use with complex shapes, and ability to cope with practical problem such as occlusion, noise of various kinds, etc.

Algebraic invariants are rather easy to implement, but they face serious problems. First, they are global descriptors. Since an entire shape has to be fitted, the problem of occlusion arises. This is a well known problem for any global method, such as moments or Fourier transforms. Second, the traditional ones are restricted to a limited repertoire of curves, mostly polynomials, such as a system of two conics. This problem has been attacked by the idea of invariant fit, in which simple shapes such as conics are fitted to more general curves invariantly. We will later describe such methods.

Differential invariants are local so the occlusion problem is less likely to be troublesome. Furthermore, they can be derived for any kind of curves, rather than just polynomials. The drawback of the method is extracting the local descriptors, such as derivatives.

It is possible to combine the advantages of the two approaches, hopefully without combining the disadvantages. We will describe a method of fitting an implicit polynomial in a window around a point of an arbitrary curve, and find the polynomial's invariants. In this way we use an algebraic, implicit method—locally.

Another issue of importance in vision is the amount of correspondence one needs to establish between elements of the observe image and the stored one. If the image is one general curve, then its invariants enable us to perform matching without any correspondence, because we can obtain an invariant signature curve that is identical for all possible views of the curve. However, this requires obtaining a large amount of data at each curve point, such as high derivatives, which reduces robustness. For pure algebraic forms such as conics, one needs correspondence between the various forms, but this is easier to achieve than simple point correspondence. A middle way is offered by "hybrid" shapes, combinations of curves with identifiable features such as points or lines. For instance, a silhouette of an airplane can have both curved and straight contours. We can use the information in the feature points or lines to reduce the order of the curve derivatives. However, the correspondence between the features has to be established.

5. Projective Geometry

From this point we specialize to projectivities. We summarize here some basic elements of projective geometry.

A projective transformation (projectivity) can be defined analytically in 1-D as

$$\tilde{x} = \frac{ax+b}{cx+d} \tag{3}$$

with a, b, c, d being arbitrary parameters. Only three of these are meaningful because an arbitrary factor can multiply both the numerator and denominator. If the projectivity has a (real) invariant point, then it can be represented geometrically as a *perspectivity*, Fig. 4. (The intersection of the two points is an invariant point.) Otherwise, the projectivity can be decomposed into a perspectivity plus translation. Unlike projectivities, perspectivities are not a group unless the fixed point is always the same. If one invariant point is at infinity, we have the *affine* sub-group, with $c = 0$. Geometrically it corresponds to a perspectivity between two parallel lines, or alternatively, a perspectivity with parallel rays (i.e. with the center of projection at infinity), plus translation.

In the plane, eq. (3) can be generalized in a straightforward way. It is convenient to write it in matrix form:

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \\ 1 \end{pmatrix} = \frac{1}{xT_{31} + yT_{32} + T_{33}} T \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{4}$$

where T is a non-singular 3×3 constant matrix, with eight significant parameters. A projectivity with an invariant line whose points are also invariant can be represented as a perspectivity. The affine sub-group has an invariant line at infinity, so it preserves parallelism in the plane (as parallel lines "meet" at infinity). A general projectivity involves combinations of perspectivities and affinities.

The matrix elements can be identified as

$$T = \begin{pmatrix} aff_1 & aff_2 & trans_x \\ aff_3 & aff_4 & trans_y \\ proj_1 & proj_2 & 1 \end{pmatrix}$$

The elements marked aff_i represent rotation, scaling in the x and y directions and shear. Together with the translation elements $trans_x, trans_y$ they make up the affine group. The $proj_1, proj_2$ elements represent tilt and slant, which are non-linear transformations.

In an affinity the $proj_i$ elements above vanish so the transformation is linear:

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} aff_1 & aff_2 \\ aff_3 & aff_4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} trans_x \\ trans_y \end{pmatrix}$$

The terms defined above should be distinguished from similar, commonly used terms such as perspective projection, or perspective camera. The latters refer to a projection from a 3-D object to a 2-D image, while the trad˙ .onal terms refer to transformations from n-D to n-D.

Homogeneous Coordinates

The form of the transformation (4) is inconvenient because the denominator leads to infinities and because of the non-linearity. One can deal with the problem by using homogeneous coordinates. The Cartesian coordinates of a point (x, y) are replaced by a triplet

$$\mathbf{x} = (x_1, x_2, x_3)$$

so that

$$x = \frac{x_1}{x_3}, \qquad y = \frac{x_2}{x_3}$$

Of course this definition is not unique, as the x_i can be multiplied by any common factor and still correspond to the same (x, y). Thus, one can express the homogeneous coordinates with the help of an arbitrary factor λ,

$$(x_1, x_2, x_3) = \lambda(x, y, 1)$$

The points with $x_3 = 0$ have no corresponding Cartesian points since the division leads to infinity. However, they are perfectly valid points of the projective space so the Euclidean infinity is now treated on an equal footing with other points. The point (0,0,0) is excluded from the space.

The general projective transformation can now be written as

$$\tilde{\mathbf{x}} = \lambda_{\mathbf{x}} T \mathbf{x} \tag{5}$$

i.e. it has the appearance of a simple linear transformation. However, when going back to Cartesian coordinates, one has to divide the new vector $\tilde{\mathbf{x}}$ by $\lambda_{\mathbf{x}} = \tilde{x}_3 = \frac{\lambda x_3}{1 + x_1 T_{31} + x_2 T_{32}}$, so the non-linearity reappears.

Lines are the projective duals of points, and homogeneous coordinates make it possible to express the duality algebraically. The line with coefficients a_1, a_2, a_3 can be expressed as the dot product

$$\mathbf{a} \cdot \mathbf{x} \equiv \mathbf{a}\mathbf{x} = a_1 x_1 + a_2 x_2 + a_3 x_3 = 0$$

Here $\mathbf{x}$ is a column vector and $\mathbf{a}$ is a row vector. It is easy to see that $\mathbf{a}$ is contragredient to $\mathbf{x}$, i.e. it transforms with T^{-1}:

$$\tilde{\mathbf{a}} = \lambda_{\mathbf{a}} T^{-1}$$

because this ensures that the new coefficients satisfy the line equation $\tilde{\mathbf{a}}\tilde{\mathbf{x}} = 0$ in the new system (describing the same line). Again, an arbitrary factor $\lambda_{\mathbf{a}}$, which depends on the line, can multiply $\mathbf{a}$ without affecting its equation or the geometrical interpretation.

In homogeneous coordinates, ordinary 2-D polynomials become homogeneous polynomials, or *algebraic forms*. Much of projective geometry involves these forms. The line is an algebraic form of the first order. The conic is an algebraic form of the second order, and it is convenient to represent it with a symmetric conic matrix A so that

$$\mathbf{x}^t A\mathbf{x} = 0$$

Again an arbitrary factor λ_A can multiply A.

Upon transformation, to preserve the above equality in the new coordinate system (since the geometrical curve is preserved), A transforms as

$$\tilde{A} = (T^{-1})^t A T^{-1} \tag{6}$$

The dual of A is the *line* conic A^{-1}, representing the tangents to A. It transforms to $TA^{-1}T^t$.

For the affine transformation, we have $T_{31}, T_{32} = 0$ so the denominator in (4) vanishes and thus we can set $\lambda_{\mathbf{x}} = 1$ for the *points*. It is still convenient to use the triplets $\mathbf{x} = (x, y, 1)$. This form is preserved by the affine transformation $T\mathbf{x}$. However, we still need to use a multiplying factor $\lambda_{\mathbf{a}}$ for the *lines* $\mathbf{a}$ because their transformation $\mathbf{a}T^{-1}$ will not preserve any normalization of $\mathbf{a}$. The duality is in fact broken here. Similarly, the coefficients of higher order forms are also multiplied by λ.

6. Overview of Projective Invariants

In this section we highlight some of the main methods of obtaining invariants and categorize them according to the domain on which they are applicable and the basic principle behind the method (Table 1). Some of the methods have applicability beyond projectivities, but we concentrate here on projective and affine invariants. The Lie prolongation method was already mentioned as a general method for obtaining general results, but we are interested here in methods that produce the invariants themselves.

The first categorization is according to the domain, or classes of objects for which the method is most useful. We distinguish here three types: 1) Local vicinities, namely points on curves (or surfaces) and their immediate vicinity. The descriptors here can be derivatives or other local quantities. 2) Whole curves or surfaces. Most methods here deal with algebraic forms such as lines and conics. Some methods are applicable to more general shapes. The descriptors here are in the form of coefficients, moments or other global quantities. 3) Hybrid shapes, that can include combinations of the previous two types.

The second criterion is the main principle on which the method is based. This can be, for instance, determinant properties or canonical frames. The principles can involve mainly differential operations, mainly algebraic, or both. A differential method is obviously better suited to local vicinities and an algebraic to global shapes, but there is some overlap. The canonical method, for instance, can use an implicit representation locally.

Accordingly, one can organize the current methods in a table as follows.

Table 1

method	type	local	hybrid	global
Wilczynski	dif	p		
Cartan	dif	a		
Canonical	dif, alg	a,p	a,p	a,p
Prolongation	alg,dif	a	a,p	a,p
Determinants	alg,dif	a	a,p	a,p
Symbolic	alg			a,p
Moments	alg			a

The first column in the table indicates whether a method is purely differential (dif), purely algebraic (alg) or a hybrid. The next column indicates whether the method can be applied for projective invariants (p), affine invariants (a), or both. Similarly for the other columns, indicating invariants of hybrid shapes and of global shapes. We will briefly highlight each method and mention some recent applications in computer vision. In subsequent sections the two most general methods, namely the determinants and the canonical methods, will be described in more detail. Other methods will only be touched on.

Wilczynski's method [1906] was the first to obtain closed form formulas for projective invariants of curves and surfaces. It was described in computer vision by this author [Weiss 1988]. While interesting mathematically it has proven difficult to implement in vision ([Brown 1991]) because of the high order of derivatives involved. (Section 7.)

Cartan's "moving frame" method [Guggenheimer 1963, Weiss 1992a] is an explicit method applicable to general transformations. However is hard to apply to transformations (such as projectivities) which do not admit a natural arc-length parameter. It is easily worked out for unimodular affine transformations, for which an affine length and affine curvature are obtained.

The canonical method [Weiss 1992a,b] is a general method that can be used locally or globally, implicitly or explicitly. It consists of defining a canonical, or standard coordinate system using the properties of the shape itself. Since this canonical system is independent of the original one, it is invariant and all quantities defined in it are invariant. Here we briefly describe the application of the canonical method to local and hybrid shapes in an implicit approach (Section 8).

The determinant method (Section 9) is based mainly on the transformation properties of determinants. If a matrix A is transformed by T to AT, then its determinant $|A|$ is transformed to $J|A|$ with J being the Jacobian of the transformation, $|T|$. Tensor dot products and traces are added in some cases to obtain complete sets of invariants. (However, much of tensor theory is not applicable because of the lack of a metric.) Invariants of a wide variety of shapes can be obtained by this simple method. Notable exception are projective invariants of curves and probably those of high order forms. For global forms such as point sets, lines and conics the method easily leads to cross ratios, to invariants of conics [Weiss 1988], etc. These invariants have been adapted, using invariant fitting, to industrial shapes [Forsyth *et al.* 1991]. For hybrid shapes, invariants were obtained by [Van Gool *et al.* 1991] and by [Brill *et al.* 1992]. They investigated general curves with known feature points. Although the method is basically algebraic, local affine invariants involving derivatives are easily obtained.

The symbolic method is the classical means by which invariants of algebraic forms were investigated. Developed by Gordan, Hilbert and others [Grace and Young 1903] it led to general theorems as well as methods of obtaining invariants for homogeneous polynomial curves of any order. At its heart it is also based on determinants, but it deals with abstract "symbols" from which the form can be built, rather than directly with the form themselves. (Section 9). For forms of order higher than two, the symbolic method is very cumbersome to implement in practice, reflecting its origin in determinants.

In the method of moments, the familiar Euclidean moments are generalized to the affine case [Taubin and Cooper 1992] and to perspectivities [Park and Hall 1987]. To find moments, one integrates over a closed shape $\xi(\mathbf{x})$ with homogeneous polynomials. In first order we have the vector $M^{[1]} = \int \mathbf{x}\xi(\mathbf{x})$, in second order we have the matrix $M^{[2]} = \int \mathbf{x}\xi(\mathbf{x})\mathbf{x}^t$, and similarly for higher order tensors in n-D. Under a linear (affine) transformation T, the moments transform in a tensor-like way, e.g. $\tilde{M}^{[2]} = TM^{[2]}T^t$. One can then find tensor invariants such as dot products, traces, eigenvalues and determinants.

This list of possibilities is not exhaustive. Invariants of areas (with unknown contours) have been studied by [Nielsen and Sparr 1990]. Affine invariant Fourier descriptors were treated in [Arbter *et al.* 1990]. Quasi-invariants, that change more slowly than the transformation, were proposed in [Binford 1981]. Euclidean curvatures have been used by many authors, e.g. [Besl and Jain 1985], [Cyganski and Orr 1985], [Stevenson and Delp 1989]. Other related papers are listed in the references.

The methods are not unrelated. The global determinantal invariants can easily be derived from the "symbolic" determinants. The canonical method is more like a computational algorithm than closed form formulas, and its relationship to the symbolic method is perhaps analogous to the relationship between methods of solving a set of linear equations. In that case we can eliminate the unknowns either by using the determinant formulas, or by Gauss elimination. The latter method brings us to a "canonical", diagonalized system and it is much more practical for higher orders. The Schwartz derivative, a 1-D invariant (Section 7), is the infinitesimal limit of the cross ratio of a line. Other relations are not yet clear.

New Geometry Challenges

The methods described above are rather invariant to the passage of time, many dating back to the 19th century. The problems of vision pose new challenges that can stimulate new developments in geometry, which in turn will benefit vision. We try here to identify such geometry challenges that are related to invariance. The value of these geometrical aspects is likely to endure beyond the specifics of the immediate applications.

In trying to find an adequate geometrical model for vision, projective geometry is only a partial answer. To improve the model, further assumptions need to be added to it, which are perhaps more context-specific. The general challenge is then to identify useful assumptions and develop the appropriate geometry based on them.

One issue is extracting the shapes (Section 10). The methods described above assume that the shapes are already given in some ideal form. In practice, of course, we are given a collection of pixels that have to be turned into curves or other shapes. In doing so some assumptions must be made that are beyond projective geometry, for instance that a curve minimizes some distances.

One important problem here is invariant fitting. It is desirable that the fitting assumption be invariant. This is a source of difficulties but also a source of new possibilities. With invariant fitting, the global methods in Table 1 can be freed of their restriction to specific forms such as conics, by fitting conics invariantly to more general curves. This has been done in the affine case by Bookstein [1979], Forsyth *et al.* [1990, 1991] and Kapur and Mundy [1992]. In the projective case, a method based on invariant segmentation into conics is due to Carlsson [1992]. Another route is opened by the canonical method, because the fitting can be done in the canonical frame [Weiss 1992b]. The method of moments does not require an invariant fit. However, high order moments are known to be sensitive to noise and occlusion.

For a local method an invariant fit is less important, but the problem arises of finding high order derivatives or fitting high order curves. Accurate derivatives were obtained in [Weiss 1991] for this purpose.

Another important problem is the connection between 2-D images and 3-D objects (Section 11). It has long been known that in general the projection from a 3-D shape to a single 2-D image does not have invariants (e.g. [Burnes *et al.* 1990]). There is simply not enough information in a 2-D image to reconstruct the missing depth information by purely geometrical methods. The invariants discussed above are 2-D to 2-D (or n-D to n-D). However, given some additional, external information, invariants can be useful here too. In [Zisserman and Mundy 1992], invariant descriptors of surfaces of revolution are derived from contours detected in a single image. Hopcroft *et al.* [1992] recover the length of three 3-D vectors using an invariant orthogonality relation. Reconstruction of 3-D invariants from multiple views, given the correspondence, is done in [Koenderink and Van Doorn 1991], [Brill *et al.* 1992] and [Barrett *et al.* 1992]. Qualitative invariants are discussed in [Weinshall 1990]. The subject is very promising but is only just beginning.

Other problems in which invariants have found use can only be mentioned here. Camera calibration, in which the invariance is to the camera parameters in addition to the geometry of the shape, was treated by [Kanatani 1990], [Mohr 1992] and others. Invariants in space-time for motion were treated in [Faugeras and Papadopoulo 1992]. Hashing methods using affine coordinates were developed by Lamdan *et al.* [1988]. Meer and Weiss [1992] studied statistical methods for point set invariants.

7. Pure Differential (Explicit) Methods

This section describes explicit differential methods. Being local, they do not suffer from the occlusion problem and can be used for an arbitrary shape. However, the parameter of the curve needs to be eliminated which reduces the robustness of the invariants. Cartan's moving frame method belongs here, but we will derive its results (local affine invariants) more simply by using the determinant method, Section 9.

A 1-D Projective Invariant

We first mention a well-known one-dimensional differential invariant, namely the Schwarzian derivative [Springer 1964]. Consider a particle moving along a straight line, with its position at a time t measured by a (non-homogeneous) coordinate $r(t)$. The Schwarzian derivative $S(r)$ is defined as

$$S(r) \equiv \left(\frac{r''(t)}{r'(t)}\right)' - \frac{1}{2}\left(\frac{r''(t)}{r'(t)}\right)^2$$

and it is invariant under projective transformations of the line, given by eq. (3) as can be checked directly. Furthermore, the differential equation

$$S(r) = g(t)$$

(where $g(t)$ is given) determines the relation $r(t)$ up to 1-D projectivity. The Schwarzian derivative is not invariant to change of the parameter t, except by a 1-D projectivity (3). It is interesting that this invariant can be obtained as an infinitesimal limit of the well-known 1-D cross ratio.

Wilczynski's Method

As described previously, a projectivity can be written in homogeneous coordinates as

$$\tilde{\mathbf{x}} = \lambda(\mathbf{x})T\mathbf{x}$$

with $\lambda(\mathbf{x})$ being an arbitrary factor, which can be different at each point $\mathbf{x}$. To find invariants, one can proceed in stages. First find invariants to the linear part T of the transformation, and from those derive invariants to λ (and also to change in parameter).

Given a plane curve $\mathbf{x}(t)$, invariants to T can be obtained by solving the linear algebraic system of equations

$$\mathbf{x}''' + 3p_1\mathbf{x}'' + 3p_2\mathbf{x}' + p_3\mathbf{x} = 0$$

for the three unknowns p_1, p_2, p_2, at each point t. It is easy to show, by multiplying the equation through by T, that these solutions p_i are invariant to T. (In fact, p_i are expressible as determinants.) However, they are not invariant to change in the arbitrary factor $\lambda(\mathbf{x}(t))$ nor to change in the curve parameter t. We can obtain functions of these p_i which are invariant to the additional transformation needed. We have the "semi-invariants"

$$P_2 = p_2 - p_1^2 - p_1' \tag{7}$$

$$P_3 = p_3 - 3p_1p_2 + 2p_1^3 - p_1''$$

These remain unchanged under multiplication of the coordinates by a factor $\lambda(\mathbf{x})$, but not under change of the parameter t.

The full invariants are

$$\Theta_3 = P_3 - \frac{3}{2}P_2'$$

$$\Theta_8 = 6\Theta_3\Theta_3'' - 7(\Theta_3')^2 - 27P_2\Theta_3^2$$

Under change of the parameter t, they transform as $\tilde{\Theta}_w = (d\tilde{t}/dt)^{-w}\Theta_w$, where $\tilde{t}(t)$ is the new parameter along the curve, and w is the weight. The subscript corresponds to the weight w.

Theorem 2. *The invariants Θ_3, Θ_8 determine all other invariants. Furthermore, they determine a plane curve except for a projective transformation.*

These two invariants still contain the unknown weight factor, which varies from point to point. To eliminate it we can use the invariant $\Theta_{12} = 3\Theta_3\Theta_8' - 8\Theta_3'\Theta_8$. We can now define the two absolute invariants [Weiss 1988]

$$I_1 = \frac{\Theta_3^8}{\Theta_8^3}, \qquad I_2 = \frac{\Theta_3^4}{\Theta_{12}}$$

These can be plotted against each other in an invariant plane with coordinates I_1, I_1. We can thus obtain an invariant signature curve identifying the original curve up to a projectivity.

Since these invariants contain the eighth derivative they are not very practical. The semi-invariant P_2 above contains the fourth derivative only. The other, P_3, contains the fifth but it can be replaced by

$$P_3^* = P_3 - P_2'$$

which again contain only the fourth derivatives.

We can clearly see the burden that the curve parameter imposes on the method. The semi-invariants P_2, P_3^* are invariant to the projectivity and contain only fourth derivatives. It is the requirement of invariance to the change of parameter t that pushes the number of derivatives needed to eight. Thus, if we can get rid of the parameter, we will need fewer local quantities and the robustness of the invariance will increase.

8. The Canonical Method

The canonical method can be used in a variety of situation, for local, global or hybrid shapes of various combinations, either in explicit or implicit ways [Weiss 1992a, 1992b]. For forms of order higher than two it is much more computationally efficient that the determinant based symbolic method. This is in analogy to the Gauss elimination method that yields a diagonal matrix in the eigenvectors frame. However, the problem here is non-linear and there is no "automatic" algorithm. Each situation has to be handled separately.

The basic idea is to transform the given coordinate system to a "canonical", or standard system, which is determined by the shape itself. Since this canonical system is independent of the original system, it is invariant. All quantities defined in it are thus invariant. The concept can be illustrated by examples of simpler transformations. If a 1-D function $x(t)$ is subject to scale transformation in x, we can obtain scale invariance by transforming to a new coordinate $\bar{x}$ in which the derivative at the origin is fixed, say $\bar{x}'(0) = 1$. We achieve this by a simple normalization $\bar{x} = x/x'(0)$. This also fixes other scale-dependent quantities such as the second derivative $\bar{x}''(0)$, so they are now scale invariant.

An important 2-D example is the Euclidean invariants. To find an invariant at a given point on a curve, we change the x, y axes so that the new x-axis is tangent to the curve at that point. We thus have $\bar{y}' = 0$, while the second derivative $\bar{y}''$ at this point is now the curvature. It is invariant since we will obtain this canonical system regardless of which system we started with. We see that by determining some of the properties of the system, the others are also determined and become invariant. We generalize this process to the projective case.

Here we use the canonical method implicitly, to avoid the parameter of the explicit differential method. To do that we fit an implicit, algebraic curve in a vicinity of a point $\mathbf{x}_0$ at which we want to find invariants. The coefficients a_i of this curve $f(a_i, \mathbf{x})$ become local descriptors at $\mathbf{x}_0$. The task now is to find the invariants of f. In the examples below, the canonical method is probably the most, if not the only suitable one.

In finding invariants, the parameter is undesirable for the following reason. The essence of finding invariants is the elimination of unknowns from the system, such as the unknown quantities describing the point of view. The parameter is also in general unknown since it can be chosen in an arbitrary way. It has to be eliminated so that the invariants will not depend on it. The more unknowns we have to eliminate, the more information we have to extract from the image, which translates in the explicit method to higher, and less reliable, derivatives. We have seen in Wilczynski's method that the need for invariance to the parameter pushes the order of derivatives from four to eight. On the other hand, the parameter is not in fact part of the geometry of the curve itself. The relation between x, y is sufficient to completely characterize the curve.

We can see the practical implication of the parameter problem in fitting a curve to the data. To obtain an eighth derivative, one has to fit eighth order polynomials to the data, for both $x(t)$ and $y(t)$. In the parameterless method we only need to fit one cubic. Lower powers are much less sensitive to noise.

The Osculating Curve

The invariants of the implicit curve are found with the help of an osculating curve at our point $\mathbf{x}_0$. We have already used the tangent to find Euclidean invariants. An osculating curve is a generalization of the tangent. A tangent is a line having at least two points in common with the curve in an infinitesimal neighborhood, i.e. two "points of contact". This can be expressed as a condition on the first derivative. Similarly, a higher order osculating curve has more (independent) points of contact, and the condition on the derivatives can be written as

$$\frac{d^k}{dt^k}(f^*(x,y) - f(x,y)) = 0, \qquad k = 0 \ldots n \tag{8}$$

with f^* being the osculating curve, f the given curve, and n the order of the osculation. Since the derivatives vanish, this condition is invariant to the parameter t. Since it has a geometric interpretation with points of contact, the condition is also projectively invariant.

In the calculation we will not need either the parameter or the above derivatives. The "data quantities" needed here are the coefficients of the given curve f, which can be obtained by fitting f to the data points. We need no more of them than in the algebraic method. Thus the robustness is increased relative to the explicit differential method. In principle, a cubic will do, having nine coefficients plus the point's position. In practice, however, we have found that a wide window is necessary for robustness to noise, and this requires a higher order curve such as a quartic

$$f(x,y) = a_0 + a_1 x + \ldots + a_{14} y^4 \tag{9}$$

(Not all its coefficients need be independent.) Fitting was done with the SVD method.

To find invariants at a point $\mathbf{x}_0$ of a given implicit curve f, we proceed in steps: i) find a simpler curve f^* that osculates it at that point. The coefficients of this osculating curve are independent of any parametrization. ii) Eliminate the factors in the projectivity by moving to a canonical coordinate system in which the osculating curve has a simple, predetermined form.

The osculating curve f^* is chosen as the simplest one that enables us to perform the factor elimination. Three factors are eliminated by moving the origin to the given point $\mathbf{x}_0$ and rotating so that the x-axis is tangent to the given curve f there. The projectivity has another five factors to be eliminated.

Local Canonical Invariants

Three of the eight projectivity parameters are eliminated by moving the origin to $\mathbf{x}_0$ and rotating so that the x-axis is tangent to the curve there. Now we need a five-parameter osculating curve f^* that passes through the origin. A suitable choice is the "nodal cubic" [Halphen 1880, Weiss 1992b]

$$f^* = c_0x^3 + c_1y^3 + c_2xy^2 + c_3x^2y + c_4y^2 + xy = 0 \tag{10}$$

This curve intersects itself at the origin so it has two tangents there, one lying along the x-axis. The other tangent is called the "projective normal" [Lane 1942]. This f^* osculates the fitted f with a seventh order contact (Fig. 5a).

Our goal is now to transform the coordinates so that this nodal cubic take on the simple coefficient-free form

$$x^3 + y^3 + xy = 0 \tag{11}$$

known as a *folium of Descartes*, Fig. 5b.

We obtain it, in a nutshell, as follows. We already have the x axis of the canonical system. The canonical y axis is now chosen as the other tangent of the nodal cubic, the projective normal. We skew the whole shape so that this projective normal becomes perpendicular to the x axis. This will eliminate the term with c_4 in the nodal cubic. Next, the coefficients c_0, c_1 are eliminated by scaling in the x and y directions. We obtain

$$\bar{x}^3 + \bar{y}^3 + \bar{c}_2\bar{x}\bar{y}^2 + \bar{c}_3\bar{x}^2\bar{y} + \bar{x}\bar{y} = 0$$

The coefficients in this system, $\bar{c}_2$ and $\bar{c}_3$, are *local affine invariants* because we have reached an affine canonical system. We have used all possible affine transformations (translations, rotation, skewing, scalings) to eliminate all the possible affine transformation factors and arrive at the above form of the cubic so the remaining coefficients are uniquely defined regardless of which system we started with.

A projective canonical system is obtained by eliminating the last two coefficients using tilt and slant. We transform the original curve f to this new system and obtain new coefficients $\bar{a}_i$ for it. Since this system is projective invariant, these $\bar{a}_i$ are invariants. We can choose some suitable combination of them as invariants I_1, I_2.

In summary, the implicit method gets rid of the parameter while the canonical method makes it practical to find invariants of the resulting cubic (and other forms).

Hybrid Shape Canonical Invariants

Some of the projectivity factors can be eliminated by using known feature points or feature lines, each of which can eliminate two factors. The remaining are eliminated as before with the help of an osculating curve, and the conic with three parameters will suffice in all cases:

$$c(x,y) = c_0x^2 + c_1y^2 + c_2xy + y = 0 \tag{12}$$

Since we need a lower order of contact of this osculating curve than we needed before (corresponding to lower derivatives), the robustness increases. However, the correspondence of the feature lines/points has to be established. We have studied [Weiss 1992b] all configurations of a curve plus one or two points or lines. We only describe here the simplest situations.

Given a feature point $\mathbf{x}_1$, we draw a line joining it with with the curve point x_0, y_0 (Fig. 6a). This is obviously a projectively invariant operation. We use this line as our new y axis. As before we skew the system so that this line becomes perpendicular to x. We thus obtain an orthogonal system. We also scale the y axis so that the distance of the feature point form the origin is unity.

To obtain the osculating conic to our fitted curve f we need only a fourth order contact, rather than a sixth as before. For an affine canonical system, we only need to scale in the x direction by eliminating one coefficient of the conic, c_0. The remaining two are affine invariants. For a projective canonical system, we use tilt and slant to eliminate the remaining conic coefficients and obtain a unit parabola $x^2 + y = 0$, Fig. 6b. The invariants are again coefficients $\bar{a}_i$ of the transformed curve $\bar{f}$.

Given a feature line, we can convert to the previous situation by finding its polar point with respect to the osculating conic, an invariant operation (Fig. 7).

9. The Method of Determinants

This is perhaps the simplest and most widely used method and can handle most common cases. However, it cannot yield the pure differential projective invariants, or high order forms in an obvious way. For hybrid forms it uses explicit derivatives with a curve parameter t. We derive here a variety of invariants in 2-D.

This method takes advantage of the transformation properties of determinants under linear transformation. Many geometrical entities can be cast in the form of determinants. In 1-D, the distance between points x_1, x_2 is

$$l_{12} = x_1 - x_2 = \begin{vmatrix} x_1 & 1 \\ x_2 & 1 \end{vmatrix}$$

while the area of a 2-D triangle can be written as

$$S_{123} = \frac{1}{2}\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} \tag{13}$$

In homogeneous coordinates, the triplets are multiplied by arbitrary factors, i.e. $\mathbf{x}_i = \lambda_i(x_i, y_i, 1)$, so the determinants are multiplied by $\lambda_1\lambda_2$ in 2-D and $\lambda_1\lambda_2\lambda_3$ in 3-D. In these coordinates, the projective transformation is linear (eq. (5)), so it only multiplies the determinants by $|T|$ and λ_i:

$$|\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \tilde{\mathbf{x}}_3| = \lambda_1\lambda_2\lambda_3|T||\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3|$$

Thus, a determinant of points/lines in homogeneous coordinates is a *relative projective invariant* of weight -1 in T and degree 1 in each λ_i.

The above properties are used to find invariants of various algebraic forms. The main trick is to find ratios of various determinants in which all the factors λ_i as well as $|T|$ cancel out, so the relative invariants become absolute. The duality of points and lines makes it possible to interchange their roles in all the formulas below. In [Bruckstein *et al.* 1991] many determinant invariants are given a geometrical meaning.

To complete the sets of invariants, dot (scalar) products and traces of tensors are useful in appropriate cases. Thus the term "dets and dots" is sometimes used. (Note that determinants are defined on square matrices, not general tensors.)

Global Projective Invariants of Forms

Here we will obtain invariants for first and second order forms, i.e. points or lines and conics, as well as combinations of such forms. In general, the configuration of forms must have a total of more than eight coefficients to eliminate the eight projective transformation factors. However, in some cases we will have an internal symmetry that reduces the number of coefficients needed.

Four collinear points (Fig. 4). The cross ratio of Euclidean distances is equal to the cross ratio of determinants in *homogeneous* coordinates, because all λ_i cancel out:

$$\frac{l_{12}l_{34}}{l_{13}l_{24}} = \frac{(\lambda_1\lambda_2 l_{12})\ (\lambda_3\lambda_4 l_{34})}{(\lambda_1\lambda_3 l_{13})\ (\lambda_2\lambda_4 l_{24})} = \frac{|\mathbf{x}_1, \mathbf{x}_2||\mathbf{x}_3, \mathbf{x}_4|}{|\mathbf{x}_1, \mathbf{x}_3||\mathbf{x}_2, \mathbf{x}_4|} \tag{14}$$

and under the transformation T this will be unchanged because $|T|$ will also cancel out.

Five points: The configuration has ten coefficients, thus can yield two independent invariants. By the same cancellation process as before, we can prove the invariance of the cross ratios of either areas S_{ijk} of triangles or the corresponding determinants in homogeneous coordinates. We have

$$I_1 = \frac{S_{423}S_{125}}{S_{124}S_{523}}, \qquad I_2 = \frac{S_{143}S_{125}}{S_{124}S_{153}}$$

The same method yields cross ratios in n-dimensional spaces.

Two points, two lines: The line coefficients $\mathbf{a}$ are contragredient to $\mathbf{x}$, i.e. they transform with T^{-1} (Section 5). Thus dot products such as $\mathbf{a} \cdot \mathbf{x}$ are invariant to T. However we still have to cancel the factors $\lambda_\mathbf{x}, \lambda_\mathbf{a}$ so that we have to use ratios again:

$$\frac{\mathbf{a}_1\mathbf{x}_1}{\mathbf{a}_2\mathbf{x}_1}\frac{\mathbf{a}_2\mathbf{x}_2}{\mathbf{a}_1\mathbf{x}_2}$$

Although there are only eight coefficients, we have an invariant because there are only seven unknowns to eliminate. This is because we have one degree of freedom of symmetry: the line joining the two points intersects the two lines, thus creating four collinear points. These points have a cross ratio, unaffected by rotation around this line.

Four points, one line: We have ten parameters, and assuming that the points $\mathbf{x}_i$ are not collinear and not on the line $\mathbf{a}$, we have three invariants

$$\frac{\mathbf{a}\mathbf{x}_1}{\mathbf{a}\mathbf{x}_2}\frac{S_{234}}{S_{134}} \qquad \frac{\mathbf{a}\mathbf{x}_1}{\mathbf{a}\mathbf{x}_3}\frac{S_{234}}{S_{124}} \qquad \frac{\mathbf{a}\mathbf{x}_1}{\mathbf{a}\mathbf{x}_4}\frac{S_{234}}{S_{123}}$$

One conic: The conic an be expressed (Section 5) as the quadratic form $\mathbf{x}^t A\mathbf{x} = 0$ with the symmetric matrix A. It transforms as $(T^{-1})^t A T^{-1}$ (eq. (6)). The matrix can obviously by multiplied by the arbitrary factor λ_A without affecting the form. Thus the *discriminant* $|A|$ is a relative invariant of weight 2 and degree 3:

$$|\tilde{A}| = |T|^{-2}\lambda_A^3|A|$$

The degree results from the fact that multiplying all the matrix elements A_{ij} by λ_A results in multiplication of $|A|$ by λ_A^3. To eliminate the factor λ_A^3 we can normalize the coefficient matrix A by the relative invariant $|A|^{1/3}$ and define a new matrix

$$\hat{A} = \frac{1}{|A|^{1/3}}A \tag{15}$$

Two conics: The configuration has ten coefficients yielding two independent invariants [Springer 1964, Weiss 1988, Forsyth *et al.* 1990]. The last authors applied them to real images (Figs. 8, 9). The joint invariants of the conics A, B can be obtained from the solutions of the invariant equation

$$|\hat{A} + \alpha\hat{B}| = 0$$

The three solutions α_i are the eigenvalues of the matrix $\hat{A}\hat{B}^{-1}$. These eigenvalues have a product of 1 due to the normalization (15). Two independent invariants are thus the sums $\sum_i \alpha_i$ and $\sum_i 1/\alpha_i$.

These invariants can be written as the traces of $\hat{A}\hat{B}^{-1}$ and its inverse:

$$I_{AB} = \text{trace}(\hat{A}\hat{B}^{-1}) \qquad I_{AB}^{\star} = \text{trace}(\hat{B}\hat{A}^{-1}) \tag{16}$$

As a direct tensor-like proof of invariance, I_{AB} is transformed by (6) to

$$\sum_{ijkl}(\hat{A}_{ij}T_{ik}T_{jl})(T_{ki}^{-1}T_{lj}^{-1}\hat{B}_{ji}^{-1}) = \sum_{ij}\hat{A}_{ij}\hat{B}_{ji}^{-1}$$

i.e. the transformation matrices cancel out. In tensor terminology, $\hat{A}$ is a covariant tensor and $\hat{B}^{-1}$ is a contravariant one, so the contraction $\sum_{ij}\hat{A}_{ij}(\hat{B}^{-1})^{ij}$ is a scalar.

A nice geometric interpretation of these invariants is given in [Mundy *et al.* 1992a]. For any point, one can find a polar line w.r.t. a given conic (Fig. 7). Given two conics, it is possible to find a polar point and a corresponding polar line that are shared by both conics (Fig. 10). The polar line coordinates in this case are an eigenvector of AB^{-1}; B^{-1} transforms the line to its polar point and A transforms the point back to the line. The cross ratio of the contact points on this line is a joint invariant.

A conic and two points: There are nine independent parameters so only one absolute invariant exists:

$$\frac{(\mathbf{a}_1 A^{-1} \mathbf{a}_2^t)^2}{(\mathbf{a}_1 A^{-1} \mathbf{a}_1^t)(\mathbf{a}_2 A^{-1} \mathbf{a}_2^t)}$$

The dual expression uses A^{-1}. The unnormalized matrix A is used here because λ_A is eliminated from the numerator and the denominator.

Global Affine Invariants of Forms

The affine transformation has only six parameters to be eliminated. From Section 5 it can be written as

$$\tilde{\mathbf{x}} = T\mathbf{x}, \qquad \text{with} \quad T_{31} = T_{32} = 0$$

Thus, unlike the projective case, we can set $\lambda_{\mathbf{x}} = 1$ for *points* and it does not need to be eliminated. Lines and conics still have $\lambda_{\mathbf{a}}$, λ_A.

The line at infinity, $\mathbf{a}^{\infty} = (0, 0, \lambda)$, remains invariant under this transformation:

$$\mathbf{a}^{\infty} = \mathbf{a}^{\infty} \lambda T^{-1}$$

Thus this line can be added to the configuration at hand to form invariants in addition to the projective ones. It can be shown [Turnbull 1928] that *all* the affine invariants can be obtained from projective invariants of the given shapes plus this line. Thus the projective methods are useful here too.

Three points with six coefficients yield one relative invariant, the area of the triangle formed by the points:

$$S_{123} = \frac{1}{2} |\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3| \tag{17}$$

transforming as $\tilde{I} = |T|I$. As $\lambda_{\mathbf{x}} = 1$, the degree is 0. In fact, the area of any shape is a relative affine invariant, and the ratio of any two areas is an absolute invariant.

Three *collinear* points yield one absolute invariant, the ratio of lengths l_{12}/l_{23}, which eliminates $|T|$. In the Euclidean case, we have $|T| = 1$ so it does not need to be eliminated. Thus any distance l_{12} between two points is a Euclidean invariant.

Four points are interesting because they show the possibility of *affine coordinates*. We can choose three points $\mathbf{x}_i$ as a basis, and any other one $\mathbf{x}$ can be expressed as a linear combination of the basis vectors (with $\mathbf{x}_1$ as an origin):

$$\mathbf{x} - \mathbf{x}_1 = \alpha(\mathbf{x}_2 - \mathbf{x}_1) + \beta(\mathbf{x}_3 - \mathbf{x}_1) \tag{18}$$

This is a linear equation system for α, β. They are invariant as the equation remains invariant to a 2-D transformation. The solution can be written as ratios of determinants which are easily shown to be equal to (13) (by subtracting the first row in (13) from the others). Thus, the affine coordinates are ratios of areas, S_{pij}/S_{123}.

One conic has one affine relative invariant (in addition to the discriminant $|A|$), resulting from the invariance of the infinite line $\mathbf{a}^\infty$

$$I_A = \mathbf{a}^\infty \hat{A}^{-1} \mathbf{a}^\infty = \hat{A}_{11}\hat{A}_{22} - \hat{A}_{12}^2 \tag{19}$$

It is relative because of the factor λ in $\mathbf{a}^\infty$. It is related to the conic area (it vanishes for a parabola).

A conic and a point. With seven coefficients, it has one absolute invariant, the *algebraic distance*

$$d = \mathbf{x}^t \hat{A} \mathbf{x} \tag{20}$$

(Again we normalize $\hat{A} = A/|A|^{1/3}$, eq. (15).)

Two conics with ten coefficients yield four absolute invariants. Two of them are identical to the projective invariants derived earlier, eq. (16). Two more relative invariants are I_A, I_B, eq (19). A joint invariant analogous to these is

$$\hat{A}_{11}\hat{B}_{22} + \hat{A}_{22}\hat{B}_{11} - 2\hat{A}_{12}\hat{B}_{12} \tag{21}$$

The last three relative invariants can form two absolute invariants.

Hybrid Shape Invariants

The determinant method can be used to find invariants of a curve combined with known reference (feature) points. Instead of determinants consisting of three points, we can use determinants of points and curve derivatives. The description here combines the results of [Van Gool *et al.* 1991, 1992] and [Brill *et al.* 1992].

Examples of relative invariants are

$$|\mathbf{x}, \mathbf{x}', \mathbf{x}''|, \quad |\mathbf{x}, \mathbf{x}', \mathbf{x}_1|, \quad |\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2| \tag{22}$$

with $\mathbf{x} = \mathbf{x}(t)$ being a curve point and $\mathbf{x}_i$ being reference points, either on the curve or not. The prime denotes differentiation w.r.t. t. They all have a weight -1 in T. Under multiplication by λ the first one transforms as

$$|\tilde{\mathbf{x}}, \tilde{\mathbf{x}}', \tilde{\mathbf{x}}''| = |\mathbf{x}, \mathbf{x}', \mathbf{x}''| \begin{vmatrix} \lambda & \lambda' & \lambda'' \\ 0 & \lambda & 2\lambda' \\ 0 & 0 & \lambda \end{vmatrix} = |\mathbf{x}, \mathbf{x}', \mathbf{x}''|\lambda^3$$

i.e. it is of degree 3. Similarly, the other invariants are multiplied by $\lambda\lambda_1^2$, $\lambda\lambda_1\lambda_2$, etc.

Under change of parameter t the first invariant transforms as (with differentiation w.r.t. $\tilde{t}$ denoted by a subscript)

$$|\mathbf{x}, \mathbf{x}', \mathbf{x}''| = |\mathbf{x}, \mathbf{x}_{\tilde{t}}, \mathbf{x}_{\tilde{t}\tilde{t}}| \begin{vmatrix} 1 & 0 & 0 \\ 0 & \tilde{t}' & \tilde{t}'' \\ 0 & 0 & (\tilde{t}')^2 \end{vmatrix} = |\mathbf{x}, \mathbf{x}_{\tilde{t}}, \mathbf{x}_{\tilde{t}\tilde{t}}|(\tilde{t}')^3$$

so it is of weight 3 in $\tilde{t}'$. The others are of weight 1, etc. In short, the degree in the λ_is is equal to the number of the corresponding $\mathbf{x}$s, and the weight in $\tilde{t}'$ is the number of differentiations.

Thus, to obtain absolute invariants, we have to find ratios of the relative ones in which all these factors cancel out. For that to happen, the total number of differentiations in the numerator and denominator has to be equal, and similarly for the number of times a particular point $\mathbf{x}$ or $\mathbf{x}_i$ appears.

It is useful to eliminate the $|T|$ and the degrees first and obtain relative invariants of weight 1 in $\tilde{t}'$. Given such an invariant, we can use it to define an invariant arc-length:

$$\tau = \int_0^t \text{abs}(I^1) dt$$

where I^1 is any invariant of weight 1 in $\tilde{t}'$, i.e. it transforms to $\tilde{I}^1 dt/d\tilde{t}$. We have

$$I_1^1 = \frac{|\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2||\mathbf{x}, \mathbf{x}', \mathbf{x}''|}{|\mathbf{x}, \mathbf{x}', \mathbf{x}_1||\mathbf{x}, \mathbf{x}', \mathbf{x}_2|} \quad I_2^1 = \frac{|\mathbf{x}, \mathbf{x}', \mathbf{x}_1||\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3|}{|\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2||\mathbf{x}, \mathbf{x}_1, \mathbf{x}_3|}$$

The first invariant above needs two reference points and second derivative, while the second invariant needs only first derivatives but three reference points. Their ratio is an absolute invariant. To find more invariants we have to allow more than one point to be on the curve. For example, with two curve points $\mathbf{x}_a, \mathbf{x}_b$ we have an absolute invariant

$$I_{ab} = \frac{|\mathbf{x}_a, \mathbf{x}_b, \mathbf{x}_b'|}{|\mathbf{x}_b, \mathbf{x}_a, \mathbf{x}_a'|} \left(\frac{|\mathbf{x}_a, \mathbf{x}_a', \mathbf{x}_a''|}{|\mathbf{x}_b, \mathbf{x}_b', \mathbf{x}_b''|} \right)^{1/3}$$

Other expressions can be found in [Brill *et al.* 1992].

These invariants can be interpreted in non-homogeneous coordinates. Fixing the third coordinate of $\mathbf{x}$ to 1, the third coordinate of $\mathbf{x}'$ becomes 0. By simple manipulations on determinants, the expressions (22) become 2×2 determinants in Cartesian coordinates:

$$|\mathbf{x}', \mathbf{x}''|, \quad |\mathbf{x} - \mathbf{x}_1, \mathbf{x}'|, \quad |\mathbf{x} - \mathbf{x}_1, \mathbf{x} - \mathbf{x}_2|$$

and the two relative invariants take the form of [Van Gool *et al.* 1991]:

$$I_1^1 = \frac{|\mathbf{x}', \ \mathbf{x}''||\mathbf{x} - \mathbf{x}_1, \ \mathbf{x} - \mathbf{x}_2|}{|\mathbf{x} - \mathbf{x}_1, \ \mathbf{x}'||\mathbf{x} - \mathbf{x}_2, \ \mathbf{x}'|}$$

$$I_2^1 = \frac{|\mathbf{x} - \mathbf{x}_1, \ \mathbf{x}'||\mathbf{x}_1 - \mathbf{x}_2, \ \mathbf{x}_1'|}{|\mathbf{x} - \mathbf{x}_1, \ \mathbf{x} - \mathbf{x}_2||\mathbf{x} - \mathbf{x}_1, \ \mathbf{x}_1'|}$$

Like the cross ratios, these expressions can be expressed in metrical terms. If t is the arc-length, then $|\mathbf{x}', \mathbf{x}''|$ above is nothing but the Euclidean curvature κ. The absolute invariant I_{ab} takes the form

$$I_{ab} = \frac{d_{ab}}{d_{ba}} \left(\frac{\kappa_a}{\kappa_b} \right)^{1/3}$$

where d_{ab} is the distance of point a from the tangent at point b.

Local Affine Invariants

Here we obtain pure local invariants by using the determinants of the derivative vectors $\mathbf{x}^{(n)}$ at a curve point only. From the transformation rules of the last subsection, the 2-D determinant $|\mathbf{x}', \mathbf{x}''|$ is a relative invariant of weight 3. The degree is 0 since $\lambda_{\mathbf{x}} = 1$ in the affine case. It can thus be used to define the *affine arc-length* [Guggenheimer 1963]

$$\tau = \int_{t_0}^{t} |\mathbf{x}', \mathbf{x}''|^{1/3} dt$$

It is absolute with respect to $\tilde{t}'$ but relative with respect to T. We will now use it as an invariant parameter for all our differentiation, denoting derivatives by $\mathbf{x}_\tau$. Higher order invariants can now be obtained either by differentiation or directly as determinants [Bruckstein *et al.* 1991]. We obtain the *affine curvature:*

$$\kappa_{af} = |\mathbf{x}_{\tau\tau}, \mathbf{x}_{\tau\tau\tau}|$$

Fig. 3 shows the affine arc-length and curvature of the curves in Figs. 1, 2. The weight in $|T|$ can be eliminated using higher order invariants.

The affine curvature is constant for conics and only conics. The conic area (a relative invariant) is $\pi\kappa_{af}^{3/2}$.

Local Euclidean Invariants

In this case $|T| = 1$ so the corresponding weight is 0 and the expression $|\mathbf{x}', \mathbf{x}''|$ now has weight only with respect to the parameter change $\tilde{t}'$. We have a new relative invariant, the length of the vector $\mathbf{x}'$, which is preserved because T is orthonormal. We can thus obtain an absolute invariant, the *Euclidean curvature*

$$\kappa = \frac{|\mathbf{x}', \mathbf{x}''|}{(\mathbf{x}'^t \mathbf{x}')^{3/2}} = \frac{x'y'' - x''y'}{((x')^2 + (y')^2)^{3/2}}$$

Choosing the parameter as the arc-length

$$\tau = \int (\mathbf{x}'^t \mathbf{x}')^{1/2} dt$$

the denominator in the curvature becomes 1.

The Symbolic Method

The symbolic method extends the determinants approach to any set of homogeneous polynomials, or algebraic forms of degree n

$$f(x_1, x_2, x_3) = \sum_{i+j+k=n} a_{ijk} x_1^i x_2^j x_3^k = 0 \tag{23}$$

They are of interest because they preserve their form under projectivities. A general method for deriving such invariants is the "symbolic" method [Grace and Young 1903, Turnbull 1928]. In this method the algebraic form f is factored formally into a power of one linear form:

$$f(x_1, x_2, x_3) = (\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3)^n$$

Of course, the factorization cannot be done if α_i are numbers, since f contains many more coefficients than three. However, it can be done with the α_i being abstract entities ("symbols") satisfying certain rules of multiplication.

The "fundamental" theorem for form invariants can now be stated in terms of these symbols as

Theorem 3. *Every invariant of a set of algebraic forms can be expressed by determinants and dot products of the symbols. All invariants can be derived from a finite number of basic ones.*

The method can be applied straightforwardly to curves of any degree. However, its complexity escalates sharply with higher degree. For the cubic, the two invariants S, T take up nearly two printed pages in [Salmon 1879]. Thus, while the symbolic method is useful for deriving general theorems, other methods are more practical.

10. Shape Extraction

To apply invariants we need to extract curves from the data, usually a noisy set of pixels. The problems arising from this are common in vision, e.g. stability and robustness. Here we concentrate on the invariant aspects of the problem.

To obtain useful shape descriptors from the raw data one has to make some assumptions about the shape and/or the noise. The shape undergoes a projective transformation but the noise does not, and this can influence our fitting strategy.

In obtaining global invariants, one fits a form such as a conic to a general shape. The noise is assumed small and the main deviation from the fit is due to the geometry of the shape. Thus here the fit has to be invariant. In local methods the main deviation from the fit is due to noise so invariant fit is less important. The problem here is to increase reliability.

Invariant Fitting

Much of the work in this area has been concerned with fitting conics to general closed curves. The method of [Bookstein 1979] and [Forsyth *et al.* 1990, 1991] uses minimization of the algebraic distance. This distance of a point $\mathbf{x}_i$ from a conic A, namely $d_i = \mathbf{x}_i^t \hat{A} \mathbf{x}_i$, has been shown (eq. (20)) to be an affine invariant. It is not necessarily positive and one wants to minimize the average square distance of n points

$$\frac{1}{n} \sum_i d_i^2 = \frac{1}{n} \sum_i (\mathbf{x}_i^t \hat{A} \mathbf{x}_i)^2$$

As before A is normalized so that $|\hat{A}| = 1$. Without normalization the expression could be multiplied by an arbitrary factor λ_A and would not be invariant. (Besides, the obvious minimum would be 0.) The goal is now to find a conic $\hat{A}$ that minimizes this distance subject to the constraint $|\hat{A}| = 1$. Since all normalized conics have an invariant algebraic distance from the data, the minimal distance and conic are also invariant.

The method was used for images containing collections of industrial objects in [Forsyth *et al.* 1990, 1991]. Conics were fitted to several objects in the image (Fig. 9). To solve this non-linear constrained minimization, they used Lagrange multipliers in an iterative method. The joint invariants (16) of pairs of conics were computed and indexed. Repeating the process from a different viewpoint, the same invariants appeared and could be used to identify the objects by looking at the index tables. Now search was needed for the identification.

The non-linear optimization does not pose a problem if the objects are conic or close to it, but for general shapes it can become more complicated. The optimization problem was studied analytically for simple objects by [Kapur and Mundy 1992]. It was shown that in most cases studied the best fitted conic was unique, but for certain "dumbbell" shapes there were two or three conics that fitted equally well (Fig. 11).

Another approach to conic descriptors is due to Carlsson [1992]. Given a closed shape, we can attempt to inscribe inside it an ellipse having five contact (tangency) points with the curve. In general, this may not be possible (we can always fit a conic to five tangents but it may not be contained in the shape.) We can settle for a one-parameter family of inscribed conics having only four contact points. The parameter can be chosen invariantly. This can be shown as follows. Fig. 12a shows a quadrilateral in which an ellipse is inscribed. The sides $\mathbf{a}_i$ have to satisfy the equation of the line conic $\mathbf{a}_i A^{-1} \mathbf{a}_i = 0$. This happens if the conic matrix has the form

$$A^{-1} = q_1 (\mathbf{a}_1 \times \mathbf{a}_2)(\mathbf{a}_3 \times \mathbf{a}_4)^t + q_2 (\mathbf{a}_1 \times \mathbf{a}_3)(\mathbf{a}_2 \times \mathbf{a}_4)^t$$

This is a family with the parameter q_1/q_2. (It can be symmetrized by $A^{-1} + (A^{-1})^t$.) Since the cross product in homogeneous coordinates is the intersection point $\mathbf{x}_{ij}$ of the sides $\mathbf{a}_i, \mathbf{a}_j$ the conic can be written as

$$A^{-1} = q_1 \mathbf{x}_{12}^t \mathbf{x}_{34} + q_2 \mathbf{x}_{13}^t \mathbf{x}_{24}$$

The expression is invariant since T and λ factor out under transformation. Therefore q_1/q_2 is invariant. In Fig. 12b a 36-side polygon was segmented. All possible quadrilaterals with inscribed ellipses were examined, and for each of them the family member with $q_1/q_2 = 1$ was selected. Three of the four resulting ellipses are seen as meaningful.

The canonical method described earlier offers a way of obtaining a local invariant fit. We want the distance to be minimal in the canonical system (previously the minimization was done in the given system). This will make the fit invariant. We proceed iteratively as follows. Starting with a non-invariant least squares fit, we obtain a curve and a canonical system corresponding to it. We thus make some progress towards the final canonical system. We transform all data points to our new system, repeat the fitting and canonization, and continue until convergence. This method has yet to be tested.

Local Curve Extraction

The local methods do not rely on invariant fit but they face the problem of high order derivatives or fitting of high order implicit curves. It is of interest to examine here what kinds of assumptions are used in the different methods.

Both the implicit and parametrized method need at least nine points to obtain two projective curve invariants. The difference shows up when fitting is done to a larger number of points. In the implicit method, the assumption is that a distance roughly *perpendicular* to the shape is minimal. In the explicit method, the minimized functions are $x(t), y(t)$, measuring distances parallel to the x, y axes. These distances are very inaccurate when the curves are close to parallel to the axes, and can introduce substantial errors. We also have to obtain two fitted functions $x(t), y(t)$ rather than one. Thus an implicit fit seems more natural. It eliminates the parameter before it enters the invariant expressions and adds to an accumulation of errors. In addition, the explicit method assumes the existence of some ordering among the data points so that a parameter can be assigned to them, which is not always the case.

The problem of high order derivatives of the explicit method was analyzed in [Weiss 1991] and it was shown that for a polynomial curve it is possible to obtain accurate derivatives if the window size is wide enough and the filter is of high order. Instead of the Gaussian $g(x)$ we used order l filters of the form

$$F_l = \sum_{0}^{l} (H(x)_i) g(x)$$

with H_i being Hermite polynomials which are orthogonal with respect to the Gaussian weight function. Finite, discrete versions are described in [Meer and Weiss 1992a].

11. 3-D Shapes

3-D shapes can undergo Euclidean motions, and it is useful to represent these objects using 3-D Euclidean invariants. This simplifies indexing and recognition. It is possible that affine transformations are also useful; if a 3-D object is projected on a screen, and the screen is viewed obliquely, one gets the impression that the object is distorted affinely. Mathematically, affine transformations are easier to handle than Euclidean because of their linearity, even though they are more than one needs.

The main problem here is how to recover the 3-D invariants from 2-D views. It is well known [Burns *et al.* 1990] that this cannot be done without some external, or "model-based", assumption, namely prior information about the shape of the object. This is easy to see if we have a 1-D view of a 2-D point set. Each point in the image can have any depth, i.e. it can be located anywhere on the line that passes through the image point and the projection center (Fig. 13). Looked at from a different viewpoint, the points can thus be projected to arbitrary locations in the second image.

However, given some information about the model, we can recover some of its characteristics using invariants, as the examples below show. Given the full model, the pose can be recovered from one view.

Recognition from One View

Single view recognition is perhaps the holy grail of vision and the original motivation for invariants. We describes examples in which some invariant prior knowledge is combined with information from the image, to obtain invariant indexing functions for recognition.

Zisserman *et al.* [1992] have derived 3-D invariants for surfaces of revolution (and their 3-D projective equivalents) from studying 2-D contour invariants. In particular, tangents of several kinds, such as bi-tangents (touching the contour in two points), are useful because tangency is invariant. These tangents can easily be detected on the image (Fig. 14).

In perspective projection of a 3-D object onto the image, the plane that passes through the projection center touches the surface of revolution at two points. (The three points determine the plane.) The line that passes through these two contact points is a bi-tangent to the surface. This bi-tangent intersects the axis of symmetry, because of the symmetry of the situation. The bi-tangent of the object projects into the image as the bi-tangent of the contour. The contour is not symmetric in the image, but we can find features indicating symmetry. Fig. 14 shows that one can match bi-tangents in the image corresponding to symmetric bi-tangents in the object. Their intersection point in the image is a projection of the corresponding intersection in 3-D. Since the 3-D intersections lie on the symmetry axis, their 2-D projections lie on the image projection of this axis.

Since we have a projection of collinear points, their cross ratio is invariant. We can measure it on the image and use it for indexing of the 3-D object. The two lamp shades of Fig. 14 were clearly distinguished by this method. The method will work for objects that are projectively equivalent (in 3-D) to a surface of revolution, such as objects with an elliptical cross section.

In another example, [Hopcroft *et al.* 1992] have used a model consisting of three orthogonal vectors of arbitrary length in 3-D. These vectors $\mathbf{X}_i = (X_i, Y_i.Z_i)$ satisfy the orthogonality conditions

$$\mathbf{X}_i^t \mathbf{X}_j = 0, \qquad i, j = 1, 2, 3$$

These relations are unaffected by the vectors' lengths or by 3-D Euclidean motions. Under orthographic projection, with the measured image coordinates x_i, y_j, we have $X_i = x_i, Y_i = y_i$. We want to find the missing depths Z_i. From the above orthogonality relations we have

$$Z_i Z_j = -(x_i x_j + y_i y_j)$$

from which Z_i can be found, e.g. $Z_1^2 = (Z_1 Z_2)(Z_1 Z_3)/(Z_2 Z_3)$.

Reconstruction from Multiple Views

Multiple views are of help in object reconstruction provided we have the correspondence. The correspondence cannot be inferred from projective geometry, and again we need model based knowledge. Thus, for the purpose of obtaining projective invariants we assume that the correspondence is given. In principle, reconstruction can be handled without invariants by simple triangulation. However, we are not really interested in the 3-D coordinates of the object's points but in its 3-D invariants. We will see that these can be recovered directly from the images. This has been done for point sets by Koenderink and Van Doorn [1991], Barrett *et al.* [1992] and others, and for curves by Brill *et al.* [1992].

For point sets, we can choose a basis of four points $\mathbf{X}_i$ and use them to define 3-D affine coordinates of any other point $\mathbf{X}$ (see eq. (18)):

$$\mathbf{X} - \mathbf{X}_1 = \sum_{i \neq 1} \alpha_i (\mathbf{X}_i - \mathbf{X}_1) \tag{24}$$

with $\mathbf{X}$ being 3-D "world" vectors. Obviously the three coordinates α_i are preserved under a linear transformation in 3-D so they are 3-D invariants.

It turns out that the 3-D invariants α_i can be recovered directly from two 2-D images obtained by an "affine camera" [Mundy and Zisserman 1992], i.e. a transformation with a linear 3×2 matrix T^*:

$$\mathbf{x} = T^* \mathbf{X} + \mathbf{t}$$

with $\mathbf{x}$ being the 2-D image coordinates. Applying this transformation to (24) eliminates the translation $\mathbf{t}$ and yields

$$\mathbf{x} - \mathbf{x}_1 = \sum_{i \neq 1} \alpha_i (\mathbf{x}_i - \mathbf{x}_1)$$

We thus have two equations for the three unknowns α_i. A second view adds two more equations, so the three invariants can be recovered.

For 3-D curves, one can consider the configuration in which the two images are in the same plane, and set apart from each other only by a horizontal distance. In viewing the 3-D point (X, Y, Z) we obtain the same y in both images, and x_l, x_r in the left and right images respectively. With both cameras using perspective projection, the projection can be represented by a transformation of quadruples of coordinates which is linear except for a factor $1/(x_l - x_r)$ [Brill *et al.* 1992]:

$$(X, Y, Z, 1)^t = \frac{1}{x_l - x_r} T^*(x_l, x_r, y, 1)^t$$

with T^* being a 4×4 matrix. Thus the transformation is similar to a 3-D projectivity, being an analog of eq. (5), and again the non-linear factor can be replaced by an arbitrary λ. Thus, many invariants of the above transformation can be obtained by methods described earlier for the projective space.

Since the above transformation contains the camera parameters, its invariants are invariant to the camera calibration. Thus the difficult calibration task becomes unnecessary. It has been shown that seven corresponding points are sufficient to obtain invariants to camera calibration. The invariants are also unchanged under 3-D projective or affine transformation, so again we have recovered 3-D invariants directly from the 2-D images.

12. Conclusion

We have seen that invariance is a very powerful tool for object recognition. It overcomes some major outstanding problems such as the need to find the correct point of view or other distortion factors. We have surveyed many of the mathematical methods involved. We have seen that the geometrical aspect of object recognition can be solved in 2-D by invariants alone. The problem of recovering a 3-D object from a 2-D image cannot be solved by geometry alone—we also need information about the object; but here too invariants are of significant help when combined with model-based knowledge. Future work will be done along several lines: 1) Developing a better fusion between invariants and model-based knowledge, for 3-D reconstruction. 2) Using robust estimation methods for more reliable extraction of the invariants 3) Developing invariants for more general transformations such as deformations. Research in these areas is just beginning and major discoveries may still be ahead of us.

References

Abhyankar, S.S. [1990], *Algebraic Geometry for Scientists and Engineers,* American Mathematical Society, Providence.

Arbter, K., Snyder, W.E., Burkhardt, H., and Hirzinger, G. [1990], "Applications of affine-invariant Fourier descriptors to recognition of 3-D objects", *IEEE Trans. PAMI*-**12**, 640–647.

Ballard, D., and Brown, C.M. [1982], *Computer Vision*, Prentice Hall, Englewood Cliffs, NJ.

Barrett, E., Payton, P., Haag, N., and Brill, M. [1991], "General methods for determining projective invariants in imagery", *CVGIP:IU*, **53**, 45–65.

Barrett, E.B., Brill, E.H., Haag., N.N., and Payton, P.M. [1992], "Invariant Linear Methods in Photogrammetry and Model-Matching", in *Geometric Invariance in Machine Vision*, eds. J.L. Mundy and A. Zisserman, MIT Press, Cambridge, MA.

Besl, P.J., and Jain, R.C. [1985], "Three-dimensional object recognition", *ACM Computing Surveys*, **17**, 75–145.

Bhaskaracharya [1150], *Beejaganit*, Ujjain.

Binford, T.O. [1981], "Inferring surfaces from images", *Artificial Intelligence*, **17**, 205–244.

Bookstein, F.L. [1979], "Fitting conic sections to scattered data", *CGIP*, **9**, 56–71.

Brill, M.H., Barrett, E.B., and Payton, P.M. [1992], "Projective Invariants in Two and Three Dimensions", in *Geometric Invariance in Machine Vision*, eds. J.L. Mundy and A. Zisserman, MIT Press, Cambridge, MA.

Brown, C.M. [1991], "Numerical evaluation of differential and semi-differential invariants", TR-393, University of Rochester Computer Science Department.

Bruckstein, A., Holt, J., Netravali, A.N., and Richardson, T.J. [1991], "Invariant Signatures for Planar Shape Recognition under Partial Occlusion", AT&T TR, October 1991.

Burns, J.B., Weiss, R., and Riseman, E.M. [1990], "View variation of point set and line segment features", *Proc. DARPA IU Workshop*, 650–659.

Carlsson, S. [1992], "Projective Invariant Decomposition of Planar Shapes", in *Geometric Invariance in Machine Vision*, eds. J.L. Mundy and A. Zisserman, MIT Press, Cambridge, MA.

Chang, S., Davis, L.S., Dunn, S.M., Eklundh, J.-O., and Rosenfeld, A. [1987], "Texture discrimination by projective invariants," *Pattern Recognition Letters*, **5**, 337–342.

Cyganski, D. and Orr, J. [1985], "Applications of tensor theory to object recognition and orientation determination", *IEEE Trans. PAMI*-**7**, 662–673.

Duda, R.O., and Hart, P.E. [1973], *Pattern Recognition and Scene Analysis*, Wiley, New York.

Faugeras, O.D., and Papadopoulo, T. [1992], "Disambiguating Stereo Matches with Spatio-Temporal Surfaces", in *Geometric Invariance in Machine Vision*, eds. J.L. Mundy and A. Zisserman, MIT Press, Cambridge, MA.

Forsyth, D., Mundy, J.L., Zisserman, A., and Brown, C.M. [1990], "Projectively invariant representations using implicit algebraic curves", *Image and Vision Computing*, **8**, 130–136.

Forsyth, D., Mundy, J.L., Zisserman, A., Coelho, C., Heller, A., and Rothwell, C. [1991], "Invariant descriptors for 3-D object recognition and pose", *IEEE Trans. PAMI*-**13**, 971–991.

Fubini and Čech [1927], *Geometria Proiettiva Differenziale*, Zanichelli, Bologna.

Gordan, P. [1885], *Vorlesungen Über Invariantentheorie*, Leipzig.

Grimson, W.E.L., and Lozano-Pérez, T. [1987], "Localizing overlapping parts by searching the interpretation tree", *IEEE Trans. PAMI*-**9**, 469–482.

Grace, J.H., and Young, A. [1903], *The Algebra of Invariants*, Chelsea, New York.

Guggenheimer, H. [1963] *Differential Geometry*, Dover, New York.

Halphen [1880], "Sur les invariants différentiels des courbes gauches", *J. Ec. Polyt.*, **28**, 1.

Hopcroft, J.P., Huttenlocher, D.P., and Wayner, P.C. [1992], "Affine Invariants for Model-Based Recognition", in *Geometric Invariance in Machine Vision*, eds. J.L. Mundy and A. Zisserman, MIT Press, Cambridge, MA.

Hilbert, D. [1893], "Über die vollen Invariantensysteme", Mathematische Annalen, **42**, 313–373.

Kanatani, K. [1990], *Group Theoretical Methods in Image Understanding*, Springer, Berlin.

Kapur, D., and Mundy, J.L. [1991], "Fitting affine invariants to curves", in *Geometric Invariance in Machine Vision*, eds. J.L. Mundy and A. Zisserman, MIT Press, Cambridge, MA.

Klein, F. [1926], *Entwicklung der Mathematik*, Berlin.

Koenderink, J.J., and Van Doorn, A.J. [1991], "Affine invariants from motion", *J. Opt. Soc. Am.* A, 377–385.

Kriegman, D.J., and Ponce, J. [1990], "On recognizing and positioning of curved 3-D objects from image contours", *IEEE Trans. PAMI*-**12**, 1127–1137.

Lamdan, Y., Schwartz, J.T., and Wolfson, H.J. [1988], "Object recognition by affine invariant matching", *Proc. CVPR*, 335–344.

Lagrange, J.L. [1773], Berlin Memoires, p. 265.

Lane, E.P. [1932], *Projective Differential Geometry of Curves and Surfaces*, University of Chicago Press.

Lane, E.P. [1942], *A Treatise on Projective Differential Geometry*, University of Chicago Press.

Maybank, S.J. [1992], "The Projection of Two Non-coplanar Conics", in *Geometric Invariance in Machine Vision*, eds. J.L. Mundy and A. Zisserman, MIT Press, Cambridge, MA.

Lowe, D. [1985], *Perceptual Organization and Visual Recognition*, Kluwer, Boston.

Meer, P., and Weiss, I. [1992], "Smoothed differentiation filters for images", *J. Visual Communication and Image Representation*, **3**, 58–72.

Meer, P., and Weiss, I. [1992a], "Point/line correspondence under 2D projective transformation", Proc. *CVPR*, 115–121.

Mumford, D. [1965], *Geometric Invariant Theory*, Springer, New York.

Mundy, J.L., and Zisserman, A. [1992], "Introduction—Towards a new framework for vision", in *Geometric Invariance in Machine Vision*, eds. J.L. Mundy and A. Zisserman, MIT Press, Cambridge, MA.

Mundy, J.L., Kapur, D., Maybank, S.J., and Quan, L. [1992a] "Geometric Interpretation of Joint Conic Invariants", in *Geometric Invariance in Machine Vision*, eds. J.L. Mundy and A. Zisserman, MIT Press, Cambridge, MA.

Nagata, M. [1963], "Complete reducibility of rational representations of a matric group", *J. Math. Kyoto Univ.*, **3**, 369–377.

Nielsen, L., and Sparr, G. [1991], "Projective area invariants as an extension of the cross ratio", *CVGIP:IU*, **54**, 145–159.

Olver, P.J. [1986], *Application of Lie Groups to Differential Equations*, Springer, New York.

Park, K., and Hall, E. [1987], "Form recognition using moment invariants for three dimensional perspective transformation", *Proc. SPIE Vol. 726, Intelligent Robots and Computer Vision,* 90–108.

Pizlo, Z., and Rosenfeld, A. [1991], "Recognition of Planar Shapes from Perspective Images Using Contour-based Invariants", TR-528, Center for Automation Research, University of Maryland.

Salmon, G. [1879], *Higher Plane Curves*, Chelsea, New York.

Springer, C.E. [1964], *Geometry and Analysis of Projective Spaces*, Freeman, San Francisco.

Stevenson, R.L., and Delp, E.J. [1989], "Invariant reconstruction of visual surfaces", *Proc. IEEE Workshop on Interpretation of 3-D Scenes,* 131–137.

Taubin, G., and Cooper, D.B. [1992], "Object Recognition Based on Moment (or Algebraic) Invariants", in *Geometric Invariance in Machine Vision*, eds. J.L. Mundy and A. Zisserman, MIT Press, Cambridge, MA.

Turnbull, H.W. [1928], *Determinants, Matrices and Invariants*, Blackie and Son, Glasgow.

Ullman, S., and Basri, R. [1991], "Recognition by linear combination of models", *IEEE Trans. PAMI*-**13**, 992–1006.

Van Gool, L., Kempenaers, P., and Oosterlinck, A. [1991], "Recognition and semi-differential invariants", *Proc. CVPR*, 454–460.

Van Gool, L., Moons, T., Pauwels, E., and Oosterlinck, A. [1992], "Semi-Differential Invariants", in *Geometric Invariance in Machine Vision*, eds. J.L. Mundy and A. Zisserman, MIT Press, Cambridge, MA.

Wayner, P.C. [1991], "Efficiently using invariant theory for model-based matching", *Proc. CVPR*, 473–478.

Weinshall, D. [1990], "Qualitative depth from stereo, with applications", *CVGIP* **49**, 222–241.

Weiss, I. [1988], "Projective invariants of shapes", *Proc. DARPA Image Understanding Workshop*, 1125–1134.

Weiss, I. [1991], "High Order Differentiation Filters that Work", TR 545, Center for Automation Research, University of Maryland.

Weiss, I. [1992a], "Noise Resistant Invariants of Curves", in *Geometric Invariance in Machine Vision*, eds. J.L. Mundy and A. Zisserman, MIT Press, Cambridge, MA.

Weiss, I. [1992b], "Local Projective and Affine Invariants", TR 612, Center for Automation Research, University of Maryland.

Weyl, H. [1939], *The Classical Groups*, Princeton University Press.

Wilczynski, E.J. [1906], *Projective Differential Geometry of Curves and Ruled Surfaces*, Teubner, Leipzig.

Wilczynski, E.J. [1908], Projective Differential Geometry of Curved Surfaces (Second Memoir), *Amer. Math. Soc. Trans.*, **79**.

Zisserman, A., Forsyth, D.A., Mundy, J.L., and Rothwell, C.A. [1992], "Recognizing General Curved Objects Efficiently", in *Geometric Invariance in Machine Vision*, eds. J.L. Mundy and A. Zisserman, MIT Press, Cambridge, MA.

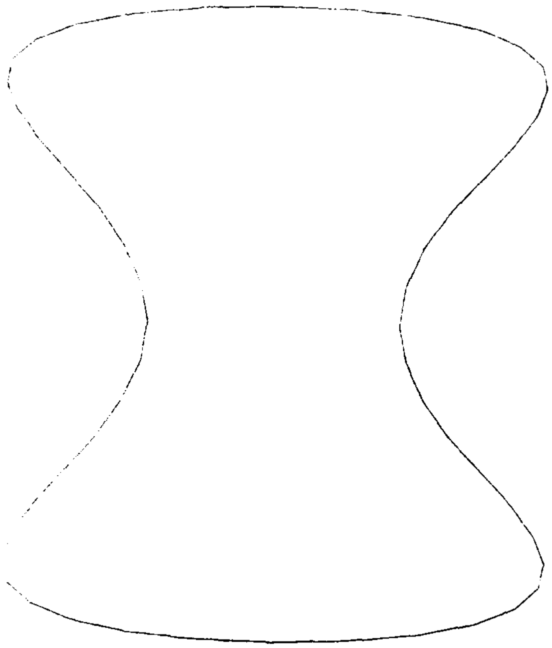


Fig. 1: Projection 1

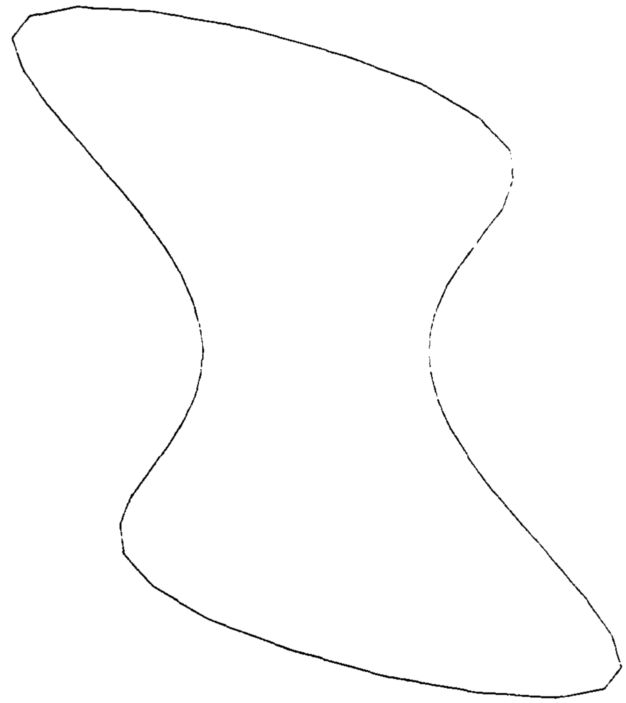


Fig. 2: Projection 2

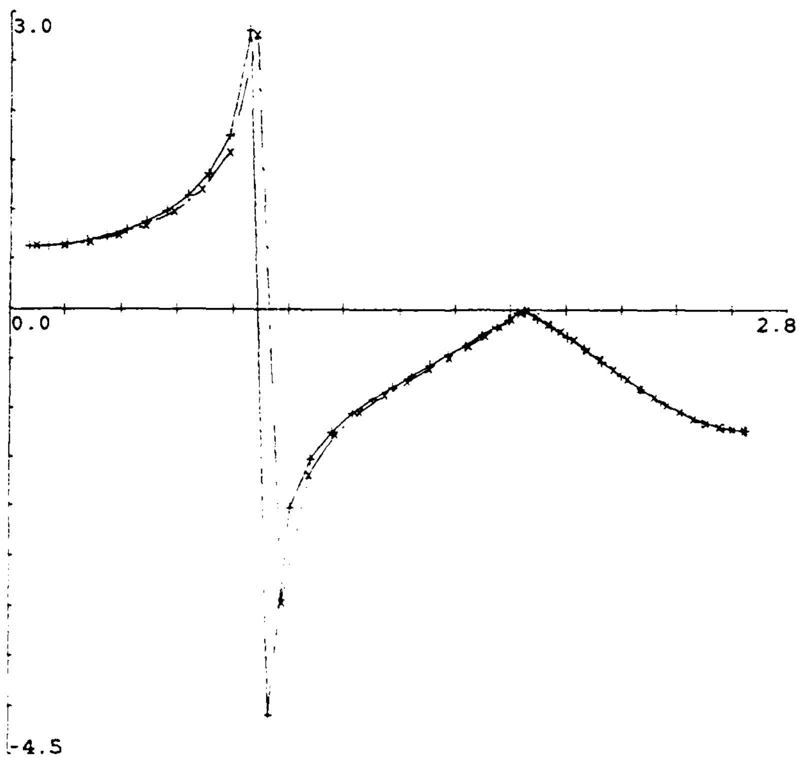


Fig. 3: Invariant signature

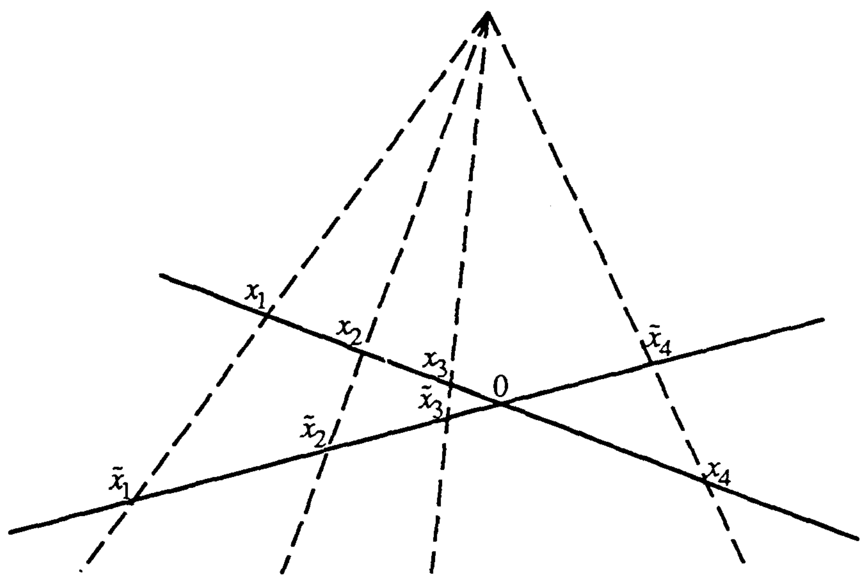
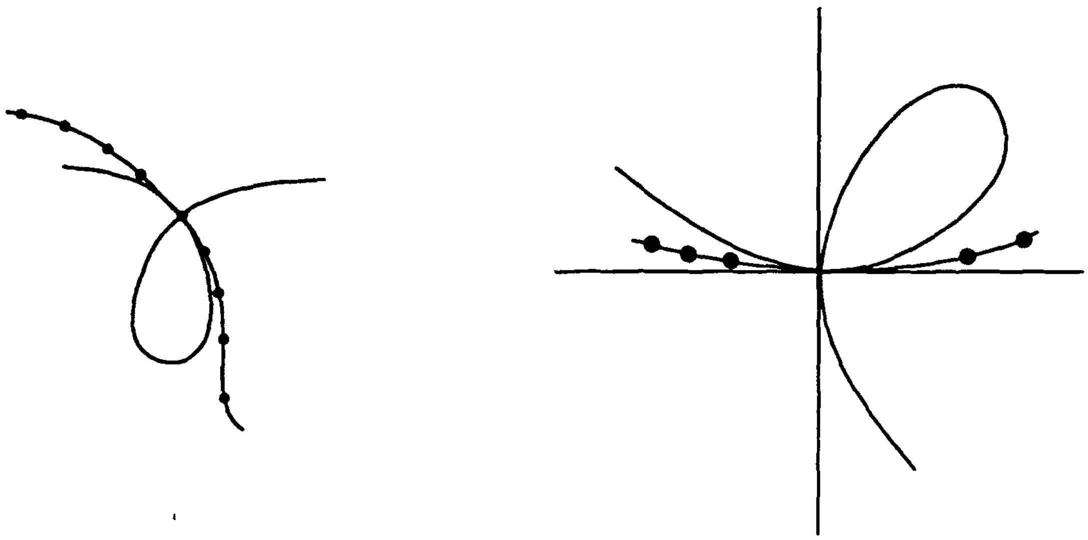


Fig. 4: Perspectivity. Cross ratio of four points is preserved.



(a) (b)

Fig. 5: (a) Osculating nodal cubic. (b) Osculating folium (leaf) of Descartes

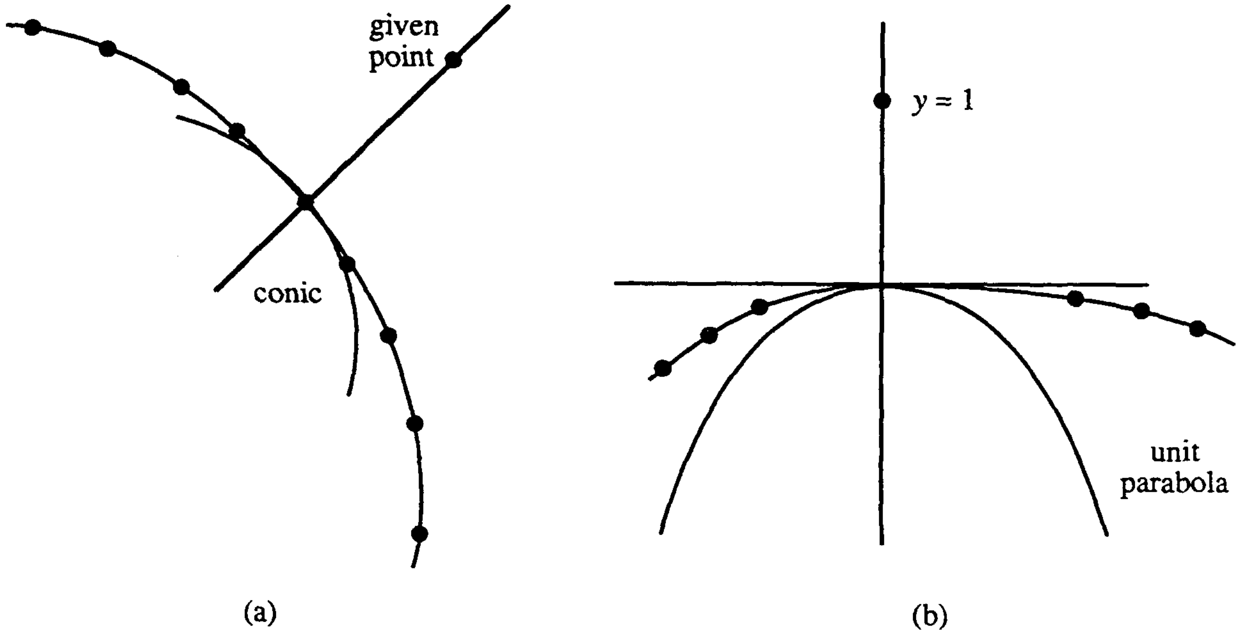


Fig. 6: (a) Osculating conic. (b) Canonical conic and point

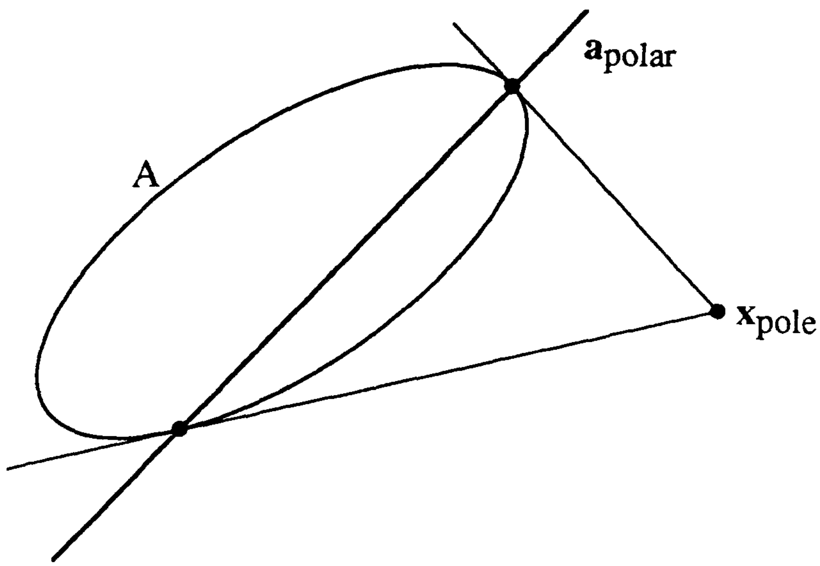
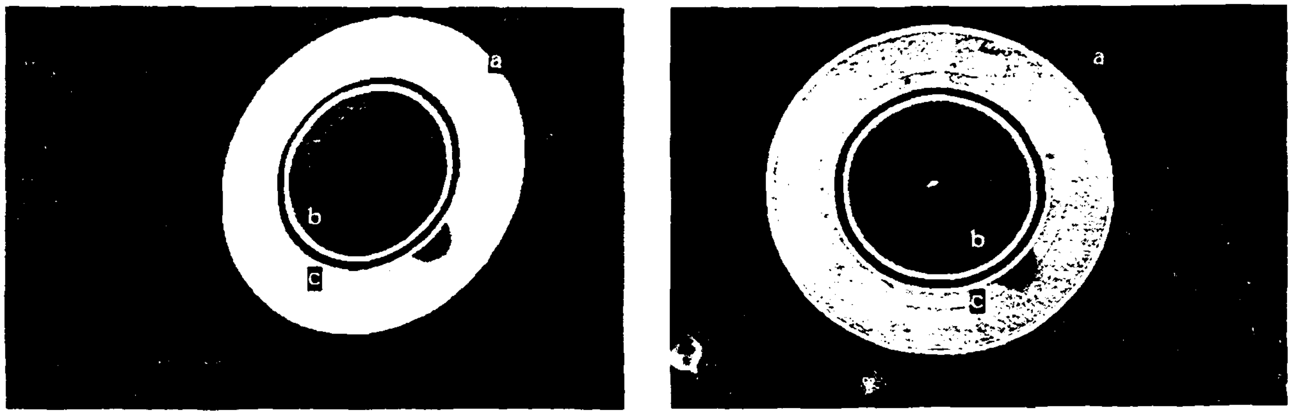


Fig. 7: A point and a line which are polar to each other with respect to a conic



a b

Fig. 8: Images of a computer tape, with two fitted conics in overlay. The data for the conics in these images was obtained by acquiring the image edges using a local implementation of Canny's edge finder, linking edges, and then choosing corresponding curves by hand. In these images, the conics have been drawn three pixels thick to make them visible. These conics were used to obtain the joint scalar invariants [Forsyth *et al.* 1990].

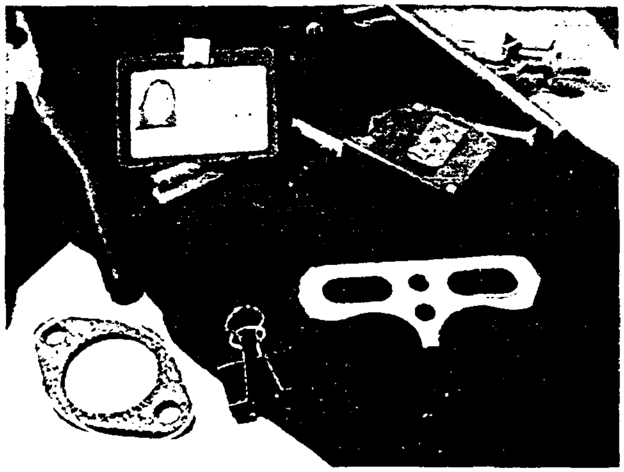


Fig. 9: The joint scalar invariants of a pair of conics can be used to find instances of models in scenes, when the objects involved have plane curves which lie on their surfaces. Here we show an instance of a gasket found in a cluttered scene by fitting conics to all of the curves using projectively invariant fitting techniques, and marking those pairs of conics with the correct joint scalar invariants. The data for the conics in this image was obtained by acquiring the image edges using a local implementation of Canny's edge finder and then linking these edges. Note that the system has ignored the wide range of distracting curves, because they do not have the right joint scalar invariants. The outside curve for this gasket is clearly not a conic, so that this result demonstrates projectively invariant fitting [Forsyth *et al.* 1990].

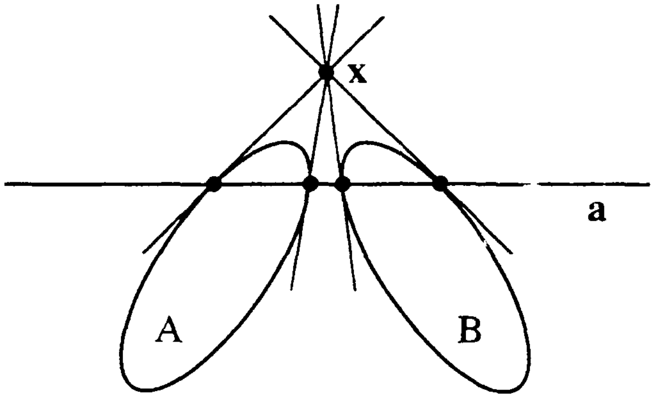


Fig. 10: The geometric configuration in which a polar line (and point) are eigenvectors of AB^{-1} [Mundy *et al.* 1992].

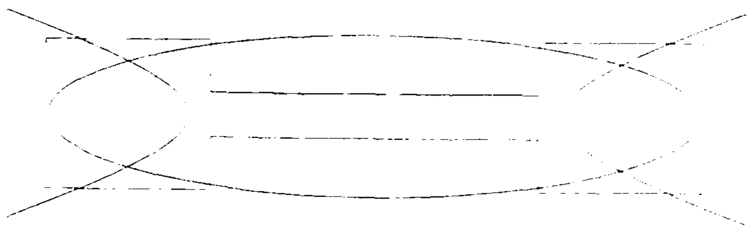


Fig. 11: Dumb-bell and the two conics which are the best fit [Kapur and Mundy 1992].

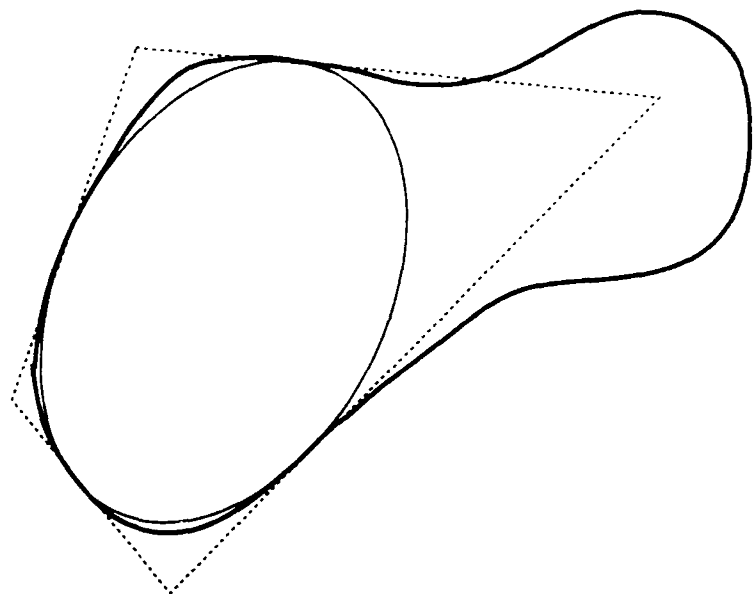


Fig. 12: (a) Four-sided polygon formed by the tangents of the contact points [Carlsson 1992].

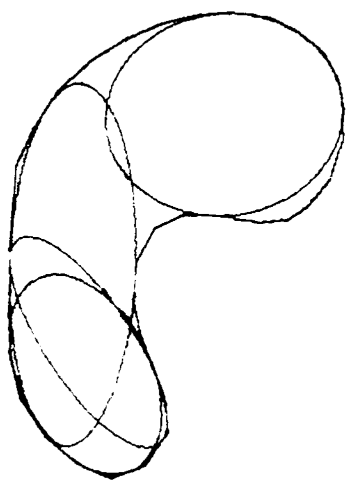


Fig. 12: (b) Inscribed ellipses with ratio $q_1/q_2 = 1$ [Carlsson 1992].

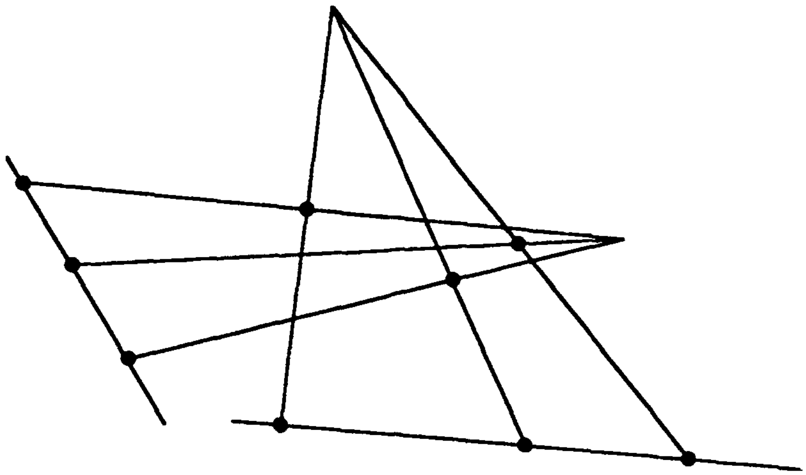


Fig. 13: Two projections from 2-D to 1-D.

Fig. 14: This figure shows two views each of two different lamp-stands. Bitangents, computed by hand from the outlines, are overlaid [Zisserman *et al.* 1992].

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
AUGUST 1992

3. REPORT TYPE AND DATES COVERED
TECHNICAL

4. TITLE AND SUBTITLE
GEOMETRIC INVARIANTS AND OBJECT RECOGNITION

5. FUNDING NUMBERS
DACA76-92-C-0009
N0004-91-J-1222

6. AUTHOR(S)
Isaac Weiss

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Center for Automation Research
University of Maryland
College Park, MD 20742-3275

8. PERFORMING ORGANIZATION REPORT NUMBER
CAR-TR-632
CS-TR-2942

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
DARPA, 3701 N. Fairfax Dr., Arlington, VA 22203-1714
US Army Topographic Engineering Center, Ft. Belvoir, VA
ONR, 800 N. Quincy St., Arlington, VA 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

In this paper we discuss the role of the general invariance concept in object recognition, and review the classical and recent literature on projective invariance. Invariants help solve major problems of object recognition. For instance, different images of the same object often differ from each other, because of the different viewpoint from which they were taken. To match the two images, common methods thus need to find the correct viewpoint, a difficult problem that can involve search in a large parameter space of all possible points of view and/or finding point correspondences. Geometric invariants are shape descriptors, computed from the geometry of the shape, that remain unchanged under geometric transformations such as changing the viewpoint. Thus they can be matched without search. Deformations of objects are another important class of changes for which invariance is useful.

14. SUBJECT TERMS
Invariance, matching, object recognition, perspective, projective

15. NUMBER OF PAGES
46

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR