Invariance - A New Framework for Vision

David Forsyth, Joseph L. Mundy, Andrew Zisserman,
and Christopher M. Brown.

Abstract

We show that curved planar objects have shape descriptors that are unaffected by the position, orientation and intrinsic parameters of the camera.

These shape descriptors can be used to index quickly and efficiently into a large model base of curved planar objects, because their value is independent of pose and unaffected by perspective. Thus, recognition can proceed independent of calculating pose. Object curves are represented using conics, attached with a fitting technique that commutes with projection. This means that the pose of an object can be determined by backprojecting known conics. We show examples of recognition and pose determination using real image data.

1 Introduction

The fundamental problem of computer vision is that shape measured in images depends not only on object shape, but also on the position, orientation and intrinsic parameters of the camera. If it is possible to define shape descriptors that are unaffected by perspective transformations, then image measurements of these descriptors can be matched to object properties regardless of camera viewpoint. Shape descriptors with these properties are known as invariants.

We believe that invariance is the essential property of a shape description.

Many properties are invariant to projection: for example, straight lines project to straight lines and intersections are preserved. The exploitation of these invariants has been responsible for the success of polyhedral model based vision. However, for smooth curves and surfaces, invariants such as zeroes of curvature, the cross ratio and Gaussian curvature do not provide a sufficiently strong set of constraints for successful model based vision. There has been a correspondingly limited success in representing and recognising curved objects. This paper shows one way of constructing and exploiting a rich invariant theory for plane curves.

In section 2, we discuss the mathematics and ideas underlying our use of invariant theory, and show a broad range of examples of invariants.

In section 3, we show the usefulness of this theory in model based vision. We demonstrate a projectively invariant representation for plane curves using conic curves. We build a simple and effective model based vision system which uses the projective invariance of this representation to recognise planar objects. Because the descriptor does not change whatever the pose of the object, this system effectively decouples the problem of identifying objects from that of determining their pose.

In section 4, we show that it is possible to recover pose from our invariant representation by solving the simple problem of backprojecting known conics.

Finally, we discuss the prospects for a wider use of invariant theory in vision.

2 Invariant theory

We adopt the notation that corresponding entities in two different coordinate frames are distinguished by large and small letters. Vectors are written in bold font, e.g. $\mathbf{x}$ and $\mathbf{X}$. Matrices are written in typewriter font, e.g. $\mathtt{c}$ and $\mathtt{C}$.

Given a group $\mathcal{G}$ and a space $\mathcal{M}$, an action of $\mathcal{G}$ on the space associates with each group element $g \in \mathcal{G}$ a map $g : \mathcal{M} \rightarrow \mathcal{M}$:

$$id(x) = x \quad (1)$$

$$(g_1 \times g_2)(x) = (g_1(g_2(x)) \quad (2)$$

where $g_1, g_2 \in \mathcal{G}$, id is the identity element of the group, and $\times$ is the group composition function. An *invariant* of a group action is defined as follows:

> **Definition** An invariant, $I(\mathbf{p})$, of a function $f(\mathbf{x}, \mathbf{p})$ subject to a group, $\mathcal{G}$, of transformations acting on the coordinates $\mathbf{x}$, is transformed according to $I(\mathbf{P}) = I(\mathbf{p})h(g)$. Here $g \in \mathcal{G}$ and $h(g)$ is a function only of the parameters of the transformation and does not depend on the coordinates, $\mathbf{x}$, or on the parameters, $\mathbf{p}$. $I(\mathbf{p})$ is a function only of the parameters, $\mathbf{p}$.

*DAF acknowledges the support of Magdalen College, Oxford. AZ acknowledges the support of the United Kingdom Science and Engineering Research Council. JLM acknowledges the support of the General Electric Coolidge Fellowship. The General Electric Corporate Research and Development Laboratory is supported in part by the following: DARPA contract DACA-76-86-C-007, AFOSR contract F49620-89-C-003. CMB acknowledges the support of the DARPA U.S. Army Engineering Topographic Laboratories Grant DACA76-85-C-0001 and the Air Force Systems Command (RADC, Griffiss AFB, NY) and Air Force OSR Contract F30602-85-C-0008, which supports the Northeast Artificial Intelligence Consortium.

In what follows, we concentrate on scalar invariants and the term invariant should be understood to mean scalar invariant, except where the context clearly indicates otherwise.

Seeing invariants in terms of the action of continuous groups (otherwise known as *Lie* groups) makes it possible to write down a system of linear first order partial differential equations in the invariant, which can then be solved using a symbolic mathematics package [14, 7, 6]. Invariants for a wide range of situations can be constructed using this machinery, which can also be used to predict the minimum number of invariants available. Table 1 shows the minimum number of independent invariants available for different systems of curves under a range of group actions. The mathematical literature contains other computational tricks for constructing invariants, of which the symbolic method appears to have the most potential value [8].

An invariant is defined in the context of a particular transformation. We have concentrated on the case of the plane projective group. This models the situation where a plane curve is subject to rigid motion in space, and projected using perspective. In fact, rigid motion and perspective projection produces a subset of the actions of the plane projective group, and the subtle but important distinctions between the two situations are the subject of active research.

A number of invariants are exploited in vision at present, with varied success. [10] demonstrated the value of using invariants of camera rotation and invariant decompositions in computing such information as optical flow. [18] has also raised the issue of invariant representations.

2.1 Examples of invariants

Example 1 - Plane rotation: The plane rotation group acts on the plane, by the mapping $\mathbf{x} = \mathrm{R}\mathbf{X}$, where R is a 2D rotation matrix. Any function of the distance from the origin to a point is invariant under the action of this group. Under the action of this group combined with the multiplicative group, $(x, y) :\longrightarrow \lambda(x, y)$, the function $x^2 + y^2$ is an invariant of weight 2. In the second case there is no scalar invariant, however.

Example 2 - Plane translation: Any element ϵ of the one dimensional translation group acts on the plane, by the mapping $\{x', y'\} = \{x + \epsilon, y\}$. The y coordinate of any point is invariant under the action of the group.

Example 3 - Homogenous polynomials: Consider the space of homogenous polynomials in n variables, $x_0, ..., x_{n-1}$. Write $\mathbf{x}$ for $\{x_0, ..., x_{n-1}\}$. The general linear group (all matrices T of non-zero determinant) acts on this space, by taking a polynomial $p(\mathbf{x})$ to $P(\mathbf{x}) = p(\mathrm{T}\mathbf{x})$. Here the coefficient of a monomial in $P(\mathbf{x})$ is determined by computing the coefficient of that monomial in the expansion of $p(\mathrm{T}\mathbf{x})$. Clearly, this action takes a polynomial of order k to another of order k, so we can see it as an action on the homogenous polynomials of order k. Furthermore, if we write $\mathbf{p}$ for the coefficients of p, we have for an invariant I, $I(\mathbf{P}) = I(\mathbf{p})|\mathrm{T}|^w$ where T is the transformation matrix and $|\mathrm{T}|$ indicates the determinant of T. Invariants of this action formed a major research topic of 19th century mathematics: an introduction can be found in, for example, [8]. A modern treatment of some of this work is given by [17] or by [3].

Example 4 - Differential invariants: Differential invariants are invariant functions of the position and derivatives of a curve at a point. Differential invariants are clearly important in vision. Curvature, torsion and Gaussian curvature, all differential invariants under Euclidean actions, have been widely applied. For example, a projective differential invariant for plane curves has been known for a long time [18, 11]. However, this invariant is an extremely large and complex polynomial in the derivatives of the curve, and it is not known how useful in practice it will be. Table 2 shows the number of derivatives required for a differential invariant of a plane curve under a range of group actions.

Example 5 - Projective invariants for pairs of plane conics: A plane conic can be written as $\mathbf{x}^t \mathrm{c}_1 \mathbf{x} = 0$, for $\mathbf{x} = (x, y, z)$ and a symmetric matrix c_1, which determines the conic. A pair of coplanar conics has two scalar invariants, which we will describe here. Given conics with matrices of coefficients c_1 and c_2, we define:

$$
\begin{aligned}
I_{\mathrm{c}_1\mathrm{c}_2} &= Trace({\mathrm{c}_1}^{-1}\mathrm{c}_2) \\
I_{\mathrm{c}_2\mathrm{c}_1} &= Trace({\mathrm{c}_2}^{-1}\mathrm{c}_1)
\end{aligned}
$$

Under the action $\mathbf{x} = \mathrm{T}\mathbf{X}$, c_1 and c_2 go to $\mathrm{C}_1 = \mathrm{T}^t\mathrm{c}_1\mathrm{T}$ and $\mathrm{C}_2 = \mathrm{T}^t\mathrm{c}_2\mathrm{T}$. In particular, using the cyclic properties of the trace, we find:

$$
\begin{aligned}
I_{\mathrm{C}_1\mathrm{C}_2} &= Trace(\mathrm{T}^{-1}{\mathrm{c}_1}^{-1}(\mathrm{T}^t)^{-1}\mathrm{T}^t\mathrm{c}_2\mathrm{T}) \\
&= Trace({\mathrm{c}_1}^{-1}\mathrm{c}_2) \\
&= I_{\mathrm{c}_1\mathrm{c}_2}
\end{aligned}
$$

A similar derivation holds for $I_{\mathrm{c}_2\mathrm{c}_1}$. Note that ${\mathrm{c}_1}^{-1}\mathrm{c}_2$ transforms to $\mathrm{T}^{-1}{\mathrm{c}_1}^{-1}\mathrm{c}_2\mathrm{T}$, which is a similarity transformation, and so its eigenvalues are preserved. This provides an alternative demonstration of invariance.

Example 6 - Projective invariants for systems of lines: A system of four coplanar lines that intersect in a single point[1], is dual to a system of four collinear points. This system has the familiar cross ratio as its invariant.

For five general coplanar lines. There are two projective invariants:

$$I_{l1} = (I_{431}I_{521})/(I_{421}I_{531})$$

and

$$I_{l1} = (I_{421}I_{532})/(I_{432}I_{521})$$

The lines are written in homogenous coordinates as $a_i x + b_i y + c_i z = 0$, and I_{ijk} is the determinant of the matrix $\{\mathbf{l}_i, \mathbf{l}_j, \mathbf{l}_k\}$ where $\mathbf{l}_i$ is $\{a_i, b_i, c_i\}^T$ (see [6]). Furthermore, because points and lines are dual in the projective plane, we have immediately that that functions are also invariants for a system of five coplanar points, which are not collinear.

[1] A useful example in vision is a set of parallel lines (which intersect in a single point at infinity).

curve (no. of d.o.f.)	plane Euclidean group (3 d.o.f.)	orthographic projection (5 d.o.f.)	Affine projection (6 d.o.f.)	projective mappings (8 d.o.f)
conic (5)	2	0	0	0
cubic (9)	6	4	3	1
quartic (14)	11	9	8	6
2 coplanar conics (10)	7	5	4	2
five lines (10)	7	5	4	2
a conic and two lines (9)	6	4	3	1

Table 1: The minimum number of functionally independent scalar invariants for plane algebraic curves under a variety of groups important in vision. By "orthographic projection", we mean that the plane on which the curve lies is subject to rigid motions in three space, and then projected onto the image plane using orthography. This case has only 5 degrees of freedom because the image curve is completely unaffected by changes in the distance to the object plane. The details of the argument used to derive these numbers appear in [6].

Example 7 - Projectively invariant measurements: If there is a distinguished conic curve in the plane (say, $\mathtt{c}$), then for two points $\mathbf{x}_1$, $\mathbf{x}_2$ that do not lie on the conic, the function:

$$\frac{(\mathbf{x}_1{}^T \mathtt{c} \mathbf{x}_2)^2}{(\mathbf{x}_1{}^T \mathtt{c} \mathbf{x}_1)(\mathbf{x}_2{}^T \mathtt{c} \mathbf{x}_2)} \tag{3}$$

is independent of the frame in which the points and the conic are measured. In turn, this can be used to define a projectively invariant metric (see [17]).

3 Recognising curved planar objects using algebraic invariants

3.1 Exact curves

Invariants provide a powerful approach to recognising objects because they provide descriptors of shape that are unaffected by object pose. The invariant for four parallel lines, discussed briefly in section 2, example 6, can be used to recognise pallets in images, because pallets possess systems of coplanar, parallel lines. Figures 1 and 2 demonstrate this approach.

Algebraic invariants can also be used to recognise objects that possess planar curves. Given a pair of coplanar conics, their joint projective invariants (section 2, example 5) are a representation that is invariant to Euclidean motion and perspective, i.e., the joint invariants calculated for image curves will have the same values as those calculated from object curves, whatever the object pose and the camera parameters. Since a conic can be represented both by $\mathtt{A}$ and the $k\mathtt{A}$, to evaluate and interpret these invariants we need to make some assumption to set the relative scale of the conic matrices. As a result of the projectively invariant fitting techniques we demonstrate below, all the conics we use will be normalised by the frame-invariant criterion $|\mathtt{A}| = 1$. In practice, these descriptors are stable and have sufficient dynamic range to be useful (see table 3, and figure 3). This invariant description enables building a model based vision system that recognises such objects quickly and efficiently, regardless of their pose with respect to the camera. A simple model based vision system that uses these invariants to recognise labels consisting of a pair of coplanar conics is demonstrated in figure 4.

plane Euclidean group (3 d.o.f.)	orthographic projection (5 d.o.f.)	Affine projection (6 d.o.f.)	projective mappings (8 d.o.f)
2	4	5	7

Table 2: The number of derivatives required for a scalar differential invariant under a variety of groups important in vision. This assumes invariance both to "geometric" actions and reparametrisation. The details of the argument used to derive these results appear in [6].

3.2 Representing general curves by conics

For both of the above systems, the object data lies (up to noise) on an algebraic curve of the right type.

If it is possible represent an arbitrary curve by a conic, the invariants resulting from this representation can be exploited. To achieve this, the representation must have this crucial frame independence property:

> Given an observation of a data set in a transformed frame, the representation computed for this set is exactly the original representation transformed according to the change of frame.

This means that the process of choosing a representation commutes with projection. In fact, it is possible to construct a representation with this property, as the following theorem, proved in [4], shows:

> **Theorem 1** Let $I(\mathbf{p})$ be an invariant of the polynomial form $Q(\mathbf{x}, \mathbf{p})$ under a group of linear transformations $\mathcal{G}$. Assume I is homogenous of degree n, with weight w. Let $< \mathbf{p} >$ be the parameter vector determined by minimizing $\sum_i Q^2(\mathbf{x}_i, \mathbf{p})$ over a set of points, $\mathbf{x}_i$, subject to the constraint, $N(\mathbf{p}) = I(\mathbf{p}) = constant$. If the point set is transformed under $\mathcal{G}$, i.e., $\mathbf{x} = \mathtt{T}_g\mathbf{X}$, let $\mathcal{T}_g$ be the corresponding transformation matrix *for the coefficients* $\mathbf{p}$. The coefficients of the polynomial fitted to the point set in the new frame are given by $< \mathbf{P} >$. Assume that n is odd or that w is even (or both). Under these conditions, we have:
>
> $$<\mathbf{p}>= k_g.\mathcal{T}_g <\mathbf{P}>$$
>
> where k_g is a scalar depending on $g \in \mathcal{G}$.

The curve chosen by this approximation process is effectively decoupled from the frame in which it is observed, and has the desired frame independence property. This is true of the

Conics	First joint invariant	Second joint invariant
Conics a and b from figure 3a	3.419	3.546
Conics a and b from figure 3b	3.418	3.543
Conics a and b from figure 3c	3.414	3.538
Conics a and b from figure 3d	3.407	3.528
Conics b and c from figure 3a	3.022	3.021
Conics b and c from figure 3b	3.023	3.021

Table 3: The joint scalar invariants computed for the indicated pairs of conics for the four different images of a computer tape from different positions and angles, shown in figure 3. Note that the joint scalar invariants for the coplanar conics a, b for the four images are effectively constant. Furthermore, the values of the invariants for different pairs of conics are different.

curve, $\{\mathbf{x} \mid Q(\mathbf{x}, \mathbf{p}) = 0\}$, and not of the *polynomial* $Q(\mathbf{x}, \mathbf{p})$. The theorem applies to algebraic curves of higher degree as well as to conics.

It is not difficult to convince oneself that this theorem must be true for the case where the normalization is a scalar invariant. It is possible to show that a solution to this fitting problem exists. We do not yet know if it is unique or not. The fitting problem presents interesting numerical difficulties: currently, it is solved using the techniques of [4], but we are investigating using the polynomial continuation (see, for example [12]).

3.3 Recognising planar objects

The fitting theorem means that the joint scalar invariants of section 2 can be used to represent curved planar objects, *even if the model does not contain plane conics.* Given an object that has a pair of coplanar curves that will both be visible at the same time, one may compute the joint scalar invariants for the conics fitted to these curves. These invariants form a projectively invariant description of this set of plane curves. It is possible to find instances of a model in an image by fitting conics to every available curve, computing the joint scalar invariants for each pair of conics, and then extracting those pairs of curves that have the appropriate values of joint scalar invariants. A model based vision system built in this way is intrinsically fast, because the invariants can be used to index into a large model base directly - there is no need to search the model base. Several examples appear in figure 5. It is sensitive to occlusion, however ([16] discusses noise issues that appear in fitting conics to small numbers of data points).

This system is different from earlier model based vision systems in a number of important ways:

- Curves are not segmented into polygonal approximations.
- It is unnecessary to search the model base. The invariant descriptors index a model directly.
- At this stage, no pose information is involved in recognition. It is therefore possible to *identify* an object without knowing *where* it is. Section 4 shows that, given that the model has been identified, pose recovery is simple.
- It is straightforward to acquire models for this system, because these models are projectively invariant. As a result, the invariants measured in any view of the model have the same value, so that model invariants can be calculated directly from conics fitted to curves extracted from *an image* of the object. The model can be imaged from any viewpoint, and no correction for aspect ratio or camera parameters is required. The model database then consists of the pair of invariants and error thresholds for each object.

4 Pose Determination

Once an object has been positively identified, the extra constraints offered by its known identity can be exploited to determine the transformation parameters between the object plane and the image. Invariant fitting allows a pair of coplanar curves to be modeled by a pair of coplanar conics. By construction, the modelling conics undergo the same projective distortion that the original curves do. Consequently, the problem of pose determination is equivalent to:

> Given a known pair of conics on the world plane, and their corresponding conics in the image, determine the transformation between the two planes.

4.1 Back projection of a conic pair

A perspective projection between the image plane and the object plane is determined by six parameters, which give the pose of the object relative to the camera. Each conic has five independent parameters, so the solution is overdetermined (10 constraints on 6 unknowns). Currently, this system is solved using the following approach:

The transformation variables are partially eliminated between the equations for the image conic coefficients to leave four equations in the $\{p, q\}$ pose variables which specify the orientation of the object plane. These equations consist of two conics, a quartic and an 8-ic (see [6] for details). The conics intersect in at most four real points - giving a possible four-fold ambiguity. However, the remaining equations must vanish at a solution for $\{p, q\}$. In general, this occurs only for one of these roots. Pose can therefore be uniquely recovered once a pair of conics has been matched. Once the orientation of the plane is determined, the other pose parameters simply scale, rotate and translate the conic pair in the plane, and are easily recovered. The details appear in [19].

4.2 Implications for model acquisition

Determining pose requires that the coefficients of the model conic *in the object plane* be known. If the object curves are known conics (this occurs in the case of the labels), then there is no difficulty. One needs to choose a coordinate system on the object plane within which to express these conics, and a sensible choice is the natural frame of one of the conics, e.g. for an ellipse the centre as origin and coordinate axes aligned with the principal axes.

However, if the object curves are not conics it is necessary to fit conics to them, using the invariant fitting method [4]. If the object curves are known in the world plane coordinate system then the fitting process produces the model directly. If the curves are not known in the world frame then they can be obtained by fitting curves in an image (using the invariant fitting method) and projecting these fitted curves back to the world plane. Backprojection requires that the transformation between the object plane and the image plane be known. This can easily be recovered by imaging the object together with a known *calibration* pair of conics (or a single circle) that is coplanar with the object curves. The calibration conics then determine the transformation between the image plane and the world plane.

4.3 Pose results

As exact pose between image and object plane is difficult to measure, we include an example where the relative motion between a views is calculated for a real object. The object curves are not conics. The object is rotated in its own plane by so motion between views is accurately known. An example is shown in figure 6. The results are given in table 3. There is good agreement between the actual and computed rotation, and the measured coordinates of the object plane remain reasonably stable.

5 Discussion

We have demonstrated a simple, efficient model based vision system that uses the principle of invariance to recognise objects, without regard to pose. We have shown that, once an object has been identified, its pose can easily be recovered. Extensive generalisation of this work is possible:

- The invariant fitting theorem works for algebraic curves of any degree. Higher degree curves have richer invariant theories in general, but the complexity of the numerical problem in fitting the curves is massively increased. It may be possible to solve these problems to create richer invariant descriptors for plane curves.
- The invariant fitting theorem applies only to point sets that are within projection. To obtain an invariant representation for *curves* requires integrating algebraic distance with respect to a projectively invariant parameter. However, simply summing algebraic distance at all the points leads to useful invariants in practice.
- The projectively invariant function associated with a single conic (equation 2.1) can be extended to a metric. It should be possible to use this function to avoid having to backproject models to verify the hypothesis that an instance has been found, by checking that pairs taken from a range of nearby features have the right function values. This has the advantage that the precise camera calibration needed to compute pose and backproject in conventional hypothesis verification, is unnecessary.

view	slant $\sigma/^\circ$	tilt $\tau/^\circ$	r/mm	$\theta/^\circ$
A	44.12	97.27	339.98	0.00
B	48.00	93.21	316.62	92.36

Table 4: Pose results for the two views of the mouse shown in figure 6. The mouse was rotated by $\sim 90^\circ$ between the views. The plane is the same in both cases.

- The representation used is tightly concentrated - 2 numbers represent any planar shape. As a result, we expect that distinct shapes will have the same representation. The projectively invariant functions (equation 2.1) associated with the representing conics can be used to distinguish between two models that have the same representation in terms of joint scalar invariants, again by checking the function values for pairs taken from a range of nearby features.
- The system as defined is sensitive to occlusion. It may be possible to overcome this difficulty by using the projective differential invariants described in section 2, example 4, or by exploiting the projectively invariant function associated with one of the conics (equation 2.1).
- The differences between plane projectivities and rigid motion with perspective projection (perspectivities) are profound and subtle. Planar projectivities have eight degrees of freedom, and form a group. Perspectivities have only six degrees of freedom, but do not form a group. However, the decreased generality of the transformations involved in perspectivities means that further tests for a model instance should be available. In section 4 we saw that pose was overdetermined - if, for example, there was *no* solution for $\{p, q\}$ that solved all four equations, this would be strong evidence that the wrong model has been matched. We believe that a better understanding of perspectivities is important to progress in vision.
- Recognising three dimensional curved objects from their outlines is an important problem. Ponce and Kriegman [15] generate outlines using elimination theory, and search over pose parameters for the best match between model and object outlines. It is possible to describe techniques based in invariant theory for direct recognition of curved objects, but these techniques require either eliminating an impractical number of variables from a system of polynomials, or solving huge dense polynomial systems [6]. It is uncertain whether these techniques can be refined to be practical.

Acknowledgements

Our thanks to Michael Brady and the Robotics Research Group, Oxford, for providing a pleasant environment in which to think and work. Our thanks to Andrew Blake, Margaret Fleck and Constantinos Marinos for a number of stimulating discussions. Charles Rothwell skilfully and efficiently implemented the model based vision systems we describe. Thanks to GE for the phone calls and a grant that made this paper possible.

References

[1] Brooks, R. A., "Model-Based Three-Dimensional Interpretations of Two Dimensional Images," IEEE PAMI, Vol. 5, No. 2, p. 140, 1983.

[2] Canny, J.F. "Finding edges and lines in images," TR 720, MIT AI Lab, 1983.

[3] Dieudonné, J. and Carrell, J.B. *Invariant Theory old and new*, Academic, 1971.

[4] Forsyth, D.A, Mundy, J.L., Zisserman, A. and Brown, C.M "Projectively invariant representations using implicit algebraic curves, " *Proc. 1st European Conference on Computer Vision*, Springer Verlag Lecture Notes in Computer Science, 1990. Revised version to appear in *Image and Vision Computing*, 1990.

[5] Forsyth, D.A., Mundy, J.L., Zisserman, A. and Brown, C.M. "Projectively invariant data representation by implicit algebraic curves: a detailed discussion," Internal report, Department of Engineering Science, Oxford University, 1990 (In preparation).

[6] Forsyth, D.A., Mundy, J.L., Zisserman, A. and Brown, C.M. "Applications of Invariant Theory in Computer Vision," Proceedings of the International Workshop on Integrating Symbolic and Numerical Computing, Saratoga Springs, 1990 (in preparation).

[7] Forsyth, D.A., Mundy, J.L. and Zisserman, A. "Transformational invariance - a primer," To appear in British Machine Vision Association Conference Proceedings, 1990.

[8] Grace, J.H. and Young, A. *The algebra of invariants*, Cambridge University Press, Cambridge, 1903.

[9] Huttenlocher, D.P., Ullman, S., " Recognizing Solid Objects by Alignment," Proc. DARPA Image Understanding Workshop, p. 1114, 1988.

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

[11] Lane, E.P. *A treatise on projective differential geometry*, University of Chicago press, 1941.

[12] Morgan, A.P. *Solving polynomial systems Using Continuation for Scientific and Engineering Problems*, Prentice-Hall, Englewood Cliffs, N.J., 1987.

[13] Nielsen, L. "Automated guidance of vehicles using vision and projectively invariant marking," *Automatica*, **24**, 2, 135-148.

[14] Olver, P.J. *The application of Lie groups to differential equations*, Springer Graduate texts in mathematics 107, Springer Verlag, 1986.

[15] Ponce, J. and Kriegman, D.J. "On recognising and positioning curved 3 dimensional objects from image contours," *Proc. DARPA IU workshop*, 461-470, 1989.

[16] Porrill, J. "Fitting ellipses and predicting confidence envelopes using a bias corrected Kalman filter," Proc. 5'th Alvey vision conference, 1989.

[17] Springer, T.A. *Invariant theory*, Springer-Verlag lecture notes in Mathematics, **585**, 1977.

[18] Weiss, I. "Projective invariants of shapes," *Proc. DARPA IU workshop*, 1125-1134, 1989.

[19] Zisserman, A.P., Marinos, C.I., Forsyth, D.A., Mundy, J.L., and Rothwell, C.A. "Relative motion and pose from invariants," To appear in British Machine Vision Association Conference Proceedings, 1990.

List of Figures

1. The model used for a pallet consists of the four parallel lines shown in the figure above. These lines yield a single invariant, equivalent to the cross ratio.

2. Two pallets found in a real image, by finding sets of four near parallel lines and marking those sets that are close to the correct value of the invariant.

3. Four images of a computer tape, with fitted conics in overlay. 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 shown in tables 1 and 2.

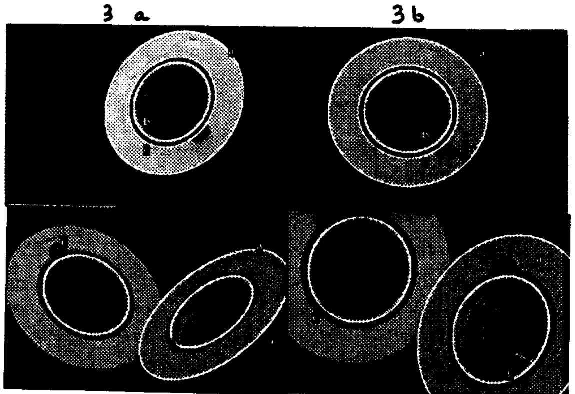
4. Four views of projectively invariant labels found in cluttered scenes. The label numbers are correct in each case. These labels are conceptually similar to Nielsen's [13] labels, but use different invariants). The labels come from a model base of 15 labels.

5. Six examples of the model based vision system working on real objects. (a) shows a gasket viewed approximately frontally. Models of the gaskets were made using images like this. (b) shows the gasket shown in (a), seen in a different view. The gasket was recognised and labelled correctly despite the large change in viewing angle. (c) shows a cluttered scene containing four gaskets, and (d) shows the gaskets correctly

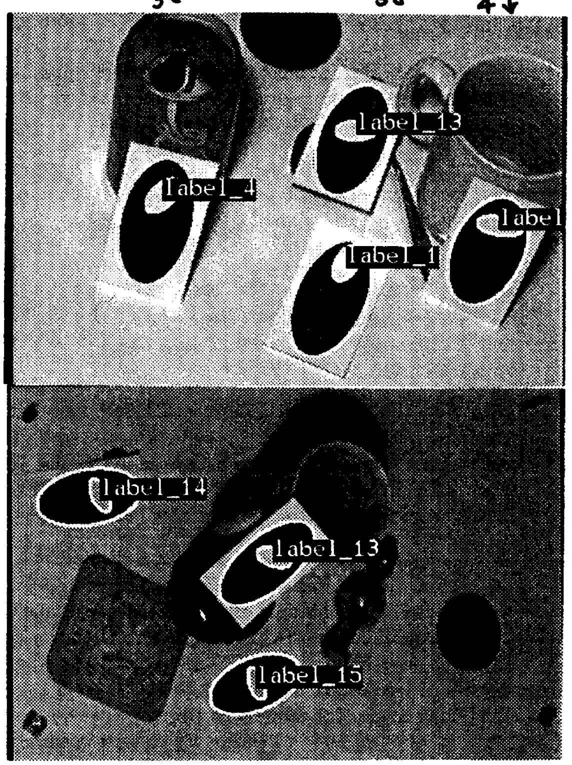
recognised and labelled. (e) shows another cluttered scene, also containing four gaskets, and (f) shows the gaskets labelled in that scene. Note that the micrometer recognised as gasket 3 can easily be dealt with by verifying the model using backprojection.

6. Images of a mouse with representing conics superimposed. The motion between views is a 90° rotation and small translation of the mouse with the camera static. The plane of the mouse buttons is approximately the same in each image. Results of the pose recovery program are shown in table 4.

3 a 3 b



3c 3d 4↓



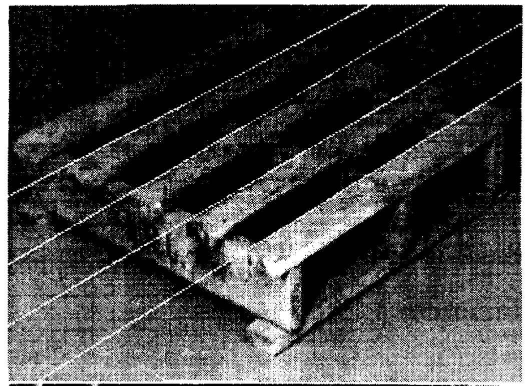
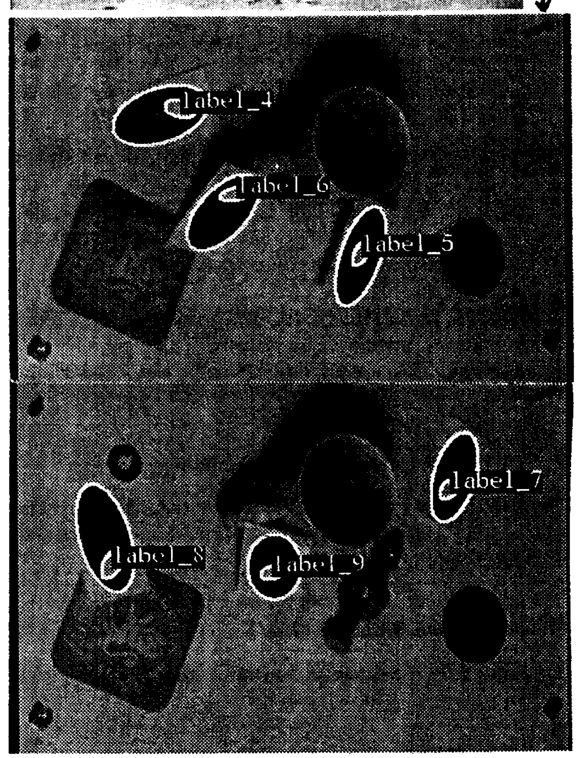
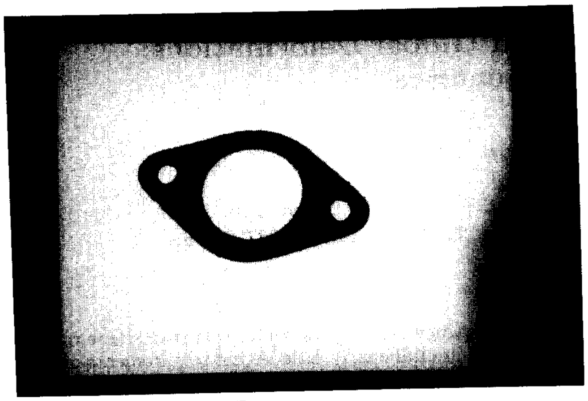


FIG 1

FIG 2

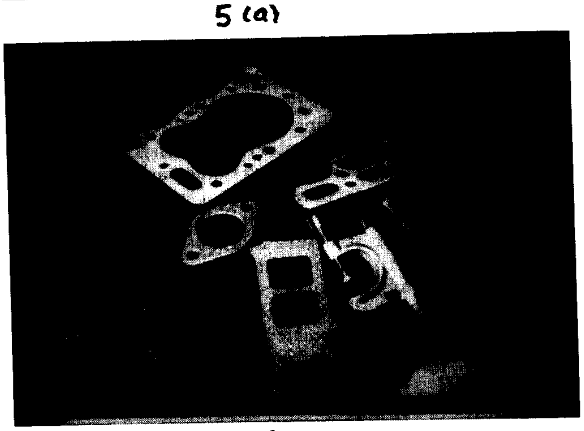
FIG 4 ↓





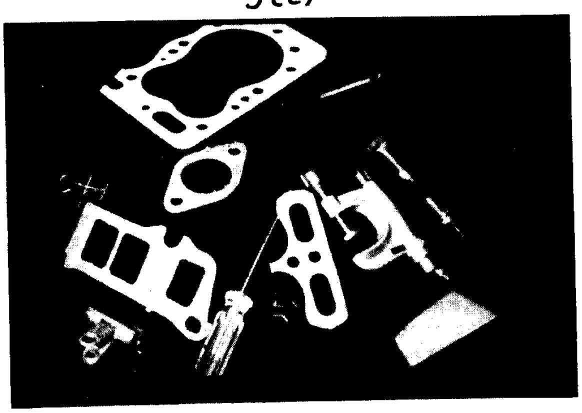
5 (a)

5 (b)

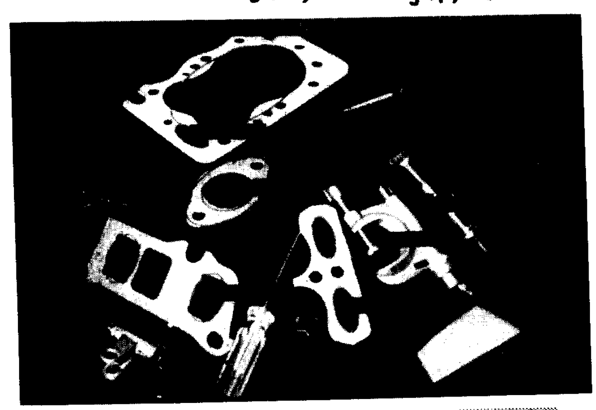


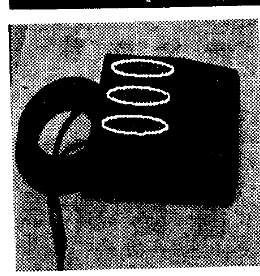
5 (c)

5(d) ↑ 5(f) ↓



5 (e)





6