PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

# Projectice subgroups for grouping

The Royal Society

*Phil. Trans. R. Soc. Lond. A* 1998 **356**, 1251-1266
doi: 10.1098/rsta.1998.0220

# Projective subgroups for grouping

By Luc J. Van Gool

*Katholieke Universiteit Leuven, ESAT-MI2, Kardinaal Mercierlaan 94, B-3001 Leuven, Belgium*

The history of grouping in computer vision stretches a long way back. One strand could be called geometry-based and focuses on shapes with special regularities such as symmetry. From a mathematical point of view, the subparts of such shapes are related by special transformations. The presented work proposes to systematically classify such 'special' transformations by studying projective subgroups that come with fixed structures. These are geometric entities such as points or lines that remain fixed under the projectivities in the subgroups. As subgroups have their own invariants, these can then be used to guide the search for the corresponding groupings.

**Keywords: computer vision; grouping; invariants; projective geometry; geometric constraints**

## 1. Introduction

Grouping is the process of combining visual information into perceptual entities. It acts as a kind of shortcut between low-level features and scene interpretation, quickly assembling parts that probably belong together. Here we focus on grouping *planar edges*, with the following goals:

1. *A principled approach.* In the literature, concepts like 'goodness' and 'non-accidentalness' (Wertheimer 1938; Kanade 1981; Lowe 1984, 1985) have been used to draw up catalogues of grouping types. As a matter of fact, the types that were included were selected on more or less intuitive grounds. Moreover, such catalogues showed little structure and typically did not hint at possible approaches to detect the groupings. Here, a more systematic classification of grouping types is propounded, albeit from a rather restrictive geometric point of view. Directly tied to the classification is an approach to detecting the grouping types.

2. *Including perspective effects.* Grouping has often been carried out under the assumption of (pseudo-)orthographic projection. This has to do with the fact that many more perceptual cues survive the corresponding affine skewing than the projective skewing that amounts from the more realistic, perspective model. Here, the full perspective nature of projection will be taken into account.

3. *Efficient grouping.* Grouping is about combining parts into larger configurations. Hence, there is a risk for combinatorial search. Here, the combined use of invariants and the Hough transform is proposed to minimize that risk.

Much of our analysis is based on the classifications of subgroups of the plane projectivities. Nevertheless, the planar shapes that are involved can be part of non-planar configurations. Moreover, sometimes the same geometrical analysis applies




TEX Paper

to the grouping of curved surfaces. Surfaces of revolution are a good case in point, as their outlines share their geometrical constraints with planar symmetric shapes (Zisserman *et al*. 1995).

The structure of the paper is as follows. Section 2 discusses the kind of subgroups that will be considered. In particular, the concept of fixed structures as a guiding principle in this analysis is discussed. The paper also recapitulates some issues of general, projective invariants, then presses on with their specialization towards the subgroups in the subsequent sections: § 3 for fixed points and lines and § 5 for fixed sets of points. Section 4 is an intermezzo, linking some of the results that are useful for recognition rather than grouping. Section 6 introduces the cascaded Hough transform to aid in the detection of the fixed structures. The results are then brought together in a strategy for geometry-based grouping, described concisely in § 7. It is based on the fixed structures and the invariants of the subgroups that they define. Section 8 concludes the paper and comments on possible future work.

## 2. Identifying subgroups for grouping

### (*a*) *Fixed structures and subgroups*

Invariants are useful tools for grouping because they allow one to find matches while avoiding combinatorial search. Using general projective invariants is not necessarily the optimal approach, however. This may be because such invariants need a minimum of contour information for their extraction, e.g. a 'bitangent segment' (Carlsson *et al*. 1996), and as grouping is about matching parts these may lack such rich local structure. A second problem is that far too many matches might result. Consider what would happen if one were to look for the other half of a mirror symmetric shape in a whole pile of identical objects. All half shapes would match under projective invariants. This may result in hundreds of possible matches which then have to be checked further whether they really represent a mirror symmetry. In such cases symmetry-specific invariants can increase efficiency considerably, as they selectively pick out half shapes that are in symmetric positions. The existence of such symmetry-specific invariants hinges on the existence of projective subgroups to which the skewed symmetries would have to belong. As the sequel shows, so-called fixed structures yield a direct route to finding such subgroups.

Consider two planar shapes in three-dimensional (3D) space. Suppose that there exists a 3D projective transformation that maps one shape to the other. This is the basic grouping configuration studied here. A special case is when the two planes coincide, as with the two halves of a mirror symmetry. The existence of the projectivity in 3D implies that in an image of such a configuration, the projections of the shapes are related by a projectivity in 2D.

Furthermore, if the 3D projectivity maps certain structures onto themselves, i.e. keeps these structures fixed in three-dimensional space, then *a fortiori* their images will remain fixed under the 2D projectivity in the image. Trivial as this observation may be, it is important to keep in mind that not too many features survive the projection onto the image. Taking mirror symmetry as an example, symmetric points have the same distance to the axis, the same curvature, etc., in 3D space, but not in the image (Glachet *et al*. 1993). Yet, the projectivity that maps the symmetric halves onto each other in the image still has a symmetry axis, i.e. a straight line all points of which are mapped onto themselves. Also, pairs of symmetric points still form fixed pairs in the sense that one point is mapped to the other and vice versa. The *fixed structures* survive in the image and hence provide a solid basis for the non-accidentalness paradigm.

Moreover, projectivities that keep the same structures fixed, e.g. a specific line or point, form subgroups of the projectivities (Van Gool *et al*. 1995). Thus, if groupings are organized according to the fixed structures of the corresponding projectivities, these projectivities belong to specific subgroups, for which specific invariants can be extracted. This is the crux of the matter. The subgroups defined by the fixed structures yield invariants. These allow one to match the parts of the grouping without combinatorial search, e.g. using hashing as for recognition (Rothwell 1993; Carlsson *et al*. 1996). These invariants are also grouping-specific, i.e. geared towards a specific type of configuration such as a symmetry, rather than being general projective invariants. This adds to the efficiency of the search and makes it possible to match curve segments that would be too small for effective projective matching.

Finally, classifying projectivities in terms of fixed structures makes explicit the equivalences between cases which might otherwise be treated separately. As an example, the detection of mirror and point symmetries in perspective views can be proved to be one and the same problem from a mathematical point of view, precisely because they have the same kind of fixed structures in the image (Van Gool *et al*. 1996).

### (*b*) *Selection of fixed structures*

When searching for structures that could remain fixed under projectivities, it stands to reason to first concentrate on the simplest kind of structures that remain qualitatively invariant. Examples are points, lines and conics, because points are mapped to points, lines to lines, and conics to conics. There are other such structures, like curves with constant projective curvature, but these are considered too intricate to be of practical use here. Most of the analysis will be carried out in the real plane.

A further distinction can be made between cases where the remaining structures—points, lines, and conics—are fixed individually or as a set. For example, under a rotational symmetry (also when viewed obliquely), only the centre of rotation is a point that is fixed individually, but other points belong to sets of points that remain fixed as a set. In addition to this distinction, it is also useful to consider combinations of fixed structures, like transformations that keep two points and a line fixed. In particular, it comes out that complete pencils of fixed structures are a particularly relevant case, as will be seen later.

All in all, the number of cases to be considered seems to become quite high. Nevertheless, not just any combination of fixed structures is possible. A complete classification of consistent combinations has not been developed yet. This paper focuses on fixed points and lines, only introducing the case of fixed sets of points. Additional results on fixed sets of points and some comments on fixed conics are given elsewhere (Van Gool 1997).

As to the related, grouping-specific invariants, previous work on semi-differential invariants is extended. These are invariants for the description of curves, that combine point coordinates with their derivatives (Van Gool *et al*. 1992). In order to keep this paper more or less self-contained, the sequel of this section gives a short overview of how semi-differential invariants can be derived for the general case of plane projectivities.

### (c) Semi-differential invariants

Consider a general projectivity with matrix $P = (p_{ij})$ acting on points $\boldsymbol{x}_k = (x_k, y_k)^{\mathrm{T}}$ as

$$x'_k = \frac{p_{11}x_k + p_{12}y_k + p_{13}}{p_{31}x_k + p_{32}y_k + p_{33}}, \qquad y'_k = \frac{p_{21}x_k + p_{22}y_k + p_{23}}{p_{31}x_k + p_{32}y_k + p_{33}}. \tag{2.1}$$

Using shorthand notation

$$N_k = p_{31}x_k + p_{32}y_k + p_{33}$$

for the denominator, the equations (2.1) can also be written as

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \frac{1}{N} P \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \tag{2.2}$$

Starting from equations (2.2) it is then easy to show that

$$|\boldsymbol{x}'_1 - \boldsymbol{x}'_2 \quad \boldsymbol{x}'_1 - \boldsymbol{x}'_3| = \frac{|P|}{N_1 N_2 N_3} |\boldsymbol{x}_1 - \boldsymbol{x}_2 \quad \boldsymbol{x}_1 - \boldsymbol{x}_3|, \tag{2.3}$$

where vertical bars indicate determinants.

Indicating between parentheses the order of coordinate derivatives with respect to a projectively invariant parameter,

$$|\boldsymbol{x}'_1 - \boldsymbol{x}'_2 \quad \boldsymbol{x}_1'^{(1)}| = \frac{|P|}{N_1^2 N_2} |\boldsymbol{x}_1 - \boldsymbol{x}_2 \quad \boldsymbol{x}_1^{(1)}| \tag{2.4}$$

and

$$|\boldsymbol{x}_1'^{(1)} \quad \boldsymbol{x}_1'^{(2)}| = \frac{|P|}{N_1^3} |\boldsymbol{x}_1^{(1)} \quad \boldsymbol{x}_1^{(2)}|. \tag{2.5}$$

At a discontinuity like a vertex, $\boldsymbol{x}_1$ say, where a left (l) and right (r) derivative can be distinguished, one has

$$|\boldsymbol{x}_1'^{(1:\ell)} \quad \boldsymbol{x}_1'^{(1:r)}| = \frac{|P|}{N_1^3} |\boldsymbol{x}_1^{(1:\ell)} \quad \boldsymbol{x}_1^{(1:r)}|. \tag{2.6}$$

One could consider (2.6) as the counterpart of (2.5) for discontinuities.

The expressions (2.3), (2.4), (2.5), and (2.6) can be considered to be building blocks for the generation of projective invariants. A possible strategy is to take products of these building blocks, raised to appropriate powers to eliminate all the factors that they produce under projective transformations (Van Gool *et al*. 1992).

In general an invariant parameter will not be available and also invariance under reparametrization has to be realized. Fortunately, the same building blocks can be used. If the left-hand sides are calculated on the basis of a parameter $t'$ and the right-hand sides use $t$, then building blocks (2.5) and (2.6) change with $(\mathrm{d}t/\mathrm{d}t')^3$ and building block (2.4) with $(\mathrm{d}t/\mathrm{d}t')$. Again these factors should be cancelled by raising the building blocks in the product by the appropriate power (Van Gool *et al*. 1992).

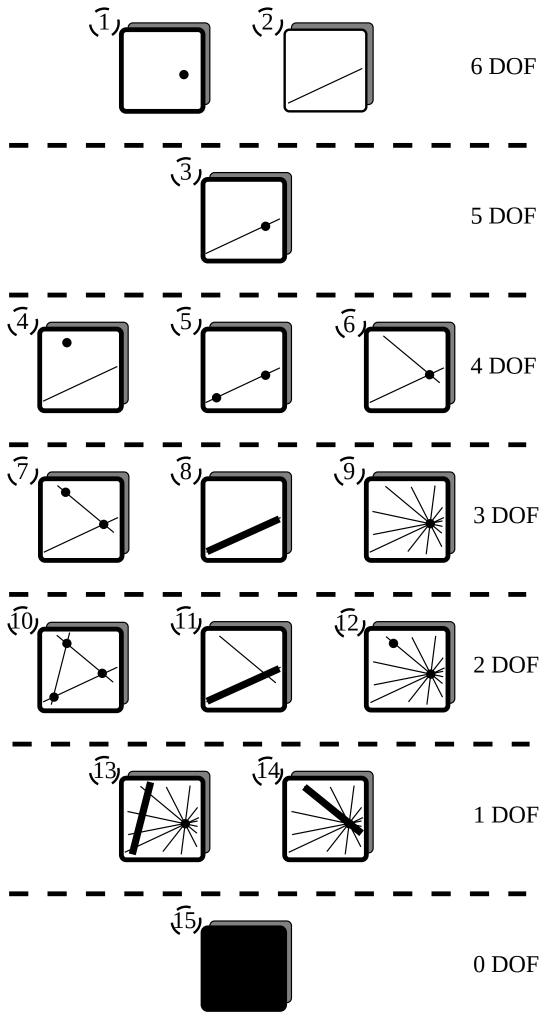


Figure 1. Classification of fixed structure subgroups for fixed points and lines.

## 3. Combinations of fixed points and lines

### (*a*) *A classification of subgroups*

The possible combinations of fixed points and fixed lines that projective transformations can share are summarized schematically in figure 1. Every square corresponds to a different type of subgroup, with a qualitatively different combination of fixed structures. A point in such a square indicates a specific (but arbitrary) fixed point; the same for a line. Note that sometimes a fixed point lies on a fixed line. In the cases indicated with numbers 8, 11, 13, and 14, a thicker line is drawn. This is to mean that every point on such a line is a fixed point and hence thick lines represent lines of fixed points. In cases numbered 9, 12, 13, and 14, a bunch of concurrent lines has been drawn. These are supposed to represent pencils of fixed lines, where all lines through a point, the so-called vertex, remain fixed. The vertex is a fixed





PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

point. Such pencils are the projective duals of lines of fixed points. The black square at the bottom is the trivial case (case 15), where all points are fixed points and the remaining subgroup only contains the identity.

Figure 1 in effect is more than an enumeration of subgroup types. Going down the classification, additional fixed structures are added, thereby gradually decreasing the dimensionality of the subgroups. The dimension of the corresponding subgroups is indicated on the right. A more detailed discussion of the subgroups and their invariants is given elsewhere (Van Gool *et al*. 1995). Note that the classes of fixed points and lines for *individual* projectivities as they are given by Springer (1964) correspond to only 7 out of the 15 classes for the subgroups. These two classifications must not be confused.

Six of the subgroup types of figure 1 are of special interest: these are the cases 8, 9, 11, 12, 13, and 14, which all contain a line of fixed points, a pencil of fixed lines, or both. These cases are of special interest because both the line of fixed points and the pencil of fixed lines fix five degrees of freedom, whereas only two parameters need to be specified to fully characterize them: the two parameters to specify the line or the vertex. This gain in degrees of freedom yields invariants that require strictly less information than that needed for the general projective invariants. One might argue that this also applies to the other cases in figure 1, but having a fixed point would, for example, lead to invariants based on the fixed point and four *additional* points, still requiring a total of five points. A similar observation can be made for all the other cases without a line of fixed points or a pencil of fixed lines.

Next it is shown how the existence of lines of fixed points or pencils of fixed lines yields invariants specific for the corresponding subgroups. Compared to the building blocks of § 2 *b* for the general projective case, these fixed structures yield additional building blocks or factors that are easier to eliminate.

### (*b*) *A pencil of fixed lines*

If there is a pencil of fixed lines, then every point is known to stay on the line of the pencil on which it lies. Denoting the pencil vertex with $\boldsymbol{x}_v = (x_v, y_v)^{\mathrm{T}}$, one therefore knows that there exists a factor $k_i$ such that for a point $(x_i, y_i)^{\mathrm{T}}$ and its image $(x_i', y_i')^{\mathrm{T}}$

$$(x_i' - x_v) = k_i(x_i - x_v),$$
$$(y_i' - y_v) = k_i(y_i - y_v).$$

Such a factor $k_i$ exists for every point $\boldsymbol{x}_i$. It immediately follows that

$$\frac{(y_i - y_v)}{(x_i - x_v)}$$

is an invariant, requiring only two points, one of which is the vertex.

In order to derive additional invariants (combinations with the different building blocks of § 2 *b*), it is important to know more about the factor $k_i$. Consider

$$|\boldsymbol{x}_1' - \boldsymbol{x}_v \quad \boldsymbol{x}_2' - \boldsymbol{x}_v| = \frac{|P|}{N_1 N_2 N_v}|\boldsymbol{x}_1 - \boldsymbol{x}_v \quad \boldsymbol{x}_2 - \boldsymbol{x}_v|.$$

This can also be written as

$$|\boldsymbol{x}_1' - \boldsymbol{x}_v \quad \boldsymbol{x}_2' - \boldsymbol{x}_v| = k_1 k_2 |\boldsymbol{x}_1 - \boldsymbol{x}_v \quad \boldsymbol{x}_2 - \boldsymbol{x}_v|$$




MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF

and therefore

$$k_1 k_2 = \frac{|P|}{N_1 N_2 N_v}.$$

From the fact that this latter equality holds for any choice of the points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, it follows that

$$k_i = \pm\sqrt{\mathrm{abs}\left(\frac{|P|}{N_v}\right)\frac{1}{N_i}}.$$

We conclude that $(x_i - x_v)$ and $(y_i - y_v)$ come as additional building blocks with the pencil of fixed lines, easing the construction of invariants. An example invariant parameter is

$$\int \mathrm{abs}\left(\frac{|\boldsymbol{x} - \boldsymbol{x}_v \boldsymbol{x}^{(1)}|}{(x - x_v)^2}\right) \mathrm{d}t.$$

### (*c*) *A line of fixed points*

If there is a line of fixed points—in the sequel referred to as *the axis*—then any point $\boldsymbol{x}_{ai}$ on it is fixed. Hence,

$$|\boldsymbol{x} - \boldsymbol{x}_{a1} \quad \boldsymbol{x} - \boldsymbol{x}_{a3}| = l|\boldsymbol{x} - \boldsymbol{x}_{a1} \quad \boldsymbol{x} - \boldsymbol{x}_{a2}|$$
$$|\boldsymbol{x}' - \boldsymbol{x}_{a1} \quad \boldsymbol{x}' - \boldsymbol{x}_{a3}| = l|\boldsymbol{x}' - \boldsymbol{x}_{a1} \quad \boldsymbol{x}' - \boldsymbol{x}_{a2}|$$

with

$$l = \frac{\|\boldsymbol{x}_{a1} - \boldsymbol{x}_{a3}\|}{\|\boldsymbol{x}_{a1} - \boldsymbol{x}_{a2}\|}.$$

It follows that

$$\frac{|P|}{NN_{a1}N_{a3}} = \frac{|\boldsymbol{x}' - \boldsymbol{x}_{a1} \quad \boldsymbol{x}' - \boldsymbol{x}_{a3}|}{|\boldsymbol{x} - \boldsymbol{x}_{a1} \quad \boldsymbol{x} - \boldsymbol{x}_{a3}|} = \frac{|\boldsymbol{x}' - \boldsymbol{x}_{a1} \quad \boldsymbol{x}' - \boldsymbol{x}_{a2}|}{|\boldsymbol{x} - \boldsymbol{x}_{a1} \quad \boldsymbol{x} - \boldsymbol{x}_{a2}|} = \frac{|P|}{NN_{a1}N_{a2}}$$

and thus $N_{a1} = N_{a2} = N_a$ where $N_a$ is one and the same value for all the points on the axis.

It then immediately follows that, for example,

$$\frac{|\boldsymbol{x}_{a1} - \boldsymbol{x}_1 \quad \boldsymbol{x}_{a1} - \boldsymbol{x}_2|}{|\boldsymbol{x}_{a2} - \boldsymbol{x}_1 \quad \boldsymbol{x}_{a2} - \boldsymbol{x}_2|}$$

is an invariant, which requires knowledge about the axis and only two additional points, hence a total of six parameters (the two points on the axis can be chosen arbitrarily). A geometrical interpretation of this invariant is that the lines $\langle \boldsymbol{x}_1, \boldsymbol{x}_2 \rangle$ and $\langle \boldsymbol{x}'_1, \boldsymbol{x}'_2 \rangle$ interesect the axis in the same point.

### (*d*) *A pencil and an axis*

If both a pencil of fixed lines and a line of fixed points exist, then the previous results can be combined. If one considers $(x_a - x_v)$, where both the point on the axis $\boldsymbol{x}_a$ and the pencil vertex $\boldsymbol{x}_v$ are fixed points now, this expression is a trivial invariant, i.e.

$$k_a = \pm\sqrt{\mathrm{abs}\left(\frac{|P|}{N_v}\right)\frac{1}{N_a}} = 1$$

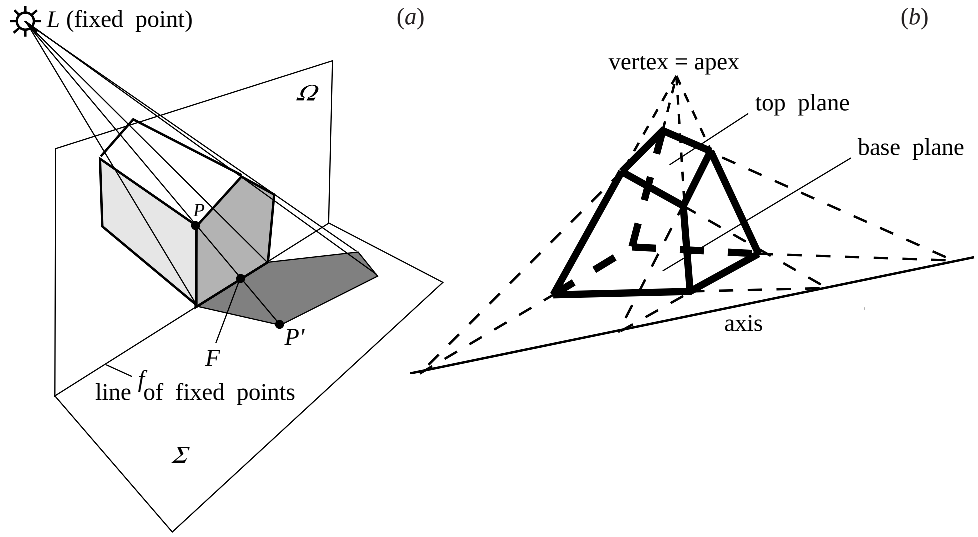


Figure 2. (*a*) Geometry of the object-shadow configuration. (*b*) Example of an extruded surface. Such shapes are formed by cutting a general cone by two planes (top and base plane).

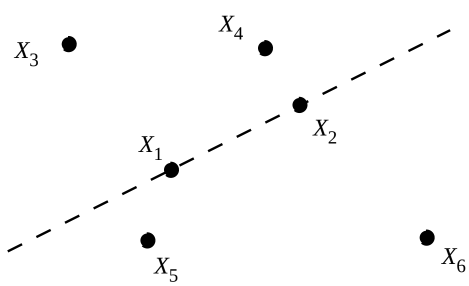


Figure 3. Butterfly configuration with labelled points.

and therefore $N_a = \pm\sqrt{\mathrm{abs}(|P|/N_v)}$, or, equivalently, $N_v = |P|/N_a^2$.

Cases with such combination of an axis and a pencil come out to be of particular, practical importance. Such *planar homologies* seem to pop up virtually everywhere in vision. Consider figure 2: the relation in the image between a planar shape and its shadow or the top and bottom plane of an extruded surface both correspond to a planar homology.

## 4. Extension to multiple views

Some of these results are not only of interest when comparing shapes within a single view, but also when different views of the same structure are available and these structures have to be matched between the images.

An example where planar homologies pop up is in determining the epipolar geometry of a pair of cameras. As has been noted before (Sinclair *et al*. 1996), the knowledge of the projectivities $P_1$ and $P_2$ for two planes between the two views suffices. What matters are the composed transformations $P_2^{-1}P_1$ and $P_1^{-1}P_2$. As a matter of fact, these correspond to planar homologies. The fixed points off the axis (the vertices) correspond to the epipoles. The lines of fixed points are the intersections of the planes as seen in each of the stereo views. Connecting the epipoles with corresponding points on the lines of fixed points yields pairs of corresponding epipolar lines.

Pursuing this line of thought a little further for the line of fixed points, i.e. the intersection of the two planes, we know that the denominators $N_a$ of the homology will be the same for all the points on this line. Following the notation of equation (2.2) we have for $P_2^{-1}P_1$, which corresponds to a planar homology,

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \frac{N_2}{N_1} P_2^{-1} P_1 \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}.$$

As all points on an axis have the same $N_a$, points on the intersection of the planes have the same $N_2/N_1$. For ease of reference, we call this the 'ratio constraint'.

From this simple ratio constraint a host of invariants can be derived that can be used for configurations consisting of planar, but not necessarily coplanar, substructures. A good example is the so-called 'butterfly configuration', as shown in figure 3.

In this figure $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, and $\boldsymbol{x}_4$ are coplanar ('first plane'), and so are $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_5$, and $\boldsymbol{x}_6$ ('second plane'). The points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ lie on the intersection of the planes, as indicated with the dashed line. The following combination of areas is invariant under changes in viewpoint:

$$\frac{|\boldsymbol{x}_1 - \boldsymbol{x}_3 \quad \boldsymbol{x}_1 - \boldsymbol{x}_4||\boldsymbol{x}_2 - \boldsymbol{x}_5 \quad \boldsymbol{x}_2 - \boldsymbol{x}_6|}{|\boldsymbol{x}_1 - \boldsymbol{x}_5 \quad \boldsymbol{x}_1 - \boldsymbol{x}_6||\boldsymbol{x}_2 - \boldsymbol{x}_3 \quad \boldsymbol{x}_2 - \boldsymbol{x}_4|}.$$

The net factor by which it would change according to the combination of factors of its building blocks (see equation (2.3)) is

$$\frac{N_1'}{N_1}\frac{N_2}{N_2'},$$

where $N$ without a prime indicates the value for the projectivity of the first plane and $N'$ corresponds to the projectivity of the second plane. This overall factor equals 1 because of the ratio constraint (i.e. $N_1/N_1' = N_2/N_2'$). The invariant, as it is given here, is directly related to the original, cross-ratio based definition of the 'butterfly invariant', i.e. the cross-ratio formed by the collinear points $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, and the intersections of the lines $\langle \boldsymbol{x}_3, \boldsymbol{x}_4 \rangle$ and $\langle \boldsymbol{x}_5, \boldsymbol{x}_6 \rangle$ with the intersection (it was in this form that the butterfly invariant was first brought to my attention by J. L. Mundy). In the meantime, Rothwell & Stern (1996) have pointed out that the three 'caging invariants' of Rothwell *et al*. (1993) can in turn be derived from the butterfly invariant. In fact, it is quite easy to derive them directly using the above constraints, which also makes it easy to derive generalizations to junctions where an arbitrary number of faces meet (Rothwell *et al*. (1993) restricted their analysis to trihedral junctions).

Another example where two planes are involved is given to show that other building blocks can be used in combination with the ratio constraint. Two planar, but not coplanar, curves touch in two points, as shown in figure 4.

Then, the combination

$$\frac{\kappa_1}{\kappa_1'}\frac{\kappa_2'}{\kappa_2}$$





PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

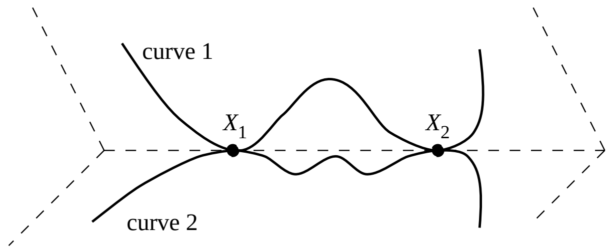


Figure 4. A pair of planar curves tangent in two points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$.

of the curvatures at the two points is an invariant, where all curvatures are measured in the image. Primes again indicate that these measurements are taken for the second curve. The invariance of the curvature based expression can be checked by rewriting it as

$$\frac{|\boldsymbol{x}_1^{(1:s)} \quad \boldsymbol{x}_1^{(2:s)}|}{|\boldsymbol{x}_1'^{(1:s')} \quad \boldsymbol{x}_1'^{(2:s')}|}\frac{|\boldsymbol{x}_2'^{(1:s')} \quad \boldsymbol{x}_2'^{(2:s')}|}{|\boldsymbol{x}_2^{(1:s)} \quad \boldsymbol{x}_2^{(2:s)}|},$$

where the specifications $:s$ and $:s'$ indicate that the derivatives are taken with respect to Euclidean arclength (for the image projections of the curves). The invariance of the above expression can again be proven by considering the net factor that results under a change in viewpoint:

$$\left(\frac{N_1'}{N_1}\frac{N_2}{N_2'}\right)^3 = 1.$$

## 5. Fixed sets of points

A set of points may, rather than being fixed individually, map onto each other. The set, not its points, is fixed. Such cases are important, because they correspond to discrete symmetries. Mirror symmetry is an example where every point belongs to a fixed pair of symmetric points. Ornamental symmetries include all cyclic and dihedral symmetry groups of different orders. Cyclic symmetry of order $n$ is synonymous to $n$-fold rotational symmetry. Dihedral symmetry groups add mirror symmetries. As a matter of fact, there is an isomorphism between the skewed symmetries as observed in the image and the 'ornamental symmetry group' of the shape. Vice versa, the existence of fixed sets of points typically are a strong indication for the presence of skewed ornamental symmetries, and in some cases it even gives a guarantee (e.g. if there is a fixed triple (Semple & Kneebone 1979)).

As in the case of a line of fixed points or a pencil of fixed lines, the presence of fixed discrete sets of points yields specialized invariants. And again, these are based on further constraints on the factors of the building blocks in § 2 *b*.

Consider a fixed $n$-tuple of points, $\boldsymbol{x}$, $\boldsymbol{x}'$, $\boldsymbol{x}''$, ..., $\boldsymbol{x}^{[n-1]}$. Consider what happens to $|\boldsymbol{x}-\boldsymbol{x}' \quad \boldsymbol{x}-\boldsymbol{x}''|$. Applying the transformation $n$ times brings all the points back to their original positions. Hence, following the factor brought about by such building blocks according to equation (2.3)

$$\frac{|P|^n}{(NN'N''\ldots N^{[n-1]})^3} = 1$$

and therefore

$$NN'N''\dots N^{[n-1]} = |P|^{n/3}. \tag{5.1}$$

A degenerate case of a point cycle is the $n$-fold repetition of the rotation centre. It follows from equation (5.1) that for this point, $\boldsymbol{x}_{\mathrm{c}}$ say, $N_{\mathrm{c}}^{n} = |P|^{n/3}$, i.e. $N_{\mathrm{c}} = \pm|P|^{1/3}$. This holds irrespective of the angle of rotation.

As an example, if one is looking obliquely at a 3-fold rotational symmetry,

$$\frac{|\boldsymbol{x}_1 - \boldsymbol{x}_2 \quad \boldsymbol{x}_1' - \boldsymbol{x}_2||\boldsymbol{x}_1 - \boldsymbol{x}_3 \quad \boldsymbol{x}_1'' - \boldsymbol{x}_3|}{|\boldsymbol{x}_1 - \boldsymbol{x}_2 \quad \boldsymbol{x}_1 - \boldsymbol{x}_3|}.$$

is an invariant under the transformation that corresponds to the 120° rotation as seen in the image. The symmetrically positioned counterparts of $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$ yield the same values, while also cycling through $\boldsymbol{x}_1$, $\boldsymbol{x}_1'$, $\boldsymbol{x}_1''$ in the appropriate way, e.g.

$$\frac{|\boldsymbol{x}_1' - \boldsymbol{x}_2' \quad \boldsymbol{x}_1'' - \boldsymbol{x}_2'||\boldsymbol{x}_1' - \boldsymbol{x}_3' \quad \boldsymbol{x}_1 - \boldsymbol{x}_3'|}{|\boldsymbol{x}_1' - \boldsymbol{x}_2' \quad \boldsymbol{x}_1' - \boldsymbol{x}_3'|} = \frac{|\boldsymbol{x}_1 - \boldsymbol{x}_2 \quad \boldsymbol{x}_1' - \boldsymbol{x}_2||\boldsymbol{x}_1 - \boldsymbol{x}_3 \quad \boldsymbol{x}_1'' - \boldsymbol{x}_3|}{|\boldsymbol{x}_1 - \boldsymbol{x}_2 \quad \boldsymbol{x}_1 - \boldsymbol{x}_3|}.$$

Although this invariant uses a total of five points as a general point-based projective invariant would, it is both simpler and more selective. This expression is not invariant under general projectivities. Note that, as usual, this symmetry-specific invariant contains information on the fixed structures of the symmetry, i.e. the fixed triple $\boldsymbol{x}_1$, $\boldsymbol{x}_1'$, $\boldsymbol{x}_1''$.

Skewed mirror symmetry deserves some special attention, both because of its special status in the grouping literature and because of its rich collection of fixed structures. In fact, it can be considered the combination of fixed pairs of points with a planar homology, with its line of fixed points (axis) and pencil of fixed lines. In the terminology of projectivities, it is referred at as a harmonic homology. For a more detailed discussion of this case, see Van Gool *et al.* (1996).

## 6. The cascaded Hough transform

Using grouping-specific invariants requires the explicit knowledge of the corresponding fixed structures on which they are based. Invariants only take the combinatorics out of the edge comparison *per se*. There still is a risk to enter a combinatorial search for fixed structures and their relations to the edges. This section introduces the 'cascaded Hough transform' (CHT) to reduce that risk.

Fortunately, however complicated the shapes to be grouped, e.g. the halves of an intricate, mirror symmetric ornament, the fixed structures remain equally simple, i.e. points, lines, and conics. Finding such simple geometric objects has been studied extensively in computer vision. A well-known technique is the Hough transform. In particular, it is very effective for finding straight lines, even if they are fragmented. Here we propose its use for the detection of some of the fixed lines and fixed points.

In the original version of the Hough transform, the lines were given a slope-intercept representation, i.e. using parameters $(a, b)$ according to the equation

$$ax + b + y = 0. \tag{6.1}$$

Using this parametrization, a pair of edge point coordinates $(x, y)$ is transformed into a line in the $(a, b)$-parameter space. Similarly, a point with coordinates $(a, b)$ in the Hough parameter space corresponds to a line in the $(x, y)$-space, i.e. the image.

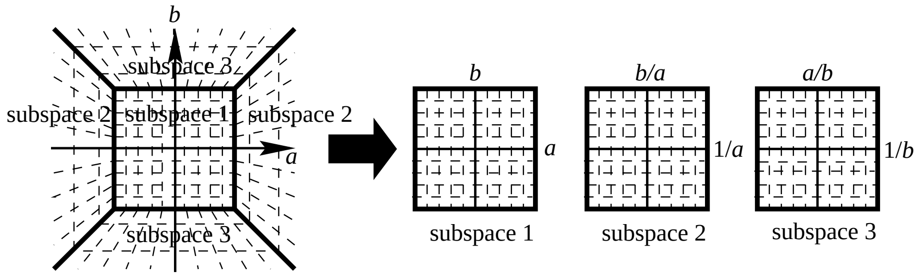


Figure 5. The original, unbounded space is split into three, bounded subspaces with coordinates $(a, b)$, $(1/a, b/a)$, and $(1/b, a/b)$, respectively.

Equation (6.1) was deliberately written in a form to emphasize the perfect symmetry between $(x, y)$ and $(a, b)$.

The CHT consists of subsequent applications of the Hough transform, with the parametrization of equation (6.1). As usual, the first Hough transform detects collinear points $(x, y)^{\mathrm{T}}$ in the image as peaks in the $(a, b)$-space. Collinear $(a, b)$s correspond to a point $(x, y)$ that is the intersection of the corresponding lines. Applying a second Hough transform to the output of the first yields such intersections as peak responses in a new $(x, y)$-space. Applying the Hough once more to this output, collinear intersections are found as peaks in a new $(a, b)$-space. It is important to note that the output of one layer is filtered before being used as input to the next. The cells that receive maximal votes are the most important structures to be passed on.

Several fixed structures can be found in this way, such as vanishing points (as line intersections) or the horizon line (as a line that contains several vanishing points). Similarly, if tangent lines are drawn at inflections, collinear intersections of these lines may indicate axes of symmetry, etc.

Of course, the $(a, b)$-parametrization is known to cause some problems as this space is unbounded. Yet, rather than going to the polar representation and thereby loosing the important symmetry, such problems can be avoided by splitting the $(a, b)$-space appropriately (Tuytelaars *et al.* 1997, 1998). This is shown in figure 5. The first subspace also has coordinates $a$ and $b$, but is used only for $|a| \leqslant 1$ and $|b| \leqslant 1$. If $|a| > 1$ and $|b| \leqslant |a|$, the point $(a, b)$ turns up in the second subspace, with coordinates $1/a$ and $b/a$. If, finally, $|b| > 1$ and $|a| < |b|$, we use a third subspace with coordinates $1/b$ and $a/b$. In this way, the unbounded $(a, b)$-space is split into three subspaces with coordinates restricted to the interval $[-1, 1]$, while a point $(x, y)$ in the original space is still transformed into a line in each of the three subspaces. As can be seen in figure 5, this can also be interpreted as an inhomogeneous discretization of the unbounded parameter space, with cells growing larger as they get further away from the origin. The same subdivision is also applied to the $(x, y)$-image space.

A more detailed discussion of the CHT and examples of its output are given Tuytelaars *et al.* (1997, 1998).

## 7. A strategy for grouping

The foregoing ideas can be combined in a grouping strategy that does away with most of the combinatorial search involved in pairing edge segments. On the one

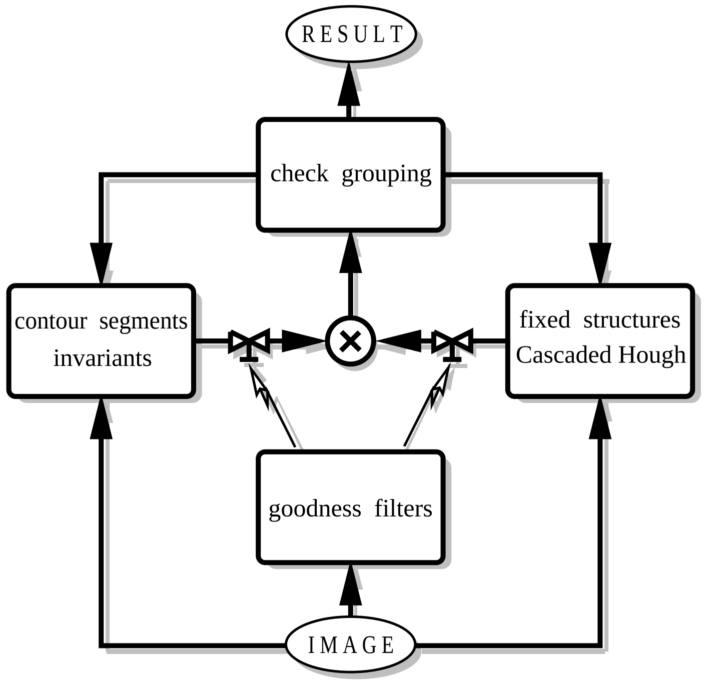


Figure 6. Proposed grouping strategy (see text).

hand, the CHT is used to quickly find potential fixed structures. On the other hand, the invariants that come with the corresponding subgroups are used to quickly group edge segments, based on hashing techniques.

Of course, no guarantee can be given that the CHT finds all fixed structures of all image groupings, but this is not necessarily required. Some groupings will be formed efficiently based on general projective invariants. In that case no fixed structures need to be known. The fixed structures of other groupings are found by reusing those already known, e.g. through the analysis of the transformations between grouped edge segments (i.e. by extracting their eigenvectors and eigenvalues). Regularly, different groupings share some of their fixed structures.

The grouping strategy can be sketched as in figure 6. The strategy, which has not been completely implemented yet, proceeds along two simultaneous tracks. On the one hand, invariants are exploited at the earliest opportunity. Initially, projective invariants are calculated for segments spanned by bitangent lines. These are matched and the matches might already yield some groupings. If there aren't many, these groupings can be analysed by considering the projective transformations that bring the segments in registration. The fixed structures of these transformations can be reused for other groupings. In parallel, the cascaded Hough scheme yields candidate fixed structures. Starting with the strongest candidates (getting the most support from image features), the invariants of the corresponding subgroups are used, mainly to those contour segments that have contributed to the extraction of the fixed structures. The strategy is to use as few fixed structures in combination as possible. Assuming a structure is fixed introduces the risk that this assumption may be wrong. Hence, if an assumption is made that several structures are fixed simultaneously, the chance of errors increases. Strong assumptions are only made as a last resort. The 'goodness filters' introduce the possibility of letting application-specific knowledge play a role and to rank hypotheses in the order in which they should be tried. Fixed structures can be given higher preference if fewer of them are combined, if they are more outspoken in the CHT, if they group longer edges, or if they do not yield too many possible matches.

Combinatorial procedures are avoided in both branches of this grouping scheme. Also note that, far from rejecting grouping based on general projective invariants, that is exactly what the system would try to achieve first.

## 8. Conclusions and future research

The use of invariants for the matching of object contours is a widespread technique. Here, grouping-specific transformation groups and invariants were derived. The major difference with more traditional groups is that some geometrical objects, fixed structures, need to be indentified for their practical use. The cascaded Hough transform was proposed as an efficient method to extract at least some of those. The actual grouping then proceeds as the interplay between hypothesizing fixed structures, but as few and least far-fetched as possible, and invariant-based matching. This is the subject of ongoing research, and much work remains to be done to arrive at a fully automatic grouping algorithm.

Future work will see a further integration of the cases discussed here: the coexistence of the three types of fixed structures—points/lines, sets of points, conics—will be considered. Another issue is the definition of the 'goodness filters' in figure 6, which determine the order of operations. The research will also be directed more strongly towards 3D patterns.

The work reported in this paper has been supported by the Flemish Fund for Scientific Research (FWO).

## References

Carlsson, S, Mohr, R., Moons, T., Morin, L., Rothwell, C. Van Diest, M., Van Gool, L., Veillon, F. & Zisserman, A. 1996 Semi-local projective invariants for the recognition of smooth plane curves. *Int. J. Computer Vision* **19**, 211–236.

Glachet, R., Lapreste, J. & Dhome, M. 1993 Locating and modelling a flat symmetric object from a single projective image. *Computer Vision Graphics Image Processing: Image Understanding* **57**, 219–226.

Kanade, T. 1981 Recovery of the 3-dimensional shape of an object from a single view. *Artif. Intell.* **17**, 75–116.

Lowe, D. 1984 Perceptual organization and visual recognition. Stanford University Technical Report STAN-CS-84-1020.

Lowe, D. 1985 *Perceptual organisation and visual recognition*. Kluwer Academic.

Rothwell, C. 1993 Recognition using perspective invariance. Ph.D. thesis, University of Oxford, UK.

Rothwell, C. & Stern, J. 1996 Understanding the shape properties of trihedral polyhedra. *Proc. Eur. Conf. on Computer Vision, Cambridge, UK*, pp. 175–185.

Rothwell, C., Forsyth, D., Zisserman, A. & Mundy, J. 1993 Extracting projective structure from single perspective views of 3D points sets. *Proc. 3rd Int. Conf. on Computer Vision, Berlin*, pp. 573–582.

Semple, J. & Kneebone, G. 1979 *Algebraic projective geometry*. Oxford: Clarendon Press.

Sinclair, D., Christensen, H. & Rothwell, C. 1996 Using the relation between a plane projectivity and the fundamental matrix. *Proc. Scand. Conf. Image Analysis*, pp. 181–188.

Springer, C. 1964 *Geometry and analysis of projective spaces*. Oxford: W. H. Freeman.

Tuytelaars, T., Proesmans, M. & Van Gool, L. 1997 The cascaded Hough transform as support for grouping and finding vanishing points and lines. *Proc. Int. Workshop on Algebraic Frames for the Perception–Action Cycle, Kiel, Germany*, pp. 278–288.

Tuytelaars, T., Van Gool, L. J., Proesmans, M. & Moons, T. 1998 A cascaded Hough transform as an aid in aerial image interpretation. *Proc. Int. Conf. on Computer Vision*, pp. 67–72.

Van Gool, L. 1997 A systematic approach to geometry-based grouping and non-accidentalness. *Proc. Int. Workshop on Algebraic Frames for the Perception–Action Cycle, Kiel, Germany*, pp. 126–145.

Van Gool, L. J., Moons, T., Pauwels, E. & Oosterlinck, A. 1992 Semi-differential invariants. In *Applications of invariance in vision* (ed. J. L. Mundy & A. Zisserman), pp. 157–192. Boston, MA: MIT Press.

Van Gool, L. J., Moons, T. & Proesmans, M. 1995 Groups for grouping: a strategy for the exploitation of geometrical constraints. *Proc. 6th Int. Conf. on Computer Analysis of Images and Patterns, Prague, Czechia*, pp. 1–8.

Van Gool, L. J., Moons, T. & Proesmans, M. 1996 Mirror and point symmetry under perspective skewing. *Proc. Conf. on Computer Vision and Pattern Recognition, San Francisco, USA*, pp. 285–292.

Wertheimer, M. 1938 Laws of organization in perceptual forms. In *A source-book of gestalt psychology* (ed. D. Ellis), pp. 71–88. London: Harcourt, Brace.

Zisserman, A., Mundy, J., Forsyth, D. & Liu, J. 1995 Class-based grouping in perspective images. *Proc. Int. Conf. on Computer Vision*, pp. 183–188.

## *Discussion*

A. Zisserman (*Department of Engineering Science, University of Oxford, UK*). Professor Van Gool has talked about fixed structures of symmetries under projectivities. Has he thought about relations where the symmetry isn't exact? Not quite a bilateral symmetry, for example.

L. J. Van Gool. Yes, I've thought about that. But the problem is that invariants don't allow too many changes of such type. That's a general problem with invariants, I think, that it's very strict, so quasi-invariants is certainly an interesting topic. What I was planning to do is look at combinations of these same building blocks. What you do for strict invariance is take these building blocks, raise them to some unknown power, and then simply solve for a system of linear equations to make sure all the factors drop out. What you could try to do for quasi-invariance is to look at variations of the logarithms of the building blocks. Using principal components of least variation can then suggest optimal exponents for quai-invariants under more general variability of shape. But I don't think it's a really good answer.

J. L. Mundy (*GE Corporate Research and Development, New York, USA*). Could Professor Van Gool say a bit more about the stability of these fixed structures? We experimented a little bit with decomposing the transform matrix into fixed points and lines and so forth, essentially, parametrizing the transform matrix with the parameters of the fixed structures, but found that the recovery was somewhat unstable. Is this our own poor numerical analytic method?

L. J. Van Gool. Is Dr Mundy suggesting trying to recover the fixed structure from the projective matrix?

J. L. Mundy. Yes, suppose you match two curves and compute the transform matrix with some kind of least-squares matching. And then compute the fixed structure from that and see if they are really fixed.

L. J. Van Gool. And they are not?

J. L. Mundy. No.

L. J. Van Gool. I don't have experience with that approach. What we have looked at so far is finding fixed structures through other means like the Hough transform and other methods. In that case, you find the fixed structure first and based on that try to find the mapping between corresponding contour segments. But obviously what you are suggesting would be part of a more complete grouping strategy. Because if you have found a grouping already, you would like to find out about further fixed structures that it has. That could be done by analysing the corresponding projectivity. We are slightly worried there because when are two eigenvalues identical? Stuff like that will really creep in and make difficulties. We haven't tried to get fixed structures from projective matrices so far. As I said, we have found them through other means like what we call the cascaded Hough transform.



PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES