Planar Homologies as a basis for Grouping and Recognition

Luc Van Gool and Marc Proesmans
(ESAT/MI2, Kath. Univ. Leuven,
Kard. Mercierlaan 94, 3001 LEUVEN, Belgium)

Andrew Zisserman
(Robotics Research Group, Dept of Eng Science, Univ. of Oxford,
Oxford OX1 3PJ, UK)

March 20, 1997

Abstract

The paper discusses a specific class of planar projective transformations, planar homologies, and illustrates their importance for geometry based grouping operations. Indeed, planar homologies keep to pop up in several areas of computer vision. Two examples are given in the paper: the analysis of planar shapes and their shadows and the detection of extruded shapes. The parameters that are needed to specify homologies are given, as well as their invariants. Since only 5 parameters are required, these invariants are simpler than general projective invariants. The work therefore further corroborates the existence of grouping-specific invariants.

*Support by Esprit BRA 'VIVA' and Esprit LTR 'IMPACT' is gratefully acknowledged. LVG and MP also gratefully acknowledges support by the Flemish government, through a project funded by the Institute for the Promotion of Scientific and Technological Research in Industry IWT. The authors would also like to thank Joe Mundy for supplying part of the image material. The authors wish to thank the reviewers for their useful comments.

1 Introduction

Invariants are now a common tool in the computer vision community for model based recognition. Their use up to this point has generally been inter-image, either for matching between a geometric model of an object in a model library and an image, or matching shapes between images. However, scenes often contain structures that are geometrically related, and these relationships lead in turn to constraints or particular transformations within a single image - i.e. intra-image invariants. One class of these transformations, planar homologies, are the central theme of this paper. They arise when two planar shapes in the scene are related by a 3-D perspectivity. The practical importance of planar homologies will be illustrated with two examples: shadows and extruded surfaces, of which typical scenes abound.

2 Properties of a Planar Homology

A plane projective transformation is a planar homology if it has a line of fixed points (the *axis*), together with a fixed point not on the line (the *vertex*) [10] (figure 1(a)). Algebraically, an equivalent statement is that the 3×3 matrix, $\mathtt{T}$, representing the transformation has two equal and one distinct eigenvalues. The corresponding ranks of $\mathtt{T} - \lambda \mathtt{I}$ have to be 1 and 2, resp. The axis is the join of the eigenvectors corresponding to the degenerate eigenvalues. The third eigenvector corresponds to the vertex. The ratio of the third to the other eigenvalue is a characteristic invariant of the homology.

The cross ratios defined by the vertex, a pair of corresponding points, and the intersection of the line joining these points with the line of fixed points, is the same for all points related by the homology. The vertex (2 dof), axis (2 dof) and cross ratio (1 dof) are sufficient to define the homology completely. One thus has 5 dof in specifying a planar homology.

Homologies and grouping Finding shapes that are related by a planar homology may start with three hypothesized pairs of corresponding points. These suffice to construct the vertex and the axis, as shown in fig. 1(b). Moreover, because three corresponding point pairs (6 dof) overdetermine the homology (5 dof), they allow critical tests to be performed to *verify* the grouping, e.g. that the vertex is on $a - a'$. Grouping related shapes based on

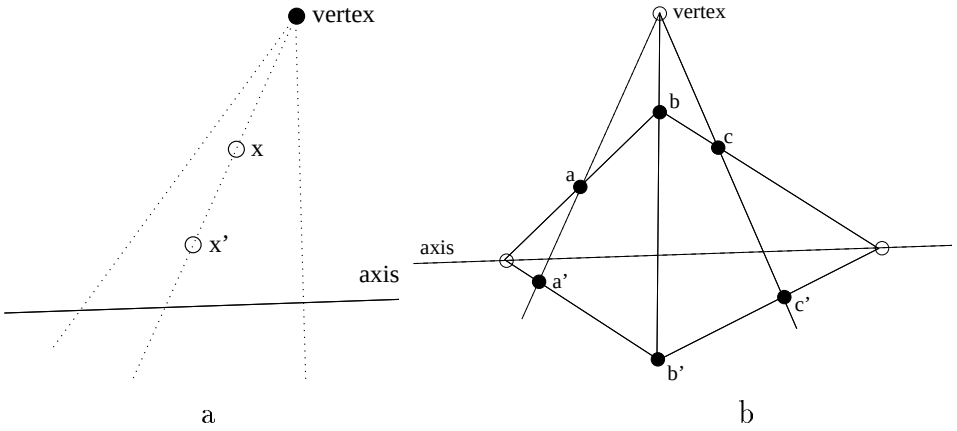


Figure 1: *(a) A planar homology is a plane projective transformation which has a line of fixed points, called the* axis*, and a distinct fixed point, not on the line, called the* vertex *of the homology. There is a pencil of fixed lines through the vertex. Corresponding point pairs* $\mathbf{x} \leftrightarrow \mathbf{x}'$ *and the vertex of the homology are collinear. The cross-ratio defined by the vertex, the corresponding points* $\mathbf{x}, \mathbf{x}'$ *and the intersection of their join with the axis, is the* characteristic invariant *of the homology, and the same for all corresponding points.*
(b) Three pairs of points that are corresponding under a planar homology suffice to find the vertex and the axis. The lines $b - b'$ *and* $c - c'$ *already yield the vertex as their intersection. The axis is found through the intersections of* $a - b$ *and* $a' - b'$*, and* $b - c$ *and* $b' - c'$*.*

projective equivalence would have required five pairs of corresponding points.

In the case that the cross ratio is harmonic, the transformation is called a planar harmonic homology. It has four dof since the cross-ratio is known. The transformation matrix $\mathtt{T}_H$ obeys $\mathtt{T}_H^2 = \mathtt{I}$, where $\mathtt{I}$ is the identity, i.e. the transformation is an involution. In a perspective image of a plane object with a bilateral symmetry, corresponding points in the image are related by a planar harmonic homology [7, 15]. The axis of the homology is the projection of the symmetry axis. Algebraically, $\mathtt{T}_H$ is a conjugate symmetry where the conjugating element is a plane projective transformation. Indeed, the conjugate symmetry is of the form $\mathtt{PSP}^{-1}$ with $\mathtt{S}$ the bilateral symmetry and $\mathtt{P}$ the projection from the object plane to the image plane. The detection of such symmetries has been discussed by several authors [11, 2, 15]. In an affine image the resulting transformation is a *skewed symmetry* [4], and the conjugating element is a plane affine transformation [6, 13].

3 Fixed sets and subgroups

Planar homologies have a fixed point (vertex) and a line of fixed points (axis). There is a direct relation between such geometrical entities that are mapped onto themselves (henceforth denoted the *fixed set*) and different subgroups, since a fixed set – a symmetry if you like – defines a subgroup. Working with smaller groups allows one to extract simpler invariants. A more general discussion on the study of projective subgroups through the specialization of fixed sets is given elsewhere [14]. In comparison to general homologies, *harmonic* homologies have additional pairs of mirror symmetric points in their fixed set. Invariants for the harmonic case and that are not invariant under general homologies are given in [2, 13, 15].

For any specialised projectivity, the major problem is to identify the fixed set in the image, for only then can the corresponding subgroup and the simpler invariants that come with it be exploited. Returning to planar homologies, the set of all such transformations does not form a group, but those with the same vertex and axis, i.e. the same fixed set, do. One has to find this vertex and axis in the image. The related shapes being part of a 3-D scene structure becomes crucial here. The detection of the fixed set can be supported by such structure in at least two ways:

1. as physical, visible features, as might happen with planar homologies

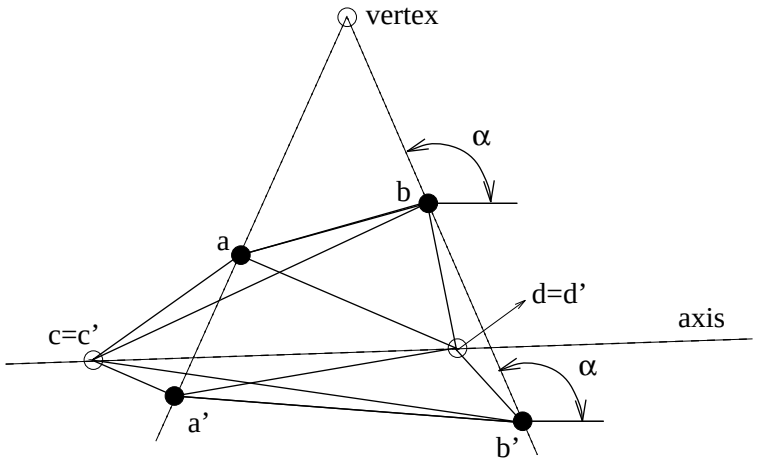


Figure 2: *Two basic types of invariants can be extracted if the vertex and axis of a planar homology are known.*

relating shadows and the planar objects that cast them, or

2. from point correspondences (cf. fig. 1(b)), e.g. for planar homologies relating sections through extruded surfaces, where physical edges may connect corresponding points.

Invariants for homologies Once the vertex and axis have been identified, the resulting planar homologies can only differ in the value of their cross-ratios. In fact, this remaining degree of freedom results in a 1-parameter subgroup of the projectivities. What are the invariants of such group?

Consider fig. 2. Looking at corresponding point pairs under a planar homology, two basic types of invariants can be used (for a more detailed discussion see [14]):

1. the angle α of the join between a point and the vertex measured with respect to an arbitary reference orientation (horizontal in the figure), and

2. the area ratio for triangle pairs, also indicated in the figure:

$$\frac{\text{Area triangle}(a,b,c)}{\text{Area triangle}(a,b,d)} = \frac{\text{Area triangle}(a',b',c')}{\text{Area triangle}(a',b',d')}$$

Note that the extraction of the first type requires knowing the vertex. The second type requires the axis. The first invariant embodies the constraint that the corresponding points and the vertex have to be collinear. The second invariant adds the constraint that lines defined by point pairs should intersect the axis in the same point if the two points are replaced by their corresponding points. This constraint also implies that it is easy to find corresponding points if one considers curves with their tangents, even if the vertex is unknown.

These invariants are close to trivial, yet not only do they support the extraction of the fixed set, they also allow simple, detailed tests on shape equivalence to make sure that shapes are really related by a planar homology.

4 Shadow Extraction and Grouping

The extraction of buildings and their structures is a first, important application where planar homologies arise naturally. The goal is to extract the line drawings of the buildings from aerial images. This is a notoriously difficult problem and, certainly in the case of buildings, missing out on a few lines or corners usually means being left with very little information. Therefore, one should bring to bear all the domain-specific aspects one can take advantage of. One of these is the presence of shadows, since aerial photographs have to be taken under good weather conditions. Another typical feature of buildings is that they are composed of planar structures. These aspects, when brought together, can turn shadows – traditionally a complicating factor due to the spurious edges they introduce (but also see [9]) – into a means of completing line drawings of buildings.

The basic observation underlying this application is that a planar object and its shadow (cast on a planar ground plane) are related by a planar homology. The vertex is the image of the light source. The axis is the image of the intersection between the object and shadow planes.

Depending on the weather conditions, the axis may be visible as a physical edge in the image. The vertex, i.e. the sun, is normally not visible in the image. However, once its position is determined, it can be used for the whole of what typically is a large image with many buildings. If for the description of a building both the vertex and the axis are available, a single shape-shadow point pair is sufficient for exploiting the transformation's special

characteristics to the full. The value of the cross-ratio becomes available and yields sufficient information to find the corresponding point for any additional point of the shape or the shadow. Therefore, if the outline of the shadow can be determined, additional points on the building can be found, even if local contrast of the edges is low or some vertices are very obtuse for the downward looking camera. Another use is checking coplanarity. If a shape-shadow pair yields a different cross-ratio, the point on the shape cannot lie in the same plane as the other points.

A shadow-based carthographic tool A carthographic tool is under development, that is intended as a flexible aid for human operators who's job it is to delineate building structures. Consider the example in fig. 3. In (a), a hangar is shown that casts a shadow on the ground plane. Some structure can easily be detected (manually) from this image. In particular, the points shown in (b) that are labeled P_1, P_1', P_2, P_2', and P_3. Also the point a indicated in (c) is easy to extract with appropriate precision. In this case the axis (intersection of the hangar front plane and the ground plane) does not clearly stand out. The tool assists in its extraction. As a first option, consider the construction of (b). The two pairs of easily detectable points P_1-P_1' and P_2-P_2' yield the fixed point as the intersection of the two solid lines. The point P_3' corresponding to P_3, which is rather difficult to extract directly from the shadow, is then determined with good precision as the intersection of the shadow boundary and the dashed line connecting the vertex and P_3. The three pairs of corresponding points then enable the determination of the axis following the construction of fig. 1(b), as shown. Alternatively, having found the point pairs (P_1, P_1') and (P_2, P_2') and the point a, the intersection of the lines P_1-P_2 and P_1'-P_2' suffices to yield a second point on the axis, and thus to fix it. Part (c) of the figure shows yet another technique, that identifies the point b based on the constraint

$$\frac{\text{Area triangle}(b, P_1, P_2)}{\text{Area triangle}(a, P_1, P_2)} = \frac{\text{Area triangle}(b, P_1', P_2')}{\text{Area triangle}(a, P_1', P_2')}$$

in order for it to lie on the axis. Finally, points P_4 and P_5 are especially difficult to find directly, whereas their counterparts P_4' and P_5' on the shadow line are easier to extract. The tool yields P_4 and P_5 upon clicking on P_4' and P_5'. This transfer is carried out using the invariant cross ratio of the planar homology.

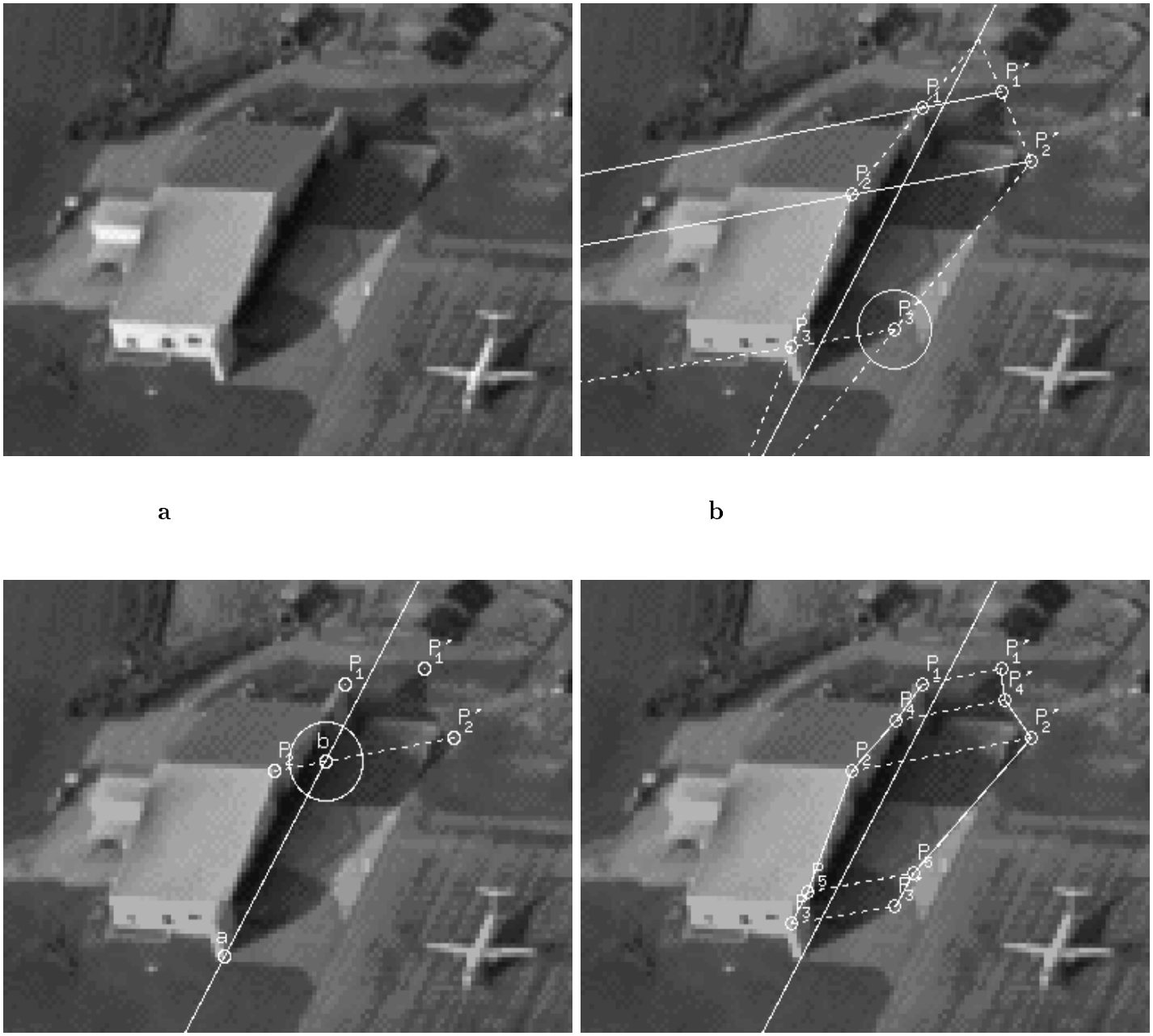


Figure 3: *(a) Scene with hangar and its shadow, (b) method for the construction of the axis using three point pairs, (c) alternative construction of this line, (d) building structure completion using the shadow.*

Affine simplifications As can be observed from the example image, the light source is distant, also in the image. In most (but not all!) aerial photographs, the shadow lines can be considered parallel. In that case the cross-ratio simplifies to a ratio of distances along these lines to the axis. This ratio will therefore be constant. Ponce's [8] criterion for skewed symmetry can e.g. easily be generalized towards such configurations assuming the distance ratio k (distance of shadow points to the axis over the distance of the object points to the axis) is known:

$$\frac{\kappa'}{\kappa} = -k \left(\frac{\sin \theta'}{\sin \theta} \right)^3$$

with κ' and κ curvatures at corresponding points and θ' and θ the angle subtended by the tangents at these points and the chord joining them. Primes are used for the shadow. Such configurations are a natural extension to the list of non-accidental properties that can be used for grouping [3, 5].

5 Extruded shapes

Extruded shapes are instances of 3D configurations with perspectively related structures. Indeed, such shapes can be modeled as the part of a cone between two planes that generate the shapes' top and base plane (see figure 4). Extruded shapes are extremely common. Examples include most tin cans, boxes, books, bottles, etc.

There is a plane projectivity $\mathtt{T}$ between the perspective structure's top and base curves in the image, which is a planar homology. The vertex $\mathbf{v}$ is the image of the apex of the cone, and the axis $\mathbf{L}$ is the image of the intersection line of the top and base planes (figure 4).

Planar invariants The line of intersection of the top and base planes can be measured in the image following the construction of fig. 1(b) (also see [1]). Figure 5(b) shows an example on a real image. Consequently, an additional line is available which is coplanar with the top and base curves. Invariants can be computed for these curves together with the line. For example, if the curve consists of four lines, then two plane projective invariants can be measured for the five lines. These are important for inter-image type of operations (e.g.

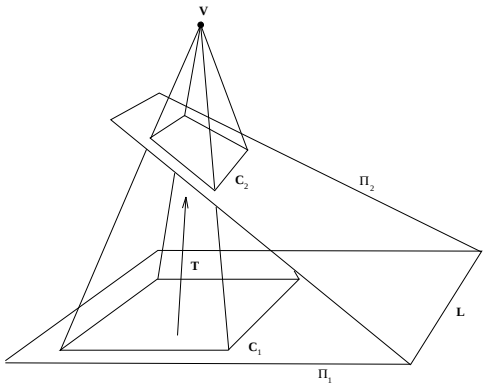


Figure 4: *The top and base image curves of a perspectively related structure are related by a planar homology* T*. Corresponding points lie on lines through* **v***, which is the fixed point of the transformation (vertex). The line* **L***, which is the image of the intersection of the two planes that cut the "cone", is a line of fixed points of the transformation (axis).*

for recognizing the object). Note that for finding corresponding top and base curves (an intra-image operation) simpler invariants using two points and the axis exist, as specified earlier. In both cases, the planar homology simplifies the analysis.

Affine invariants In the case that the top and base planes on the 3D object are parallel, the line of intersection is at infinity, and **L** is the vanishing line of the planes. Consequently, affine invariants of the top and base curves can be measured in the image (since the line at infinity has been identified for their planes). For example, if the curve consists of four lines there are two plane affine invariants. Affine invariants of each curve (which are the same for the top and base) are computed by removing the perspective distortion, i.e. transforming the curve so that the vanishing line is sent to the line at infinity.

It is worth noting that in the case of the vertex lying on the plane at infinity, the top and base planes are then related by a translation. That affine information is available in this case follows from the fact that under pure translation affine 3D structure can be recovered from uncalibrated cameras [12].

For the example shown in figure 5, the eigenvalues of the transformation



Figure 5: *(a) Original image, the polyhedral prism is a perspectively related structure. The correspondences used to compute the planar homology are highlighted in white. (b) At the top of the figure the line of fixed points, and at the very bottom the fixed point (the line intersection) computed from the eigenvectors of the planar homology (the other polyhedron is included for reference).*

object	raw	affine
1.000	1.121	1.000
1.000	0.828	0.999

Table 1: *Affine invariants measured from the four vertices on the top face of the polyhedral prism of figure 5. The invariants are the ratios of triangles which divide the rectangular face. In the table:* 'Object' values *are computed from the object;* 'raw' values *are computed directly from the image coordinates of the vertices, these are not unity due to the perspective distortion of the rectangle;* 'affine' values *are computed after transforming the points so that the vanishing line of the plane is sent to infinity.*

T are

$$\lambda = \{1.1946, 1.1946, 0.89482\},$$

and the equation of the computed line $\mathbf{L} = 0.061x + 0.625y + 805.721 = 0$ (x and y measured in pixels on a 772 by 575 image). The point transformation matrix which transforms this measured line to the line at infinity is:

$$\mathbf{x} = \mathtt{T}\mathbf{X} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_1 & l_2 & l_3 \end{bmatrix} \mathbf{X}$$

where $\mathbf{L} = (l_1, l_2, l_3)^\top$. Results are given in table 1. Clearly, the measured invariants are very accurate indicating that the perspective distortion has been very effectively removed.

6 Conclusions

The paper has illustrated the practical importance of planar homologies for intra-image but also inter-image analysis. The reported work underlies several directions of current investigations. On the theoretical side, ongoing research aims at extending the results towards different types of subgroups of the projectivities, based on their fixed sets. From a practical point of view, these subgroups offer strong potential for grouping. This is all the more so because related shapes quite often share part of their fixed sets. As examples,

shape-shadow pairs share the same vertex (the sun) in aerial photographs and the dominant class of extruded surfaces with sections perpendicular to their "spine" share the same axis if they are resting on parallel planes.

References

[1] D. Forsyth and C. Rothwell, Representations of 3D objects that incorporate surface markings, in *Applications of Invariance in Computer Vision*, eds. Mundy, Zisserman, & Forsyth, LNCS 825, Springer-Verlag, 1994, 341-357.

[2] R. Glachet, J. Lapreste, and M. Dhome, Locating and modelling a flat symmetric object from a single projective image, *Computer Vision, Graphics, and Image Processing*, **57**(2), 1993, 219-226.

[3] T. Kanade, Recovery of the 3-Dimensional shape of an object from a single view, Artificial Intelligence 17, pp.75-116, 1981

[4] T. Kanade and J. Kender, Mapping image properties into shape constraints: skewed symmetry, affine-transformable patterns, and the shape-from-texture paradigm, CMU-CS-80-133, july 1980

[5] D. Lowe, Perceptual organization and visual recognition, Stanford University, Dep. of Computer Science, Report No. STAN-CS-84-1020, sept. 1984

[6] D. Mukherjee, A. Zisserman, and M. Brady, Shape from symmetry – detecting and exploiting symmetry in affine images, *Phil. Trans. Royal Soc. London A*, 351, pp. 77-106, 1995

[7] J. Mundy and A. Zisserman, Repeated structures: image correspondence constraints and 3D structure recovery, in *Applications of Invariance in Computer Vision*, eds. Mundy, Zisserman & Forsyth, Springer, 1994, 89-106

[8] J. Ponce, On characterizing ribbons and finding skewed symmetries, *Proc. Int. Conf. on Robotics and Automation*, 1989, 49-54

[9] J. Shufelt and D. McKeown, Fusion of monocular cues to detect man-made structures in aerial imagery, *Computer Vision, Graphics, and Image Processing: Image Understanding*, **57**(3), 1993, pp. 307-330

[10] C. Springer, *Geometry and analysis of projective spaces*, Freeman, 1964

[11] F. Ulupinar and R. Nevatia, Constraints for interpretation of line drawings under perspective projection, *Computer Vision, Graphics, and Image Processing: Image Understanding*, **53**(1), 1991, 88-96

[12] L. Van Gool, T. Moons, M. Proesmans, and M. Van Diest, Affine reconstruction from perspective image pairs obtained by a translating camera, *Proc. Int. Conf. Pattern Recognition*, Jerusalem, 1994, A/290-A/294

[13] L. Van Gool, T. Moons, D. Ungureanu, and A. Oosterlinck, The characterisation and detection of skewed symmetry, *Computer Vision, Graphics, and Image Processing: Image Understanding*, **61**(1), 1995, 138-150

[14] L. Van Gool, M. Proesmans, and T. Moons, Groups for grouping, SPIE Int. Symp. on Optical Science, Appl. of Digital Image Processing XVIII, Vol.2564, pp.402-413, 1995

[15] L. Van Gool, T. Moons, and M. Proesmans, Mirror and point symmetry under perspective skewing, IEEE Conf. Computer Vision and Pattern Recognition, pp. 285-292, San Francisco, june 1996