Model Construction from a Single Perspective View using Shape from Symmetry.

C. I. Attwood, G. D. Sullivan and K. D. Baker

Intelligent Systems Group,
Department of Computer Science,
University of Reading, RG6 2AY.
Charlie.Attwood@reading.ac.uk

Abstract

A new method is described - Structure from Symmetry (SFS) - in which objects possessing bilateral symmetry may be reconstructed from a single perspective view. The algorithm exploits the groundplane constraint [9, 14] and assumes a camera model, to allow 3D recovery up to scale. The SFS algorithm has been implemented in the form of an interactive tool. The user marks a set of symmetric object points in an image, and a complete facet model description is generated as output. The tool has been extensively used to create models for the vision system developed in the VIEWS project. The algorithm has potential for automatic structure recovery if methods for identifying symmetry points can be improved.

1 Introduction

Geometrical models are widely used to encode the structural knowledge necessary to solve vision problems, and yet methods for acquiring models have largely been ignored by the computer vision community [1]. This has in turn, lead to systems which are limited by small object databases. The model requirements for vision are often different from those of other modelling application areas, such as computer graphics and CAD. Vision models require economy of representation; they must make the position of model features explicit under any transformation; and they should distinguish between the image characteristics of different model features such as crease, fold, or extremal edges, surface marks etc.

Manual construction methods are time-consuming, and require direct physical measurements of the object to be obtained, which is not always possible. *CAD* systems whilst providing some of the required properties, contain much detail which is unnecessary to the vision task and must therefore be filtered to extract the salient features [2]. Techniques using construction from multiple views, require that objects be viewed from several different view points, or, where structure from motion is concerned, that several images are captured from the same view point over time. Data used is either intensity, or more commonly, actively sensed range or depth data [3-8, 10]. These techniques are powerful but

BMVC 1993 doi:10.5244/C.7.16

require integrating data from several sources, which can lead to increased error due to the correspondence problem.

Many objects in the world possess a particular form of symmetry - bilateral symmetry. This is especially true for man-made objects, but also occurs frequently in nature. Kanade [11, 12] demonstrated recovery of objects from a *single* view. He proposed a variety of techniques, involving the use of geometrically invariant properties of objects under projection. Of particular relevance here was his use of skewed symmetry. Skewed symmetry refers to a class of 2D shapes in which symmetry is found along lines at a fixed angle to an axis. Such shapes can be described as 2D-Affine transforms of real symmetries.

Recent work by Glatchet et al. [13], has shown that hand segmented images of *flat* bilaterally symmetric objects of known height may be located from a single perspective projection, and modelled up to a scaling factor. We describe a new Structure from Symmetry (SFS) algorithm. Using the groundplane constraint [9, 14], and certain geometrical properties of symmetrical points, it allows the full 3D structure of solid objects possessing bilateral symmetry to be recovered up to a scaling factor, from a *single* perspective view. The SFS algorithm is at present embodied in an interactive tool, which requires the user to identify symmetrical control-points.

2 The Symmetry Constraint

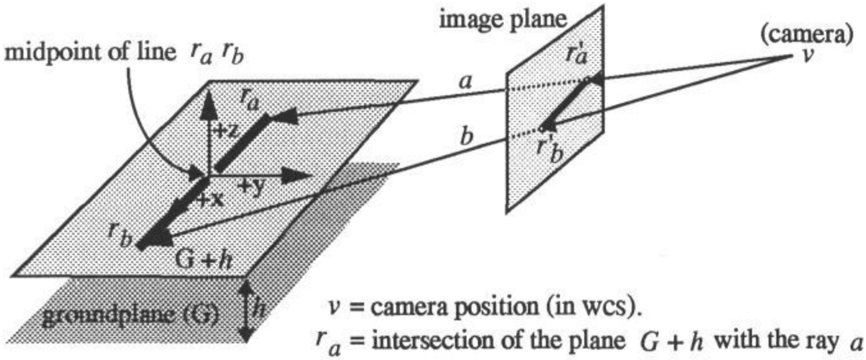
The SFS algorithm assumes a reasonably accurate camera model, giving the camera's position in the world coordinate system. The method used to estimate the 3D position of a point from its 2D image coordinates involves two stages. Firstly, the position of one pair of symmetry points is used to construct a model-centred coordinate system (mcs) up to an arbitrary scale. Secondly, further point pairs are recovered, by making use of their symmetry relationship within the mcs. The user specifies the approximate height of one point of a symmetrical pair of vertices, on the object. This estimate determines the scale of the recovered structure, which may be corrected later (an accurate value gives the correct scale).

2.1 Recovery of a model centered coordinate system

The position of the camera (v) in world coordinates is known from the camera model. A point on the image (r'_a) identifies a ray a, on which the object vertex, r_a lies (see Figure 1). Under an assumption of the height (h) this point is, its 3D position may be found by intersecting the ray with the ground plane + h. This point has a symmetry pair, r_b, of the same height, so its position may be found in a similar manner. These two points now define an axis, with its mid-point lying on the object's symmetry plane. Since the ground plane is known, a second vertical axis may be defined (the groundplane normal), extending through the mid-point. These two axes and their cross product, define a full local cartesian coordinate system for the object, in which reflection about the symmetry plane is represented as a change of sign, of one coordinate. The transformation linking the wcs and the mcs is obtained and may now be used to convert known point positions between the two coordinate systems, as required in the next stage.

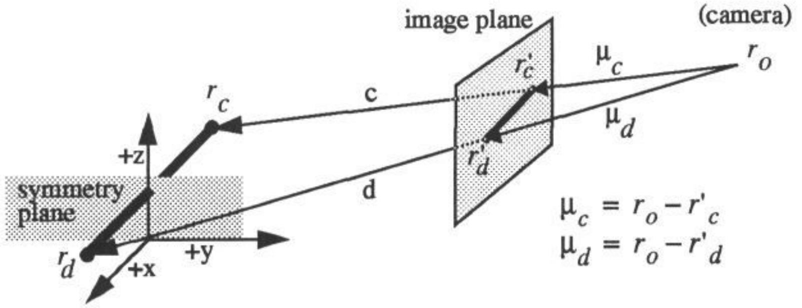
2.2 Structure from symmetry

Consider two unsolved symmetrical object points r_c and r_d (see Figure 2). The wcs to mcs transformation obtained in the previous section allows us to express the position of the camera and image points, in the mcs. The benefit of working in the mcs is that it allows the symmetry property, relating point pairs, to be expressed as simple constraints on their 3D location (see EQ 5-7 below). The



v = camera position (in wcs).
r_a = intersection of the plane $G+h$ with the ray a
r_b = intersection of the plane $G+h$ with the ray b

Figure 1. Recovery of a model-centred coordinate system.



$$\mu_c = r_o - r'_c$$
$$\mu_d = r_o - r'_d$$

Figure 2. Structure from Symmetry from the labelled points r'_c and r'_d.

position of an unknown object vertex must lie on the ray formed between the camera and the image position of that vertex. The problem is to find the distance along the ray at which the 3D point lies.

The unique solution to the 3D positions of the points r_c and r_d may be found by solving for λ_c and λ_d in the following line equations:

$$r_c = r_o + \lambda_c \mu_c \quad \text{(EQ 1)}$$

$$r_d = r_o + \lambda_d \mu_d \quad \text{(EQ 2)}$$

r_o is the position of the camera, and:

$$\mu_c = r_o - r'_c \quad \text{(EQ 3)}$$

$$\mu_d = r_o - r'_d \quad \text{(EQ 4)}$$

where r'_c and r'_d are the corresponding image points expressed in the mcs (see Figure 2.).

As the problem is set in a model-centred coordinate system, certain properties of the symmetry point pair can be utilised:

Let $\hat{x}$, $\hat{y}$, and $\hat{z}$ be the unit vector axes of the model centred coordinate system. From the symmetry constraint the following three properties can be derived; the x coordinates of the points are a simple reflection of each other (EQ 5), the y coordinates are identical (i.e. the points must both lie at the same distance along the symmetry plane) (EQ 6), and the points are at the same height (EQ 7).

$$r_c \cdot \hat{x} = -r_d \cdot \hat{x} \quad \text{(EQ 5)}$$

$$r_c \cdot \hat{y} = r_d \cdot \hat{y} \quad \text{(EQ 6)}$$

$$r_c \cdot \hat{z} = r_d \cdot \hat{z} \quad \text{(EQ 7)}$$

From the line equations, substituting for r_c and r_d into EQ (5-7) gives:

$$(r_o + \lambda_c \mu_c) \cdot \hat{x} = -(r_o + \lambda_d \mu_d) \cdot \hat{x} \quad \text{(EQ 8)}$$

$$(r_o + \lambda_c \mu_c) \cdot \hat{y} = (r_o + \lambda_d \mu_d) \cdot \hat{y} \quad \text{(EQ 9)}$$

$$(r_o + \lambda_c \mu_c) \cdot \hat{z} = (r_o + \lambda_d \mu_d) \cdot \hat{z} \quad \text{(EQ 10)}$$

Re-arranging produces a set of three simultaneous equations in the two unknowns, λ_c and λ_d:

$$\lambda_c (\mu_c \cdot \hat{x}) + \lambda_d (\mu_d \cdot \hat{x}) = -2 (r_o \cdot \hat{x}) \quad \text{(EQ 11)}$$

$$\lambda_c(\mu_c \cdot \hat{y}) + \lambda_d(\mu_d \cdot \hat{y}) = 0 \qquad \text{(EQ 12)}$$

$$\lambda_c(\mu_c \cdot \hat{z}) + \lambda_d(\mu_d \cdot \hat{z}) = 0 \qquad \text{(EQ 13)}$$

Equations 11-13 are now solved for λ_c and λ_d, using a linear least squares solution. Substituting back into the line equations 1 and 2, gives r_c and r_d; the positions of the symmetry points in the mcs.

3 Model-building Tool

The SFS algorithm has been implemented in the form of a model building tool, and has been used to produce over twenty models of vehicle types used in Esprit project P-2152 (VIEWS). A basic requirement of the model tool is that the object to be recovered is at least slightly oblique to the camera, so that three sides are in view including the plane orthogonal to the symmetry axis.

The operator uses a simple vertex labelling notation. This indicates the approximate order and handedness of points from the nearest end of the object to its furthest end. For example, 'op1' and 'np1' might define the lower edge of the front spoiler of a car, where the 'o' and 'n' prefixes refer to offside and nearside points.

Prior to using the system the user must specify a simplified facet structure for the object to be acquired. The facet structure is necessary for three reasons: (i) to aid in the regularisation of model planes, (ii) to minimise errors in the recovered points, by constraining points belonging to a facet to be co-planar and (iii) to allow the solution of the position of points without a visible symmetric partner (see below). Each point is interactively positioned on the object's vertices in the image by the user (see Figure 3.1). The SFS algorithm described in Section 2.0 is applied iteratively to the list of facets, on a per-facet basis, using the (u,v) point values assigned by the user. As the process of symmetry pair recovery is repeated, facets are progressively completed and then deleted from the list of facets to be solved. Objects with similar shape (e.g. some different makes of cars) may be reconstructed from the same facet definition structure, allowing rapid development of new models.

When a facet contains a point without a visible symmetric partner, the facet is skipped until at least three points in the facet to which the points belong, have been recovered. The point is then recovered in 3D by intersecting the ray passing through its image projection with an estimate of the plane of its facet formed from the three known points. The new point is now regularised with respect to its facet, and its occluded partner's position is found directly from reflection of the visible point in the symmetry plane. Any other occluded points in the facet are processed in the same manner, and the facet is deleted from the list. The algorithm terminates when all the facets have been processed.

Figure 3.0 shows a car selected as the ROI. In Figure 3.1, the labels assigned to the car vertices and tie-lines are shown marked on the scaled ROI image. Figure 3.2 shows two views of the recovered vehicle model. Figure 4.0 shows a more

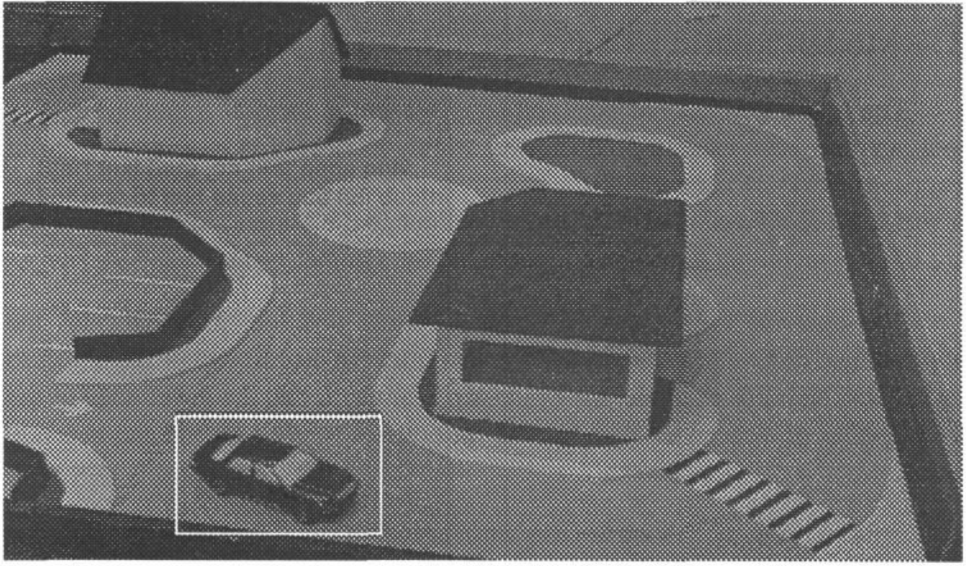


Figure 3.0. A "Toy-town" scene with a vehicular ROI selected.

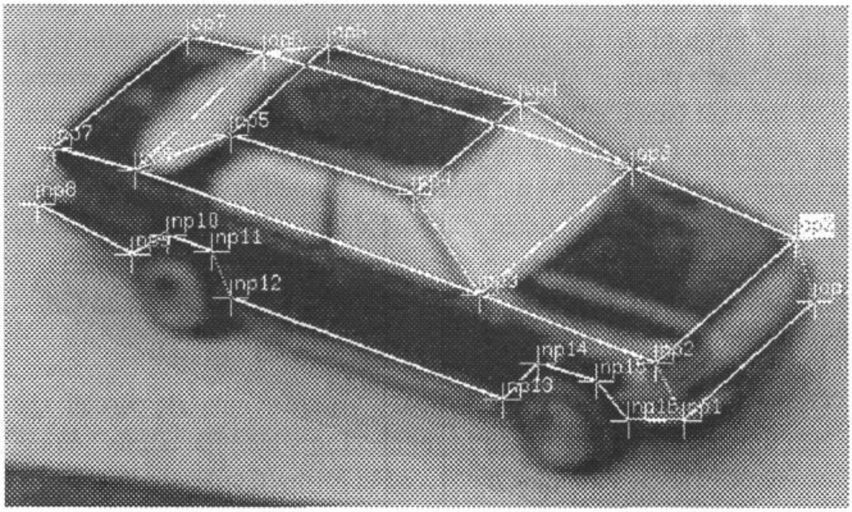


Figure 3.1. The ROI with symmetric point labels (tie-lines link visible 2D points to aid correct positioning of vertices).

difficult scene with the object of interest a long way from the camera. Despite this, an acceptable model of the aircraft tractor was extracted and is shown in Figure 4.1.

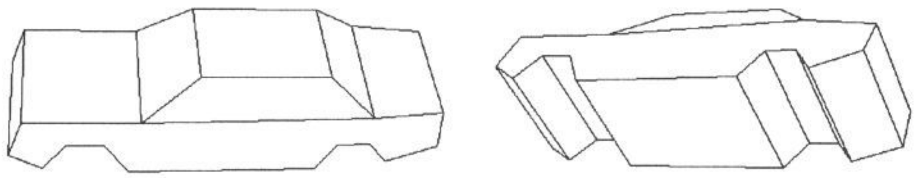


Figure 3.2. Two views of the recovered vehicle model (Note the successful recovery of underbody facets despite there being only two visible points in the original image per facet).

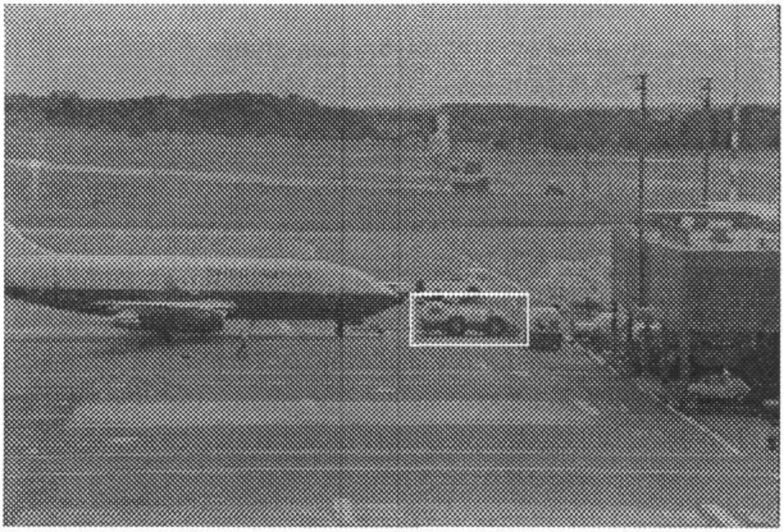


Figure 4.0. A difficult scene (Newcastle airport) where the ROI (a towing tractor) is a long distance from the camera.

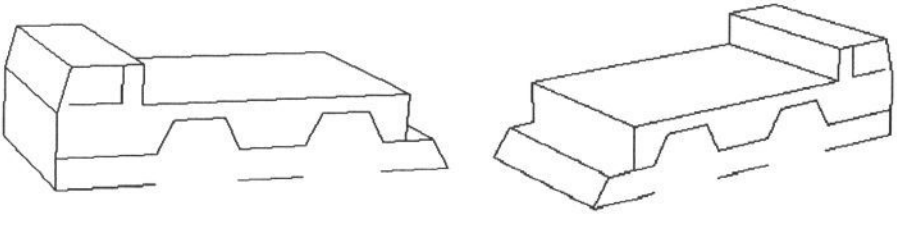


Figure 4.1. The recovered facet model, displayed at two new orientations (with some minor surface markings added).

4 Accuracy of Structure Recovery - An Experiment

The accuracy of the SFS algorithm was assessed using a test cuboid, 100x250x250mm in size, positioned 3.4 metres from the camera. The correct height of one point was given, so that recovered scale should be correct. In order to assess *structure* recovery accuracy a measure of percentage error was used - the difference between a recovered lines length and the true line length, divided by the true line length. The average percentage error for line length was 1.58% (range of 0.99% - 2.01%). This equates to just under 4mm error on the 250mm length lines, which is close to the maximum possible accuracy, given the uncertainty over vertex positioning in the image.

In addition to assessing the *structure* recovery accuracy it is also of interest to measure the algorithm's ability to recover absolute positional information. This stricter test - the Euclidean distance between the actual and recovered vertex positions - gave an average error of 8.71mm (range 6.19mm - 12.06 mm). Equivalent to an 0.26% error in depth along the ray.

The accuracy of recovery of points without a symmetry pair visible in the image (and which thus had to be recovered by fitting to a facet-plane estimate), was assessed separately. In the test cuboid image, one point was occluded. The error values for this point and its visible symmetry partner were 11.83mm and 10.76mm respectively. The adjusted average including these points was 9.36mm.

The results are quite satisfactory, and were obtained with a less than perfectly calibrated camera model. Experience to date shows that the SFS algorithm degrades gracefully with increasing camera model error.

5 Discussion and Further Developments

The SFS algorithm requires the correct identification of pairs of points in the image, which are known to lie on either side of an object's plane of bilateral symmetry. As currently implemented, an operator identifies the symmetry property relating point pairs, however it is possible to consider the automatic location of such point pairs.

Although edge detectors are notoriously bad at correctly detecting line-endings, methods such as corner-detection have improved the accuracy and reliability with which such features may be found. Objects such as vehicles contain a number of edges orthogonal to the symmetry plane, producing clusters of parallel lines in the image. By selecting a set of such lines, and their end-points, multiple hypotheses, about the object's local coordinate system may be obtained (by the method described in section 2.1), and some statistical choice made as to the most representative mcs. The end-points of remaining parallel lines in the image cluster would then be recovered in depth.

The requirement, in the implementation described above that a facet structure be pre-specified, is a consequence of the particular object model representation required by the VIEWS model system, rather than a requirement of the SFS algorithm itself, except in the role that facets play in solving for occluded points. The interactive tool has reduced model construction time from days to hours (compared with the manual construction methods previously used). The output from the tool is an ASCII model definition file in a format ready for use in the VIEWS vision system. A variety of object models have been built using the tool, including; cars, lorries, fuel tankers, baggage trucks, transit vans, water carriers, a small plane and a moving staircase.

Positioning of the model vertices markers by the user could be improved using a technique described in Rothwell *et al.* [16]. The lines passing through the first two symmetric point pairs intersect at an epipole. Subsequent symmetric pairs should both lie on a line passing through the same epipole. This property could be used to aid the accurate positioning of symmetric points by the user, helping to ensure that the assumptions of symmetry are met, and thus improving recovery accuracy.

6 Conclusion

Symmetry is an important cue to structure, and may play a significant role in human vision. Stevens [15] found that shapes with the property of skewed symmetry provide sufficient information for human observers to perceive surface orientation correctly.

An algorithm has been introduced, that utilises knowledge about the symmetry relationship between points on an object's surface, in order to allow accurate recovery of its surface shape from a single perspective view. The algorithm is currently restricted from use as a stand-alone technique by the difficulty of identifying symmetry points in the image, prior to recovery. It has however demonstrated that it can be incorporated into a practical tool for model acquisition. Future work may lead to its use in an autonomous form, if knowledge of symmetry can be extracted from the image prior to full object recognition.

Acknowledgements

This research was funded under ESPRIT P-2152 (VIEWS). We would also like to thank Anthony Worrall for his advice on the mathematical exposition of this work.

7 References.

[1] J. Y. Weng, T. S. Huang, and N. Ahuja. Motion and Structure from Two Perspective Views: Algorithms, Error Analysis, and Error Estimation. *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.11, no.5, 1989, pp.451-477.

[2] P. J. Flynn and A. K. Jain. CAD-Based Computer Vision: From CAD Models to Relational Graphs. *IEEE Trans. on Pattern Anal. Mach. Intell.*, vol.13, no.2, 1991, pp.114-132.

[3] C. H. Chien and J. K. Aggarwal. Volume/Surface Octrees for the Representation of Three-Dimensional Objects. *CVGIP*, vol.36, 1986, pp.100-113.

[4] C. Connolly *et al.* Matching from 3-D Range Models into 2-D Intensity Scenes. *Proc. of First Int. Conf. on Computer Vision*, London, June 1987, pp.65-72.

[5] F. Solina and R. Bajcsy. Recovery of Parametric Models from Range Images: The Case for Superquadrics with Global Deformation. *IEEE Trans on Pattern Anal. Mach. Intell.*, vol.12, no.2, 1990, pp.131-147.

[6] Y. R. Tabak and R. Jain. Building an Environment Model Using Depth Information. *Computer*, vol.22, no.6, 1989, pp.85-90.

[7] M. Potmesil. Generating Models of Solid Objects by Matching 3-D Surface Segments. *Proc. 1983 Int. Joint Conf. Artificial Intelligence*, Aug. 1983, pp.1089-1093.

[8] T. N. Tan, G. D. Sullivan, and K. D. Baker. Structure from Constrained Motion, *VIEWS project internal report*, RU-03-WP•T411-01, March 1991.

[9] T. N. Tan, G. D. Sullivan, and K. D. Baker. 3D Structure and Motion Estimation from 2D Image sequences. *Proc. of British Machine Vision Conference*, 69-78, 1992, Springer-Verlag, 1992.

[10] J. R. Stenstrom and C. I. Connolly. Constructing Object Models from Multiple Images. *International Journal of Computer Vision*. 9:3, 185-212, 1992.

[11] T. Kanade. Recovery of the Three-Dimensional Shape of an Object from a single View. *Artificial Intelligence*, 17, 409-460, 1981.

[12] T. Kanade. Geometrical Aspects of Interpreting Images as a Three-Dimensional Scene. *Proc. I.E.E.E*, vol. 71, 798-802, 1983.

[13] R. Glatchet, J. T. Lapreste, and M. Dhome. Locating and Modelling a Flat Symmetric Object from a Single Perspective Image. *Computer Vision Graphics and Image Processing: Image Understanding*, vol. 57, no. 2, 219-226, 1993.

[14] T. N. Tan, G. D. Sullivan, and K. D. Baker. Linear Algorithms for Object Pose Estimation. *Proc. of British Machine Vision Conference*, 600-609, 1992, Springer-Verlag, 1992.

[15] K. A. Stevens. *Surface Perception from Local Analysis of texture*, Ph.D. Thesis, Tech Report 512, MIT, Cambridge, MA, 1979. Cited by T. Kanade (1981), op.cit.

[16] C. A. Rothwell, D.A. Forsyth, A. Zisserman & J.L. Mundy. Extracting Projective Structure from Single Perspective Views of 3D Point Sets. *Proc. Fourth Int. Conf. on Vis.* Berlin, Germany, 1993