The Study of 3D-from-2D using Elimination

Michael Werman

Hebrew University of Jerusalem
Institute of Computer Science
91904 Jerusalem, Israel
e-mail: werman@cs.huji.ac.il

Amnon Shashua*

Hebrew University of Jerusalem
Institute of Computer Science
91904 Jerusalem, Israel
samm@cs.huji.ac.il

Abstract

This paper unifies most of the current literature on 3D geometric invariants from point correspondences across multiple 2D views by using the tool of Elimination from algebraic geometry. The technique allows one to predict results by counting parameters and reduces many complicated results obtained in the past (reconstruction from two and three views, epipolar geometry from seven points, trilinearity of three views, the use of a priori 3D information such as bilateral symmetry, shading and color constancy, and more) into a few lines of reasoning each. The tool of Grobner base computation is used in the elimination process.

In the process we obtain several results on N-view geometry, and obtain a general result on invariant functions of 4 views and its corresponding quadlinear tensor: 4 views admit minimal sets of 16 invariant functions (of quadlinear forms) with 81 distinct coefficients that can be solved linearly from 6 corresponding points across 4 views. This result has non-trivial implications to the understanding of N-view geometry. We show a new result on single-view invariants based on 6 points and show that certain relationships are impossible.

One of the appealing features of the elimination approach is that it is simple to apply and does not require any understanding of the underlying 3D from 2D geometry and algebra.

1 Introduction

The area of 3D-from-2D geometry, namely the body of results and methods on recovering and manipulating general 3D information from 2D projections, is often fragmented into several fundamental questions that are loosely related. When corresponding points are identified across the projections of a 3D object, one may ask the following questions:

1. Can the camera positions be recovered? If yes, how many views are necessary and how many points? How many solutions are there? What can recovered from uncalibrated cameras?
2. Given more views than necessary, and/or more points than necessary, do we obtain fundamentally different results? For example, can we tradeoff views with points? Do fundamentally non-linear processes become linear when we add views/points?
3. Are there general results for N views? I.e., is there a "view" manifold that captures the equivalence class of views of a 3D object?
4. Are there functions of views that are object-invariant? If yes, how many? What is their general form?
5. What can we tell from a single view in general? What kind of a priori information on the 3D world permits general results from a single view?

Partial answers exist for each of these items, some of the answers have been known for a while in the computer vision community, some can be traced back to the mid 19th century and the beginning of this century, and some answers are very recent. But overall, the answers are fragmentary and ranging over a wide spectrum of tools and notations.

For example, it is known that 5 corresponding points across two views provide a solution to the relative (calibrated) camera positions up to a 10-fold ambiguity [7]; 7 corresponding points provide a solution in a projective setting (uncalibrated cameras) up to a 3-fold ambiguity [7]. The case of 6 points has not been addressed in the past. A unique *linear* solution can be obtained from 8 points [15], and likewise in the projective case [6, 11, 27]. The "epipolar geometry" is captured by an invariant function of two views having a bilinear form [15, 6] — whether this is the only invariant function of two views is an issue that has not been satisfactorily resolved.

Results in the case of N ($N > 2$) views are sporadic. For $N = 3$, 6 corresponding points are sufficient for a projective reconstruction of camera position, up to a 3-fold ambiguity [25]; 7 points provide a unique *linear* solution [8, 30]. Invariant trilinear functions of 3 views come in linearly independent sets of 4 functions, have 27 distinct coefficients, and can be recovered linearly from 7 points [28]. The minimal number

*Starting from fall 1995, A. Shashua will be at the Computer Science Department of the Technion, Haifa 32000, Israel.

of algebraically independent functions of 3 views is 3 [37]. The notion of "view surfaces" that captures the N-view geometry can be found in special cases — 1D images in [34] and images under parallel projection in [13].

In this paper we approach 3D-from-2D geometry issues in a general and unified manner. The approach relies on tools taken from symbolic algebra — specifically, *elimination* theory using Grobner bases (cf. [33, 5], and for review of applications of Grobner bases for geometric theorem proving, computer graphics and robotics, see [12, 2]). With this approach we have been able to capture most of the previous results in the field of geometric invariants, extend some of them to more general cases and obtain a new general result on invariant functions of 4 views. We believe this approach has much potential and would be very useful in future progress of many other problems in computer vision. One of the appealing features of this approach is that, for the most of it, very little geometric and algebraic intuition is required in order to obtain these results.

2 3D-from-2D Geometry and Elimination Theory

The general idea behind elimination is to start with the equations describing the projection of 3D points onto corresponding 2D points over n-points and N-views; then to identify the variables we are not interested in (those we want to eliminate), variables we that we know or treat as constants, and variables we care about (those that will participate in the implicit form of our original equations); then we perform the elimination process using the tool of Grobner bases. The various questions we described above, and their solutions, differ simply in the decision of what variables we assign to each of the three sets.

Specifically, Let $P = (X, Y, Z, 1)$ denote the homogeneous coordinates of 3D points, and $p = (x, y, 1)$ the homogeneous coordinates of the projection of P, then a pin-hole camera model gives rise to a linear projection described by a matrix A:

$$p \cong AP,$$

where $\cong$ denotes equality up to scale. If nothing is known about A we are in a projective setting (i.e., given enough views and points we can recover the object up to an unknown projective transformation of 3D space). Since the projection equation is up to scale, there are 11 distinct variables to A, thus for N views we have $11N$ camera parameters.

Let n be the number of 3D points. Since we are dealing with a projective setting, there are 15 free parameters we can choose, i.e., any five of the points, which are in general position, can be assigned the standard coordinates (i.e., form a projective basis of 3D space). Thus, we have $3n-15$ space variables. Finally, we have $2Nn$ image coordinates. Taken together, we have:

1. $11N$ camera variables.
2. $3n - 15$ space variables.
3. $2Nn$ image variables.

We have $2Nn$ equations, where each equation is a rational function that expands readily into a polynomial in all three classes of variables. The zero-set of the $2Nn$ polynomials defines an algebraic manifold (variety), and the questions of interest often boil down to finding ways to project the variety onto a lower dimensional space. In other words, an invariance relation holds under the general variability of a subset of our original variables — say, for example, reconstruction of the 3D object is an invariance relation under the variability of the camera parameters. This readily suggests the use of *elimination*, via Grobner bases, as the general tool for achieving projections onto the desired lower dimensional spaces. The final result that the Grobner base computation provides is a basis for the ideal that includes all possible relationships that follow from the original equations and that include only the required variables.

Since we can eliminate at most m variables from $m + 1$ algebraically independent equations, we readily obtain a parameter counting tool for dealing with geometric issues. This method of parameter counting is similar to counting degrees of freedom of a system, with the difference that elimination arguments provide a more general counting framework, and without the subtleties (that often lead to mistakes) typically associated with degrees-of-freedom counting. As a final remark, it is important to note that the main contribution of this framework is its generality, not merely the principle of eliminating variables as a way to achieve invariance relations. Particular uses of elimination appeared in the literature in specific problems, such as in [20, 21, 1, 28, 14, 9], but not as a general methodology with the appropriate tools from symbolic algebra.

The following examples will demonstrate these simple ideas:

3 Reconstruction of 3D objects

Given N views and n points, we can achieve reconstruction by eliminating the $11N$ camera parameters and all space parameters except one, say the Z coordinate of the n'th point. This is possible if:

$$2Nn > 11N + 3n - 16 \qquad (1)$$

(the left hand side being the two equations per image point and the right hand side being the camera parameters and all the 3D points except for 5 points and one Z coordinate of a single point). In the case of two views, $N = 2$, we see that the minimal number of points is $n = 7$, where we get 28 algebraically independent equations and 27 variables to eliminate (note the lack of subtleties: because each equation introduces a new variable, we are guaranteed to have an algebraically independent set of equations). This immediately tells us there is only *one* invariant function of the form:

$$f(Z, x_i, y_i, x'_i, y'_i) = 0,$$

where $i = 1, ..., 7$, and $(x_i, y_i), (x'_i, y'_i)$ are corresponding image points. The degree of the polynomial indicates the number of solutions (or at least an upper bound) (which we know is 3 from [7]). Furthermore, running the Grobner bases program (using Macaulay or Maple, with a sufficient amount of memory) we get a third-order polynomial in Z which is a closed-from solution to the reconstruction problem (for $n = 7, N = 2$).

We can ask what happens with $n = 6$ points (a problem not addressed in the past). We see that we have 24 equations and 24 variables to eliminate. This means that a solution is not possible (we cannot eliminate all the variables we need), but there exists a polynomial with two space coordinates (one polynomial for each combination of two space coordinates), say:

$$g(Z, Y, x_i, y_i, x'_i, y'_i) = 0,$$

where $i = 1, ..., 6$. Thus if one coordinate is given, we can recover the other two space coordinates.

So far we have re-derived the results of [7] on reconstruction from 7 points and discussed the case of 6 points across two views. In a similar straightforward manner we can address the case of $N = 3$ views, as follows. To satisfy Equation 1 we see immediately that $n \geq 6$ points are needed in order to have enough equations for eliminating all the undesired variables. For $n = 6$ points we have 36 equations and 35 variables to eliminate, thus there is a unique reconstruction polynomial (for each coordinate) of the form:

$$f(Z, x_i, y_i, x'_i, y'_i, x''_i, y''_i) = 0,$$

where $i = 1, ..., 6$ and (x''_i, y''_i) are the corresponding image coordinates in the third view. The degree of $f()$ tells us the number of solutions (which we know is 3 from [25]). How many invariant reconstruction polynomials exist per coordinate for 6 points and $N > 3$ views? We saw that for $N = 2, 3$ the minimal number of points gave rise to a single invariant polynomial; however, we see that in general we will have $N - 2$ invariant functions for $N > 2$ views.

Finally, when do we obtain linear solutions? Extra points and/or extra images add algebraically independent equations — each with a new variable. Thus, the basis B of the ideal grows with the new equations. Following the elimination stage we can once again compute a basis but this time of B with respect to a single variable, thus, producing a single polynomial. If there is a linear solution, then we are guaranteed to have it as the basis.

4 Invariant Functions of Views

Here we are interested in functions that are invariant to object structure, i.e., functions of image coordinates across a number of views that vanish for all tuples of corresponding points (i.e., matching constraints). For example, the fundamental matrix (epipolar geometry) is a matching constraint (bilinear function) of two views, and the trilinear tensor of [28] (trilinear functions) is a matching constraint(s) of three views.

This goal is very similar to the reconstruction goal with the difference that we treat the $11N$ camera parameters as *constants* (i.e., combinations of them will form the coefficients of the invariant functions). In reconstruction the role of the n points is to eliminate the $11N$ camera parameters, thus for the task of invariant functions of views we need to consider only $n = 1$ points.

The number of equations is $2N$, we wish to eliminate all 3 space coordinates ($n = 1$), and the camera parameters are constants. Therefore, the inequality is:

$$2N > 3.$$

We see that for two views we have a single invariant function which is the epipolar geometry. Computing a Grobner base we indeed obtain the bilinear function of image coordinates (whose coefficients are combinations of the 22 camera parameters), as expected. If we want to get this relationship from the images without computing or knowing the camera parameters we have to eliminate them using more image points;

$$2 \times 2n > 11 \times 2 + 3n - 15$$

so that we need at least 8 pairs of image points. Note that the epipolar geometry (location of epipoles) can be computed (non-linearly) from 7 points (classic problem of Chasles [7]) up to a three-fold ambiguity, yet the bilinear invariant function of two views (the matching constraint) as a function of image pairs requires at least 8 pairs.

The advantage of this derivation over the geometric derivation of the fundamental matrix (whose entries are the coefficients of the bilinear invariant function), is that here we have a simple and rigorous proof that the epipolar constraint is the only invariant function of two views — a statement that is well accepted in the community but has not been adequately proven before.

From here we obtain also a general result for N views which is that the minimal number of algebraically independent functions of N views is $2N - 3$ (in other words, $2N - 3$ is the minimal generating set of the ideal defined by the original camera equations). For example, we have 3 independent functions for $N = 3$ views — a result previously noted in [37]. Note that these functions are not necessarily the epipolar constraints.

With a small amount of extra calculation (after we get the Grobner basis) we get the following new result; It has been shown that $N = 3$ views admit sets of *four* trilinear invariant functions with overall 27 coefficients, that can be solved *linearly* from 7 corresponding points across the three views [28]. The 27 coefficients sit in a tensor whose relation to 3D invariants, epipolar geometry and intrinsic structures of 3 views has been recently derived in [30, 8]. We see that a linear solution requires an extra point, but also an extra equation. In other words, the four trilinear equations are not algebraically independent but they are linearly independent. This raises the question of how does the linear situation grow with the number of views? How do the number of coefficients grow? Do we always need to add an extra point to obtain a linear solution? The answers for $N = 4$ are described below:

Four views admit a quad linear tensor with 81 coefficients described by 16 quadlinear functions. The coefficients of the tensor can be solved linearly from 6 corresponding points across the four views.

5 Alternate View of Elimination: Mappings

The elimination process can also be viewed as a mapping. For example, the case of matching constraints induces the mapping $\mathcal{R}^3 \mapsto \mathcal{R}^{2N}$, i.e., a 3D algebraic manifold (variety) in a $2N$ dimensional space. In the case of parallel projection we immediately have the following results:

- The rank of the span of the points in R^{2N} (the 2N coordinates of each projected 3D point) is 3 [35].
- The invariant functions (matching constraints) across views are linear [36, 23].
- Segmentation of separately moving bodies can be done by segmenting the points in R^{2N} into 3D hyperplanes, for example, using a Hough transform.

The elimination approach can be applied also to shading and color domains as well. In fact, any vision domain that has an algebraic form benifit from this approach. For example, using any model of reflection with the usual assumption of a light source at infinity (or equivalently parallel light rays) the gray level at a point in the image is a function (among other parameters) of the direction of the light source, which is defined by two parameters (often noted p and q). From this it immediately follows that all the images of a given object are in a 2D manifold regardless of direction of the light source. In the Lambertian case, the 2D manifold becomes a 3D linear subspace by describing the light source (intensity and direction) by three parameters. Thus, given three different images I_1, I_2 and I_3 of a scene taken under different lighting conditions any other picture I of this scene is linearly spanned by I_j, i.e., $I = \alpha_1 I_1 + \alpha_2 I_3 + \alpha_3 I_3$, for some coefficients $alpha_1, alpha_3, alpha_3$ [29].

Color, both of the surface and of the illumination, requires many parameters to be described exactly. Therefore, in practice a reduced set of parameters is typically used for an approximate model. For example, in the linear combination color model [16], the color of an object or a light source can be described as a linear combination of a set of basis color functions. It has been argued that 3 basis functions can approximate well the color of natural objects and light sources. The color measured at a point is alao a function of the illumination. We have immediately that the colors of a set of points are in a 3D manifold regardless of the illumination. In case the illumination model is linear, then the manifold is a linear subspace.

6 What can be obtained from a single view?

It has been stated in the past that there are no invariants from a single view [3, 4, 19]. There are still invariants over camera parameters that are not invariant to the 3D structure of the object. However, the issue of extracting information from a single image remains largely an open problem (see also [38, 32]).

Similarly to the approach for obtaining invariant functions of views, we treat a subset of our variables as constants. In this case the space coordinates are constants, which means that additional views would be required for eliminating those constants. We eliminate the 11 camera parameters, and we have $2n$ equations, thus for n points we have the following condition to satisfy:

$$2n > 11,$$

which means that for $n \geq 6$ there exist $2n - 11$ invariant functions of image coordinates whose coefficients are polynomials of the space coordinates. In other words, the functions are invariant to camera transformations (because we eliminated them), but not necessarily to the point-configuration in space, i.e., there may be several, not projectively equivalent, configurations of n space points that share the same invariant function. One possible use of these single-view-object-non-specific invariant is for indexing into object data bases.

The invariant function for 6 points has a quadlinear form and is shown Appendix A.

One way to overcome the the lack of general invariants from a single view is to employ a priori knowledge on the class of objects. This knowledge can be represented by E equations describing properties of the object. The relevant counting argument is $2n + E > 11 + 3n - 16$ for reconstruction and $2n + E > 11 + 3n - 15$ for recognition. One case that has been treated is to assume existence of bilateral symmetry of our 3D point set [19, 24, 26, 18].

A fairly complete explanation of model based invariants, model free invariants, indexing and their complexity that is based on this work can be found in [40].

7 The correspondence problem: The Issue of Symmetric Functions

It would be desirable to to have some way of computing invariants without having to solve the correspondence problem. In its most generality this means that we would like to find invariant functions that are also invariant to permutations of the image point sets. Therefore, the invariant function must be a polynomial of symmetric functions of the image points. We have

$$2Nn + Ns > 11N + 2Nn,$$

where s is the number of symmetric functions used in each image. This implies that in order to have a correspondence free invariant one must use on the average more than 11 symmetric functions per picture, leaving a complicated and high degree (thus noise sensitive) invariant. In other words, the general solution to the correspondence problem is not practically feasible, unless we make further assumptions.

8 The Case of Multiple Objects

The framework we have described so far can be easily extended to deal with multiple objects each moving independently in three space. Assume we have p objects projecting onto each view and we have correspondence of all points across the views, but we have not segmented the points into object sets. As we have seen above, each object maps onto a 3D manifold (variety) in $2N$ dimensional space. The ideal describing the multiple objects simultaneously is the intersection ideal of each of the separate ideals describing the 3D manifolds — which corresponds to the union of the 3D varieties. The dimension of the intersection ideal is the sum of the dimensions of the individual ideals. Thus, in order to have a large enough space to contain the intersection ideal we need to increase the number of images N:

$$2N > 3p.$$

For example, for $p = 2$ two objects we need $N = 4$ four images in order to find an invariant function (matching constraint) that applies to *both* objects *simultaneously*. Note, that in the special case of parallel projection the intersection ideal is still linear.

9 Motion of Lines

The framework can be easily applied to lines in space. It is known that, first, three views are necessary for reconstruction of lines, and 9 corresponding lines are required in the process [17, 37]. Second, the linear methods require 13 matching lines across three views [10, 39, 31, 22].

A line in 3D requires 4 parameters, and its projection in 2D requires two parameters. Let l denote the number of lines. We can readily state the following observations:

- For $N > 2$ we have $2N - 4$ algebraically independent relationships between the parameters of a projected line over N images.
- In a single image there are $2l + 2n - 11$ algebraically independent invariant functions (invariant over camera transformation, but not to the 3D configuration of lines and points).
- If, $2Nl > 11N + 4l - 15 - 1$, then 3D reconstruction is possible. For example, for $N = 3$ views we need at least $l \geq 9$ lines; for $N = 4$ views we need $l \geq 8$ lines; for $N = 4$ views, $l \geq 7$; for $N \geq 8$ views, $l = 6$ lines are sufficient.
- If $2N(n + l) > 11N + 4l + 3n - 16$ we can also reconstruct 3D (points and lines). For example, for $N = 2$ views, the contribution of lines drops, which means that lines do not help at all in reducing the number of required points for 3D recovery from two views. With $N = 3$ views one can recover 3D from 5 points and 2 lines.

10 The Case of Parallel Projection

An interesting special case of 3D-from-2D geometry is the case when the projection is parallel, i.e., we are in an affine setting. The appealing part of this assumption is that the elimination process reduces to Gaussian Elimination instead of the heavy tool of Grobner bases. Moreover, the elimination process always results in affine functions.

Consider the following examples. In the case of reconstruction we have the following condition to satisfy:

$$2Nn > 8N + 3n - 13$$

as there are 12 free parameters in a 3D affine projection, which shows that 4 points are needed for 2 views. The example of epipolar geometry follows from the same condition as in perspective projection:

$$2N > 3,$$

but now instead of a algebraic manifold we have a 3-space of a $2N$ dimensional vector space. For example, for $N = 2$ we have a hyperplane in $\mathcal{R}^4$, and thus the invariant function is unique and obtained by the inner product between $\mathcal{R}^4$ and the hyperplanes dual space. The case of $N > 2$ readily shows the existence of linear functions of views — which is the result of [36] (also [23] contains an alternative derivation of the same result).

The case of invariants from a single view, we have $n \geq 5$ which yields two linear functions of the form:

$$\sum_i \alpha_i x_i = 0 \qquad \sum_i \alpha_i y_i = 0$$

where the two sets of coefficients are the same, and their sum vanishes (for the case $n = 5$, each function is a hyperplane of $\mathcal{R}^5$ that passes through the origin).

Taking the case of $n = 5$ further we have that the vector of coefficients α is the solution of the system $A\alpha = 0$ where $A = [P_1, ..., P_5]$. This provides another route to epipolar geometry and the linear combination of views of [36].

11 Summary

We have seen that by recasting the camera equations as an elimination problem (looking for projections of the general variety defined by the original equations onto lower dimensional varieties), many of the current and previously solved problems related to geometric invariants become very simple.

The main ingredient of the approach is the division of our parameters into three classes — those to be eliminated, those that are constants, and those that would participate in the implicit functions we want to derive. Most, if not all, of the questions of interest reduce to a proper choice of dividing the parameters into these classes.

In addition to unifying most of the current results in the field of geometric invariants from points and lines sets, we have obtained few general results on N views, and a general result on the invariant functions of 4 views that contains several non-trivial consequences on the general understanding of linear solutions to the N-views n-points reconstruction and representations of 3D space from 2D views.

A Invariant Functions From a Single View

Six image points (x_i, y_i), $i = 1, , , 6$, satisfy a single quadlinear function of image coordinates whose coefficients are a function of the sixth space point whose coordinates are denoted by (p, q, r, s). The function is invariant to camera transformations from 3D to 2D.

$$
\begin{aligned}
pq & \begin{vmatrix} x_1 & x_2 & x_6 \\ y_1 & y_2 & y_6 \\ 1 & 1 & 1 \end{vmatrix} \begin{vmatrix} x_3 & x_4 & x_5 \\ y_3 & y_4 & y_5 \\ 1 & 1 & 1 \end{vmatrix} \\
-pr & \begin{vmatrix} x_1 & x_3 & x_6 \\ y_1 & y_3 & y_6 \\ 1 & 1 & 1 \end{vmatrix} \begin{vmatrix} x_2 & x_4 & x_5 \\ y_2 & y_4 & y_5 \\ 1 & 1 & 1 \end{vmatrix} \\
+qr & \begin{vmatrix} x_2 & x_3 & x_6 \\ y_2 & y_3 & y_6 \\ 1 & 1 & 1 \end{vmatrix} \begin{vmatrix} x_1 & x_4 & x_5 \\ y_1 & y_4 & y_5 \\ 1 & 1 & 1 \end{vmatrix} \\
+ps & \begin{vmatrix} x_2 & x_3 & x_5 \\ y_2 & y_3 & y_5 \\ 1 & 1 & 1 \end{vmatrix} \begin{vmatrix} x_1 & x_4 & x_6 \\ y_1 & y_4 & y_6 \\ 1 & 1 & 1 \end{vmatrix} \\
-qs & \begin{vmatrix} x_1 & x_3 & x_5 \\ y_1 & y_3 & y_5 \\ 1 & 1 & 1 \end{vmatrix} \begin{vmatrix} x_2 & x_4 & x_6 \\ y_2 & y_4 & y_6 \\ 1 & 1 & 1 \end{vmatrix} \\
+rs & \begin{vmatrix} x_1 & x_2 & x_5 \\ y_1 & y_2 & y_5 \\ 1 & 1 & 1 \end{vmatrix} \begin{vmatrix} x_3 & x_4 & x_6 \\ y_2 & y_4 & y_6 \\ 1 & 1 & 1 \end{vmatrix} = 0
\end{aligned}
$$

Acknowledgments

This research was supported by ARPA through the U.S. Office of Naval Research under Grant N00014-93-1-1202, R&T Project Code 4424341-01 We thank H. Furstenberg, M. Ben-or, D. Weinshall and S. Peleg for comments throughout this research.

References

[1] E. Barrett and P. Payton. Projectively invariant structures in multi-sensor imagery. In *SPIE Applications of Digitial Image processing XV*, volume 1771, pages 236–251, 1992.

[2] B. Buchberger, G. Collins, and B. Kutzler. Algebraic methods for geometric reasoning. *Annl. Reviews in Comp. Science*, 3, 1988.

[3] J.B. Burns, R. Weiss, and E. Riseman. View variation of point-set and line segment features. In *Proceedings Image Understanding Workshop*, pages 650–659. Morgan Kaufmann, San Mateo, CA, April 1990.

[4] D.T. Clemens and D.W. Jacobs. Space and time bounds on indexing 3-D models from 2-D images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:1007–1017, 1991. Also in DARPA IU Workshop, pp. 604–613, 1990.

[5] D. Cox, J. Little, and D. O'Shea. *Ideal, Varieties, and Algorithms.* UTM. Springer-Verlag, 1992.

[6] O.D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *Proceedings of the European Conference on Computer Vision*, pages 563–578, Santa Margherita Ligure, Italy, June 1992.

[7] O.D. Faugeras and S. Maybank. Motion from point matches: Multiplicity of solutions. *International Journal of Computer Vision*, 4:225–246, 1990.

[8] R. Hartley. Lines and points in three views — a unified approach. In *Proceedings Image Understanding Workshop*, Monterey, CA, November 1994.

[9] R. Hartley. Projective reconstruction and invariants from multiple images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(10):1036–1040, 1994.

[10] R. Hartley. Projective reconstruction from line correspondences. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Seattle, WI, June 1994. IEEE Computer Society Press.

[11] R. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 761–764, Champaign, IL., June 1992.

[12] C.M. Hoffman. *Geometric and Solid modeling.* Morgan-Kaufmann, 1989.

[13] D. W. Jacobs. Generalizing invariants for 3-D to 2-D matching. In *Proceedings of the 2nd European Workshop on Invariants*, Ponta Delagada, Azores, October 1993.

[14] D.J. Kriegman and J. Ponce. On recognizing and positioning curved 3-d objects from image contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:1127–1137, 1990.

[15] H.C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, 1981.

[16] L. T. Maloney and B. Wandell. A computational model of color constancy. *Journal of the Optical Society of America*, 1:29–33, 1986.

[17] S.J. Maybank. The critical line congruence for reconstruction from three images. *Applicable Algebra in Engineering, Communication and Computing*, 1995. To appear.

[18] H. Mitsumoto, S. Tamura, K. Okazaki, N. Kajimi, and Y. Fukui. 3-D reconstruction using mirror images based on plane symmetry recovering method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:941–946, 1992.

[19] Y. Moses and S. Ullman. Limitations of non model-based recognition systems. In *eccv*, pages 820–828, 1992.

[20] J. Mundy and A. Zisserman. *Geometric invariances in computer vision*. MIT Press, Cambridge, 1992.

[21] J.L. Mundy, A. Zisserman, and D. Forsyth. *Applications of Invariance in Computer Vision*. Springer-Verlag, LNCS 825, 1994. Proc. of the 2nd European Workshop on Invariants, Ponta Delagada, Azores, Oct. 199.

[22] N. Navab, O. D. Faugeras, and T. Vieville. The critical sets of lines for camera displacement estimation: A mixed euclidean-projective and constructive approach. In *Proc. 4th International Conference on Computer Vision*, Berlin, Germany, May 1993. IEEE.

[23] T. Poggio. 3D object recognition: on a result of Basri and Ullman. Technical Report IRST 9005-03, May 1990.

[24] T. Poggio and T. Vetter. Recognition and structure from one 2D model view: Observations on prototypes, object classes and symmetries. A.I. Memo No. 1347, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, February 1992.

[25] L. Quan. Invariants of 6 points from 3 uncalibrated images. In *Proceedings of the European Conference on Computer Vision*, Stockholm, Sweden, May 1994. Springer-Verlag, LNCS 801.

[26] C.A. Rothwell, D.A. Forsyth, A. Zisserman, and J.L. Mundy. Extracting projective structure from single perspective views of 3D point sets. In *Proceedings of the International Conference on Computer Vision*, Berlin, Germany, 1993.

[27] A. Shashua. Projective structure from uncalibrated images: structure from motion and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(8):778–790, 1994.

[28] A. Shashua. Algebraic functions for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1995. in press.

[29] A. Shashua. On Photometric Issues in 3D Visual Recognition from a Single 2D Image. *International Journal of Computer Vision*, in press.

[30] A. Shashua and M. Werman. Fundamental tensor: On the geometry of three perspective views. ICCV 1995.

[31] M.E. Spetsakis and J. Aloimonos. Structure from motion using line correspondences. *International Journal of Computer Vision*, 4(3):171–183, 1990.

[32] P.F. Stiller, C.A. Asmuth, and C.S. Wan. Invariants, indexing, and single view recognition. In *Proceedings Image Understanding Workshop*, Monterey, CA, November 1994.

[33] B. Sturmfels. *Algorithms in Invariant Theory*. TMSC. Springer-Verlag, 1993.

[34] C. Tomasi. Pictures and trails: A new framework for the computation of shape and motion from perspective image sequences. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 913–918, Seattle, Washington, June 1994. IEEE Computer Society Press.

[35] C. Tomasi and T. Kanade. Shape and motion from image streams – a factorization method. *International Journal of Computer Vision*, 9(2):137–154, 1992.

[36] S. Ullman and R. Basri. Recognition by linear combination of models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-13:992—1006, 1991.

[37] T. Vieville, N. Navab, and Q.T. Luong. Motion of lines in the uncalibrated case: algebra versus geometry. Technical report, INRIA, fall 1993.

[38] D. Weinshall. Model based invariants for 3-D vision. *International Journal of Computer Vision*, 10(1):27–42, 1993.

[39] J. Weng, T.S. Huang, and N. Ahuja. Motion and structure from line correspondences: Closed form solution, uniqueness and optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(3), 1992.

[40] M. Werman and D. Weinshall. Complexity of indexing: how to choose efficient indices. Technical Report TR-95-7, The Hebrew University of Jerusalem, Institute of Computer Science, 1995.