Multiple View Vision

Kalle Åström,
Centre for Mathematical Sciences, Lund University
P O Box 118, SE-221 00 Lund, Sweden
kalle@maths.lth.se

Abstract

The goal of this paper is to give a short historical overview of multiple view vision and in particular the estimation of both camera geometry and scene models using only images as input. This problem (the structure and motion problem) can be seen as the mathematical inverse of the computer graphics problem. The basic structure and motion problems for different feature types are discussed as well as some recent methods that estimate not only scene geometry, but also radiance, irradiance and illumination properties.

1. Introduction

The problem of creating virtual three-dimensional models of the environment and calculate the image of these objects from different viewpoints, is one of the ideas behind *computer graphics*, see [23, 49]. Already in 1961 a computer game 'Space war' was developed at MIT using vector graphics. In 1963 computer graphics (the hidden line algorithm) was developed for rendering simple objects such as points and lines. As the field of computer graphics developed, more advanced models were created (hidden surface (1969), diffuse lighting (1971), curved, textured surfaces (1974), ray tracing (1980), radiosity (1985), etc) and used for machine design, architecture, games and special effects in movies. Today the three-dimensional models can handle not only lines and points but also surfaces, texture, specularity, shading and different lighting conditions. It is now possible to make full-length movies like 'Toy Story' using only computer graphics. Episode I of the star wars saga, 'The Phantom Menace', uses digital manipulation in nearly every single frame. The three-dimensional models are most often created manually, which is difficult and time-consuming.

Just as part of the field of computer graphics can be seen as the generation of images of known three-dimensional models, some problems within the field of *computer vision* are their inverses in the mathematical sense, i.e. to calculate the three-dimensional models from their two-dimensional images. These images could be virtual, but most applications of computer vision use a camera as a sensor for real world objects. Another central problem is to calculate the position and orientation of the camera with respect to the scene. These problems are combined in the **structure and motion problem**, where structure refers to the three-dimensional models and motion to the position and orientation of the camera. The field of computer vision is still at its early stages. Most of the theory and the algorithms focus on the reconstruction of simple geometrical features like points and lines, similar to the field of computer graphics in 1960s, but there have been recent attempts to deal with more complicated objects. The purpose of this paper is to give an overview of such attempts and to present some applications of multiple view vision.

1.1. Examples of multiple view vision

Interaction between computers and the physical world is an increasingly important area. Cameras are becoming universal sensors, cheap, widely available. Proper use of these sensors offer many possibilities in many fields, medicine, industry, human machine interfaces. There are many such applications. Here I will only give two examples.

One such example is the use of digital cameras for precision measurements in industry. One example is the company IMETRIC in Switzerland that develops algorithms for structure and motion estimation for points in several images. Their system is based on placing retroreflective circular targets on industrial parts. During measurements they take several images with high-resolution cameras (up to 7168×4048 pixels) with flash. They measure the image positions of these targets with high precision and solve for structure and motion. Measurement accuracies reach from 1 part in 30,000 to 200,000 of the object size depending on the camera and the application

Another example is automatic model building from image sequences. Many companies try apply computer vision in the world of film-making. The idea is to automate a variety of labour-intensive tasks in special effects etc. An example of is the company REALViZ, [58] that delivers new 'software tools that dramatically reduces the time required to produce high-quality computer-generated images, animation and 3D models'.

1.2. Overview of the paper

It is difficult to give an overview of structure and motion estimation. Systems for automatic structure and motion estimation are complex and involve research in feature detection, matching, geometry, optimization, etc. Here I have chosen to focus on the geometrical problems in structure and motion estimation.

Many results in computer vision is an analysis of a special case. These solutions are important building blocks for solution of structure and motion. An overview over some of these methods are given in Section 2. In Section 3 are described so called factorization methods. In Section 4 methods for calibration and auto-calibration are described. Nonlinear optimization of image residuals or bundle adjustment methods is discussed in Section 5. Section 6 is devoted to structure and motion using apparent contours and Section 7 describe some recent advances on how to improve your scene models when you have estimated motion.

2. Special case methods

2.1. Minimal case methods

The classical minimal case problems are for points. These were formulated and solved already in 1910 by the German photogrammetrist Erwind Kruppa, [43]. With 5 point correspondences in 2 images (with known internal calibration), he showed that there are 11 solutions to the structure and motion problem. Later Demazure [18] showed that there are in fact only 10 solutions. The uncalibrated case (7 points in 2 views) has an even older history. A related problem was formulated by Chasles already in 1855 [13]. Several authors responded to that question and showed that there are in general 3 solutions to the problem, cf. [29]. See also [61]. Both of these approaches use the so called epipolar constraint. The projection of scene points can be modeled with a matrix equation

$$\underbrace{\lambda}_{scale}\underbrace{\mathbf{x}}_{image} = \underbrace{P}_{motion}\underbrace{\mathbf{X}}_{structure} , \tag{1}$$

where $\mathbf{x}$ is a 3 vector representing an image point in homogeneous coordinates, $\mathbf{X}$ a 4 vector representing a scene point in homogeneous coordinates and P is a 3×4 matrix representing camera position, orientation and internal calibration. The multilinear approach is a way to eliminate the scale factors and the scene structure. There is a essentially one-to-one mapping from pairs of camera matrices to the fundamental matrix F

$$\underbrace{(P_1, P_2)}_{motion} \longleftrightarrow \underbrace{F}_{3\times 3 \text{matrix}} \tag{2}$$

Each image point correspondence $(\mathbf{x}_1, \mathbf{x}_2)$ gives a linear constraint on the fundamental matrix:

$$\mathbf{x}_1^T F \mathbf{x}_2 = 0, \tag{3}$$

The basic idea is to use image points to calculate F using (3) and the fact that $\det(F) = 0$, and then calculate motion (P_1, P_2) using (2) and finally structure $\mathbf{X}$ from (1).

For uncalibrated cameras there is another interesting minimal case (6 points in 3 views) that was solved independently in [55, 30]. There are in general 3 solutions. Both can be seen as using the trilinear constraint. Later it was shown that these two problems (7 points in 2 views and 6 points in 3 views) are dual problems in a sense that was discovered in [12].

For lines there are several minimal cases. The affine camera cases (6 lines in 3 affine views and 5 lines in 4 affine views) were solved in [4]. The case 9 lines in 3 uncalibrated views is still unsolved.

2.2. Quasi-linear methods

Besides the minimal cases there is great interest for so called quasi-linear methods for close to minimal cases. Many times a minimal case, which might have multiple solutions and whose solutions involve solving polynomial equations, is to difficult to implement or impractical to use. Quasi linear methods use the extra equations to find simple linear methods for finding the solutions by ignoring some of the non-linear constraints. The most common quasi-linear method is the eight-point-algorithm [64, 45, 26]. It is based on the epipolar geometry. Estimate the fundamental matrix. The 8 point correspondences give 8 linear constraints (3). Ignoring the non-linear constraint ($det(F) = 0$) this determines F uniquely up to scale.

For three views a similar approach is used to solve the case of 7 points in 3 views. The idea is based on the trilinear tensor. The same tensor can be used for the case of 13 lines in 3 views or certain combinations of lines and points, cf. [59, 63, 60, 20, 28, 27]. For four view the quadrilinear tensor can be used to solve the case of 6 points in 4 views [32].

2.3. Resection and intersection

If you have more images or features, one common practice is to solve for a minimal case and then extend the solution to more features and views by so called intersection and resection. In 1759, Johan Heinrich Lambert, in a treatise "Perspectiva Liber" (The Free Perspective), developed the mathematical principles of a perspective image using space resection to find a point in space from which a picture is made. Resection is thus the method of estimating the camera geometry in a new view using image features and estimated scene features. Intersection is the method of estimating the location of additional features using known camera positions.

3. Factorization methods

A completely different approach to the structure and motion problem is to solve for all views and all features simultaneously using a factorization approach. The original idea was presented in [69] which solved the case of points projected with the affine camera model. The key here is that under ortographic camera model, a measurement matrix containing all image feature measurements

$$W = \begin{pmatrix} x_{11} & \dots & x_{1n} \\ \vdots & \vdots & \vdots \\ x_{m1} & \dots & x_{mn} \\ y_{11} & \dots & y_{1n} \\ \vdots & \vdots & \vdots \\ y_{m1} & \dots & y_{mn} \end{pmatrix} \tag{4}$$

is of rank 4 (or rank 3 if the image coordinates are centered around the center of mass). Using the singular value decomposition the matrix W can be factored as $W = PX$ where P contains the motion and X the scene structure. A more thorough overview is given in [39] This factorization method has been extended to the other camera models, cf. [50, 67, 35, 62]. Some of these methods use the notion of affine shape, which can be seen as dual projective space. It includes in its method a way of normalization the coordinates, similar to what is proposed in [26]. The factorization approach has also been extended to lines [56, 48] to combinations of points, lines and conics, [37] and also to 3D curves [9, 10].

4. Auto-calibration methods

Recently there has been an intensive research on the possibility to obtain reconstructions up to an unknown similarity transformation (often called *Euclidean reconstructions*) without using fully calibrated cameras. Since it is only possible to make reconstruction up to an unknown projective transformation (often called *projective reconstruction*) when nothing about the intrinsic parameters, extrinsic parameters or the object is known, it is necessary to have some additional information about either the intrinsic parameters,

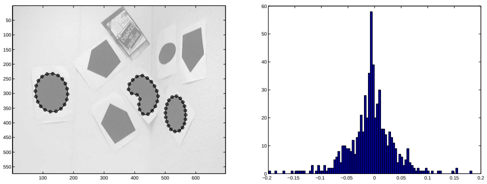

Figure 1. One of five images used in experiment. The extracted curves and 20 reprojected points on each space curve are shown. To the right is shown a histogram of the residuals (in pixels) between extracted image curves and reprojected points of the reconstructed curve. The standard deviation of these residual is 0.03 pixels.

the extrinsic parameters or the object in order to obtain the desired Euclidean reconstruction.

A priori information about the object can be used in a fairly straight-forward manner, see [21]. A priori information about the extrinsic parameters, i.e., the camera orbit, has been investigated by several researchers, e.g. [77].

The camera matrix can be decomposed as $P = K[R\,|\,(-Rt)]$, where t is the camera center, R is a rotation matrix representing camera orientation and K is contains the internal calibration parameters:

$$K = \begin{bmatrix} \gamma f & sf & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{bmatrix} . \tag{5}$$

Here f is camera constant (focal length), s represents skew, γ represent aspect ratio and (x_0, y_0) the coordinates of the principal point.

One common situation is when the intrinsic parameters are constant during the whole (or a part) of the image sequence. Auto-calibration from the assumption of constant intrinsic parameters is traditionally known as self-calibration and was the first auto-calibration problem studied in computer vision. This problem leads to the so called Kruppa equations. Several attempts to solve this problem have been made, see [46], [22]. More recent approaches to self-calibration, using more robust methods, can be found in [3] (using the more general *Kruppa constraints*), [53] (using the so called *modulus constraints*) and [70] (using the classical formulation with the *absolute conic* combined with a robust estimation method). The absolute conic is a complex conic at the plane at infinity. Basically the calibration problem boils down to a number of matrix equations:

$$P_i \Omega P_i^T = \mu_i \omega_i, \forall i = 1, \dots m, \tag{6}$$

one equation for each image i with camera matrix P_i. One wants to solve for the absolute conic Ω and the image of the absolute conic ω_i which in turn is related to the internal calibration matrix K in (5) according to

$$\lambda_i \omega_i = K_i K_i^T \quad . \tag{7}$$

During the last years, several attempt has been made to develop auto-calibration techniques under less restrictions on the intrinsic parameters of the camera. The first step in this direction was made in [31] and another in [52], where the self-calibration method presented in [53] is extended to allowing changing focal length. In [33] was given for the first time an existence proof (along with a computation method) for the possibility to make Euclidean reconstruction under the assumption of known skew and aspect ratio, from so called *Euclidean image planes*.

The next step in the development of flexible calibration techniques was to weaken the assumptions further. Simultaneously, it was proven in [34] and [51] that it is sufficient to know the skew. In fact, it was even shown in the former paper the more general result that it is sufficient to know any one of the intrinsic parameters. Observe that all other intrinsic parameters are unknown and allowed to vary between the different imaging instances.

All auto-calibration techniques are based on the assumption of generic camera motion. However, if the camera motion is not sufficiently general, e.g. pure translation or circular motion, then one might run into trouble. An interesting and important subjects in its own is to study these so called *critical motion sequences* (CMS). The pioneering work on CMS is [65] (for self-calibration) followed up by [66] and [38].

5. Bundle adjustment methods

The problem of obtaining a statistically optimal solution to the structure from motion problem, have traditionally been solved using the technique of so called bundle adjustments, see [8]. However, it is only possible to use points in the traditional approach to bundle adjustments.

In [75] the problem of estimating structure and motion from line correspondences is analyzed. They discuss both the question of obtaining an initial estimate and the question of improving these estimates.

A generalization to combinations of points, lines and conics has been made in [6]. Also for reconstruction of 3D curves it has been possible to extend the bundle adjustment method, [10].

6. Apparent contours

The **apparent contours** of three-dimensional surfaces are known to be a rich source of information. Under *known* viewer motion, reliable descriptions of curved surfaces can be recovered from the apparent contours, cf. [25, 15, 14, 72, 68]. Inference of 3D shape is even possible from a single image, cf. [71].

Motion estimation using only deformation of the apparent contour on the other hand is much more difficult. One reason for this is that the apparent contour changes as the viewpoint changes. The contour on the surface that generates the apparent contour is actually different for different viewpoints, see Figure 2. Nevertheless, techniques have been developed for estimating camera motion using only the deformation of the apparent contours. The key observation is that for pairs of images, there exists points on the respective apparent contours that are actually the projection of the same physical point. These so called **epipolar tangency points** can be found and used for estimation of apparent contours.

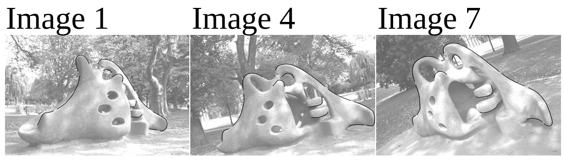

Figure 2. Image sequence of a statue in a park. Extracted contour segments are marked in black. Notice how the contours change as the camera moves.

The special case of frontier points under orthographic projection and object rotation around a single axis was considered in [57, 24]. In [54], although primarily concerned with stereo calibration from 3D space curves, it was noted that the intersection of two contours from two discrete viewpoints generated a point, visible in both images. This is also mentioned in [72], where the apparent contours are used for object modeling. This constraint was exploited in [11] in the analysis of the visual motion of space curves. An approach for parallel projection has been presented in [73, 74]. Another approach using trinocular stereo has been presented in [36].

For the general case of motion estimation this was solved by iterative, optimizing techniques in [5, 2]. For the special case of circular motion see also [47]. In two recent papers excellent experimental results have been obtained both for the case of circular motion [47] and also using a dominant plane for simplifying the situation [16].

7. Structure estimation with known motion

The resulting scene structure after running for example the eight point algorithm on a bunch of points is crude and

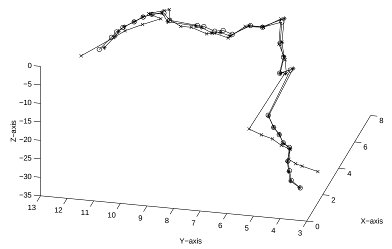

Figure 3. A comparison between the motion computed with the apparent contours, (before and after bundle adjustment) marked with crosses and the motion computed with special points, marked with rings.

uninteresting to look at. The camera motion on the other hand is completely represented. When the camera motion has been estimated it is possible to use this camera motion to further enhance the scene model. This is an important application of theory for scene modeling using known camera motion.

There has been interesting developments in this directions in the last couple of years. Three main trends are described here:

- Dense reconstruction — Methods for estimating dense textured scene structure by finding corresponding points in other images.
- Space carving — A method of building up textured voxel models of the scene by carving away transparent voxels.
- Illumination models — Methods for estimating illumination, reflectance and albedo models for the scene.

These methods are very different in character but each give strikingly realistic scene models.

7.1. Dense reconstruction

One popular method of obtaining a better (more realistic) scene model is based on finding dense feature correspondences using the estimated camera motion. These methods were originally developed for calibrated 3D systems. [51].

For stereo rigs a typical method is to use the known motion (epipolar geometry) to re-map the image to the so called standard geometry where each horizontal line is an epipolar line. The correspondence search is then reduced to a matching of points on the same scan-line. One also use constraints like order preserving, bidirectional uniqueness and detection of occlusion, cf. [40]. A dynamic programming scheme can be used to guide the search to the most probable match [19]. A similar algorithmic idea, but for more than 2 images has also been developed [41]. These algorithms yield reasonably good results, but have their shortcomings when it comes to more general camera geometry. In practice different 3D reconstructions obtained from different subsequences of images are fused in a later step.

7.2. Space carving

Another fascinating method for generating 3D models from images with known camera geometry is the space carving method, cf. [44]. The method is based on representing the scene as colored voxels. One starts with a set of voxels (large enough to include the actual scene. By examining voxels of the scene models one can determine if the voxels colors are consistent with the images. If not these voxels are removed (carved) from the scene model. This procedure is repeated until the whole model is consistent with the images. The final model is then the hull of all voxel models consistent with the images. The approach has several nice theoretical and practical properties: It is provably correct for arbitrary camera positions, No constraints on camera viewpoints are imposed, photo-realistic models are obtained. There are some practical problems. With real images it might sometimes happen that a voxel is erroneously removed. The algorithm then continues to carve away (non-consistent) voxels of the interior resulting in big holes in the reconstruction. Since the whole approach is voxel based, the final modes has a edgy look.

7.3. Illumination models

A third interesting trend is to model not only geometry and intensity (texture maps) of the model, but also to include illumination aspects of the model.

In [76, 17] is presented a method for recovering reflectance properties of all surfaces in a real scene from a sparse set of photographs. The approach is based on knowing the geometry of both the cameras and the scene and the light sources (with high accuracy). Using high dynamic range photographs of the scene one estimates both radiance and irradiance of the scene. The reflectance properties of a surface point is often described by bi-directional reflectance distributions (BRDF). These describe how much light incoming from one direction to the point is reflected to the other directions. Using only a sparse set of images it is not possible to estimate arbitrary BRDF's. A solution is to limit the estimation to low-parameter reflectance models, see also [42]. A further simplification is made by assuming that the scene is divided into areas with similar reflectance properties. Since geometry and reflectance models are estimated it is possible to render not only new views not only from new view-points but also with novel lighting conditions, e.g. by inserting virtual lights or virtual objects with interesting reflectance properties.

8. Conclusions

In this paper we have given a very brief overview over methods and tools for solving some aspects of the structure and motion problem. It has been shown that, although much work still is needed, impressive progress has been made during the last decade. Much is understood about how to estimate geometric and photometric properties of a static scene given only images as input. With such impressive progress we can hope to solve more difficult structure and motion problems with more automatic methods and more realistic scene models in the future.

The structure and motion problem is one of the fundamental problems of computer vision. Successful solution of this problem has many interesting applications where scene modeling and measurements are needed. A successful solution can also be used as a basis for the solution of other difficult computer vision questions, e.g. recognition, segmentation.

Will the structure and motion problem become so important in the future that special hardware will be built, as the scene modelers in graphic cards of today? Will they ever be as widespread as the graphic card?

9. Acknowledgments

I have tried to give a personal view of the field of multiple view vision. Obviously the results presented are the due to many researchers. I thank all of these.

References

[1] K. Åström. *Invariancy Methods for Points, Curves and Surfaces in Computational Vision*. PhD thesis, Dept of Mathematics, Lund University, Sweden, 1996.

[2] K. Åström, R. Cipolla, and P. J. Giblin. Generalised epipolar constraints. *Int. Journal of Computer Vision*, 33(1):51–72, 1999.

[3] K. Åström and A. Heyden. Euclidean reconstruction from constant intrinsic parameters. In *Proc. International Conference on Pattern Recognition, Vienna, Austria*, pages 339–343, 1996.

[4] K. Åström, A. Heyden, F. Kahl, and M. Oskarsson. Structure and motion from lines under affi ne projections. In *Proc. 7th Int. Conf. on Computer Vision, Kerkyra, Greece*, 1999.

[5] K. Åström and F. Kahl. Motion estimation in image sequences using the deformation of apparent contours. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21(2):114–127, 1999.

[6] K. Åström, F. Kahl, A. Heyden, and R. Berthilsson. A statistical approach to structure and motion from image features. In *Proc. International Workshop on Statistical Techniques in Pattern Recognition, Sedney*, 1998.

[7] K. Åström and M. Oskarsson. Solutions and ambiguities of the structure and motion problem for 1d retinal vision. *Journal of Mathematical Imaging and Vision*, 12:121–135, 2000.

[8] K. B. Atkinson. *Close Range Photogrammetry and Machine Vision*. Whittles Publishing, 1996.

[9] R. Berthilsson and K. Åström. Extension of affi ne shape. *Journal of Mathematical Imaging and Vision*, 11(2):119–136, 1999.

[10] R. Berthilsson, K. Åström, and A. Heyden. Reconstruction of curves in r^3 using factorization and bundle adjustment. *Int. Journal of Computer Vision*, 2000. accepted for publication.

[11] S. Carlsson. Suffi cient image structure for 3D motion and shape estimation. In J.-O. Eklundh, editor, *Proc. 3rd European Conf. on Computer Vision, Stockholm, Sweden*, volume I, pages 83–91. Springer–Verlag, 1994.

[12] S. Carlsson and D. Weinshall. Dual computation of projective shape and camera positions from multiple images. *Int. Journal of Computer Vision*, 27(3):227–241, 1998.

[13] M. Chasles. Question 296. *Nouv. Ann. Math.*, 14(50), 1855.

[14] R. Cipolla. *Active Visual Inference of Surface Shape*. PhD thesis, University of Oxford, 1991. also published in Springer-Verlag as LNCS 1016 (1995).

[15] R. Cipolla and A. Blake. Surface shape from the deformation of apparent contours. *Int. Journal of Computer Vision*, 9(2):83–112, 1992.

[16] G. Cross, A. W. Fitzgibbon, and A. Zisserman. Parallax geometry of smooth surfaces in multiple views. In *Proc. 7th International Conference on Computer Vision, Kerkyra, Greece*, pages 323–329, September 1999.

[17] P. Debevec. Rendering synthetic objects in to real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *in Proc. SIGGRAPH 98*, 1998.

[18] M. Demazure. Sur deux problemes de reconstruction. Technical Report 882, INRIA, Rocquencourt, France, 1988.

[19] L. Falkenhagen. Hierarchical block-based disparity estimation considering neighbourhood constraints. In *Proc. International Workshop on SNHC and 3D imaging, Rhodes, Greece*, 1997.

[20] O. Faugeras and B. Mourrain. About the correspondence of points between n images. In *IEEE Workshop on Representation of Visual Scenes, MIT, Boston, MA*, pages 37–44. IEEE Computer Society Press, 1995.

[21] O. D. Faugeras. Stratifi cation of three-dimensional projective, affi ne and metric representations. *Journal of the Optical Society of America A.*, 12(3):465–484, March 1995.

[22] O. D. Faugeras, Q.-T. Luong, and S. J. Maybank. Camera self-calibration: Theory and experiments. In G. Sandini, editor, *ECCV'92*, volume 588 of *Lecture notes in Computer Science*, pages 321–334. Springer-Verlag, 1992.

[23] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes. *Computer Graphics, Principles and Practice*. Addison-Wesley Publishing Company, 1990.

[24] P. J. Giblin, F. E. Pollick, and J. E. Rycroft. Recovery of an unknown axis or rotation from the profiles of a rotating surface. *J. Opt. Soc. America*, 11(A):1976–1984, 1994.

[25] P. J. Giblin and R. Weiss. Reconstruction of surfaces from profiles. In *Proc. 1st Int. Conf. on Computer Vision*, pages 136–144, London, 1987.

[26] R. Hartley. Minimizing algebraic error in geometric estimation problems. In *Proc. 6th Int. Conf. on Computer Vision, Mumbai, India*, pages 469–476. IEEE Computer Society Press, 1998.

[27] R. I. Hartley. A linear method for reconstruction from lines and points. In *Proc. 5th Int. Conf. on Computer Vision, MIT, Boston, MA*, pages 882–887, 1995.

[28] R. I. Hartley. Lines and points in three views and the trifocal tensor. *Int. Journal of Computer Vision*, 22(2):125–140, 1997.

[29] O. Hesse. Die cubische gleichung, von welcher die lösung des problems der homograpie von m. chasles abh ngt. *J. Reine Angew. Math.*, 26:188–192, 1863.

[30] A. Heyden. Consistency and correction of line-drawings, obtained by projections of piecewise planar objects. In J.-O. Eklundh, editor, *Computer Vision – ECCV'94*, volume 800-801 of *Lecture Notes in Computer Science*, pages 411–419. Springer Verlag, 1994.

[31] A. Heyden. *Geometry and Algebra of Multiple Projective Transformations*. PhD thesis, ISBN 91-628-1784-1, Lund University, Sweden, 1995.

[32] A. Heyden. Reduced multilinear constraints: Theory and experiments. *Int. Journal of Computer Vision*, 30(1):5–26, 1998.

[33] A. Heyden and K. Åström. Euclidean reconstruction from image sequences with varying and unknown focal length and principal point. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 438–443, 1997.

[34] A. Heyden and K. Åström. Minimal conditions on intrinsic parameters for Euclidean reconstruction. In *Proc. 2nd Asian Conf. on Computer Vision, Hong Kong, China*, 1998.

[35] A. Heyden, R. Berthilsson, and G. Sparr. An iterative factorization method for projective structure and motion from image sequences. *Image and Vision Computing*, 1999.

[36] T. Joshi, N. Ahuja, and J. Ponce. Structure and motion estimates from dynamic silhouettes under perspective projection. In *Proc. 5th Int. Conf. on Computer Vision, MIT, Boston, MA*, pages 290–295, 1995.

[37] F. Kahl and A. Heyden. Structure and motion from points, lines and conics with affine cameras. In *Proc. 5th European Conf. on Computer Vision, Freiburg, Germany*, pages 327–341, 1998.

[38] F. Kahl and B. Triggs. Critical motions in Euclidean structure from motion. In *Proc. Conf. Computer Vision and Pattern Recognition*, 1999.

[39] T. Kande and D. Morris. Factorization methods for structure from motion. *Phil. Trans. R. Soc. Lond.*, A(356):1153–1173, 1998.

[40] R. Koch. *Automatische oberflachenmodellierung starrer drei-dimensionaler objekte aus stereoskopischen rundum-ansichten*. PhD thesis, University of Hannover, Germany, 1996.

[41] R. Koch, M. Pollefeys, and L. Van Gool. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. In *Proc. 5th European Conf. on Computer Vision, Freiburg, Germany*, 1998.

[42] J. J. Koenderink, A. J. vad Doorn, and M. Stavridi. Bidirectional reflection distribution function expressed in terms of surface scattering modes. In B. Buxton and R. Cipolla, editors, *Proc. 4th European Conf. on Computer Vision, Cambridge, UK*, pages 28–39. Springer-Verlag, 1996.

[43] E. Kruppa. Zur Ermittlung eines Objektes Zwei Perspektiven mit innerer Orientierung. *Sitz-Ber. Akad. Wiss., Wien, math. naturw. Kl. Abt*, IIa(122):1939–1948, 1905.

[44] K. N. Kutulakos and S. S. M. A theory of shape by space carving. In *Proc. 7th Int. Conf. on Computer Vision, Kerkyra, Greece*, 1999.

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

[46] Q.-T. Luong. *Matrice Fondamentale et Calibration Visuelle sur l'Environnement-Vers une plus grande autonomie des systèmes robotiques*. PhD thesis, Université de Paris-Sud, Centre d'Orsay, 1992.

[47] P. Mendonca, K.-Y. Wong, and R. Cipolla. Circular motion recovery from image profiles. In *ICCV Vision and Algorithms Workshop: Theory and Practice, Corfu, Greece, 21-22 September*, 1999.

[48] D. D. Morris and T. Kanade. A unified factorization algorithm for points, line segments and planes with uncertainty models. In *Proc. 6th Int. Conf. on Computer Vision, Mumbai, India*, 1998.

[49] W. N. Newman and R. F. Sproull. *Principles of Interactive Computer Graphics*. McGraw-Hill Book Company, 1979.

[50] C. J. Poelman and T. Kanade. A paraperspective factorization method for shape and motion recovery. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(3):206–218, march 1997.

[51] M. Pollefeys, R. Koch, and L. Van Gool. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. *Int. Journal of Computer Vision*, 32(1):7–25, 1999.

[52] M. Pollefeys, L. Van Gool, and A. Oosterlinck. Euclidean reconstruction from image sequences with variable focal length. In B. Buxton and R. Cipolla, editors, *ECCV'96*, volume 1064 of *Lecture notes in Computer Science*, pages 31–44. Springer-Verlag, 1996.

[53] M. Pollefeys, L. Van Gool, and M. Oosterlinck. The modulus constraint: A new constraint for self-calibration. In *Proc. International Conference on Pattern Recognition, Vienna, Austria*, pages 349–353, 1996.

[54] J. Porrill and S. B. Pollard. Curve matching and stereo calibration. *Image and Vision Computing*, 9(1):45–50, 1991.

[55] L. Quan. Invariants of 6 points from 3 uncalibrated images. In J.-O. Eklund, editor, *Proc. 3rd European Conf. on Computer Vision, Stockholm, Sweden*, pages 459–470. Springer-Verlag, 1994.

[56] L. Quan and T. Kanade. Affine structure from line correspondences with uncalibrated affine cameras. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(8):834–845, Aug. 1997.

[57] J. H. Rieger. Three dimensional motion from fixed points of a deforming profile curve. *Optics Letters*, 11:123–125, 1986.

[58] L. Robert. Computer-vision for the post-production world: Facts and challenges through the realviz experience. In *ICCV Vision and Algorithms Workshop: Theory and Practice, Corfu, Greece, 21-22 September*, 1999.

[59] H. Schubert. Die trilineare beziehung zwischen drei einstufigen grundgebilden. *Mathematische Annalen*, 17, 1880.

[60] A. Shashua. Trilinearity in visual recognition by alignment. In J.-O. Eklundh, editor, *Computer Vision - ECCV94*, pages 479–484. Springer Verlag, Lecture Notes in Computer Vision, 1994.

[61] A. Sparr and G. Sparr. On a theorem of M. Chasles. Technical report, Dept of mathematics, Lund Institute of Technology, Sweden, 1994.

[62] G. Sparr. Simultaneous reconstruction of scene structure and camera locations from uncalibrated image swquences. In *ICPR'96*. IEEE, 1996.

[63] M. E. Spetsakis and J. Y. Aloimonos. A unified theory of strucure and motion. In *Proc. DARPA Image Understanding Workshop*, 1990.

[64] P. Stefanovic. Relative orientation - a new appraoch. *ITC Journal*, 3:417–448, 1973.

[65] P. Sturm. Critical motion sequences for monocular self-calibration and uncalibrated Euclidean reconstruction. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 1100–1105, 1997.

[66] P. Sturm. *Vision 3D non calibrée: Contributions à la reconstruction projective et étude des mouvements critiques pour l'auto-calibrage*. PhD thesis, Institut National Poytechnique de grenoble, 1997.

[67] P. Sturm and B. Triggs. A factorization based algorithm for multi-image projective structure and motion. In B. Buxton and R. Cipolla, editors, *Computer Vision – ECCV'96*, volume 1065 of *Lecture Notes in Computer Science*, pages 709–720. Springer Verlag, Lecture Notes in Computer Science, 1996.

[68] R. Szeliski and R. Weiss. Robust shape recovery from occluding contours using a linear smoother. In T. Terzopoulis and C. Brown, editors, *Real-Time Vision*. Cambridge University Press, 1994.

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

[70] B. Triggs. Autocalibration and the absolute quadric. In *Proc. Conf. Computer Vision and Pattern Recognition*, 1997.

[71] F. Ulupinar and R. Nevatia. Perception of 3-d surfaces from 2-d contours. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 15(1):3–18, 1993.

[72] R. Vaillant and O. D. Faugeras. Using extremal boundaries for 3D object modelling. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 14(2):157–173, 1992.

[73] B. Vijayakumar, D. Kriegman, and J. Ponce. Invariant-based recognition of complex 3D curved objects from image contours. In *Proc. 5th Int. Conf. on Computer Vision, MIT, Boston, MA*, pages 508–514, 1995.

[74] B. Vijayakumar, D. Kriegman, and J. Ponce. Structure and motion of curved 3D objects from monocular silhouettes. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 327–334, 1996.

[75] J. Weng, T. Huang, and N. Ahuja. Motion and structure from line correspondances: Closed-form solution, uniqueness, and optimization. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 14(3), 1992.

[76] Y. Yu, P. Debevec, J. Malik, and T. Hawkins. Inverse global illumination: Recovering reflectance models of real scenes from photographs. In *in Proc. SIGGRAPH 98*, 1998.

[77] A. Zisserman, P. Beardsley, and I. Reid. Metric calibration of a stereo rig. In *Proc. IEEE Workshop on Representation of Visual Scenes*, 1995.