

Evaluation of Corner Extraction Schemes Using Invariance Methods*

Anders Heyden[†] Karl Rohr[‡]

[†] Dept of Mathematics, Lund University, Box 118, S-221 00 Lund, Sweden

[‡] FB Informatik, University of Hamburg, Vogt-Kölln-Str. 30, 22527 Hamburg, Germany

email: heyden@maths.lth.se rohr@informatik.uni-hamburg.de

Abstract

We describe a new method to evaluate corner extraction schemes using invariance methods. Since the locations of corners in an image depend both on the intrinsic parameters of the camera and the relative position and orientation of the object with respect to the camera, the exact positions of corners in an image are generally not known. To circumvent the need for this knowledge, we here use sets of points (instead of individual points) extracted from images of polyhedral objects and projective invariants to calculate a manifold of constraints on the coordinates of the corners. We then estimate the variance of the detected corners from the distance of the coordinate vector to this manifold. This is independent of the camera parameters and the relative position and orientation between the camera and the object. Five different kinds of corner extraction schemes are investigated. The purpose of the paper is to show that invariance methods can effectively be used to make this comparison rather than to make a thorough comparison of different corner extraction schemes.

1. Introduction

Corner extraction schemes are important in computer vision to provide input to other kind of operations, such as recognition and reconstruction. Most approaches to recognition and reconstruction are based on the locations of extractable points, usually corners, in one image or in several images, see for example [18], [6], [4], [10] and [8]. Since the reliability of the overall approach heavily depends on the accuracy of the extracted points it is necessary to have good performance of the corner extraction schemes as well as having an estimate of how accurate the locations of corners are. Such estimates can be used in a subsequent step for hypothesis testing, for example, in recognition tasks.

However, it is difficult to evaluate the performance of corner extraction schemes because the true positions of the corners in an image are generally not known. This knowledge would be at our hand only if we knew exactly the intrinsic parameters of the camera as well as the relative position and orientation of the object with respect to the camera. However, since this knowledge is in general not available it is hard to judge the performance of corner extraction schemes.

Invariants have been an area of extensive research in computer vision during the last few years (e.g. [15], [17]). Invariants tell us relations between the object and image that are invariant under the camera position, orientation and geometry, that is intrinsic parameters and extrinsic parameters. These invariants constrain the locations of the corners in the image to lie on a certain manifold, here called the manifold of constraints. This manifold can be used to estimate the variance of the corner extraction schemes using statistical methods.

In the present paper we will do this for five different kinds of corner extraction schemes. The first is manual selection, the second is the local operator of Kitchen and Rosenfeld [9], the third is the local operator together with nonlinear diffusion, the fourth is the model fitting approach of Rohr [13] and the fifth is model fitting in conjunction with nonlinear diffusion.

There are different invariants that can be used depending on how much a priori knowledge we have of the object. If the only knowledge is that the object consists of planar surfaces, that is a polyhedral object, the only constraints of the locations of the corners are obtained from the so called second order syzygies, see [7]. If we know the shape of the different planar surfaces, for example if they are rectangular, this can be used to constrain the coordinates further, see [18]. This can be effectively done using the theory of shape and depth and the so called S -matrix, see [18] and [7]. If we have a complete model of the object we can estimate the 11

*This work has been done within the ESPRIT project BRA EP 6448, VIVA.

parameters in the uncalibrated projective camera model and then calculate the errors.

In previous work, the performance of different corner detectors on real images has mainly been judged by visual inspection of experimental results (e.g. [9]). Analytic studies of corner detectors have been done by Giraudon and Deriche [5] and Rohr [14]. Giraudon and Deriche use the L-corner model of Berzins [2] for specific aperture angles of 45° and 90° to compare the localisation of three local (differential) corner detectors. In Rohr [14] an analytic study of ten local (differential) corner operators has been performed. This study is based on the parametric L-corner model in [12],[13] and evaluates the localisation of corners for aperture angles in the whole range of 0° to 180° . For each corner detector an implicit equation is derived which characterises the localisation of the corner points in dependence of all model parameters (aperture angle, contrast, image blur).

In Coelho et al. [3] planar invariants (e.g. the cross ratio) have been used to compare the performance of three indirect corner detectors. Indirect corner detectors first extract edge points and then operate on this data to determine the corner positions.

In the present study, we use projective invariance methods to investigate the performance of five direct schemes for extracting corners. In contrast to our previous study for single points in [14] we here analyse sets of points. The considered points are corners of 3D polyhedral objects, e.g. L-, T-, Y-, and Arrow-corners. The efficiency of the different approaches is judged w.r.t. consistency conditions which the extracted point sets have to fulfil for a valid 3D interpretation.

2. Corner extraction

We investigate five direct schemes for extracting the corner positions of 3D polyhedral objects. Direct schemes can be classified into local and semi-global approaches. From each class we have selected one representative. Additionally, each of the two approaches has been investigated in conjunction with a nonlinear diffusion scheme. The results are compared with manually selected positions.

1. Manual selection: The image is displayed on the screen and the pixel (integer-)coordinates are picked out by hand.

2. Local operator (Kitchen and Rosenfeld 82): To detect corners Kitchen and Rosenfeld [9] proposed a differential operator which consists of the first and second partial derivatives of the image $g(x, y)$:

$$KR(x, y) = \frac{g_x^2 g_{yy} - 2g_x g_y g_{xy} + g_y^2 g_{xx}}{g_x^2 + g_y^2} \quad (1)$$

In our implementation we have used 5×5 Beaudet masks [1] to compute the partial derivatives.

3. Local operator and nonlinear diffusion: One problem with local operators is that they are relatively noise sensitive. Often points are detected which do not correspond to corners. Nonlinear diffusion approaches, on the other hand, reduce the amount of noise while at the same time preserve the systematic intensity variations. To reduce the influence of the noise (and the influence of possibly other distortions) we therefore suggest to apply the local corner operator on nonlinear diffused images (instead of using the original images). In our case, we investigate the nonlinear diffusion approach of Schnörr [16] in conjunction with the local operator of Kitchen and Rosenfeld [9].

4. Model fitting (Rohr 90,92): In Rohr, [12], [13] a model-based approach for recognising edges and corners has been introduced. This approach is based on a parametric model, which represents the systematic intensity variations of corners resulting from polyhedral 3D objects. The model is a superposition of L-corner model functions and comprises arbitrary complex corners in terms of the number of intersecting edges, for example, L-, T-, Y-, and Arrow-corners. Corner localisation is achieved by fitting the model directly to the image intensities within semi-global regions (e.g. 25×25 windows) yielding subpixel positions. Under reasonable assumptions the exact geometric positions of the corners can be determined independently of the amount of image blur. Denoting the parametric model with $g_M(x, y, p)$ and the image intensities with $g(x, y)$ the parameters p are estimated by a least-squares fit:

$$\min_{p \in \mathbb{R}^n} \sum_{i=1}^m (g(x_i, y_i) - g_M(x_i, y_i, p))^2 \quad (2)$$

5. Model fitting and nonlinear diffusion: To reduce the influence of the noise and of possibly other distortions we here apply the model fitting approach as described above also on nonlinear diffused images obtained with the approach in [16].

3. Invariance methods

In this section we will briefly describe the different methods used to check the consistency of the detected corners. We remark that the ‘true’ location of the corners are not known. This means that we can only check the consistency of the image with respect to the a priori information of the object. We remark that the camera model used is the standard uncalibrated perspective camera and we have assumed that no nonlinear deviations are present.

The image of the truncated cube, see the left part of Figure 1, contains 4 planar surfaces, one on top, one on the left, one on the right and one triangular in the middle. The only assumption made on the truncated cube is that these four surfaces are planar. When this assumption is made there is

only one consistency condition on the location of the corners. Consider the lines between the top surface and the left surface, between the top surface and the right surface and between the left surface and the right surface. These lines are intersections between the surfaces and are seen as lines connecting some of the located points in the image. The lines are determined by the coordinates of the six points on the lines. The three surfaces defining the lines, the top surface, the left surface and the right surface, meet somewhere in 3D space. In general three planes meet in a single point. Since the three lines mentioned above are intersections between pairs of these planes, these three lines must also intersect in one point.

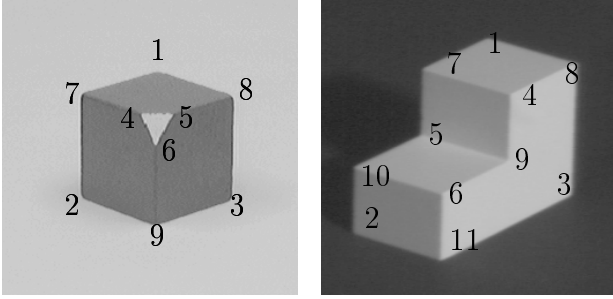


Figure 1. Image of a truncated cube and an L-shaped object.

This condition can be formulated mathematically by the so called second order syzygies of the S -matrix or shape matrix, see [18], [7]. Let $S(\mathcal{Y})$ denote the S -matrix corresponding to the point configuration \mathcal{Y} . The S -matrix can be used to determine whether the line-drawing is correct or not by using the following theorem, see [7].

Theorem 3.1. \mathcal{X} is an impossible planar configuration if and only if

$$\text{rank } S(\mathcal{A}) \geq |V(\mathcal{A})| - 3 \quad (3)$$

for some collection of faces \mathcal{A} of \mathcal{Y} , where $|V(\mathcal{A})|$ is the number of points in \mathcal{A} .

In this case the image is correct if and only if $\text{rank } S(\mathcal{Y}) = 5$. In order to check this condition we have to examine the rank by using a singular value decomposition. If the sixth singular value is close to zero we can detect the corresponding linear combination, w , of the columns that is closest to zero. Then it is possible to adjust the entries in the S -matrix such that the linear combination, w , of the columns are exactly zero, see [7] again. This is a linear algorithm and it decreases the rank of the new S -matrix, \tilde{S} , to 5. Finally, we must find a point configuration as close as possible to the original one, in the Euclidean metric, but with the new S -matrix, \tilde{S} , as its S -matrix. This problem is

again a quadratic goal function with linear constraints and the solution is

$$\tilde{x} = x - \tilde{S}(\tilde{S}^T \tilde{S})^{-1} \tilde{S}^T x \quad , \quad (4)$$

where x is a vector with components x_i and similarly for y .

In order to test the consistency of the detected corners for the L-shaped object in the right part of Figure 1 we must use more a priori information than just planarity of the faces. We will use two different levels of a priori information.

First, if we know the shape of the faces, we can calculate the S -matrix, $S(\mathcal{X})$ of the 3D object, \mathcal{X} , see [18]. The following equation, see [18] again, is useful

$$\text{diag}(\alpha)S(\mathcal{X}) = S(\mathcal{Y}) \text{diag}(c) \quad , \quad (5)$$

where \mathcal{X} is the object, \mathcal{Y} is the image, α is a vector containing the relative depths and c is a vector needed to adjust for different scale factors. From this equation it is possible to solve for α and c , using linear methods. Then the new S -matrix can be calculated as $\tilde{S} = \text{diag}(\alpha)S(\mathcal{X})$. Finally we use (4) again which gives a corrected image.

Secondly, if we have a complete model of the object we can do a little more. The camera equation can be written, using a camera matrix, P :

$$\lambda_i \mathbf{x}_i = P \mathbf{X}_i \quad , \quad (6)$$

where $\mathbf{x}_i = [x_i \ y_i \ 1]^T$, $\mathbf{X}_i = [X_i \ Y_i \ Z_i \ 1]^T$, and $\lambda_i = \alpha_i^{-1}$ are the inverse relative depths. Here it can also be seen that these equations are linear in λ and P . Thus it is possible, using only a singular value decomposition, to compute the relative depths and the camera matrix. Finally, the new corrected coordinates in the images can be calculated from (6) now when both the relative depths and the camera matrix are known.

4. Statistics

It is possible to estimate the variance without bias if we know the degree of freedom, or dimension, of the manifold that constrains the correct corner positions, see [11] for further details. If we measure the error as the distance to some manifold with dimension d it can be shown that the variance can be estimated, in a first order approximation, as

$$\bar{\sigma}^2 = \frac{1}{n-d} \sum_{i=1}^n e_i^2 \quad . \quad (7)$$

In the truncated cube case we have just one constraint on the image coordinates and thus $d = n - 1 = 18 - 1 = 17$. In the first consistency test of the L-shaped object we have $n - d = 11$ constraints on the corner positions. In the second consistency test we also have $n - d = 11$ constraints on the coordinates, which is the same as the number of coordinates

minus the degree of freedom in the projection matrix for an uncalibrated camera.

The results are shown in Table 1. The average standard deviation have been calculated using all points, in the three tests, that is $18 + 22 + 22 = 62$ coordinates and $n - d = 1 + 11 + 11 = 23$ as the total number of constraints.

We remark that these estimates are obtained from a nonexact calculation of the shortest distance to the manifold of constraints and this will of course bias the estimates, making them bigger. We have also used a linear approximation of the manifold of constraints which will also bias the estimates.

5. Experimental Results

Our approach has been investigated on the camera recorded images shown in Figure 1. The corner positions for the truncated cube image have been extracted using the five schemes described in Section 2. Note, that the quality of the image is not very good. The noise is relatively large and also several other distortions are present. Therefore, the application of the local operator resulted in many misdetections. When using the local operator on the nonlinear diffused image the number of misdetections was reduced largely. To obtain the right number of corners, in either case we manually selected those detected points which were closest to the actual corner positions. The model fitting approach also has been assisted interactively to obtain good starting values for the minimisation procedure. As with the local operator the robustness improved when using the nonlinear diffused image.

Figure 2 shows the resulting errors (in pixels) of the consistency test described in Section 3 applied to the corner points of the five extraction schemes. We have also marked the three points with largest deviations. Note, that our test only checks the consistency of point sets and therefore the corrected points may well be displaced from the true positions. It can be seen that the smallest values of the estimated standard deviation are obtained with the model fitting approach and with model fitting in conjunction with nonlinear diffusion. The local operator yields the largest value. Regarding the three points with largest errors we see a remarkable consistency among the different extraction schemes. Nearly all approaches rate the corner points 6, 9, and 5 (in this order) to be farthest away from the corrected positions. Also, we have about the same relative errors in horizontal and vertical direction.

In Table 1 are the estimated standard deviations (in pixels) for the different corner extraction schemes. The best result is obtained with the model fitting approach followed by model fitting in conjunction with nonlinear diffusion, manual selection, local operator together with nonlinear diffusion, and local operator. It appears that nonlinear diffusion

as a preprocessing step increases the accuracy of the local operator by a factor of three while the accuracy of the model fitting approach decreases slightly.

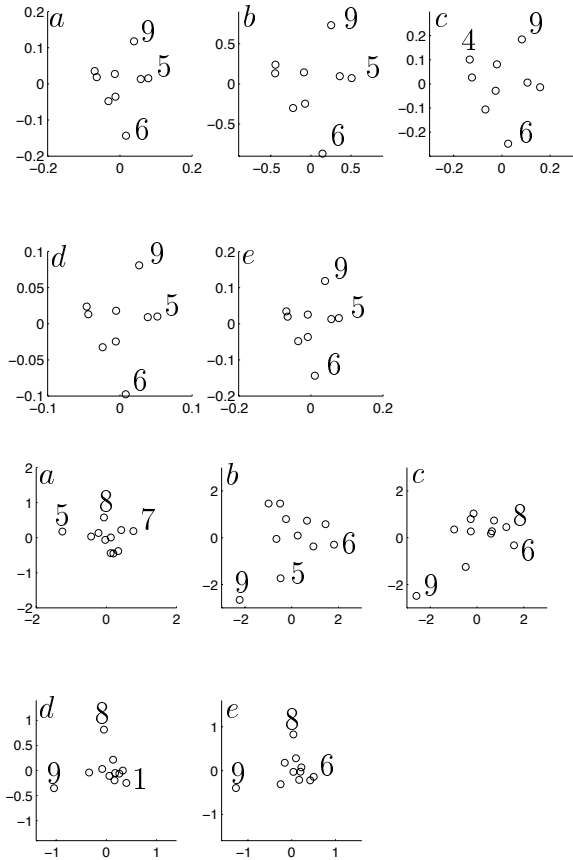


Figure 2. Graphical representation of the errors (in pixels) for the truncated cube(above) and for the second consistency test applied to the L-shaped object (below); a) manual selection, b) local operator (Kitchen/Rosenfeld), c) local operator & nonlinear diffusion, d) model fitting (Rohr) and e) model fitting & nonlinear diffusion. The three ‘worst’ points are indicated with their point number from Figure 1.

For the L-shaped object in Figure 1 the resulting errors (in pixels) of the second consistency test described in Section 3 are shown in Figure 2. Table 1 shows the estimated standard deviations (in pixels) for the different corner extraction schemes using both consistency tests. Also with this image and the two consistency tests we qualitatively obtain the same result as before.

Taking all measurements for the two images and the corresponding three consistency tests we obtain the average (corrected) estimates of the standard deviation of the different corner extraction schemes as shown in the last row of Table 1. For the model fitting approach the standard deviation

turns out to be about 0.5 pixels. Model fitting in conjunction with nonlinear diffusion and manual selection are somewhat worse and the accuracy of the local operator is about three times worse than the model fitting approach.

6. Conclusions

We have investigated the performance of five different corner extraction schemes using three different types of invariants to constrain the locations of the coordinates of the set of corners. The variance of the corner extraction schemes has been estimated using the distance to the manifold of constraints. It turns out that the model fitting approach of Rohr [13] yields the best result, manual selection is somewhat worse than model fitting and the local operator of Kitchen and Rosenfeld [9] turns out to be a factor of three worse. Nonlinear diffusion as a preprocessing step seems to make the local operator more robust but slightly decreases the accuracy of model fitting.

The use of invariance methods to evaluate corner extraction schemes seems to be successful and easy to use, since the camera does not need to be calibrated and the relative position and orientation between the camera and the object need not be known in advance. However, in order to make a more reliable estimate of the performance of the different corner extraction schemes one has to use more images and statistical methods.

Acknowledgement

This work has been done as a collaborative project within the ESPRIT-BRA Project VIVA, sponsored by the European Union (EU). We also thank Kalle Åström for interesting discussions of the statistics involved.

References

[1] P. R. Beaudet. Rotationally invariant image operators. *Proc. Intern. Joint Conf. on Pattern Recognition*, (Kyoto/Japan):579–583, 7-10 Nov. 1978.
 [2] V. Berzins. Accuracy of Laplacian Edge Detectors. *Computer Vision, Graphics, and Image Processing*, (27):195–210, 1984.

[3] C. Coelho, H. A., J. L. Mundy, D. A. Forsyth, and A. Zisserman. An Experimental Evaluation of Projective Invariants. *Geometric Invariance in Computer Vision*, (J.L. Mundy and A. Zisserman (Eds.)):87–104, 1992.
 [4] O. D. Faugeras. *Three Dimensional Computer Vision, A Geometric Viewpoint*. The MIT Press, Cambridge, Massachusetts, London, England, 1993.
 [5] G. Giraudon and R. Deriche. On Corner and Vertex Detection. *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pages 650–655, 3-6 June 1991.
 [6] A. Heyden. Reconstruction from Image Sequences by means of Relative Depths. *Proc. ICCV'95*, pages 1058–1063, 1995.
 [7] A. Heyden. On the Consistency of Line-Drawings, Obtained by Projections of Piecewise Planar Objects. *J. of Mathematical Imaging and Vision*, (to appear), 1996.
 [8] A. Heyden and K. Åström. A Canonical Framework for Sequences of Images. *Proc. IEEE Workshop on Representation of Visual Scenes*, pages 45–42, 1995.
 [9] L. Kitchen and A. Rosenfeld. Gray-level corner detection. *Pattern Recognition Letters*, (1):95–102, 1982.
 [10] S. Maybank. *Theory of Reconstruction from Image Motion*. Springer Verlag, Berlin, Heidelberg, New York, 1993.
 [11] C. R. Rao. *Linear Statistical Inference and Its Application*. John Wiley & Sons, Inc. 1965.
 [12] K. Rohr. Über die Modellierung und Identifikation charakteristischer Grauwertverläufe in Realweltbildern. *12. DAGM-Symposium Mustererkennung*, (Oberkochen-Aalen, Informatik-Fachberichte 254):217–224, 24-26 Sept. 1990.
 [13] K. Rohr. Recognizing Corners by Fitting Parametric Models. *Intern. J. of Computer Vision*, 9(3):213–230, 1992.
 [14] K. Rohr. Localization Properties of Direct Corner Detectors. *J. of Mathematical Imaging and Vision*, 4(2):139–150, 1994.
 [15] C. Rothwell, D. A. Forsyth, A. Zisserman, and J. L. Mundy. Extracting Projective Structure from Single Perspective Views of 3D Point Sets. *ICCV'93*, pages 573–582, 1993.
 [16] C. Schnörr. Unique Reconstruction of Piecewise-Smooth Images by Minimizing Strictly Convex Nonquadratic Functionals. *J. of Mathematical Imaging and Vision*, 4(2):189–198, 1994.
 [17] G. Sparr. Projective Invariants for Affine Shapes of Point Configurations. *ESPRIT/DARPA Invariants Workshop*, 1991.
 [18] G. Sparr. Depth-Computations from Polyhedral Images. *ECCV'92, Lecture Notes in Computer Science*, 588:378–386, 1992.

Type	Manually	Local op. (Kitchen/R.)	Local op. & nonlin. diff.	Model fitting (Rohr)	Model fitting & nonlin. diff.
Truncated cube	0.2571	1.5750	0.4564	0.1690	0.2486
L-shaped object, I	0.7290	1.9395	1.7015	0.6057	0.6710
L-shaped object, II	0.5869	1.6451	1.4799	0.4852	0.5587
Average	0.6493	1.7892	1.5624	0.5379	0.6060

Table 1. Estimated standard deviations (in pixels) for the different corner extraction schemes.