POLYOC: Photogrammetric Measurements of Buildings and Architectural Models via Multi-View Stereovision

Volker Steinhage

Universität Bonn/Germany
Institut für Informatik

Abstract

We present a new method to derive a reliable 3D description of buildings and architectural models based on stereovision using multiple camera views. This multi-view stereo technique has been successfully applied to many objects and has proved to be more reliable and more accurate than the well known binocular and even trinocular stereo techniques. The method has the following attractive features:

- *Flexibility: Arbitrary positions of an arbitrary amount of different camera views are allowed.*
- *Automatic Calibration of each camera view is obtained by a simple procedure.*

- *Reliable Stereo Analysis: The matching of corresponding image features of different images relies on a basic paradigm of Hypothesis Prediction and Validation.*

- *Knowledge Base: Heuristics can be applied to correct the object reconstruction.*

- *CAD-Interface: The 3D description of the reconstructed object can automatically transferred via DXF to a CAD-system.*

- *Low Costs: No special and expensive technical equipment like a laser or an UV range finder is necessary. A camera, a PC with CAD-system and a digitizing device like a digitizing tablet with a pointing pen are sufficient.*

Abstract

Das System POLYOC akzeptiert als Eingabe drei oder mehr Foto- oder Videoaufnahmen eines Bauobjekts oder auch eines Architekturmodells, z.B. eines Tragwerkmodells.

Zunächst sind relevante Bildmerkmale zu extrahieren; d.h. hier : Extraktion relevanter Raumkanten und Eckpunkte fiir jede Aufnahme. Dieser Prozeß kann wahlweise automatisch durch Algorithmen der Bildverarbeitung oder durch Anklicken der Eckpunkte mit einem Fadenkreuz auf dem Monitor oder mit einem Sensorstift auf dem Digitalisierungstablett erfolgen.

Der Hauptschritt zur automatische Berechnung des dreidimensionalen Gittermodells basiert auf einer erweiterten Stereoanalyse. In diesem Schritt erfolgt eine automatische Bestimmung der Kameraparameter für jede Aufnahme, eine Zuordnung der Eckpunkte zwischen allen Bildaufnahmen sowie die Bestimmung der dreidimensionalen Koordinaten mittels Triangulisation.

In einer Nachbearbeitung kann der Benutzer eine automatische Modellkorrektur durchführen. Hier stehen z.B. die folgenden wissensbasierten Prozeduren zur Verfügung : Heuristiken zur Vertikalstellung von Raumkanten mit einer Winkelabweichung von bis zu 50 von der Raumvertikalen, zur Horizontallegung von Raumkanten mit einer Winkelabweichung von bis zu 50 von der Raumhorizontalen sowie eine Heuristik zur Winkelkorrektur von nicht horizontalen oder vertikalen Raumkanten, die eine bekannte Neigung haben (z.B. Dachneigung).

Im letzten Schritt erfolgt ein Transfer der dreidimensionalen Objektrekonstruierten im DXF-Format in ein CAD-System.
Das System POLYOC ist flexibel. Es arbeitet mit beliebig vielen Bildern aus beliebig unterschiedlichen Aufnahmepositionen. Die Verwendung mehrerer Bilder erhöht die Sicherheit und Genauigkeit der Rekonstruktion. Die Korrespondenzen zwischen den extrahierten Bildmerkmalen werden auf verläß1iche Weise durch ein generate-and- test- Verfahren verifiziert.

Abstract

We present a new method to derive a reliable 3D description of buildings and architectural models based on stereovision using multiple camera views. This multi-view stereo technique has been successfully applied to many objects and has proved to be more reliable and more accurate than the well known binocular and even trinocular stereo techniques. The method has the following attractive features:

- *Flexibility: Arbitrary positions of an arbitrary amount of different camera views are allowed.*
- *Automatic Calibration of each camera view is obtained by a simple procedure.*
- *Reliable Stereo Analysis: The matching of corresponding image features of different images relies on a basic paradigm of Hypothesis Prediction and Validation.*
- *Knowledge Base: Heuristics can be applied to correct the object reconstruction. -*
- *CAD-Interface: The 3D description of the reconstructed object can automatically transferred via DXF to a CAD-system.*
- *Low Costs: No special and expensive technical equipment like a laser or an UV range finder is necessary. A camera, a PC with CAD-system and a digitizing device like a digitizing tablet with a pointing pen are sufficient.*

Binocular and Trinocular Stereovision

Stereo analysis has been developed to measure positions and shapes of 3D objects for applications in aerial cartography, automatic surveillance, parts identification in industrial automation, and passive robot navigation. The main problem in stereo analysis is the matching of stereo images to ascertain the corresponding points, the so called correspondence problem by [Ullman79].

In conventional binocular stereo analysis. various matching techniques have been used to avoid ambiguous matches: relaxation in [Barnard&Thompson801, coarse-to-fine in [Marr82], [Grimson851, and [Posch881, and dynamic programming in [Otha&Kanade851. Although these schemes are useful as stereo analysis strategies they cannot avoid ambiguous matches which are attributed to the appearance of scene points with similar features or to local vanishing of true match points caused by occlusions or due to noises in video signals.

In trinocular stereovision a third camera view is added to solve the correspondence problem geometrically. While in binocular stereovision the search for corresponding points is a search along conjugated epipolar lines, in trinocular stereovision the search for corresponding points between two images can be reduced to a simple verification at a precise location in the third image. A survey about approaches in trinocular stereovision is given in [Steinhage89].

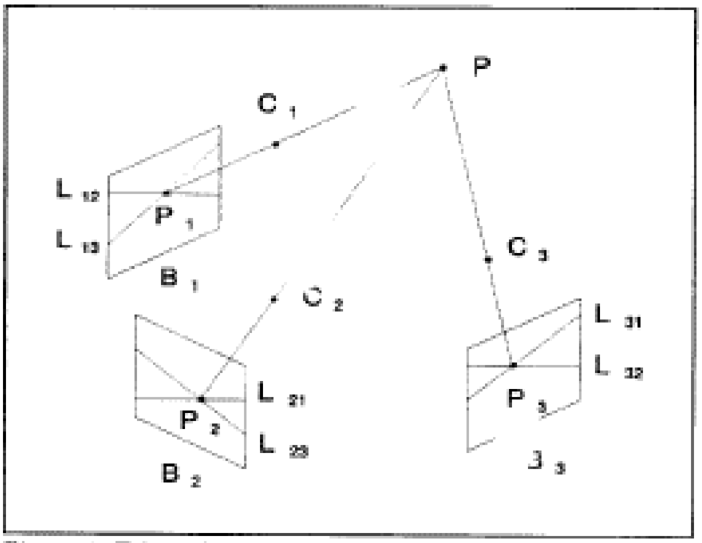


Figure 1: Trinocular stereovision.

Figure 1 shows the epipolar geometry for trinocular vision: each camera is modeled by its optical center Ci and its image plane Bi. An image point Pi of the image plane Bi defines together with the optical center Ci and the second optical center Cj the epipolar plane (Pi, Ci, Cj). The intersection of this plane with the image plane Bj yields the epipolar line Lji on which the corresponding image point Pj in the image Bj must lie. To verify the correspondence of the two image points Pi and Pj the corresponding image point Pk in the third image Bk has to lie on the intersection point of both epipolar lines Lki and Lkj.

While the use of a third view increases the reliability of the correspondence analysis in general, there are certain cases remaining where correspondence cannot be established. These cases are often not mentioned in other references or occluded by definition like the so called general position assumption in [Binford81]. and [Marr82]. But our experiences in trinocular vision show that these cases quite often occur and are not negligible because they result in wrong matches. We have examined the following error cases:

(1) The visible point in two images does not appear in the third image. The third image cannot verify a true correspondence found in the two images.

(2) Two distinct pairs of corresponding points in two images are both verified by the same point in the third image. This happens whenever one object point is hidden by another point in one image. The computation of the D position will yield a wrong value for the hidden object point.

(3) A false correspondence between two points in two images is verified by a third false correspondence in the third image. Ibis case appears if the lines of sight from the three cameras to three distinct object points intersect at one point. A non-existing point will be computed.

The Multi-View Algorithm

The three error cases described above indicate that additional information is needed to solve correspondences more reliable. So we propose in [SteinhageEtAl91] a multi-view stereovision system POLYOC which derives the correspondences in principle via three images. But the algorithm takes in account a user-defined number of additional images. Each additional camera view provides an additional measurement which increases the reconstruction reliability and accuracy of each 3D point. The three error cases can he solved much more reliable. POLYOC includes the following steps:

(1) calibration of the camera images.

(2) image processing and feature extraction.

(3) object reconstruction via multi-view stereovision.

(4) object correctures via knowledge based heuristics.

(5) transformation of the 3D description into DXF

Calibration

Theoretically six non coplanar points are sufficient for determining the perspective transformation matrix T uniquely. T is a matrix of dimension 3 x 4, but it is defined up to a scale factor. One can specify T uniquely assuming that

$$t_{34} = 1.$$

Each time an image point Pi = (u,v)t is identified as the image of a scene point P = (x,y,z)t, this provides the following two linear equations on the eleven unknowns remaining for determining the transformation matrix T:

$$P^t * t_1 + t_{14} - u * (P^t * t_3 + 1) = 0$$

$$P^t * t_2 + t_{24} - v * (P^t * t_3 + 1) = 0$$

where tjk is the element of rank (j,k) in T, and tj is the 3-vector obtained from the first 3 elements of the jth row of T:

$$t_i = (t_{i1}, t_{i2}, t_{i3})t$$

The 3D coordinates (xc, yc, zc) of the optical center C of each camera view is then obtained by solving:

$$(0, 0, 0)t = T * (x_c, y_c, z_c)t$$

which is a system of three linear equations in three unknowns.

Practically several points of the object which has to be measured are available as these six calibration points. The user of POLYOC has to determine at least six points of the object with known 3D coordinates just once to initialize the calibration procedure for all camera images.

Feature Extraction

The reconstruction process via the multi-view algorithm is driven by comers and edge line drawing. For extracting these image features two procedures are available in POLYOC:

(1) Automatic picture processing and feature extraction via algorithms for extracting and thinning of edge segments and determining the line equations of the edges and the coordinates of the corner points in each image.

(2) Manual feature extraction via a digitizing device like a digitizing tablet with a pointing pen.

Automatic Feature Extraction

If the surfaces of the object which has to be measured are mostly uniform in texture the procedure for automatic feature extraction is worth to be applied. This procedure includes the following steps:

(1) Reduction of the noise signal by a 5x5-Median operator. This operator replaces each pixel intensity by themedian value of its neighbours within a 5x5 - neighbourhood.

(2) Extraction of edge points by the Sobel operator. The Sobel operator estimates the first derivative of the intensity values of the center pixel within a 3x3-neighbourhood. If this value exceeds a preset threshold the center pixel is classified as an edge pixel.

(3) Thinning of the extracted edge picture by an algorithm described in [PavIadis821. This algorithm cheeks the3 x 3 -neighbourhood of an edge pixel extracted by the Sobel operator and eliminates edge pixels in such a way that all remaining edge pixels constitute edge segments which are only one pixel thick.

(4) Computation of connected edge lines by a weighted polarized Hough transform proposed in [Lie/Chen901. The standard Hough technique detects connected edge lines by transforming all edge pixels into a two-dimensional parameter space where every edge pixel votes for all that lines on which this pixel potentially can lie on. The peaks in the parameter space determine the line equations of the edge lines which actually exist in the original image. Unlike the standard Hough transform, the weighted polarized Hough transform restricts the transform mapping to within a polarizing zone of the local fitting line. thus eliminating noisy votes on parameter space.

(5) Computation of the 2D corner points. The comer points are derived by investigating the intersections of the edge lines determined by the weighted polarized Hough transform.

The following figure 2 shows the steps of the automatic feature extraction procedure working on a simple test object: (a) the original video image, (h) the edge picture extracted via the Sobel operator, (c) the Hough space of the thinned edge picture, (d) seven peaks found in the Hough space corresponding the seven line equations of the connected edge lines of the Sobel picture. (c) the computed corner points projected in the original video image, and (f) the complete object reconstruction derived from four video images via the multi-view stereo analysis.

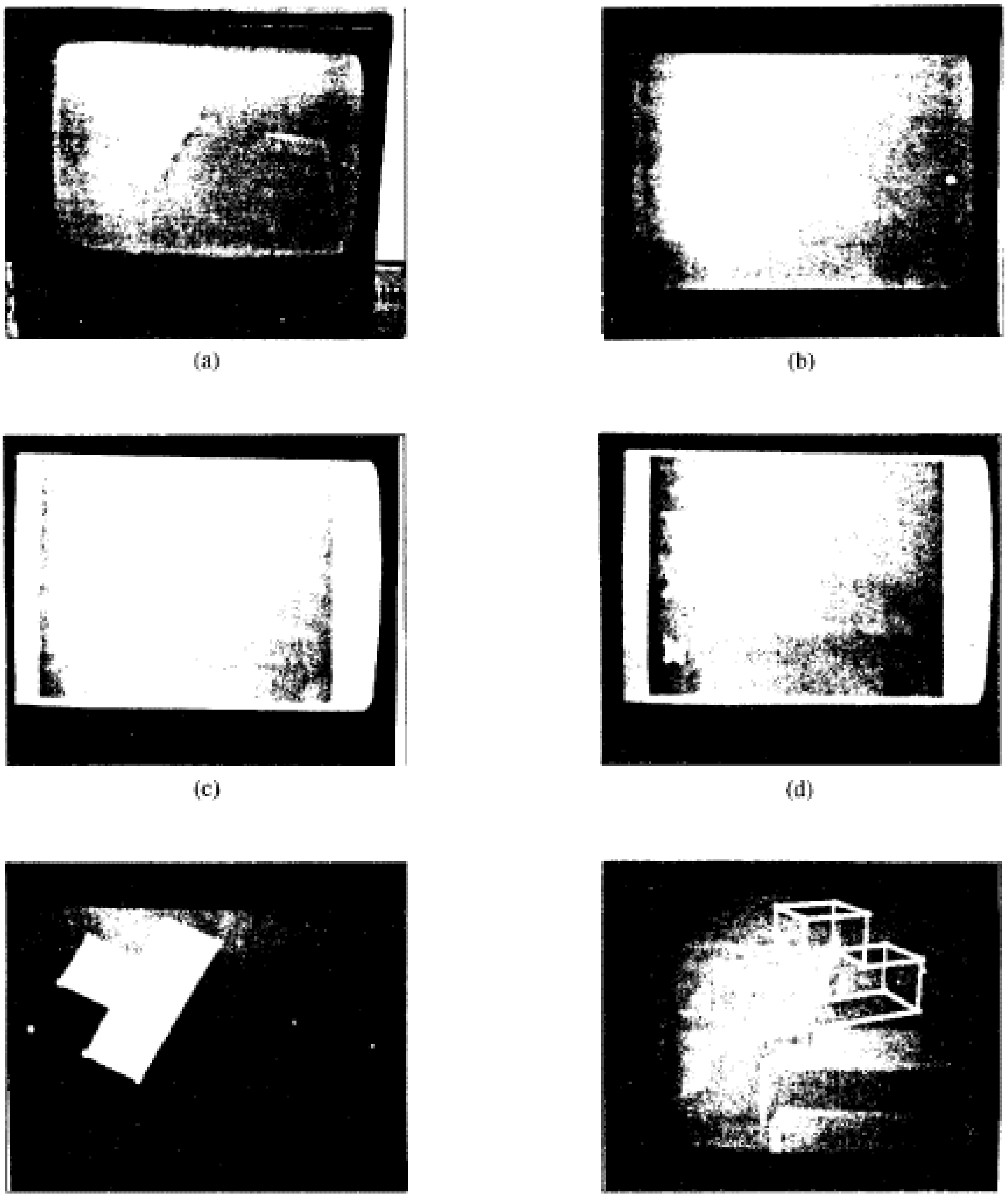


Figure 2 : (a) the original video image of a simple test object, (b) the edge picture extracted via the Sobel operator, (c) the Hough space of the thinned edge picture, (d) the peaks found in the Hough space corresponding the line equations of the connected edge lines of the Sobel picture, (e) the computed corner points projected in the original video image, and (f) the complete object reconstruction derived from four video images via the multi-view stereovision algorithm.

Manual Feature Extraction

If the surfaces of the object which has to be measured contain too many unimportant details or the surrounding of the object shows to much structures the procedure for automatic feature extraction would extract too many irrelevant information. The consequence would be a slow down of the whole reconstruction process and unnecessary sources of error. In this case the manual feature extraction procedure is advantageous.

The manual feature extraction procedure allows the user of POLYOC to digitize the images of the object of interest by hand. Via a digitizing device the user can mark significant points of the object and connect them with the edges found in the image. For this purpose POLYOC offers a lot of helpful procedures for drawing and erasing comer points and connected loops of edges etc. So the user of POLYOC can easily extract all significant corner points and edges of the building or model which has to be measured.

Determining the Calibration

As in section 2 stated at least six points with known 3D coordinates are necessary for the calibration of each image. The user of POLYOC has to determine these calibration points out of the list of extracted corner points. For this purpose a catch procedure is available to determine the calibration points in the extracted and thinned edge pictures in an elegant fashion.

3D Reconstruction via Multiple-View Stereovision

The result of the feature extraction are two lists for each image: the first list contains all extracted corner points of the image, the second list holds the edges connecting the comer points. Furthermore the camera parameters obtained from the calibration points via the calibration procedure are attached to each image. The reconstruction process via the multi-view stereo algorithm is driven by the extracted comer points. The matching of corresponding comer points of different images relies on a basic paradigm of Hypothesis Prediction and Validation.

Hypotheses Prediction

The hypotheses for corresponding corner points are derived via a generalized epipolar constraint which requires that each pair of comer points along the conjugated epipolar lines between two images has to he confirmed by a corner point lying at the intersection of the pair of conjugated epipolar lines in a preset number of the other images.

Hypotheses Validation

The hypotheses for corresponding corner points are verified via a neighbourhood constraint over the edges which connect the corner points. This constraint requires for each set of three connected comer points in one image that the corresponding three comer points in a preset number of the other images are also connected in the same fashion. Furthermore the difference between the 3D positions for a vertex computed from each pair of images may not exceed a preset small threshold.

Manual 3D Completion

In general the multi-view stereo algorithm terminates with a partial 3D object reconstruction. Because of the strong constraints the algorithm uses in the prediction and validation of the matching hypotheses some correct 3D edges are not accepted by the stereo matcher. But all rejected 3D edges are hold in a separate list and can be tested by the user of POLYOC. For this purpose POLYOC offers a procedure which shows projections of the object reconstruction from arbitrary viewing points. The rejected 3D edges are also subsequently projected into the reconstruction. So the user can easily decide whether a 3D edge is correct and can take it over into the 3D object reconstruction. Knowledge based Correction

The user of POLYOC can choose several knowledge based heuristics, for example: (1) a heuristic which sets up vertical all those edges whose direction differs not over a preset threshold from the orthogonal direction of the ground plane, (2) a heuristic which lays all those edges horizontal whose direction is nearly parallel to the ground plane. (3) heuristics which relay on the lengths, symmetries and parallelisms of the 3D edges are in work.

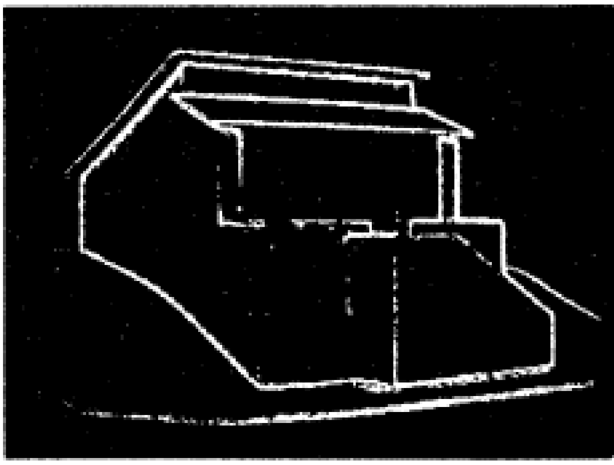
Converting the object reconstruction into DXF

The result of the multi-view stereovision algorithm are two lists again: the first list holds all reconstructed 3D corner points of the object, the second list contains all 3Dedgesof the object. The result of the knowledge based correction procedures are a second version of the original reconstruction: again two lists as mentioned above. These two list are converted into DXF format in such a way that two layers are created in the target CAD-system: one layer for the original and uncorrected reconstruction and a second layer for the corrected object reconstruction.

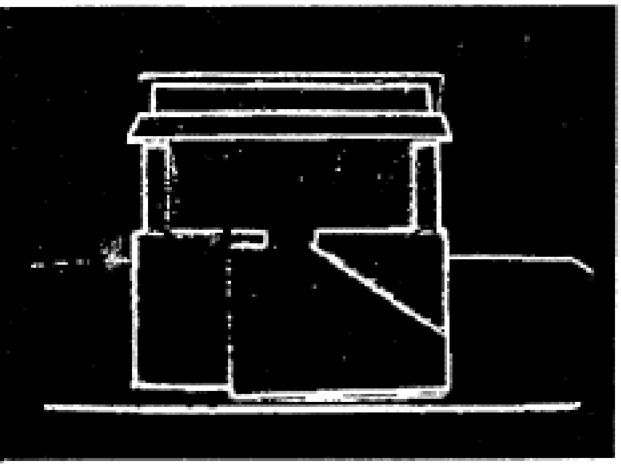
Experimental Results

The multi -view stereovision system POLYOC that we have sketched has been successfully applied to many scenes and has proved reliable and accurate results. Reliability and accuracy increase with each additional camera view. Using four or more images within the multi-view stereo algorithm an accuracy of the reconstructed 3D positions within 1 mm is reachable.

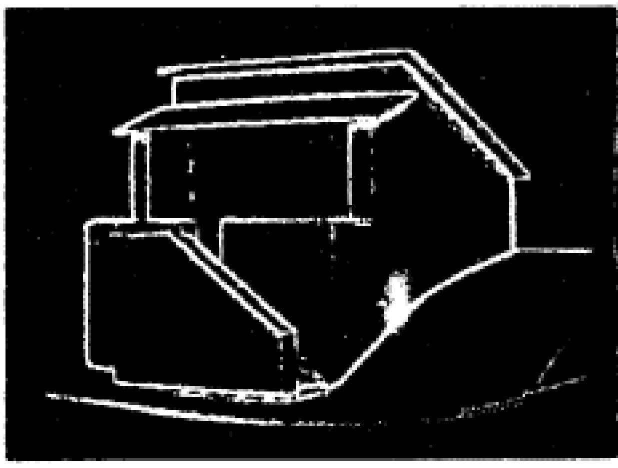
Figure 3 shows the object reconstruction process of an architectural model via POLYOC: (a) - (c) the Sobel-pictures of three video images, (d) the partial object reconstruction obtained by POLYOC from these three images, (c) the PC-interface of POLYOC for selecting the digitized images: beside the three images of (a) - (c) now three additional images are selected for the object reconstruction process, (f) the object reconstruction obtained from these six input images, (g) the nearly complete object reconstruction obtained using again six additional input images.



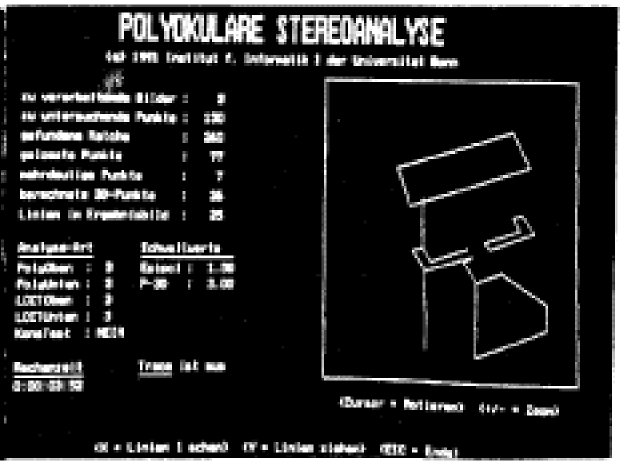
(a)



(b)

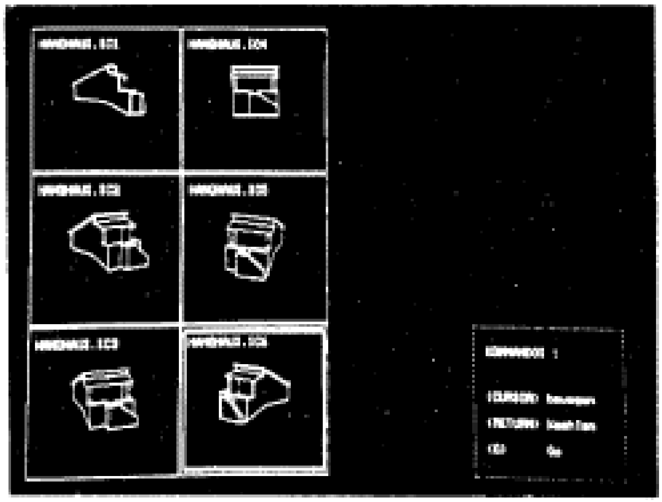


(c)



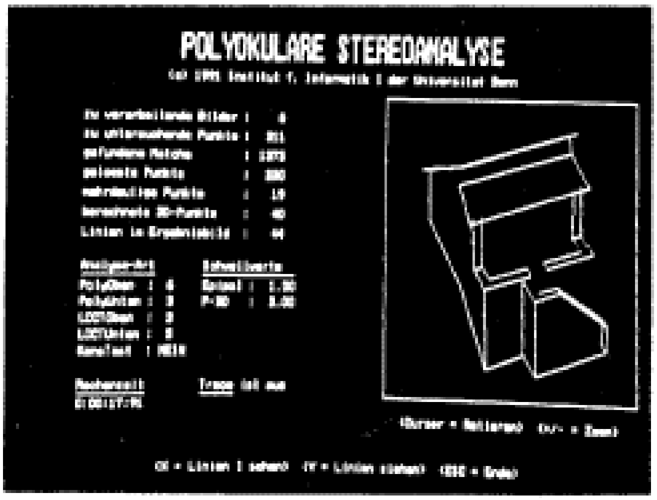
(d)

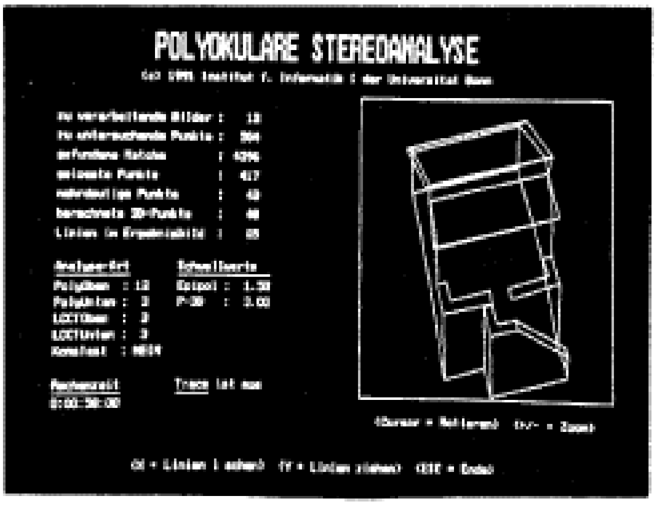
Figure 3 : (a) - (c) the Sobel-pictures of three video images, (d) the partial object reconstruction obtained by POLYOC from these three images.



(e)

(f)





(g)

Figure 3 : (e) the PC-interface of POLYOC for selecting the digitized images, (f) the object reconstruction obtained from the six selected images in (e), (g) the nearly complete object reconstruction obtained using again six additional input images.

References

[Ayache&Lustman87] N. Ayache, F. Lustman, "Trinocular Stereovision: Recent Results", in Proc. 10th IJCAI, 826-828, 1987.

[Ayache&Hansen88] N. Ayache, C. Hansen, "Rectification of Images for Binocular and Trinocular Stereovision", in IEEE 9th Intern. Conf. on Patt. Recog., 11- 16, 1988.

[Barnard&Thompson80] : S. T. Barnard, W. B. Thompson, "Disparity Analysis of Images", in IEEE Transact. Pattern Analysis and Machine Intelligence, Vol. PAMI-2, No. 4, 333-340, 1980.

[Binford81] : T.O. Binford, "Inferring Surfaces from Images", in Artificial Intelligence 17,205-244, North Holland, Amsterdam, 1981.

[GerhardEtAl86] : A. Gerhard, H. Platzer, J. Steuer, R. Lenz, "Depth Extraction by Stereo Triples and a Fast Correspondence Estimation Algorithm", in IEEE 8th Intern. Conf. on Pattern Recognition, 5 12-5 15, 1986.

[Grimson85]: W. Grimson, "Computational Experiments with a Feature Based Stereo Algorithm", in IEEE Trans. Pattern Anal. Machine Intell., Vol. PAMI-7, No. 1, 1985.

[GurewitzEtAI86] : E. Gurewitz, I. Dinstein, B. Sarusi, "More on the Benefit of a Third Eye for Machine Stereo Perception", in IEEE 8Lti Intern. Conf. on Pattem Recognition, 966-968, 1986.

[HungEtA189] : K,C. Hung et al., "Polyhedron Reconstruction Using Three-View Analysis", in Pattern Recognition, Vol. 22, No. 3, 231-246, 1989.

[Ito&lshii86] : M. Ito, A. Ishii, "Three-View Stereo Analysis", in IEEE Trans. Pattern Anal. Machine Intell., Vol. PAMI-8, No. 4, 524-532, 1986.

[Marr82] : D. Marr, "Vision", Freeman, San Francisco, 1982.

[Posch88] : S. Posch, "Hierarchische linienbasierte Tiefenbestimmung in einem Stereobild", in Informatik Fachbericht 181 der 12. GWAI, Springer Verlag, Berlin et al., 275-285, 1989.

[Ohta&Kanade85] : Y. Ohta, T. Kanade, "Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming", in IEEE Trans. Pattern Anal. Machine Intell., Vol. PAMI-7, No. 2, 1985.

[OthaEtAI86] : Y. Ohta, M. Watanabe, K. Ikeda, "Improving Depth Map by Right-Angled Trinocular Stereo", in IEEE 8th Intern. Conf. on Patt. Recogn., 519-521, 1986.

[Steinhage89] : V. Steinhage, "Objektrekonstruktion durch trinokulare Stereoanalyse", KI-Bericht Nr. 12, Universität Bonn, 1989.

[SteinhageEtA191] : V. Steinhage, I. Daniels, A. Schick, "A Multi-View Stereovision Approach for Polyhedron Reconstruction", in 4th Intern. Conf. on Comp. Analysis of Images and Patterns CAIP '91, Dresden 1991.

[Thorpe&Shafer83] : C. Thorpe, S. Shafer, "Correspondence in Line Drawings of Multiple Views of

Objects-, in Proc. 9th IJCAI, 959-965, 1981.

[Ullman79] : S. Ullman, "The Interpretation of Visual Motion", MIT Press, Cambridge, Mass., 1979.

[YachidaEtAI86] : M. Yachida, Y. Kitamura, M. Kimachi. "Trinocular Vision: New Approach for Correspondence Problem", in IEEE 8th Intern. Conf. on Pattern Recognition, 1041-1044, 1986.

Order a complete set of
eCAADe Proceedings (1983 - 2000)
on CD-Rom!

Further information:
http://www.ecaade.org