INTERPRETING SINGLE IMAGES OF POLYHEDRA

Carola Braun
Institute of Photogrammetry
University of Bonn, FRG

abstract

The paper presents a concept of a rule-based system for deriving the 3D structure of polyhedral objects in perspective projections using known geometrical constraints. The system may supply a complete reconstruction of objects, especially of buildings in aerial images, by matching the results of individual image interpretation in space.
As buildings can be simplified to polyhedra, the procedures first are developed for polyhedral objects. Many real objects have groups of parallel and orthogonal lines and rectangular corners. The presented system uses this information for reconstructing polyhedral objects with the aim of a geometrically correct shape description. The geometric reasoning process, implemented in the system, is able to automatically determine the structure of the object step by step, asking the operator to provide information about geometrical constraints if necessary. Examples illustrate the applicability of the approach.

keywords

rule-based system, inverse perspectivity, perceptual grouping, geometric reasoning, image analysis

1 Introduction

The concept of a rule-based system is presented for deriving the geometrical shape of polyhedral objects in perspective projection using known geometrical constraints. The system is intended to be used for a complete reconstruction of buildings in aerial images, matching the results of individual image interpretations in space.
The system is part of a long term project for the development of procedures for the recognition and reconstruction of buildings.
As man-made objects like buildings may reasonably well be described by their geometric structure and this structure can be approximated by a polyhedra (skyscrapers, terrace houses), the procedures of the system first are developed for polyhedral objects in single images. It is also possible to extend the interpretation process for buildings, consisting not only of planes but of cylindrical surfaces or cones (castles, churches, congress halls).
There are two possibilities for deriving the 3D information of an object. The classical approach first applies point or edge detection and possibly also grouping procedures to get a symbolic description of the images (Fig. 1.a). This description is either based on pixels, attributes and/or relations. Matching the image descriptions of several images and determining the object points in space by a spatial intersection allows a 3D reconstruction of the object. The other approach applies edge detection and grouping procedures in order to get sketches of the object (Fig. 1b). A reasoning process then infers from the 2D features of the sketches to single 3D models of the object which can easier be matched in space in order to get a 3D description of the object. The advantage of this aproach is that the problems involved in the step "matching in 2D" are avoided, however replaced by the additional interpretation step of sketches. Thus, in contrast to most existing approaches, the matching process will take place in object space after the 3D interpretation of the individual images.
Researchers have used various techniques for interpreting sketches or line drawings. In their Mosaic image understanding system HERMAN and KANADE (1984) implemented several strategies to extract 3D information about houses from single aerial images by exploiting the special structure of the houses (vertical lines, flat roofs) and the special orientation of aerial images. MULGAONKAR and SHAPIRO (1985) have presented a PROLOG-based reasoning system which allowed to interprete perspective line drawings. It contains a large number of rules of the inverse perspectivity and of grouping processes. The system is able to interpret even images basing only on extracted line segments by means of these rules, but seems to be inefficient for interpreting sketches of even rather simple objects. HARALICK (1989) has collected a number of

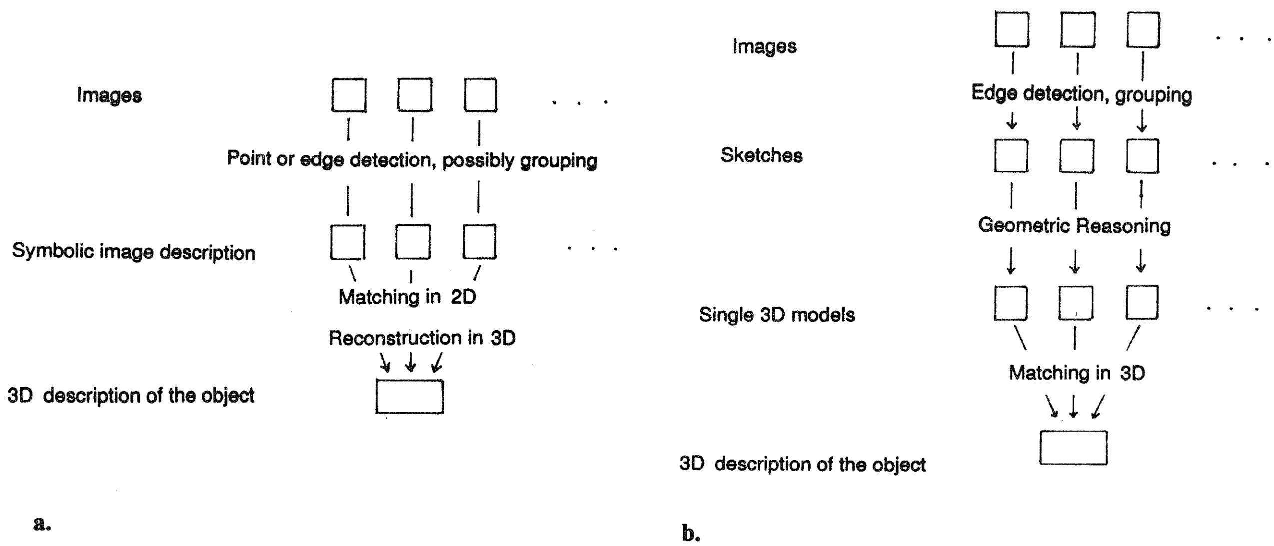


Fig. 1 Deriving 3D information of an object

a. Classical approach

b. The matching in 2D is replaced by two steps

rules of the inverse perspectivity for deriving 3D attributes from image data and hypothesis about the object's shape. SUGIHARA (1986) probably has presented the most consistent theoretical framework for the geometrical interpretation of line drawings. The reasoning is based on line drawings which are labeled according to HUFFMANN (1971) and CLOWES (1971), deriving a set of linear constraints for the parameters of the object's faces with the exception of a few form parameters (>=4) which have to be derived by other means. The approach is not able to handle incomplete sketches or even non linked edge images, thus assumes the labeling of edges to be complete.

The approach presented in the following aims at the geometrical interpretation of perspective and orthogonal images. The sketch, derived from an image, may be incomplete and inaccurate, without any apriori information about the object's orientation or about the exterior orientation of the images as well. The approach tries to link the features of the procedures mentioned above. The process of interpreting single images is devided into several steps. As the kind, number and order of these steps of the process are dependent on the input data, on the kind of object and on the image, there are several ways to solve the task. Therefore the steps are formulated as rules, organized by the control modul of a rule-based system, finding one possible short way to the solution.

Chapter 2 gives an overview of the approach, chapter 3 contains the geometric reasoning process. The rule-based system is described in chapter 4, illustrated by some examples.

2 Overview of the approach

The vision system for Interpreting Single Images of Polyhedra (ISIP) uses four levels for the representation of data (Fig. 2). The first level is the original image, a single black and white or color one. Methods of noise cleaning or image restauration being edge preserving may be applied for preprocessing.

As primitives, like extracted straight line segments and detected regions with similar intensity (blobs), the data

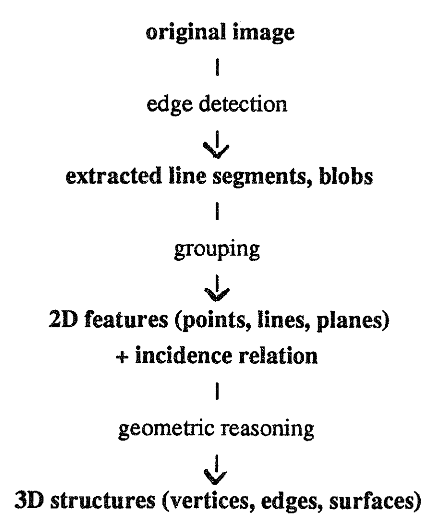
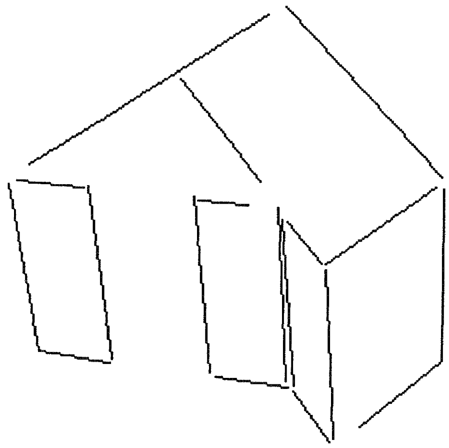
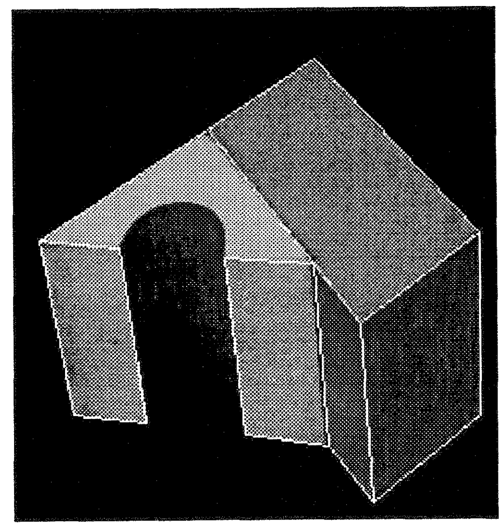


Fig. 2 Four levels for the representation of data



a.



b.

Fig. 3 Image of a toy block

a. Extracted line segments

b. Sketch

structure of level 2, are not sufficient to be connected with an object model, it is necessary to get a relational description between these primitives.

Methods for finding relationships or groups of primitives with common properties are perceptual grouping procedures (MOHAN/NEVATIA 1987, STRAFORINI et al. 1990). These groups supply hypothesis about the possible appearance of an object, their common properties might be parallelity, collinearity, connectivity, symmetry or texture.

The data structure of level 3 are 2D features, consisting of points, lines and planes, including a list of incidence relation (line is in plane), which is introduced as a hypothesis (VOSSELMANN 1991). In this case the result of the grouping is represented by groups of lines, each group belonging to a special plane. Sketches of the objects, being represented by the data of level 3, can be derived automatically or semiautomatically by the system in dependence of the input data. Approximately digitizing the toy block of Fig. 3 for example supplies initial values for a model. A best estimator then fits the model in the extracted line segments in order to get an input for the interpretation of sketches of polyhedra (SCHICKLER 1992). The geometric reasoning, being the internal part of the rule- based system, is then able to automatically determine the 3D shape of the object step by step, asking the operator to provide information about further geometrical constraints if necessary.

The 4th level finally consists of 3D structures, i.e. vertices, edges and surfaces, derived from the 2D features of the 3rd level. The result of the interpretation process is the reconstructed object, represented by a 3D geometric model composed by the 3D structures and the incidence relation.

3 Geometric Reasoning

3.1 Motivation and Task

The human visual system is able to interpret the geometry of an objet from one perspective line drawing without any additional information about the object (Fig. 4). For a computer line drawings are only simple collections of lines in a plane, therefore special algorithms are needed for deriving the scene structure. It is shown how some aspects of the human interpretation process can be transferred to a computer by applying some of the assumptions the human visual system obviously has implicitly made about the object, e.g. assumptions about the existence of parallel and perpendicular lines of the house in Fig. 4.

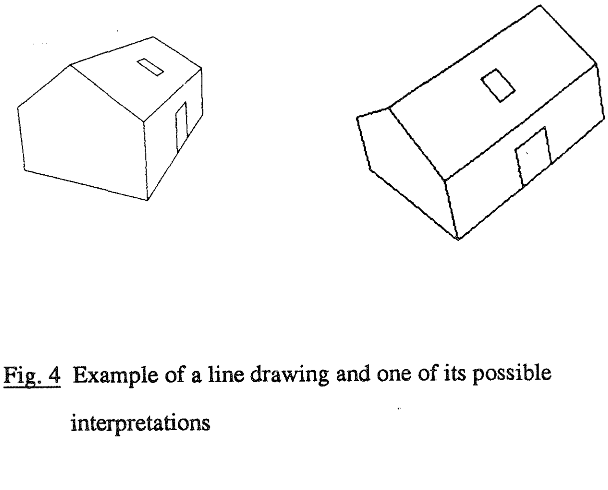


Fig. 4 Example of a line drawing and one of its possible interpretations

Inferring from the 2D features in the image to the 3D structures of the object is the task of the geometric reasoning process which is solved by building up hypothesis about relations and geometric properties of the object in order to compensate the information lost in central projection. The assumptions mentioned above belong to these relationships.

Especially the regular structure of polyhedral objects supplies several relationships, e.g. parallel and perpendicular lines as well as horizontal and vertical lines or a trihedral corner.

The geometric reasoning process, implemented in the system, is devided into two parts:

First relationships between 3D entities are determined by the system either automatically by grouping procedures or by an operator. Rules of the perspective geometry supply the mutual orientation of camera and object. In the second step the inverse perspective problem is solved by using the relationships as geometrical constraints. The shape of the object is determined by the system, calculating the planes of the object in space step by step.

3.2 Relationships between 3D entities

The following relationships between 3D entities are used in the reasoning process:

<u>given hypothesis</u>

- Faces of the object are assumed to be planes (polyhedral object).
- Incidence relation: line is in plane, is_in(line, plane)

<u>determined by grouping</u>

- Parallel relation
- Perpendicular relation
- Collinear relation

<u>provided by an operator</u>

- Trihedral corner
- Features with known length -> scale factor
- Features with known angle

The first two relations are introduced as hypothesis to the system before the reasoning process starts. The next three relations can automatically be determined by grouping procedures. Observe, that all the grouping procedures are not based on any knowledge about the 3D shape of the scene, any how they are able to supply the basis to build up hypothesis. In case of not finding enough relationhips, the system asks the operator to provide information about geometrical constraints.

3.2 Rules of the perspective geometry

The relationships between 3D entities, introduced as hypothesis about the shape of the object, are components of several rules of the perspective geometry used in the system. If a hypothesis is built up, the rule calls routines for calculating real values for attributes of spatial objects or relations, e.g. the direction vector of parallel lines.

In the following some rules of the perspective geometry are presented:

<u>1. Vanishing points</u>

If the grouping procedure has found at least three lines in the image assuming them to be parallel in space, the vanishing point and the corresponding direction vector is calculated (Fig. 5).

Procedures for locating vanishing points are described in (BARNARD 1984, MAGEE/AGGARWAL 1984 and BRILLAULT O'MAHONY 1991).

<u>2. Two parallel lines</u>

If two lines can be assumed to be parallel, the direction vector of the parallel lines can directly be calculated (NEVATIA/ULUPINAR 1991).

<u>3. Rectangular corner</u>

A rectangular corner consists of three lines, being perpendicular to each other and meeting in one point. As polyhedra or buildings often have a rectangular corner, this information can be applied as a starting rule, in case the corner is assigned by an operator. A known rectangularity allows to determine the mutual orientation of object and camera as the rotation matrix just represents the direction vectors of these three lines in space. There exist two solutions for the problem because the rectangular corner can be considered to be in front of the lines or back as well (KANATANI 1990, PAN 1990).

<u>4. Two parallel lines intersecting one perpendicular line</u>

If these three lines are located by an operator or a grouping procedure, the direction vector of the normal of the plane can be determined (Haralick, 1989).

<u>5. Three known vanishing points</u>

If three vanishing points are located in the image and their direction vectors enclose three angles of 90 degrees, the rotation matrix, the focal distance and the position of the principle point can be estimated.

<u>6. Two known vanishing points</u>

If two vanishing points are located in the image and their direction vectors enclose an angle of 90 degrees, the rotation matrix and the focal distance can be estimated.

<u>7. Plane of the object</u>

If two direction vectors of lines, not being parallel, and one point in space are known, the parameters of this plane of the object can be determined in space.

<u>8. Point in a known plane</u>

If a point, given in the image, belongs to a plane known in space, the 3D coordinates of this point can uniquely be calculated by intersecting the image ray of the point with the plane in space.

<u>9. Known scale factor</u>

In case of a distance between two object points is given, it is possible to transform the 3D model to an object of the real world.

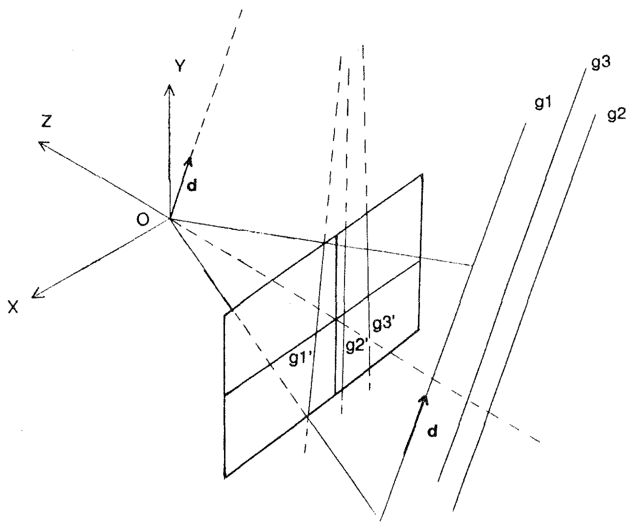


Fig. 5 Representation of the direction vector of a vanishing point and its corresponding lines in the image and in space

There are numerous other possibilities to reason about 3D structures from image data and adequate hypothesis about the object's form which may also be used (HARALICK 1989).
The rules 1-6 are used to solve the first task of the geometric reasoning to determine the direction vectors of lines in space, while the rules 7-9 are used to derive the object's shape.

4 Rule-based System

4.1 Concept

The task of the vision system for interpreting images of polyhedra (ISIP) is to derive the object's shape in space.
A large variation of input data is admissable. There is no limitation in the complexitiy of the object as long as the planes of the object are connected and the object fullfills the assumption being approximable as a polyhedra. The system is able to evaluate perspective images as well as orthogonal ones. The exterior orientation of the image may vary, also the used camera, the input images may be terrestrical or aerial ones or images with short or long distances to the object. Because of this variety of admissable input data and the resulting big number, kind and order of procedures to be applied, there exist several strategies to solve the task. Therefore the vision system ISIP is organized in components of a rule-based system, finding one short way to the solution (NIEMANN 1981).
The information, involved in the 2D sketch, and the information provided optionally by an operator is collected as the intermediate data of the system (Fig. 6).

The different rules within the interpretation system consist of a condition and a routine. The first conponent of the rule may either be a hypothesis about relationships or an assumption, resulting from the input data or the strategy of the interpretation process. The rules are based on apriori knowledge about the problem and their form of representation is known, so the operator could define knowledge in additional rules. The rules in the knowledge base are listed without a defined order, only the two components of each rule are connected together. The rules in the knowledge base of the system are devided into four classes. There are rules of the perspective geometry (cf. chap. 3.3) and rules, resulting from the strategy of the interpretation process or the used camera or the exterior orientation of the image. It will be shown how rules of the other three classes are applied to derive the object's shape (cf. chap. 4.4).
The sequence of steps in which the relationships between 3D entities and the planes of the object are calculated, is determined by the control modul using the given 2D sketch and applying procedures of algebra and rules of the knowledge base. This central component of the system contains a special strategy and structure of order, handling the given knowledge, facts and rules, bottom up. The results of the system are declarations concerning the interpretation process and a 3D model of the object.

4.2 Input of the system

The data of the input images is stored in lists, according to the data structures of level 3 (cf. chap. 2).
P : { {pnr x' y' } ... }
L : { {lnr b e } ... }
P is a list, consisting of n points with point number and 2-D coordinates (x',y'). L is a list consisting of m lines. Each of these lines consists of a line number, the beginning and ending point of that line.
E : { {enr l1 ... ln} ... }
E contains the incidence relations used as hypothesis by the system.

4.3 Control modul

The system sucessively calculates the parameters of the planes of the object, stopping the process if all planes of the list E are determined in space respectively the points and the lines. The geometric reasoning depends on the geometrical constraints, which are either found automatically or given manually, on the actual status of knowledge, which is obtained so far, and on the parameters to be searched for. It is important to keep manual inputs by an operator as small as possible though in some cases inputs given by an operator may be necessary for the solution.

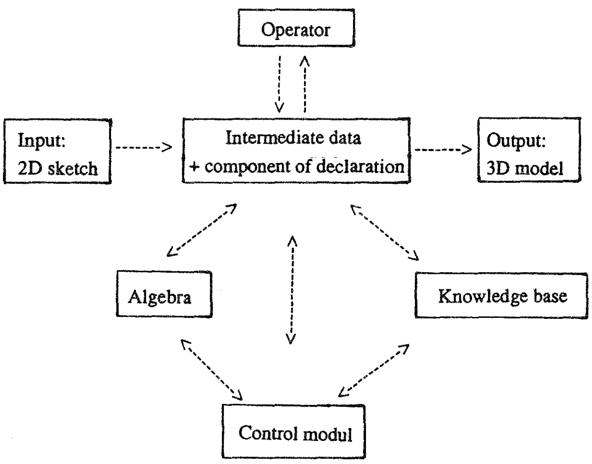


Fig. 6 Components of the rule-based system

As there is no apriori knowledge about the object at the beginning except the incidence relations or optionally about the camera or exterior orientation, the reasoning can only be started by building up hypothesis about geometrical relationships. In case of redundant information the results can be improved by calculating the values for the attributes in a simultanous evaluation process. If an operator is asked to indicate geometrical constraints, the system might have more information available as necessary.
The following strategy is implemented as one part of the control modul, solving the 2nd task of the reasoning process (cf. chap. 3.1):

```
Initially E contains all unknown planes.
repeat
   repeat ( *internal reasoning* )
      for all planes in list E do
         if plane determinable then perform
            • plane determination
            • point_in_plane determination
            • line_in_plane determination
            • delete plane from E
         endif
      endfor
   until no plane in E is determinable any more or E is empty
   if E is not empty then (* operator request *)
      ask the operator for a spatial relation
   endif
until all planes of the object are determined in space
```

There is freedom in the sequence of planes in line 4 of the algorithm which may be used to optimize the interpretation with respect to stability. The operator request may be assisted by the system, providing a list of candidate hypothesis from which the operator has to choose one. This may again be used for optimization. Finally all not yet used hypothesis may be combined to determine best estimates for the spatial coordinates of the object.

4.4 Examples

The reasoning for a collection of houses is discussed using the example in Fig. 7. The bold lines are classified as lines whose direction cosines in space can automatically be derived by locating vanishing points (Fig. 7a). These are the three vanishing points pointing downwards, backwards and to the left. The system now calculates the parameters of all those planes which fullfill the criteria for determination. These are all planes with at least two bold lines not being parallel (cf. rule 7 in chap. 3.2). As during the first pass the direction cosines of the undetermined lines in space (thin lines) could not be calculated, the parameters of the remaining planes were determined by the system in a second pass, still within the internal reasoning (cf. control modul, line 3). Therefore in this example it is not necessary for an operator to indicate any geometrical relations. The result of the collection of buildings, generated by an alternative view, is shown in Fig. 7b.

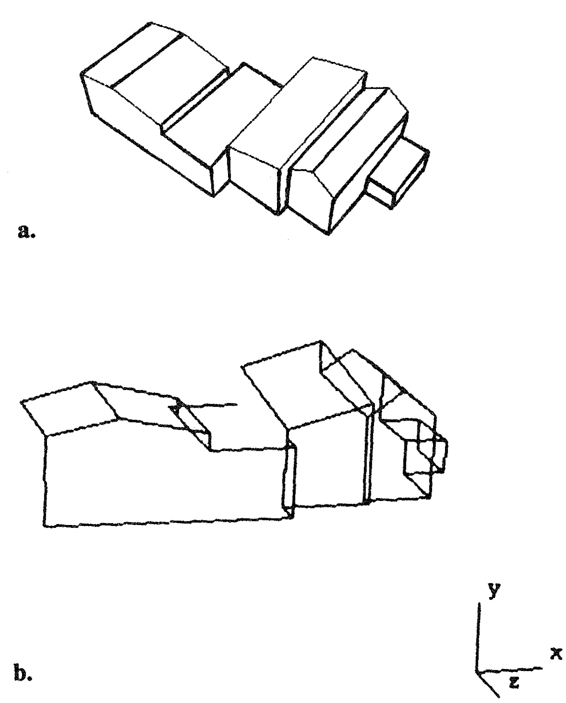
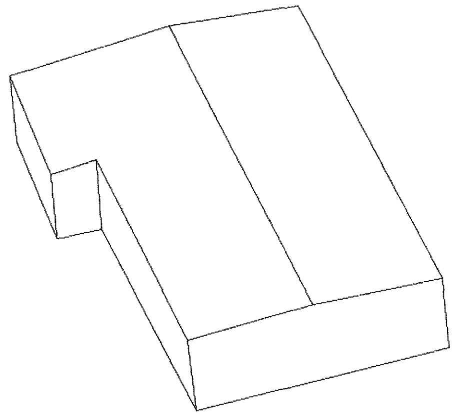


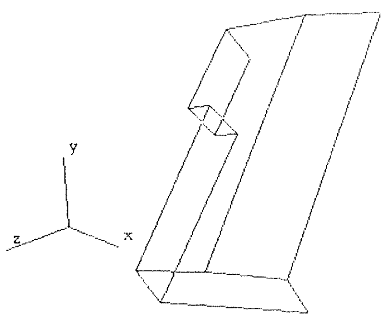
Fig. 7 A collection of houses

a. Perspective line drawing

b. Representation of the reconstructed collection of houses



a.



b.

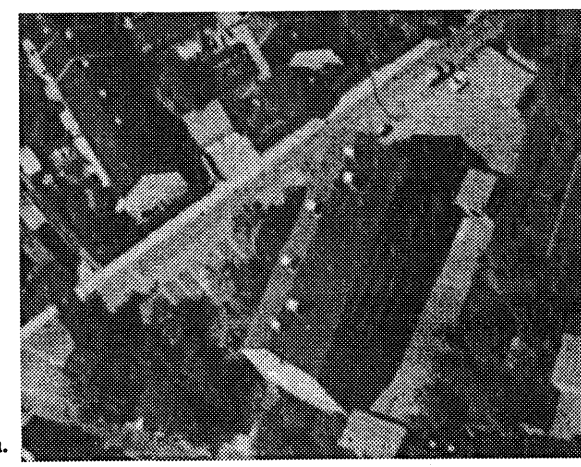
Fig. 8 Building

a. Orthogonal Sketch

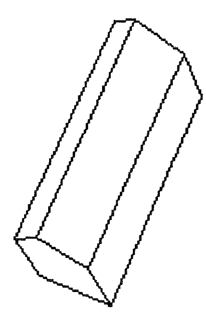
b. Representation of the building in space

In case an orthogonal sketch is given (Fig. 8a), the operator is asked to provide information about a trihedral corner. This starting rule is sufficient for the reconstruction of this building in space (Fig.8b). The problem of the trihedral corner can only be solved if the two neighbour legs of one leg of the trihedral corner are situated in two different quadrants (SUGIHARA 1986). Especially sketches, being manually drawn, and parts of aerial images nearby the image border can be regarded as an orthogonal projection, not fullfilling this criteria. The system then is able to automatically correct the position of the principle point in the image in order to determine the direction vectors of the three lines in space.
Fig. 9b represents an orthogonal sketch of a single house being derived from a part of an aerial image. As the exterior and interior orientation of the input image in Fig. 9a is given, the input data can be corrected with the coordinates of the principle point, before the reasoning starts. Again only by indicating a trihedral corner at the object, the operator gets the reconstructed house in space (Fig. 9c). Here the supply of information is even reduced to the indication of a trihedral corner. In case of the position of the nadir point, the vanishing point of the vertical lines, is known in the image, the system is able to use this information for choosing the correct solution.
Fig. 10 finally represents the result for interpreting the line drawing of the toy block in Fig. 3b, showing that the input data for the reasoning process need not to be a complete line drawing.

a.



b.



c.

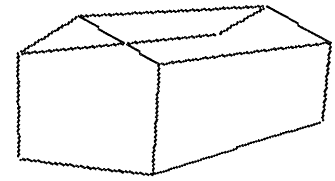


Fig. 9 House

a. Aerial image

b. Sketch derived from a part of an aerial image

c. Result of the reasoning process

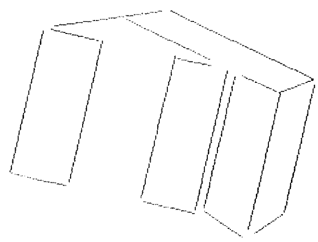


Fig. 10 Representation of the reconstructed toy block of Fig.3

5 Discussion

A concept of a vision system with the aim to interpret images of polyhedral objects has been presented. It could be shown that an interpretation of single images is feasible.

For solving the interpretation problem, assumptions are made about planar faces of the object and the incidence relation between the 2D features and hypothesis about spatial relationships of the object, being derived either by grouping procedures or by the operator.

The result of the vision system is the 3D reconstruction of a polyhedral object from one single image. The input image may contain real and artificial objects and graphical sketches as well.

An automatic derivation of buildings or groups of buildings may represent an important contribution for the aquisition and actualization of information-systems. Applications are city-planning in simulated parts of a town and the analysis of the climate in cities (BECKRÖGE/JUNIUS 1991).

An interpretation of single images of buildings in real-time may be implemented in systems for automated navigation of cars or planes. The results of the vision system presented here are also suited as models for the feature based object location in aerial images (SCHICKLER 1992).

For further applications the system can be transferred to all kinds of objects being approximately a polyhedra, e.g. when grasping of machine parts in robotik.

Especially the interpretation of scanned line drawings (inverse computer graphics) can be regarded as a part of a man-machine interface. In all cases the result of the interpretation of single images may used as input for procedures of 3D-CAD graphics.

Future work will concern the extension to non planar surfaces and the fusion of single 3D models.

6 Literature

BARNARD, St. (1982): Interpreting Perspective Images. Technical Note 271, SRI Project 1009, Menlo Park, 1982

BECKRÖGE, W. and JUNIUS, H. (1991): Klimatologische Begutachtung bestimmter stadtplanerischer Untersuchungsbereiche im Dortmunder Südwesten unter Einsatz eines Geo-Informationssystems. VR 53/3+4, Juni 1991

BRILLAUT-O'MAHONY, B. (1991): New Method for Vanishing Point Detection. CVGIP:IU, Vol. 54, No. 2, 1991, pp. 289-300

CLOWES, M. B. (1971): On Seeing Things. Artificial Intelligence 2, 79-116, 1971

HARALICK, R. M. (1989): Monocular Vision Using Inverse Perspective Projection Geometry. Analytic Relations, Conf. on Computer Vision and Pattern Recognition, San Diego, 1989, pp. 370ff

HERMAN, M. and KANADE, T. (1986): The 3-D Mosaic Scene Understanding System. From Pixels to Predicates, Ablex Publishing Corporation, New Jersey, 1986

HUFFMANN, D. A. (1971): Impossible Objects as Nonsense Sentences. Machine Intelligence, Vol.6, ed. by B. Meltzer, D. Michie (Edinburgh University Press, Edinburgh, U.K. 1971), pp.295-323

KANATANI, K. (1990): Group-Theoretical Methods in Image Understanding, Springer, Berlin 1990

MAGEE, M. and AGGARVAL, J. (1984): Determining Vanishing Points from Perspective Images. CVGIP, Vol. 26, 1984, pp. 256-267

MOHAN, R. and NEVATIA, R. (1987): Perceptual Grouping for the Detection and Description of Structures in Aerial Images. IRIS, Univ. of South Calif., Rep. 225, Los Angeles, 1987.

MULGAONKAR, P. G. and SHAPIRO, L. G. (1985): Hypothesis-Based Geometric Reasoning about Perspective Images. Proc. 3rd Workshop on Computer Vision, Bellaire, Michigan, 1985, pp.11-18.

NEVATIA, R. and ULUPINAR, F. (1991): Constraints for Interpretation of Line Drawings under Perspective Projection. CVGIP:IU, Vol.53, No.1, 1991, pp.88-96

NIEMANN, H. (1981): Pattern Analysis, Springer, Berlin 1981

PAN, H. P. (1990): A Spatial Structure Theory in Machine Vision and its Application to Structural and Textural Analysis of Remotely Sensed Images, PhD Thesis, ITC, 1990

SCHICKLER, W. (1992): Feature Matching for Outer Orientation of Single Images using 3D Wireframe Controlpoints. ISPRS Congress, COM. III, Washington D.C., 1992

SUGIHARA, K. (1986): Machine Interpretation of Line Drawings. The MIT Press, Cambridge, MA, 1986.

STRAFORINI, M., C. COELHO, M. CAMPANI and V. TORRE (1990): On the Understanding of Indoor Scenes. Intern. Workshop on Robust Computer Vision, Seattle, WA, 1990

VOSSELMANN, G. (1991): Relational Matching. Promotion Universität Bonn, 1991