# Modelling strategies in photogrammetry

MONIKA SESTER, Stuttgart

**ABSTRACT**

In order to extract object information from digital images, models of these objects have to be available. The paper investigates the modeling problem in photogrammetry, where different types of models have to be provided for varying tasks. To solve the question of the model aquisition, a procedure is sketched, which extracts hierarchical model descriptions from given examples.

## 1. INTRODUCTION AND OVERVIEW

The primary task of photogrammetry is to extract information from images. This is conventionally done using photographs and measuring the interesting data in it. Going from analog to digital images, the task remains the same. The main difference being, that the measuring process is now transferred to a computer. Relying on the computational power of machines, in the early days of ComputerVision this automation seemed to be close at hand. What wasn't clear however is the importance of knowledge, namely knowledge humans intuitively apply confronted with a problem. In order to measure, the object of interest has to be known and identified in the image. The identification is a cognitive process, which humans do mostly unconsciously - which is perhaps the reason why it has been seriously underestimated. Identification consists of mainly two steps, namely a model of the object and a strategy specifying how to apply the models in order accomplish the task. The paper concentrates on the modeling aspect by presenting and describing different model-types which are applied in digital photogrammetry - depending on various tasks in this domain. The model building process, especially in natural domains, is a non-trivial task. Therefore a semi-automatic modeling strategy is demonstrated, which tries to use methods similar to those humans apply when acquiring new knowledge: Models, i.e. internal descriptions of objects are derived from examples using learning techniques.

## 2. TASKS IN PHOTOGRAMMETRY

The conventional task of photogrammetry is to measure objects in images in order to be able to use this information primarily for mapping processes. Thus the objects involved are the earths surface with its interesting objects, or industrial and architectural objects in close range applications, respectively. In the following the main tasks belonging to the photogrammetric domain are summarized:

- Orientation tasks: in order to be able to measure 3-D-coordinates from two-dimensional images, two overlapping images have to be oriented relatively to each other and absolutely to the world coordinate system.
- Aerial triangulation in order to determine point coordinates in a whole block of images: the measuring problem is confined to ground control points and to points that merely serve as tie-points in order to link adjacent images.
- Derivation of digital elevation models from stereo images: points that represent the surrounding landscape are measured in a stereo-image-pair.
- Extraction of 'semantic information' from the images, namely distinct physical objects which have to be projected into maps or geographic-information systems (GIS). To this end, all the

data that might be relevant for a GIS has to be measured. This involves both the exact geometry and features and attributes of the objects.Another application is the notification of changes in the environment. This involves a comparison of the GIS data and the actual data in the image and results in an update of the database.

## 3. THE CENTRAL PROBLEM: IDENTIFICATION

The measurement process is divided into an identification step and the measurement itself. Humans mostly have no problem to identify an object, they do it automatically and mostly unconsciously. In order to automate this process, a system has to be provided which knows what to measure and how to find it. An object for which a description is provided, has to be found in an image using a certain strategy - thus a matching problem has to be solved.

Usually some object features - called model of the object - have to be identified in the image. In order to accomplish this, the same object representation has to be provided both in the image and the model. All features found in the image potentially belong to the searched object. This can be a great amount, depending on the given features. Using the grey-value as a feature will give a lot of possible candidates, whereas a more specific description will restrict the number of matches. In any case a strategy of how to find and evaluate the matches has to be provided.

Thus identification needs a model and a recognition strategy in order to find a consistent transformation of the model into the image. These two components depend on each other in a complementary way: using simple methods on one side, presumes elaborate techniques on the other. In this way, a simple model has to rely on a powerful matching strategy, which incorporates a huge amount of knowledge. If, e.g., points are used as features, there are many potential matching candidates in the image. A search procedure, which normally has exponential complexity, has to be applied to find the correct transformation. In order to reduce the ambiguity, additional knowledge about image and object has to be introduced into this search process: approximate position, overlapping area, orientation of the images, type of object, etc. This information keeps the range of possible candidates in a manageable size. On the other hand, using sophisticated models, i.e. a better object description, leads to a small number of possible matches. In this case, the matching strategy degrades to a comparison of a few representations, which can be accomplished very efficiently. Here the computational burden lies in the derivation of the model features in the image. Generally speaking, increasing the amount knowledge on one side, eases the processes involved on the other side. This paper focuses on the modeling aspect of the interpretation task.

## 4. MODELS

Modeling is a topic which can be approached in different ways. Basically, any interior representation or idea we have of objects and processes of the world around us, can be called a 'model'. These can be distinguished into object models and process models. The first category defines e.g. our knowledge about houses, books or friends, while the second stands for ideas like 'how to come from one point in the city to another ?' or 'how to identify a partly occluded object ?' (e.g. by looking around the occluding object or by removing it). In the following, however, there is a concentration on object models - process or functional models will not be considered. Specifying object models, the following distinctions are possible. This enumeration reflects a successive specialization going from one item to the next:

**very general models** serve for mere perception purposes: these models are general criteria like gestalt-phenomena [Metzger 1950] to distinguish what is figure and ground in an image. These models are basis for any object identification. These criteria include concepts of stability, smoothness, or good continuation. E.g. the notion, that 'an object is, what moves together' [Uexküll cited after [Lorenz 1977]]belongs to this type of models. This criterion is used in image sequence analysis to determine, which image parts have moved.

**general/generic models** describe not a single, but a class of similar objects and thus help to globally categorize objects.

**specific models** exactly describe a single object: using this model allows to identify a unique object.

Each of the enumerated model types has to define the object basing on itsproperties and attributes. These descriptions basically vary in the following ways:

**attribute based description** The description consists of a collection of attributes of the object. E.g. in multispectral classification, a pixel is classified to a specific landuse class basing on its color, thus 3 numeric values; the determination of an interest point depends on derivatives of the gradient. Attributes can be symbolic as well: a simple house in a 2D-blocks world can be described as a set of 2 regions, namely a triangle and a rectangle.

**relational description** In a relational description not only attributes are specified, but also dependencies between the object parts. In the case of the toy house, a specification would be the relation that the triangle is on top of the rectangle.

**hierarchical description** Hierarchical models describe an object in an structured way in form of parts and relations: each level of hierarchy defines a specific level of detail of the description, where the steps go from specific to more general. A hierarchy thus represents a fine-to-coarse structure of an object: at the basic level, the object is described in terms of elementary features (e.g. points and lines). These features can be grouped to form new objects in the next level, which themselves define object parts of the 'final' object. The hierarchical description of the house is therefore the collection of points and lines (with their annotated attributes and relations) on the first level, the two areas on the following level (again with their properties), and the finally the object 'house' in the last level.

All the model representations named rely on object attributes and relations between them. There are different types of attributes, which can be used to describe an object. Depending on which features are used, the modeling can be performed along several lines:

**Geometric models** The traditional approach in Computer Vision is to describe objects in terms of geometric properties, namely points, lines, regions and their relations. This is adequate as long as objects with rigid shapes and known form are to be modeled. In industrial applications, e.g., objects on the conveyor belt, which have to be identified or grasped, can be described very effectively basing on geometric features.
There are other approaches to the modeling which do not solely rely on geometric features of the objects, but furthermore include physical or functional aspects. This is important for situations, where the shapes to be identified are not rigid, or the objects cannot be described basing on their shape at all.

**Contextual models** [Strat and Fischler, 1991] have developed a system, that relies on contextual models. The basic idea is to include the whole scene context into the model of an individual object. The procedure is applied to outdoor scenes, where the task is to distinguish e.g. sky, ground and trees.In order to recognize these kinds of objects, e.g. 'sky' can be defined as being of color blue, of certain texture, in the neighborhood of trees and above the ground.

**Functional models** Another method to describe objects is use its functional behavior. [Stark and Bowyer,1991] give an example of a description of a chair as something to sit on. The primary request is that the object has to have a flat surface to somehow stand stable on the ground. All objects which share these properties can be used as chairs.

**Physical models** A rather new approach is to include knowledge about physical properties into the model. [Birnbaum et al. 1993] presented a program which uses knowledge about causality, gravity and balance for segmentation of a blocks world image: an image of a stack of blocks is segmented basing on the question, why this block is stable. This involves e.g. the consideration that the upmost block has to be supported by another object in order to remain stable.

Another aspect concerns the representation of these models in a computer. Popular description schemes for high level knowledge are semantic nets, frames, if-then rules.

## 5. SPECIALITIES IN DIGITAL PHOTOGRAMMETRY

### 5.1 Advantages ....

In photogrammetry there is a huge amount of knowledge about the observation geometry and the flight parameters available. Due to standardisations in aerial photogrammetry, there are certain schemes, along which photographs are taken. These schemes provide (approximate) knowledge about:

- the orientation of an image,
- the overlap of images: image are taken in a strip-wise manner with a given overlap inside the stripes and along the stripes,
- the aspects of objects in the image (objects are seen from above).
- Furthermore, there are additional sensors available, which provide new information: GPS-sensors measure the position of the camera projection center with an accuracy of less than a decimetre and the orientation in the range of 0.1 gon [Schade 1993]; Laser-ranging-techniques give spotwise height information; multispectral scanners extract also non-geometric object features.

### 5.2 State of the Art in Automation in Digital Photogrammetry

Using the advantages stated above, there has been great success in certain domains of photogrammetry, especially in surface reconstruction.

1. Least squares area based matching uses grey-values as elementary objects. This method requires good approximate values for the transformation parameters between the two images.
2. Using interest points as model features has successfully been applied to different tasks, all of them dealing with the matching of two images. The principal assumption of this approach is, that interest points bear more information then a mere gray-value. The meaning of a interest point is a mathematical definition basing on gray-value-gradients. Which physical process however produced the gradients is not defined: ideally, it is invoked by a distinct point in the real world, however texture or shadows also may lead to interest points.
   The traditional way of deriving digital elevation models form stereo images is to measure homologue points in both images. A human operator is choosing only a few points, namely those which represent a whole area. In addition he chooses points with a physical meaning, namely breaklines. Furthermore he uses 'common sense' knowledge in the way that he does not measure terrain surface on trees or houses. Here it is obvious, that the cognitive abilities of

humans are important in order to distinguish ground points from other objects. Nevertheless there are successful implementations of automatic DEM derivation: the idea is basing - like in the human ideal - also on a point based matching. The cognitive ability of the human being, who carefully chooses typical terrain points, is replaced by a procedure which measures a lot of points. The terrain surface is then computed with the help robust algorithms. These procedures guarantee that trees and objects (as long as they are small enough) and false correspondences are treated as outliers and eliminated, with the result of an averaged terrain surface [Match 1991].

The same approach has been applied to the automatic tie-point-transfer in aerial triangulation [Tsingas 1991], in navigation problems [Hahn 1992], and for relative orientation of two images [Haala and Hahn,1993]. All these tasks have in common, that the meaning of the individual points used is of minor importance, only the fact that a point can be found in the next image is counting. The computational load lies in the matching procedure which has to eliminate false correspondences based on heuristics and additional knowledge. The knowledge this approach uses is the fact that in overlapping (stereo) images there are corresponding points that can be found (in a certain restricted search area, given in the knowledge about the overlap of the images) and matched with the help of robust algorithms.

Another method to reduce the number of false correspondences and furthermore to diminish the amount (or accuracy) of prior information is to use higher order features. This can be accomplished by combining points in a certain neighborhood (spatial and scale) to form new structures [Zimmermann 1992]. These structures bear more information than the single points, since the relations about neighborhood are included. Using this technique can reduce the amount of search in the matching procedure, since there are fewer correspondences. Still however, the features need not have a physical meaning, which makes them hard to understand for humans.

These point based technique are working, as long as the individual object (the point) is of no interest, it only contributes to a common aim in connection with other measurements. If however individual measurements contradict the general trends, there are problems. In the case of the DEM-generation, a huge house or tree, with many measured points will lead to a 'terrain' surface, that lies on the house or tree respectively. In this case, additional knowledge about the semantics of the measured object has to be included.

3. [Schickler 1992] presents a system where model based matching is used to detect houses in aerial images. In order to accomplish this task, detailed knowledge about the position and orientation of the image and of the 3D-model (houses) is used, so the search for the house can be concentrated on a distinct part of the image.

   [Stilla and Jurkiewicz, 1991] use a generic description of houses, streets and settlements in order to find these structures. The individual descriptions are quite simple (a roof is a set of two parallelograms). The search is accomplished with the help of a powerful blackboard system. Since however the model is quite general, the geometry of the found structures, especially the houses, is detected only vaguely.

### 5.3 Difficulties ....

The problem in photogrammetry on the other hand is, that it deals with objects in the images that can not be standardized. Furthermore they belong to the class of natural objects for which descriptions are hardly available - in contrast to industrial applications, which mostly have to do with objects of fixed look and geometry. For objects like trees, houses or streets it is impossible to provide specific, individual descriptions due to their countless number. Also additional factors like illumination cannot be controlled in advance.

### 5.4 .... and Issues

In order to encompass the combinatorial explosion of possible matches when no or only poor prior knowledge is given, there is a new tendency in Computer Vision, called active vision: the scope of this approach is to reduce the ambiguity of possible candidates by reducing the field of view to a region of interest and not scanning the whole image. This region of interest can e.g. be a moving object, which is to be traced. Once found, not the whole image has to be interpreted, but only a certain range around the proposed location of the object (e.g. [Pahlavan et al. 1993]). Active vision is an issue which focuses at improving the strategic side of the interpretation task.
On the modeling part, as shown in section 4, there are different models with different representations available. The question is, which model to use for which purposes. In general, using all information which is known about a scene supports the interpretation task to a great deal. This means to include allsensors and context available, and use a hierarchical modeling scheme. This hierarchy supplies descriptions on various levels of detail, going from general to specific, and also the functional dependencies between the levels.
The question remains, which technique to use in order to define such models. Normally, they are formed by hand by selecting the seemingly relevant features and relations. In complex scenes and with complex natural objects, it is often not clear in advance, what the relevant object features are. Thus in the following a procedure is sketched, which derives models from examples in a learning process.

## 6. LEARNING MODELS FROM EXAMPLES

An analysis of human perception and knowledge acquisition reveals, that these processes proceed on a level of near unconsciousness. That makes it hard to understand the underlying processes and to transform it to a machine. Physiological studies demonstrate, that humans have a bundle of techniques available of how to perceive objects in their surrounding. These techniques have developed during the evolution history and are now part of the inborn structure of man [Lorenz 1977]. These techniques and schemes enable humans on one hand to perceive individual objects in a agglomeration of gray-values sensed by the eyes (separation of figure and ground). On the other hand, they help to derive new knowledge basing on known structures. Without these basic structures no perception and no cognition is possible, since only that can be known and thought, of which the structural elements are known.
In order to transfer this knowledge acquisition technique to a machine, a perceptive apparatus, and a set of schemes has to be provided. In this paper this has been done for a restricted area, namely for a two-dimensional world of geometric objects. In this world, there is a fixed set of objects, which have some geometric properties. Furthermore the system has to know a couple of methods in order to manipulate these objects. Basing on this information (which is further specified in the next sections) a technique to acquire new knowledge with the help of supervised learning is presented.

### 6.1 Automatic Derivation of a Generic Parcel Model

The actual application is in the domain of landuse classification, namely the modeling of parcel structures. This example shall demonstrate the need for a flexible modeling process and for generic models for object recognition and image interpretation. A field of land is a 'generic' object, since it describes a whole class of (possibly different shaped) objects. To describe each parcel individually might be possible in a restricted area, is however impossible for all fields. A possible generic description of could be: **normally, a parcel is a rectangular area.** This is a valid

description. The first problem is, that it has to be transferred to software (what is 'normal' ?). The second is that it is not a unique description: it holds true for many other structures in our environment. What is important for the description of parcels is the knowledge about the surrounding of a single parcel. In that way, both the description of the individual object and the typical relations it has to its neighbors have to be included into the model.
The system knows about a simple 2D-world. This world consists of objects of three different types: points, lines and regions. The system has programs to handle and manipulate these objects. This knowledge defines the expressive power of the system. It can be subdivided into basic objects, basic relations and basic functions for object recognition and manipulation.

### 6.2 Basic Objects and Relations

Basic objects are all the objects the system can perceive directly. The system knows points, lines and regions, which it can extract from given data with the help of special functions. These functions include simple geometric routines like classification of points, determination of the length and angle of a line, the derivation of closed regions from a list of lines, the area size, type and form. The system knows about fundamental relations between objects, namely point-relations (distance between two points, distance between point and line, point-in-polygon-test), line-relations (distance of 2 lines, angle between two lines, parallelity, orthogonality) and relations between two regions (points (of different types) two regions have in common, difference in size, form or type, common lines, etc.).

### 6.3 Basic Functions

These functions are stored in the system database in a flexible rule structure. A rule specifies, which function can be applied for a specific concept. This form of knowledge organization aims at a separation of knowledge strategies and data. Thus the way the data is handled can be fixed, even if the database is extended by new rules. If e.g. a color image is available, simply a new rule of how to handle color, can be added to incorporate the new functionality, while the strategies remain the same. In addition to the mere object-manipulation functions, there are metafunctions, for basically two functionalities:

1. Generalization is done by grouping object parts to form a new, a more general object. To this end, a structural clustering algorithm known from numerical statistics is applied, which successively merges objects which are spatially proximate. The definition of 'proximity', however is depending on the task and the domain. E.g. a set of edge pixels can be combined to form a line. In this case proximity is defined in terms of a common orientation and spatial proximity of the edge pixels, which are both numerical values. Using the clustering procedure for symbolic data requires the definition of symbolic proximity or neighborhood in a numeric way.
2. In order to derive such a proximity measure for different tasks, a learning algorithm is provided. ID3 [Quinlan 1989] is a learning system that starts from a collection of attributes of some classified objects and tries to find out, what the discrimination attributes for the classes are. This functionality is used especially incases, where a huge amount of examples is available, but the inner relations among the data are unknown. One example is to derive the causes for plant diseases from a collection of observations of infected and sound plants and their attributes (like color of leaves, size, speckles on leaves, etc.). A classification into infected and non-infected plants is given implicitly in the data - the learning step is deriving an explicit representation of this (hidden) relation, by revealing the dependencies between plant attributes

Photogrammetric Week'93 Eds. D. Fritsch / D. Hobbie, Wichmann, Karlsruhe, 1993.



and an infection. ID3 works basing on the information theoretic principle. The data set is successively divided with the help of the most discriminating attribute, the one giving the maximal information gain, to finally end up with a decision tree. ID3 is a powerful method. However it requires the input to be provided in form of attribute lists, so no higher level relations can be treated. Furthermore the program assumes the data to be consistent.

### 6.4 Learning Parcel Model with Learning Techniques

The method is now demonstrated with the example of the modeling of a parcel structure. Starting point is a 'parcel object' represented as a collection of points and lines in the database[1] (Figure 1.a). This input has to be treated with the systems' functions, namely the procedures to recognize points, lines and regions, and their attributes and relations. Using all the knowledge leads to a model of the parcel in terms of a collection of areas with their attributes and specific neighbor-relations (Figure 1.b).

```
[poi poi1 [126.955 112.322]]
[poi poi2 [123.684 20.4973]]
[poi poi3 [74.1003 19.1632]]
...
[poi poi11 [27.7588 176.564]]
[poi poi12 [29.7168 17.969]]

[line lin1 poi11 poi9]
[line lin2 poi7 poi5]
[line lin3 poi4 poi2]
[line lin5 poi9 poi7]
...
```

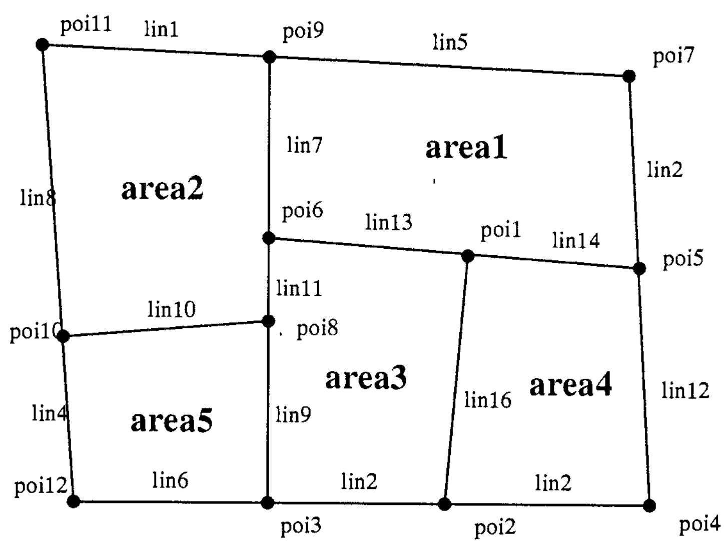


Figure 1: Starting point of analysis (a) and derived regions (b).

In order to generalize this description, possibilities to merge these elementary objects into higher level structures are sought. If no such functions are given for this specific object, the learning capability can be invoked. A 'teacher' is asked to give examples for a possible merging of objects, and to name this procedure. One possibility is to start from an intuitive idea of neighborhood in parcel structures, and to give examples for a typical neighborhood of two fields of land, e.g. fields 1 and 4. The question is, what is the feature (or the features) these two objects share, which allows for the merging ? In the learning phase, examples (and counterexamples) for this concept 'neighbor' have to be specified by the teacher, along with a classification. The system looks up or calculates all the attributes, the given concept could have and stores them in a datafile. In the case of learning a concept that concerns the relation of two regions, all the functions the system has available for 2 regions can be applied. So all the attributes that can possibly be extracted for the two regions are calculated, since it is not known in advance, which function (or which attribute the function

[1]The program is written in POP11, a flexible programming environment, which eases list-bases procedural programming. The following extracts are written in POP11-syntax.

derives) will result in the searched functionality. Thus the routines for the extraction of information about the two regions are called:

- distance of the individual regions,
- length of common line,
- number of common junction-types: T-junctions, T-junctions with a common stem, L-junctions,..
- etc.

In that way, all features of the examples are gathered.

```
% neighbor common_ell common_tee common_stem common_frk common_arw common_njn5
                        form_diff size_diff distance common_sides  %
[neighbor    0 2 2 0 0 0 same_form 1534.19    75.0233     1.0 ] ;;;  [2 5]
[no_neighbor 0 2 1 0 0 0 diff_form 1184.86    68.509      0.7 ] ;;;  [1 2]
[neighbor    0 2 2 0 0 0 same_form 118.404    54.2336     1.0 ] ;;;  [4 3]
[no_neighbor 0 0 0 0 0 0 diff_form 2719.05    112.4       0 ]   ;;;  [5 1]
[no_neighbor 0 0 0 0 0 0 same_form 1492.14    98.7988     0 ]   ;;;  [4 5]
```

After enough examples have been carefully selected and presented, the collected data is treated by ID3. As a result, automatically a function 'neighbor' in form of a decision tree is derived, which gives a classification depending on a set of attributes.

```
define NEIGHBOR ( area1,area2 ) -> klasse ;
   vars klasse,area1,area2
   undef -> klasse;
   if ( COMMON_STEM (area1,area2) ->>> val) <= 1 then
      'no_neighbor' -> klasse;
   elseif ( COMMON_STEM (area1,area2) ->>> val) > 1 then
      'neighbor' -> klasse;
   endif;
enddefine;
```

The function calculates whether two areas are neighbored in the defined sense and reads as follows: the neighborhood of two regions can be solely defined basing on the number common special points, namely those T-junctions, which share a common stem. If the number of these points (calculated by routine COMMON_STEM) is less than 1, then the two areas are not neighbored, else they are.

It should be stressed, that the information about neighborhood was given implicitely in the data - however in order to use it as a function, it has to be turned into an explicit form.

This new function is now available in the function database. It enables the system to use this new property for the clustering process. This results in a successive clustering of neighboring parcels - using the notion of neighboring just learned and defined in the new function. The clustering stops when no more objects are available, to which the function can be applied.

Region **area2** and **area5** are merged in the first step, since they are neighbors in the sense of the new function. Merging them results in a new region **area2'**.After merging **area3** with **area4**, this new region **area3'** is coming into the neighborhood of **area1**.

[area2 area5] -> area2'
[area3 area4] -> area3'
[area1 area3'] -> area1'
[area1' area2'] -> area1"

In this case, a single function was sufficient to 'explain' the whole object 'parcel structure'. This is a hint, that a parcel of land is a recursive structure: a partition of a parcel results in parcels again.

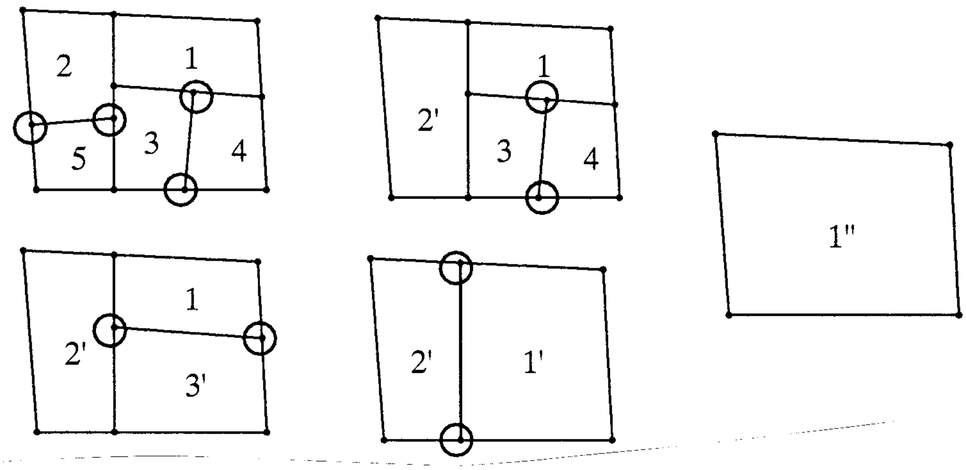


Figure 2 : Successive clustering of parcels sharing relation 'neighbor' (indicated with circles).

This fact is obvious, when considering the formation of such a structure during reallotment:a big parcel is divided into smaller ones which themselves may be divided further on.
As a result of the clustering the partitioning process of a big parcel into its smaller parts is reconstructed. In this way, a hierarchical description of a parcel structure is gained. On the lowest level, a parcel of land consists of a set of points and straight lines. The following level describes it as a closed region (of a set of possible forms), while the newly learned relation specifies the connections to neighboring parcels.

### 6.5 Evaluation

A semi-automatic procedure for model generation has been presented with the example of two-dimensional objects. The ideal has been the human ability to perceive, recognize and acquire new models. Starting with a set of basic elementary objects, relations and functions, the aim was to construct a hierarchical and general model, which is described in the 'language' of the system. A fundamental prerequisite for the result of a good model is to provide a 'rich environment', namely a thorough background knowledge of the system. If the system is lacking some basic perceptive functionalities, the models it can derive will be restricted. If e.g. the notion of the T-junctions and stems were unknown, the system could not detect the 'neighbor'-relationship.
The purpose of the modeling process is to derive models in order to apply them for interpretation tasks, namely to recognize and identify unknown parcels. To accomplish this, all different representation levels which evolved during generalization in the modeling procedure can be used: starting from straight lines, looking for closed areas, combining them pairwise when sharing the property of neighborhood, to finally end up with a complete parcel structure. The application of such models to real data has not yet been investigated. In principle, all kinds of matching procedures can be applied, depending on the type of features the model bases on.
The expressive power of the system presented is limited by the type of learning procedure applied. As stated above, only unary relations, namely attribute lists, can be treated, so no higher order relations between objects can be introduced. In these relations have to be transferred into the attribute-list form, which is however not adequate in some cases. Describing e.g. a house, the facts that the windows are inside the wall, orthat the chimney is on top of the roof cannot be transferred into the form requested. In this case, other types of learning strategies have to be applied.

## 7. CONCLUSIONS

After an overview of the tasks photogrammetry has to deal with, the fundamental problem of automation, namely object identification as a cognitive process has been named. The matching problem between model and image can be approached on different levels, depending on the task. Mere matching between images (relative orientation, aerial-triangulation, Sensor-location) can be accomplished with general 'point'-models with great success. The photogrammetric task of object extraction for mapping purposes however basically needs object models. In the case of natural objects, which exist in a variety of types, the need for a generic description is obvious. These models have to be described in a suitable way. Thus a procedure has been presented which allows for model generation basing on learning techniques. The result is a generic, hierarchical model description. The idea was to imitate the human knowledge acquisition capabilities. In such a complicated domain like the interpretation of natural scenes, a fully automatic system is unrealistic to date. It is clear, that a lot of additional knowledge has to be incorporated into the system, but often unknown, what kind of knowledge it has to be. Thus a procedure has been developed, which uses all the knowledge available in order to explain a given input. If the interpretation stops due to missing facts or functions, there is the possibility to extend the knowledge base by learning new concepts, presented by a teacher.

## 8. REFERENCES

Birnbaum, L., M. Brand and P. Cooper (1993): Looking for Trouble: Using Causal Semantics to Direct Focus of Attention. Proceedings of the ICCV, 1993.

Haala, B. and M. Hahn (1993): Quality and Performance Analysis of Automatic Relative Orientation. SPIE Conference on Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision, Orlando, Florida, 1993 (to be published approx. August 1993).

Hahn, M. (1992): 3-D Egomotion from Long Image Sequences. High Precision Navigation 91, Linkwitz, K., Hangleiter, U. (ed.), Dümmler Verlag, Bonn, 1992.

Krzystek, P. (1991): Fully Automatic Mensuration of Digital Elevation Models with MATCH-T. Proceedings of the 43rd Photogrammetric Week, 1991.

Lorenz, K. (1977): Die Rückseite des Spiegels, dtv, 1997.

Metzger, W. (1953): Gesetze des Sehens, Kramer Verlag, Frankfurt, 1953.

Pahlavan, K., T. Uhlin and J.-O. Eklundh (1993): Dynamic Fixation. Proceedings of the ICCV, 1993.

Quinlan J.R. and R.L. Rivest (1989): Inferring Decision Trees Using the Minimum Description Length Principle. Information and Computation, vol. 80, 1989.

Schickler, W. (1992): Feature Matching for Outer Orientation of Single Image Usinge 3-D Wireframe Controlpoints. International Archives of ISPRS, Comm. III, Washington D.C., 1992.

Stark, L. and K. Bowyer (1991): Achieving Generalized Object Recognition through Reasoning about Association of Function to Structure. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, no. 10, 1991.

Stilla, U. and K. Jurkiewicz (1991): Objektklassifikation mit einem Blackboardorientierten Inferenzmechanismus. Forschungsinstitut für Informationsverarbeitung und Mustererkennung (FIM)-Bericht Nr. 230, Ettlingen, Juni, 1991.

Strat, T.M. and M.A. Fischler (1991): Context-Based Vision: Recognizing Objects Using Information from Both 2-D and 3-D Imagery. IEEE Transactions on Pattern Analysis and Machine Intelligence vol. 13, no. 10, 1991.

Tsingas, V. (1991): Automatische Aerotriangulation. Proceedings of the 43 rd Photogrammetric Week, 1991.

Zimmermann, G. (1992): Automatic Acquisition of Optical Landmarks from Image Sequences. High Precision Navigation 91, Linkwitz, K., Hangleiter, U. (ed.), Dümmler Verlag, Bonn, 1992.