

STRUCTURAL ANALYSIS OF COMPLEX AERIAL PHOTOGRAPHS

Makoto Nagao
Department of
Electrical Engineering
Kyoto University
Sakyo, Kyoto, 606
JAPAN

Takashi Matsuyama
Department of
Electrical Engineering
Kyoto University
Sakyo, Kyoto, 606
JAPAN

Hisayuki Mori
Department of
Electrical Engineering
Kyoto University
Sakyo, Kyoto, 606
JAPAN

A new system for the structural analysis of complex aerial photographs is presented. This system has the ability of focussing its attention of the analysis on the limited local areas where objects are highly supposed to exist. Several kinds of strong and typical features are extracted, and these primary features of objects are combined to extract rough areas of the objects. This focussing mechanism saves the total processing time and facilitates the detailed analysis. The recognition process of the system is implemented according to the 'production system'. The knowledge sources in this system are object-detection subsystems which analyse their individually focussed local areas and recognize specific objects respectively. All the results of the analysis are written in the common blackboard, and the system finds out conflicts and recovers errors by backtracking to feature extractions and low level processings. This architecture enables us to organize smoothly the diverse knowledge required to describe the complex structure on the ground surface.

1. INTRODUCTION

Recently several systems for the analysis of aerial photographs have been developed to locate objects on the ground surface[1],[2],[3]. When we are going to get a description of the structure on the ground surface by analysing an aerial photograph, we find several difficulties which are not encountered in other image analysis areas:

- (1) The size of a picture is very large.
- (2) As it is impossible to control photographing conditions, the quality of a picture is apt to change.
- (3) The sizes and the textural properties of objects vary quite widely.
- (4) Variation of objects in a scene entails calculation of many different features and the diverse knowledge of the world.
- (5) There are too many situations on the ground surface to build a general model which can represent all possible spatial arrangements of objects.

In order to overcome these difficulties and to realize an efficient and reliable analysis, we have developed a new system based on the 'production system'[4]. The system consists of a group of object-detection subsystems which individually search for specific objects by communicating with each other via a common 'blackboard'[5],[6].

The analysis process of this system is divided into the following steps.

- (1) Segmentation: After noise removal, an aerial photograph is segmented into elementary regions according to the multispectral property.
- (2) Global survey of the whole scene: Regions with characteristic properties are extracted. These are used to confine the spatial domains of objects where object-detection subsystems work.
- (3) Detailed analysis of focussed local areas: Each object-detection subsystem focuses its attention on a specific local area and checks the existence of a specific object.
- (4) Communication among object-detection subsystems: All the information about the properties of regions and recognized objects is stored in the blackboard. Object-detection subsystems interface with it in a uniform way. The system controls the overall flow of the analysis by managing the information in the blackboard. It solves conflict among object-detection subsystems, and corrects errors by backtracking.

This paper mainly describes the control structure of the system, and demonstrates its performance with experimental results. The detailed algorithms of picture processing routines are described in [7].

2. FOCUS OF ATTENTION IN THE ANALYSIS OF AERIAL PHOTOGRAPHS

Aerial photographs we want to analyse are 4 band multispectral images of complex suburban areas taken from the low altitude(Fig. 1).

As the size of a picture is very large, the analysis should be done as efficiently as possible. If each object-detection subsystem, which is very sophisticated and time-consuming, were applied to the whole picture area, it would take a prohibitive time to complete the analysis. In order to solve this problem, we adopted the focussing mechanism which human being seems to use when he interprets a complex scene. When he sees a scene, at first he globally surveys it to find the characteristic areas which attract his interest. Then, he goes into the detailed examination of some local area to find out objects using his knowledge of the world. The more intensively he focuses, the more specialized knowledge he comes to use.

In our system, several kinds of regions with prominent features(characteristic regions) are extracted by simple picture processing programs. Then, object-detection subsystems apply sophisticated programs only in the local areas where specific objects are supposed to exist. As the time-consuming processings are applied only in small local areas, the total processing time can be reduced very much. Fig.2 shows the schematic drawing of this focussing process.

In order to specify the spatial domains of objects as correctly as possible, we utilize such features as size, shape, brightness, multi-spectral properties, texture and spatial relations among regions. We extract seven types of characteristic regions, i.e. large homogeneous regions, elongated regions, shadow regions, shadow-making regions, vegetation regions, water regions and high contrast texture regions.

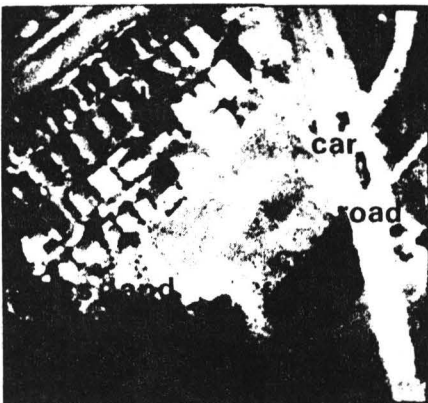


Fig.1 A picture of a complex suburban area

All the features used here are stable against the change of photographing conditions, and all parameters are automatically and adaptively determined from the picture data under analysis. Therefore extracted characteristic regions can be the reliable basis for the subsequent detailed analysis.

Fig.3 shows the characteristic regions extracted from the picture of Fig.1.

3. OBJECT-DETECTION SUBSYSTEM

An aerial photograph contains a variety of objects such as crop fields, forest areas, grass lands, roads, rivers, houses and so on. The diverse knowledge should be incorporated to describe the structure on the ground surface. In addition, these objects, especially in a small country as Japan, are intricately arranged without definite spatial relationships.

Taking these conditions into account, it seems to be natural to divide the system into a group of object-detection subsystems. Each of them is designed to find out specific objects using the knowledge of their intrinsic properties and the environments in which they are embedded. The diverse knowledge of the world is distributed in object-detection subsystems. This facilitates the implementation of the system.

The type of object-detection subsystems can be classified into two categories according to the information they use in selecting the candidate regions of objects.

(1) Picture data-driven subsystem

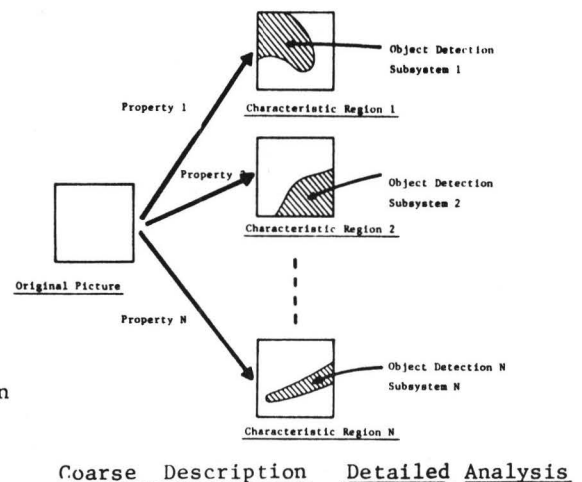


Fig.2 Focussing mechanism of the system

The subsystems of this type check the existence of the local areas with specified properties by combining characteristic regions. If there are such local areas, they examine them in detail by activating special purpose feature extraction programs.

(2) Model-driven subsystem

The subsystems of this type pick up the candidate regions by using the spectral and spatial relationships with already recognized objects. For example, it is very difficult to detect cars by a simple knowledge that they are rectangular unless we impose the condition that they are on the roads or the parking lots. Therefore the subsystem for the detection of cars entails the recognition of roads or parking lots.

The present system has thirteen object-detection subsystems for eight kinds of objects: crop field, bare soil(crop field without plantation), forest, grass land, road, river, car and house. Fig.4 shows intermediate results of detection of these objects.

4. THE STRUCTURE OF THE BLACKBOARD

Each object-detection subsystem interfaces with the common data base, the blackboard, in order to test conditions for activation and to write

in the result of the analysis. All the communications among object-detection subsystems are made via this blackboard. Fig.5 shows the schematic drawing of the structure of the blackboard.

The blackboard contains global parameters as well as the properties of regions and objects. They denote the global properties of the picture data, that is, photographing conditions of an aerial photograph such as the direction of the sun, threshold values for checking the similarity of multispectral properties and so on. From these global parameters, each object-detection subsystem gets the information of the quality of the picture data under analysis. Therefore, it can successfully find out objects in spite of unstable photographing conditions of aerial photographs.

In this system, elementary regions, which are segmented according to the multispectral property in segmentation, are considered as the basic units for all higher level processings. Their fundamental properties(the average gray level in each spectral band, size, location and some basic shape features) are stored in the property table in the blackboard(Fig.5).

We store LABEL PICTURE in the blackboard in order to denote the spatial relationships among elementary regions. It is the symbolic

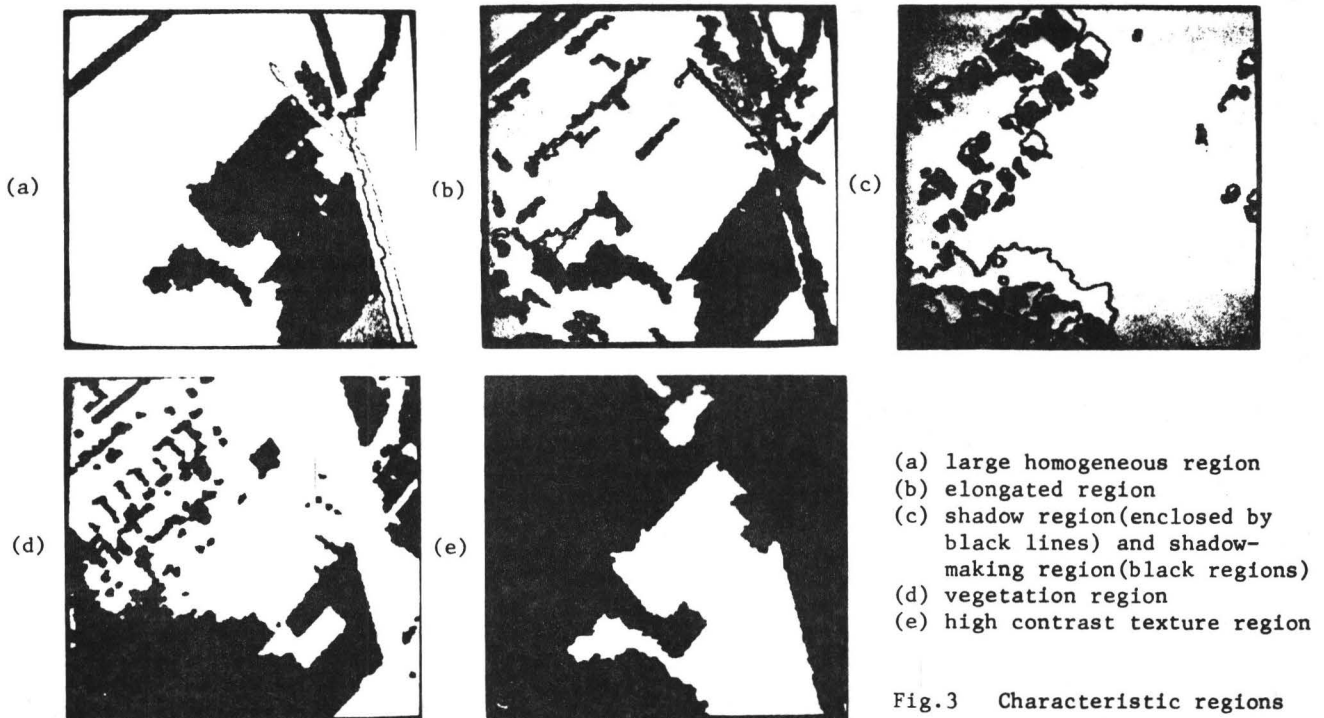


Fig.3 Characteristic regions

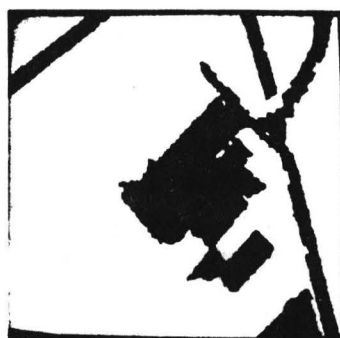
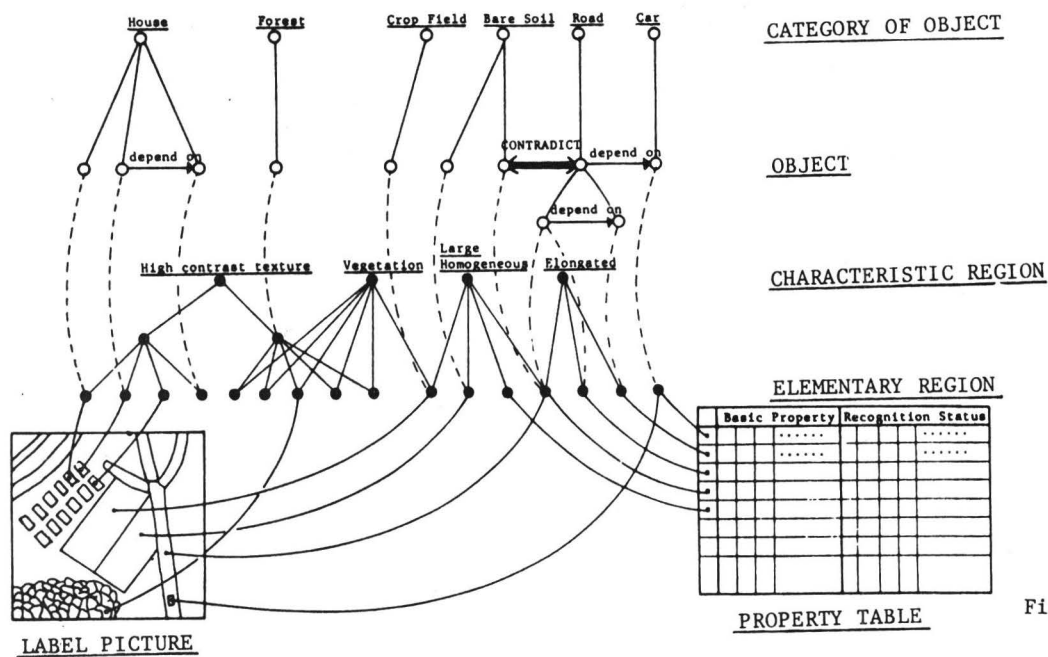


Fig.4 Intermediate results of object detection:
Several regions are recognized as different objects.
These conflicts are solved by the system(see Sec.5).



picture where each pixel in an elementary region is labeled with the same unique region number(Fig.6). The reason for this is:

The spatial relationships among regions used by object-detection subsystems are very different and depend on the properties of objects they want to find out. Therefore it is not economical, and is even sometimes impossible to calculate all spatial relationships in advance. For example, when we want to estimate the location of a house by using the regularity of the spatial arrangement of already recognized ones, it will take a long time to find a candidate region of a house if we do not use the two-dimensional image.

As each pixel in LABEL PICTURE contains an unique region number, we can easily get the properties of the region to which it belongs. On the other hand, the location of a region in LABEL PICTURE is denoted by the two coordinate pairs, (BX, BY) and (EX, EY), in the property table(Fig.7). When one wants to calculate a new feature of a region, one only has to scan within the rectangle area specified by these coordinate pairs(Minimum Bounding Rectangle). This is very useful to save the processing time of picture processing programs.

Each characteristic region in the blackboard is represented by a characteristic region node which denotes a group of elementary regions. Two-dimensional images of characteristic regions can also be constructed from LABEL PICTURE. These are very useful in locating clusters of elementary regions such as forest and residential areas.

When each object-detection subsystem recognizes an object, it generates an object node in the blackboard, and the object node and its constituent elementary regions are connected by part/whole relations. If the recognition of an object depends on the properties of already recognized ones, the node of a new recognized object is



Fig.6 LABEL PICTURE: each pixel in an elementary region(enclosed by black lines) is labeled with an unique region number.

linked with those of old ones(Fig.5). By storing the history of the recognition process, the system can give back the state of the blackboard to the correct one when it finds out an error in it.

5. CONTROL MECHANISM OF THE SYSTEM

As each object-detection subsystem recognizes objects independently of the others, the system incorporates the mechanism to solve conflicts between these subsystems.

The property table has the field where each object-detection subsystem returns one of the following recognition statuses: "unanalysed", "recognized", "irregular shaped" and "rejected". the system checks this field, and if some elementary region is recognized as different objects by multiple subsystems, it evaluates the reliability value of each object to which that region belongs. Then the objects nodes except the most reliable one are deleted, and the corresponding recognition statuses of the region are changed from "recognized" to "rejected".

If there are other objects which have been recognized in connection with rejected objects, the system also deletes those object nodes by traversing the dependency links from rejected objects. In this case, the recognition statuses of their constituent regions are given back to "unanalysed", for they might be recognized by using the properties of the other objects of the same kind.

Errors in segmentation are also repaired by the system. Object-detection subsystems checks various properties of regions to recognize objects. They returns the recognition status, "irregular shaped", if the shape of a region is not suitable for an object while all the other properties

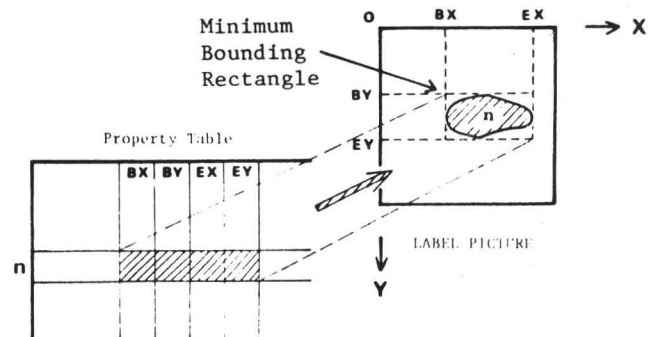


Fig.7 Specification of the location of an elementary region in LABEL PICTURE

are satisfactory. Then, the system re-analyses that region by applying the split/merge program. If the region has the bottle-neck as shown in Fig. 8, it is divided into small regions. If the boundary of the region is very rough and irregular (Fig. 9), neighboring small regions with the similar color are merged with it. These newly generated regions are added in the blackboard as temporary regions and analysed by object-detection subsystems. If a temporary region is successfully recognized, it is registered as a new elementary region in connection with its related object, and the original region is deleted from the blackboard. If the result of the recognition of the new region contradicts with that of the original one, the system deletes one of them depending on the reliability. When the temporary region is not recognized, it is removed from the blackboard, and the corresponding recognition status of the original region is changed to "rejected".

In the case of Fig. 8, the bottle-neck results from the error of mismerging two adjacent houses. When this region is splitted into two regions, the house detection subsystem comes to recognize them successfully. The region in Fig. 9 corresponds a crop field, but it is not recognized because of the irregularity of the boundary. After the merging process, it comes to take the smooth boundary, and is recognized as a crop field.

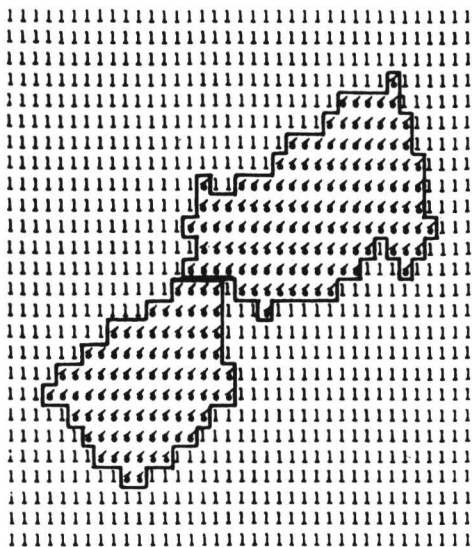


Fig.8 Splitting a region with a bottle-neck

By the benefits of the above-mentioned control mechanisms of the system, each object-detection subsystem can use the properties of already recognized objects without considering the results of recognition by the others. The system stops the analysis when no new objects are recognized.

Fig. 10 shows the final result of the analysis. Although white regions with no marks are left unanalysed, we can see that almost all objects which belong to clear semantic categories are successfully recognized. (As we do not have the ground truth data, the evaluation of the result is done by visual inspection.)

6. CONCLUSIONS

The system for the structural analysis of complex aerial photographs has been presented. The major conclusions on this system are the following:

- (1) The focussing mechanism realizes the efficient analysis and successfully isolates objects. Table 1 shows the efficiency of this focussing mechanism.
- (2) The production system architecture gives us the following benefits.

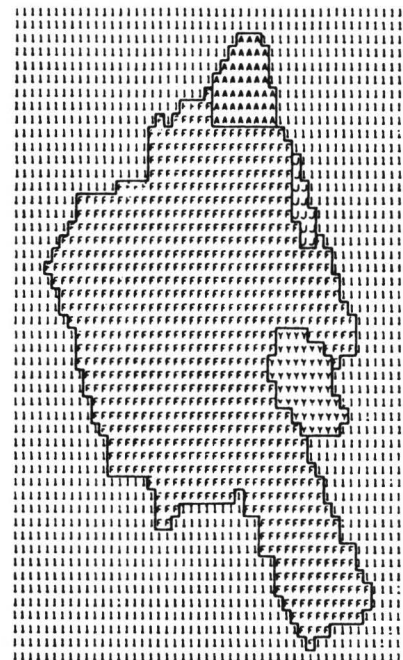


Fig.9 An irregular shaped region: the central large region is merged with neighboring small regions.

- A. Modularity: Diverse knowledge required to describe the structure of the ground surface can be individually implemented in object-detection subsystems.
- B. Model-driven analysis: We can use the properties of already recognized objects in order to analyse unrecognized areas and to find out contextual-dependent objects.
- C. General control mechanism: The system takes the full responsibility for the maintenance of the blackboard. It solves the conflicts among object-detection subsystems and corrects the errors in segmentation. If necessary, it undoes the content of the blackboard to remove the effects of errors. These mechanisms of the system integrate mutually independent object-detection subsystems.

This system is implemented on a large computer, FACOM M-190. The total CPU time was about 150 sec. in the picture of Fig.1. Experiments on several different aerial photographs have shown that this system works fairly well.

SCAN AREA	CANDIDATE REGION	TIME(sec.)
in MBR*	large homogeneous regions	0.14
whole area	large homogeneous regions	0.39
in MBR*	all regions	0.84
whole area	all regions	28.95

Table 1 Efficiency of the focussing mechanism:

Time denotes the processing time for the crop field detection subsystem to calculate straightness of region boundaries. As we focus only on large homogeneous regions, we can save processing time very much.

* MBR denotes a minimum bounding rectangle(see Fig.7)

REFERENCES

- [1] Barrow, H.G. et al "Interactive Aids for Cartography and Photo Interpretation: Progress Report, October 1977" In Proc. of the Workshop on Image Understanding, Oct., 1977, pp. 111-117
- [2] Quam, L.H. "Road Tracking and Anomaly Detection in Aerial Imagery" In Proc. of the Workshop on Image Understanding, May, 1978, pp. 51-55
- [3] Nevatia, R. and Price, K. "Locating Structures in Aerial Images" In Proc. IJCPR-78, Nov.,1978, pp. 686-690
- [4] Davis, R. and King, J. "An Overview of Production Systems" AIM-271, Stanford A.I.Lab., 1975
- [5] Erman, L.D. and Lesser,V.R. "A Multi-level Organization for Problem Solving Using Many Diverse Cooperating Sources of Knowledge" In Proc. IJCAI-75, 1975, pp.483-490
- [6] Lesser, V.R. and Erman, L.D. "A Retrospective View of the HEARSAY-II Architecture" In Proc. IJCAI-77, 1977, pp. 790-800
- [7] Nagao, M., Matsuyama, T. and Ikeda, Y. "Region Extraction and Shape Analysis of Aerial Photographs" In Proc. IJCPR-78, pp. 620-828

Fig.10 Final result of the analysis:

White regions without mark are unrecognized(20% of the whole picture area)

BS: bare soil
(crop field without plantation)

CF: crop field

G : grass land

F : forest

RD: road

R : roof of a house

C : car

r : misrecognized roof
(these regions have similar multi-spectral properties to that of true roof)

