*MVA '96* IAPR Workshop on Machine Vision Applications, November. 12-14, 1996, Tokyo, Japan

# An Image Retrieving System Based on User–Specified Recognition Model

Wei Wu *
CyberNet CO. LTD.

Hiroshi Mo, Masao Sakauchi †
Institute of Industrial Science
University of Tokyo

## Abstract

In this paper, we shall discuss an extension of MTDM model for image retrieval based on user's specification. A MTDM model is composed of objects which describes the static and the matching control methods for recognition targets. Characteristics of the describing targets are declared by abstract members of the MTDM model objects. In this way, the differences of data description can be easily hidden. In our proposed image retrieval system, user specifies a MTDM model to each keyword for image retrieval that may be not exist in the image database. The retrieval acquisition is structured and represented in the form of MTDM model. When a keyword is not found in the image database, the related reasoning model(MTDM model) will be fired to match corresponding objects form each image. By the extension of MTDM model, the differences of data are integrated. The image retrieval system can treat objects in a general way and the recognition model and the retrieving model can be integrated by the MTDM model.

## 1 Introduction

For more flexible information retrieval from image database, the automatic acquisition of keywords and the technique to describe keywords are important. Limited by the image understanding technology, some researches aim at techniques of describing the contents of image in multiple abstract levels[1, 2, 3, 4]. By this approach, the input images are first analyzed by an image understanding system. Then the image retrieving system uses as keywords the unperfected analysis results which correspond to the internal level of the contents structure[5, 6, ?, 7]. In these researches, description of keywords, integration of different type of data and integration of recognition model and retrieval model remains unsolved. As a result, the keywords themselves are not structured; the describing ability is weak. Further more, it is still difficult for image understanding system to integrate different types of input data by model and to describe in a general way the recognition results for different types of input data description. To solve the above mentioned problems, we have proposed an active model for drawing understanding system named MTDM[8]. In this paper, we shall discuss the extension of MTDM model for image retrieval based on user's specification.

A MTDM model is composed of two types of objects. One is called model-object, which describes the static attributes of the recognition targets. The other one describes the matching control method for reasoning process and is organized as a tree structure named matching-tree. When applied to object recognition, a matching-tree is declared according to the recognition purpose and the characteristics of the objects. The model-objects refer to the recognition target and its components are independent of the matching-tree and can be reused in different matching-trees. In both types of describing objects, characteristics of the describing targets are declared by their abstract members. In this way, the differences of data description can be easily hidden. This ability has been proven by our previous experimentation of distortion absorbing for drawing recognition[9].

A prototype of drawing understanding system which base on the MTDM model has been developed[9]. In this paper, we shall discuss the extension of MTDM model for image retrieval based on user's specification. House-images are used as retrieval targets.

## 2 Declaring Retrieval Query by MTDM Model

In general, a retrieval acquisition is declared by a logical expression composed of a set of keywords and logical operators. With the description contents of the keywords, there are two kinds of keywords exist:

1. Entities : The entities are Objects which are described in the input image. For example, in a house-image, the entities may be tree, roof, veranda and so on.

*Address: 3–2–5 Higasinakano, Nakano-ku, Tokyo 164 Japan. E-mail: wuwei@sak.iis.u-tokyo.ac.jp

†Address: 7–22–1 Roppongi, Minato-ku, Tokyo 106 Japan. E-mail: mo@sak.iis.u-tokyo.ac.jp

2. Abstract concept: Abstract concept describes the state of contents represented by an input image. It depends to the attribute value of entities. For example, in a house-image, type of the house, comfortableness of environment and so on are abstract concepts. They depend to the shape of the house entity, area and composition of the background entities and so on.

In our propotype system, Each entity or concept is related to a model-object of MTDM model which describes the static characteristics and the matching control for the object. The retrieval query then can be given as a MTDM model which is composed of some model objects according to their sub-concepts, entities and relational operations. In this way, the recognition model and the retrieving model can be integrated by the MTDM model.

### 2.1 Integrating Different types of data description

Edge vectors and color segments data which extracted from a image data are used as input data to the retrieval system. In the system, the color segment object is defined based on polygon object and specialized to hold a color attribute. The following model-object description declaras the color segment as a object which derived form polygon object.

```
element{
  name : colorSeg;
  kind_of : polygon;
  member : color;
  method : hasColor, display;
};
```

With the declaration, the color segment object inherits the characteristics of polygon and adds a color attribute. By this way, the color segment can also be treated as a object composed of vectors.

### 2.2 Declaring Retrieval Query by Matching-Tree

In general, the retrieval acquisition in concept level will not directly relate to entities in the input image. By the static view of declaring, the acquisition can be structed as following.

1. Building a concept relating to the acquisition. For example, "comfortable environment", "Japanese style" and so on.

2. A retrieval acquisition is composed of a set of keywords and relational operators. It can be decomposed as sub-concepts which are entities or abstract concepts, so a acquisition can be represented as a tree structure.

3. Repeating the dividing process, the tree structure will at last has entities as its leaf nodes. An internal node of the tree are related to a sub-concept or relational operator.

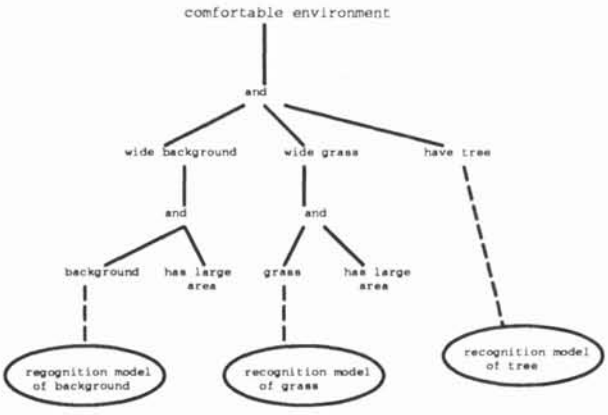


Figure 1: Matching-tree for retrival of "comfortable environment"

The structure tree then can be transformed to a matching tree easily by relating a model-object for each node in the tree.

Fig.1 shows an example of retrieval query described by a matching-tree. The MTDM model declares a concept of "comfortable environment" for a house-image. The leaf nodes in the matching tree correspond to the components of the retrieving target and the internal nodes of describe the relationship between the component objects. The query goal is matched according to the related matching-tree. An entity is related to a model-objects which declares the static characteristics of the entity. For matching an object, a set of matching-tree corresponding to different recognition purposes can be specified. The matching-tree of retrieval query can be flexibly bound to any specific matching-tree of entities. As a result, The variation of matching-tree of components will not cause variation of the matching-tree related to the retrieving acquisition.

## 3 Experimental Results

A prototype system which is based on the abovementioned approach has been developed. In this system, a set of models about component entities of house-image is defined. Roof, veranda, tree and so on are some examples of the model set. An input image is described by using the color segment data and edge data extracted from the bitmap data of the image. A matching-tree related to an entity is designed to match the entity from the edge data, the color segments or both of them according to the characteristics and recognition purpose. Fig.2 and Fig.4 show examples of matching-trees for entities described in a house-image image. A roof model matches shape attributes from edge data and uses color segments as supplementary information . A tree model matches color segments which include many noise (The noise is defined as a set of short edges enclosed by its model). In spite of the differ-

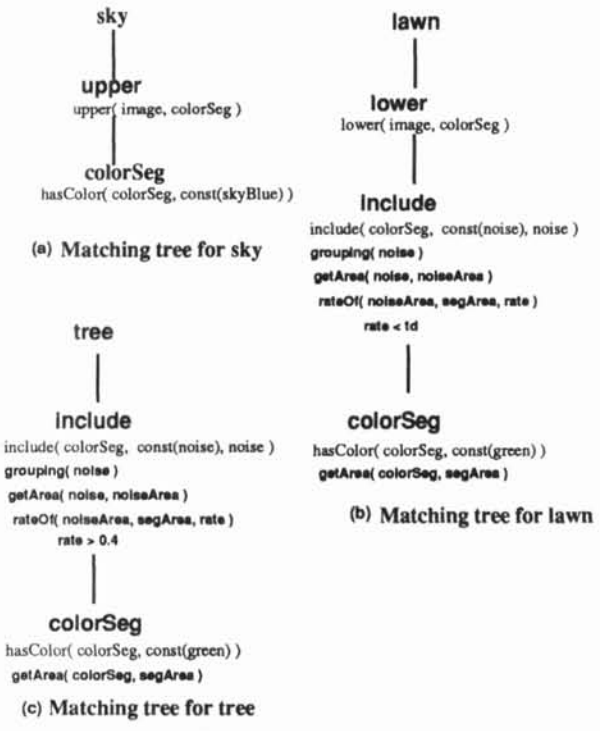


Figure 2: Matching-trees for background entities

ences in data representation, the matching-tree can treat the models corresponding to its construction components in the same way.

A house-image is composed of two kinds of information, one is the building and another is the background. For the building object, shape, materials(represented by color) are considered important. In the other way, the color and area parameters may be considered important for the components of background kind of objects. In our system, four types of components are defined as background object, they are sky, tree, grass and pond. It is difficult to declare these components by their shape parameters. In the case of representing the environment of a house, we shall not evaluate the detail shape parameters of the background objects, but extract them by evaluating the information of color and position. The matching trees for the background components are shown in Fig.2.

The matching trees are fired[8] by color segment objects. The matching tree for matching sky matches a sky object if a color segment has "large" area and in the "upper" location of the image. The matching tree only evaluates the color segment data. The lawn and tree both have green color, they are recognized by evaluating number of shade objects included in the green color segment.

When the matching tree of lawn is fired by a color segment with "green" color, it firstly matches shade objects(represented by "noise" objects in our model declaration) included in the area of the firing element. If the extent area of matched shades is less then a given threshold, a lawn is matched. A tree is matched in a opposite way.

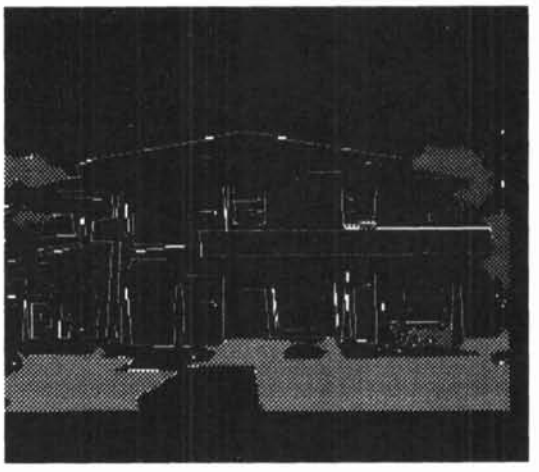

Figure 3: Recognition result of "comfortable environment"

In the input color segment data, shades are represented by segments with small area. By the segmentation preprocessing, small segments are deleted. In our system, model for matching shades is built to match shades from edge vectors. As shown in Fig.??, in the declaration of matching tree, the matching control of members of model-objects are only declared, it is no need to classify the difference between objects matching from color segments and objects matched from edge vectors. All model-objects described with different types of input data can be treated in a general way.

The retrieval acquisition for "comfortable environment" is matched by firing the matching tree shown in Fig.1. When there is no "background object" exist in the input image, the reasoning MTDM model related to the "background" object will be firstly fired. The operation of retrieval and reasoning are integrated in the same form of matching tree. Fig.3 shows a recognition result of "comfortable environment".

The matching trees for components of a house are built in the similar way(Fig.4. The following is a parts-model declaration for roof objects.

```
element{
  name : roof;
  kind_of : colorSeg;
  member : ;
  method : ;
};
```

In the Fig.4(a), the matching tree for roof with triangle shape will be fired by a "stableLine" object. The "near" node searches "stableLine" objects around the end points of the firing object. The "isTriangular" node tries to match a "triangle" object

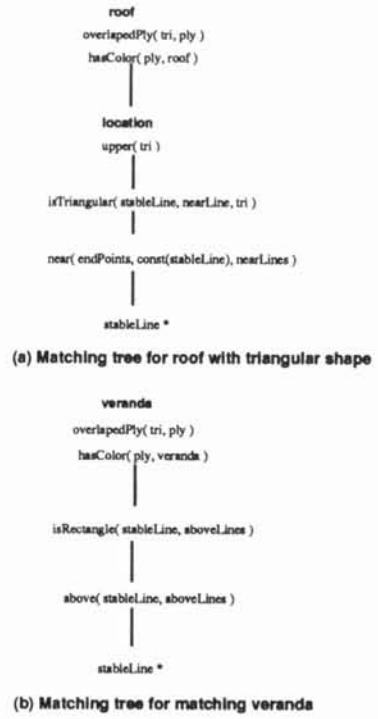


Figure 4: Matching trees for roof and veranda

from the firing object and searched "stableLine". The "location" node checks the "upper" relation for the matched triangle. At the last, in the root node of the matching tree, the color attribute of color segments included in the matched triangle are evaluated to color of roof. Fig.4(b) shows a matching tree for matching a veranda. It works in a similar way to that of roof.

In each of matching step, the matching targets may be represented by edge vectors or color segments. The differences of data representation are hidden and do not represent in the matching tree declaration. Fig.5 shows a recognition result for roof and veranda.

## 4 Conclusion

In this paper, we have discussed an extension of MTDM model for image retrieval. The goal of a retrieval acquisition is related to a abstract concept. The relationship among goal and sub-goals are organized in the form of MTDM model. The leaf nodes of the matching tree represent the retrieval keywords and are related to entities in the image. The propotype system retrieves images by firing the matching tree relating to the retrieval goal. When an unexisting keyword be found, the system fires the related reasoning model to match the keyword object. In this way, The retrieval acquisition can be represented in a structured form. Matching of key-

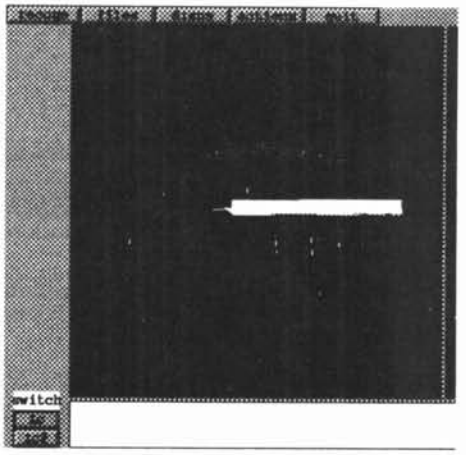


Figure 5: Recognition result of roof and veranda

words can be implemented in appropriate algorithm by defining corresponding matching trees that are suitable to the retrieval target and recognition objects.

## References

[1] S.WW.Smoliarand H.Zhang, "Content-Based Video Indexing and Retrieval", IEEE MultiMedia, pp62-72, Vol.1, No.2, 1994

[2] J.Kang and A.D.Narasimha, "Identifying Faces Using Multiple Retrievals", pp27-38, IEEE MultiMedia, Vol.1, No.2, 1994

[3] M.Akaishi and Y.Tanaka, "Search for Every Composite Object that Satisfies Ambiguously Specified Relative Locations of Its Components", pp1916-1925, Trans. of IPSJ. No.8, 1995

[4] K.Ozaki and Y.Yano, "The Development of a 3-Tuple-based Art Museum Database System ArtFinder ",pp2690–2701, Trans. of IPSJ., No.11,1995

[5] M.Sakauchi, "Database Vision and Image Retrieval", IEEE MultiMedia, pp79–81,Vol.1, No.1, 1994

[6] Y.Gong, "An Image Database System with Content Capturing and Fast Indexing Abilities", First International Conference on Multimedia Computing and System, 1994

[7] J.Yamane, M.Sakauchi, "A Construction of a New Image Database System Which Realizes Fully Automated Image Keyword Extraction", p1216-1233, IEICE. Trans., E76-D, 10, 1993

[8] W.Wu, W.Lu, M.Sakauchi, "An Object-Oriented Model For Drawing Understanding and Its Ability of Noise Absorption", ICDAR '95, 1995

[9] W.Wu, S.Satoh and M.Sakauchi, "A Multipurposes Drawing Understanding System with Object-Oriented Model Description", Trans. IPSJ, Vol.37, No.6, 1996