Reasoning About Images:
Using Meta-knowledge in Aerial Image Understanding

Peter G.Selfridge
Kenneth R.Sloan, Jr.

Department of Computer Science
University of Rochester
Rochester, New York 14627

Abstract

Image Understanding (I.U.) shares with Artificial Intelligence the need for mechanisms of using knowledge to control computation. In I.U., this knowledge takes the form of prior knowledge about objects and knowledge gained from doing the computations themselves. This paper presents an approach to the general I.U. problem and a specific program for locating buildings in aerial photographs. Prior knowledge is stored as an Appearance Model, which represents the appearances of possible buildings. A three stage program generates operator and parameter sequences to achieve recognition. Each level uses partial results from below to 1) search parameters for the best match, 2) infer obscuring image conditions and deal with them, and 3) infer high-level conditions such as the presence of occluding objects. This use of partial results is a kind of "metaknowledge", enabling the program to evaluate its progress and change hypotheses and computations accordingly.

Introduction

Image Understanding (I.U.) shares many problems with the field of Artificial Intelligence (A.I.). The use of knowledge to control processing is one such general problem. In an I.U. domain, this can be formulated as the problem of using both prior knowledge about the task domain, and knowledge acquired as processing proceeds, to choose a sequence of computations which will achieve recognition and location of the desired objects.

It is a tenet of this research that a program to do this must be prepared to use partial results, such as parual matches, to change the sequence of operators and their parameters. In doing so, it must be able to evaluate the results of its processing and be prepared to infer, at several levels, what image conditions are impeding recognition of the desired objects.

This paper describes a program to locate buildings in aerial photographs using the above approach. The program has three parts, shown in figure 1, each of which embodies different levels of knowledge and inference ability.

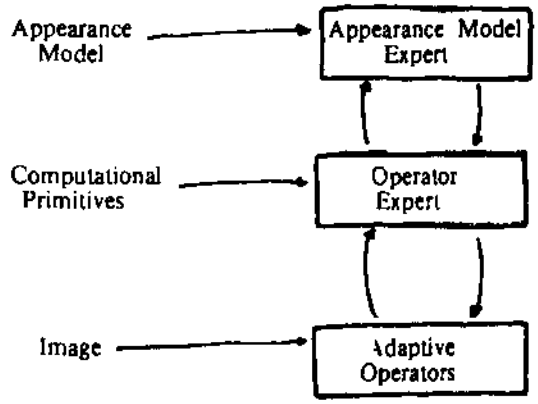


Figure 1: A Building Location Program

An Appearance Model Expert chooses subgraphs of the building Appearance Model. A Region Description is derived from this sub-model and passed to the Operator Expert. This Expert chooses operators to compute candidate regions and calls other operators to try to match the region characteristics with the Region Description. One kind of operator is an Adaptive Operator, which can vary a single parameter to optimize the match with a specific value or range of a feature. The Operator Expert can infer the existance of a possible image problem and take appropriate action. The Appearance Model Expert can infer high-level conditions and check spatial relations between regions to recognize buildings complexes.

Appearance Models and The Appearance Model Expert

An Appearance Model is a data structure encoding the expected appearances of a class of objects. This model could be derived by a "smart" modeller from many sources of prior knowledge such as a light model, a camera model, and a 3D domain model. The "appearance primitive" in our model is a region.

Our Appearance Model for buildings is a hierarchical graph encoding building properties and relations between buildings and other entities, such as roads or shadows. The model also stores verified and candidate regions in appropriate places. The hierarchy allows the Appearance Model Expert to choose a subgraph, representing the current hypothesised building type, on the basis of accumulated region evidence as well as knowledge of the Appearance Model itself.. From that subgraph, a Region

Description is created by combining all the region property values from that subgraph. This Region Description is a representation of the desired object in terms of expected region properties, and is passed to the Operator and Image Problem Expert.

Operator Expert

The Operator and Image Problem Expert is given a Region Description from the Appearance Model Expert, and attempts to locate a region with the desired properties. To do this, it uses knowledge to invoke appropriate operators to generate candidate regions, and calls Adaptive Operators to try to achieve good matches on specific properties.

Adaptive Operators

An Adaptive Operator is a program that is given a candidate region and a desired range of property values for a single property. This represents a goal to the operator of computing a derived region from the candidate with its property as close to the desired range as possible. Each Adaptive Operator has knowledge of the effects of varying a single parameter of a single operator. An Adaptive Operator uses this knowledge to vary the parameter, creating derived regions from the candidate, to gel a region that is the best match possible.

Eor example, Adaptive-Match-On-Size knows that for a dark region created by thresholding, raising that threshold will shrink the size of the region, and lowering it will increase the size. It uses that knowledge to get the region size as close as possible to a given size range. Other Adaptive Operators include one that maximizes region contrast, one that adaptively finds the best polygonal representation for a region, and one that maximizes the "liny-ness", or percentage of a region boundary that is made up of line segments.

Evaluations and Meta-knowledge

Each level in this system uses evaluations to do reasoning. These evaluations judge the match between the region description of the desired region and the current candidate region in various ways. Metaknowledge in each level enables that level to make hypotheses about the image and object under consideration. These hypotheses further affect processing by indicating different operators and different location strategies.

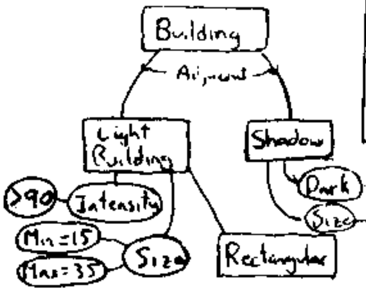
Adaptive Operators use evaluations of a single region property to change a single operator parameter. The Operator Expert uses evaluations, indications of either full or partial failure of a particular Adaptive Operator, to make hypotheses about the object. Infered hypotheses can be global, like hypothesising noise in the image, or local, like hypothesising a merging condition of the desired object with a neighboring one. These hypotheses then guide further processing by generating new constraints and indicating new operators.

Finally, partial matches at the Appearance Model Expert level can indicate high-level conditions such as object occlusion, which may lead the Appearance Model Expert to try to locate the occluding object.

An Example

This section illustrates a single example of the execution of our system, applied to a window of an aerial photograph. The partial Appearance Model used is one hypothetically derived, perhaps from knowledge of this particular site.

Appearance Model: Image:

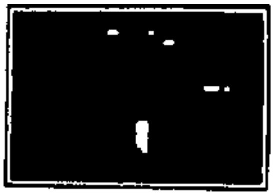


To instantiate this model, the Appearance Model Expert first attempts to locate the building by choosing the left sub-model and creating the following Region Desription:

Region Description: Light
Intensity > 100
20 < Size < 40
Rectangular:
4 corners tolerance 5.0
4 sides tolerance 5.0

This RD is given to the Operator Expert, which calls operators to try to find a region matching this description. Threshold (100) is called, creating this image:

The lower region is choosen as an initial candidate because its size is closest to that of the RD. Adaptive-Match-On-Size varies the threshold, creating the following series of images, with the threshold and size of the derived region below each figure:



Threshold = 172
Size = 11

Threshold = 100
Size = 65

Threshold = 150
Size = 14

Threshold = 131
Size = 19

This candidate is now the right size. Adaptive-Rectangle attempts to see this region as a rectangle, and it succeeds. This result is passed to the Appearance Model Expert, which now tries to find the shadow, generating the following Region Description:

Region Description: Dark
10 < Size < 20
Attention: 20, 5, 35, 20

The last attribute indicates to look in a rectangle around the located building. The Operator Expert calls Adaptive-Match-On-Size and again succeeds. The adjacency condition is checkea, ana succeeds as well. The two regions, corresponding to the located building and shadow, are shown here:

Discussion

The example above illustrates two features of our system. First, it shows as Adaptive Operator at work generating derived regions to satisfy some criterion (region size). Second, it shows the Appearance Model Expert locating an object by parts, using a constraint (adjacency of parts) to limit search for the second part. It does not illustrate a current ability of the Operator Expert: the detection of random noise and its application of an operator (local averaging) to alleviate it.

Other workers in this domain have written programs embodying some of the approachs of our program, and have achieved some good results[l, 2, 3, 4]. The analog of our Appearance Model is usually encoded as a fixed set of region properties, rather than a graph of different alternatives amenable to intelligent interpretation[l, 2]. Information from located objects is used in [1, 2] to refine a property table, which can result in a new segmentation. However, little use of partial matches is made, and none with the flexibility of our system.

The most novel feature of our approach is the use of Adaptive Operators. While the hill-climbing techniques they embody are classic, indeed, amoung the oldest in A.I. [5], their application to an I.U. domain is new, as is their use in a framework of our kind. It is important to recognize that adaptive techniques, as all other computation, must be used in conjunction with an appropriate control structure.

The framework of our program provides a flexible rule-based system that applies operators when needed and relies on the Adaptive Operators to fine tune candidate regions. Each level in our program can evaluate and use the results from the level below, and deal with partial successes in a manner appropriate to that level. This meta-knowledge enables each level, and the program as a whole, to adapt to unknown and unexpected image and object conditions in a robust way.

Conclusion

Our program is currently being run on large (256 by 256) • aerial photographs of various complexity. Further details will be presented in [6].

References

1. Kestner, W., Bohner, M., Sharf, R., and M. Sties, "Object Guided Segmentation of Aerial Images", in Proceedings of the Fifth International Conference on Pattern Recognition, pp. 529-531, 1980

2. Nago, M., Matsuyama, T., and H. Mori, "Structured Analysis of Complex Aerial Photographs", in Proceedings of the Sixth International Joint Conference on Artificial Intelligence, pp. 610-616, 1977

3. Nago, M., Matsuyama, T., and Y. Ikeda, "Region Extraction and Shape Analysis of Aerial Photographs", in Proceedings of the Fourth International Conference on Pattern Recognition, pp. 620-628, 1978

4. Nevatia, R. and K. Price, "Locating Structures in Aerial Images", in Proceedings of the Fourth International Conference on Pattern Recognition, pp. 686-690, 1978

5. Selfridge, O., Howland, B., and M. Minsky, "Hill Climbing: Some Remarks on Multiple Optimization", Information and Control, 1959

6. P. Selfridge, "Reasoning About Images: Application to Aerial Image Understanding", forthcoming PhD thesis, 1981