COMPUTER INTERPRETATION OF PEANUTS CARTOONS

Mark R. Adler
Department of Artificial Intelligence
University of Edinburgh
Edinburgh, Scotland

We describe some features of a program which interprets cartoons which are two-dimensional representations of events in the three-dimensional world. The main problems confronting the recognition procedures stem from the difficulties in describing the irregular ourved line drawings which comprise this domain:

1) Shane Descriptions Curved objects are more difficult to describe than polyhedra. The difficulty is compounded when curved objects are also irregular, and vary each time they are hand-drawn.

2) Occlusion Hidden contours cannot be found by boundary extension for the irregular curved cartoon world.

The strategy we adopt (based on ideas of Guzman [1971]) is to reduce the importance of precise shape descriptions by exploiting context informaion provided by the use of models.

Models

Individual components of our simple scenes are locally ambiguous because each shape can have several possible referents, e.g. a rectangular region may be part of a leg, a sock, or a baseball bat, etc. Models are proposed on rather gross features and once invoked provide context information to determine the interpretation. The model hierarchy we employ has two main roles:

1) To provide shape information, i.e. a description of the regions which occur in the scene at a local level -- its appearance.

2) To provide information about relationships between regions. This is a more global view. It explains how small regions fit together to form larger, more meaningful structures.

The relational, models of objects in the Peanuts world guide the analysis by applying information about how regions may be connected to represent people and objects. Sophisticated techniques are applied to handle complications arising from occlusion and variations in orientation. The shape models guide us in the matching of regions to shape descriptions. Each shape model has an associated series of feature tests (varying with the complexity of the shape) which are applied to the data defining the region. Guided by the relational context, the shape tests only need to verify the region identity rather than generate it.

The model information is structured into a hierarchy based on functional groups of regions. This deals with self-occlusion as well as with objects which have flexible joints and may appear in various attitudes. These sub-structures are represented by a variety of sub-models. Each reflects a possible region configuration accounting for differences among the characters and for the variety of orientations and body positions.

Occlusion

In our cartoon domain, it is important to discover which objects are occluded since this may alter their perceived shape to the point where they can no longer be matched to a shape model. With proof of occlusion, the system is prepared to lower its threshold for acceptance; otherwise, the current model must be rejected in favour of a more appropriate one, retaining previously computed information when possible.

A T-joint may be used as a powerful tool to confirm occlusion. T-joints may be interpreted as the place where an object s boundary disappears beneath another, or else where it reappears. The pairing heuristic finds the correct pair of such T-joints, thus providing information about region occlusion. Unfortunately, not all T-joints indicate occlusion; they also may occur when one region merely abuts another. The pairing heuristic is only applied when occlusion is the suspected cause of a region-to-model match failure. It can determine the proper pairing of T-joints to any depth of occlusion. The implemented system also has the capability of dealing with "curved" T-joints whose orientation may be ambiguous.

Experiments and Extensions

We have tested the reliability of the system by systematically distorting the presented data. By using drawings created by interpolation between two characters, we have shown that the shape models based on the presence of key features can distinguish between valid representations and "nonsense" scenes. By enlarging the heel of a shoe in stages, we observed the strong context effect of the hierarchy that forced acceptance of the distorted shape. Further experiments have led to the refinement of the control structure to use intermediate-level knowledge to overcome some problems of model instantiation. This work is seen as the first stage of a multi-scene analysis, using one scene to predict the next.

References

Adler, M.R. (1977) Forthcoming Ph.D. Thesis. University of Edinburgh.

Adler, M.R. (1976) Recognition of Peanuts cartoons. In Proceedings of AISB Summer Conference. July 1976. University of Edinburgh.

Guzman, A. (1971) Analysis of Curved Line Drawings Using Context and Global Information. In Machine Intelligence 6, eds. B. Meltzer & D. Michie. Edinburgh: University Press.