Proceedings of the Third International Conference on Autonomous Agents
Seattle, WA, May 1999

Perceptual Grouping and Attention in a Multi-Agent World

Randall W. Hill, Jr.
University of Southern California
4676 Admiralty Way, Suite 1001
Marina del Rey, CA 9092-6695
(310) 822-1511
hill@isi.edu

1. INTRODUCTION

For some time now, we have been building Soar agents that model the behavior of a military helicopter crew [2]. The crew functions performed by the pilot and gunner are simulated in a single agent, which we will hereafter refer to as the pilot agent, or pilot for short. Teams of these pilots perform tactical operations in a synthetic battlespace inhabited by thousands of other agents. A synthetic battlespace is a high-resolution, distributed simulation environment, which models combat situations at the level of the individual entity, which can be a vehicle such as a tank, a truck, an airplane, or an individual soldier. The synthetic environment includes a model of terrain, weather, daylight, time of day, dust, smoke, and craters.

Two perceptual problems arose when we deployed Soar-based pilots in this dynamic, highly populated, multiagent world. The first problem was perceptual overload. The pilot was unable to cope with a large number of entities in its field-of-view: it lacked a focus of attention. The second problem was that the outputs of the pilot's perceptual system did not provide the level of abstraction needed by its reasoning system. Percepts were resolved as individual agents, yet, some tasks require reasoning about and understanding groups of other. We addressed the latter issue by introducing perceptual grouping, whereby the pilot can perceive groups of agents as a coherent whole rather than as individuals. This provided the right level of detail to the reasoning system while also providing a way of focusing attention. Attention can still be oriented toward individuals, but groups are the primary focus.

2. PERCEPTION IN A VIRTUAL PILOT

The pilot agent is implemented in Soar, which is both an architecture for constructing intelligent systems and a unified theory of cognition [5]. The perception of other agents is integrated with reasoning and action in the Soar decision cycle. Since the synthetic battlespace is a highly dynamic environment, the pilot cannot afford to spend too much time on any given part of the decision cycle. When perceptual processing dominates reasoning and acting then the pilot flies erratically and crashes into the terrain. Without a way of focusing attention, the perceptual system slows the decision cycle to a point where the agent cannot keep up with events in the world.

The perception of agents (known as entities in the synthetic battlespace) is driven by the arrival of a stream of entity-state updates. Each update characterizes the momentary state of an entity: it provides information about the identity, location, and velocity of an entity, such as a tank. These updates are filtered through models of the helicopter's visual sensors to determine what information is potentially perceptible by the crew. Entities that are too far away will be imperceptible. Entities within the perceptible range of the model may still be rendered imperceptible if they are occluded by a terrain feature or an environmental factor such as smoke or dust. Given that the entity state information is directly available to the pilot agent, many of the standard vision problems can be finessed. For instance, understanding whether a new update refers to a known entity, or whether it represents a newly perceived object is not a problem. Each entity has a unique identifier in the simulation that can be used to associate the entity state information with visual objects in the agent's memory. Several other important problems remain, however. We address two of these problems in this paper.

3. GROUPING AND ATTENTION

3.1 Automatic Grouping

Group objects are formed by the pilot's perceptual system in one of two ways: *automatically* or *voluntarily*. Automatic grouping is initiated by the perceptual system based on proximity and force-id (i.e., friendly or foe). One at a time, agents perceived by the pilot are compared to existing groups. If an agent is within K meters of the center of mass of a group and it has the same force attribute value as the group, then it is clustered with that group. Otherwise a new group is formed.

3.2 Voluntary Grouping

The second method of grouping is voluntary: it is invoked by the virtual pilot for tasks such as tracking, where it helps to perceive a set of visual objects as a group instead of as individuals. The pilot identifies the individual agents it wishes to track and perceptually groups them with a command to the perceptual. For instance, when the task is to escort a division of transport helicopters, it is more effective for the virtual pilot to orient itself on the center of mass of a group rather than on a particular helicopter. To accomplish this the pilot visually identifies the helicopters it intends to escort and the perceptual system forms an exclusive group object for them, which can then be used for the escort task.

3.3 Automatic versus Voluntary Grouping

The need for two methods of grouping arises from the nature of the tasks in which the pilot needs to reason about groups of visual objects. Proximity-based groups are formed in an automatic fashion: the pilot does not control their formation. By tracking groups, the pilot can more easily understand a scene. While proximity-based groups are sufficient for tasks like evading and targeting, they are not sufficient for tasks like escorting. This is because the escorting task involves tracking a specific set of agents (e.g., a specific group of transport helicopters), and it is likely that this set will not be exclusively contained in one proximity-based group. In practice, two situations must be addressed. When the agents the pilot desires to track together are not located close enough to one another to be automatically clustered by proximity, they are clustered into disparate groups. The second case occurs when task-irrelevant agents are proximally mixed with the agents being tracked. These situations raised the need for a voluntary, or top-down, method of grouping. It provides the pilot with a method for forming exclusive groups of agents in service of a specific task to be performed, and such groups can be voluntarily disbanded.

3.4 Dynamic grouping

Groups are dynamic in nature: they form, split, merge, move, and change shape. In operational terms, splitting occurs when a sub-unit of vehicles maneuvers away from its group. For example, a scout platoon may break away and move ahead of the main body of a force in order to scout the front. For this reason, perceptual grouping must be dynamic. The attributes of group objects are updated every input phase. If members of a group have changed their positions, then the center of mass has to be re-computed along with all the other geometric relations. Splitting occurs when a member of a group moves beyond a distance of K + □ from the group's center of mass. K is the threshold defining the radius of the group and □ is a small increment of K used for making the split. Once an agent is split from a group, it is not orphaned for long, however. The clustering algorithm checks the orphan's distance to the center of mass of each of the other proximity-based groups—it is added to the first group within K meters. If there are no groups within K meters of the orphan, a new group is formed with the orphan as its only member, and other members may subsequently be added to it.

1.5 Perceptual attention

There is evidence that humans process perceptual information in two stages, preattentive and attentive [3], and that grouping may occur during the preattentive stage [4]. Furthermore, attention can be driven top-down and bottom-up [1]. Inspired by these hypotheses about human perception, the perceptual processing in the pilot is also divided into two stages: attentive and preattentive.

When attention is given to an agent in the perceptual field, then all of its attributes are computed, including the expensive geometric relationships. Preattentive processing only updates the perceived agent's basic state information, which involves very little computation. The pilot controls whether an agent receives attentive or preattentive processing by setting filters on the perceptual system. A filter specifies what kind of visual object should be made cognitively accessible (i.e., admitted to the pilot's working memory), providing the pilot with a way of focusing its visual attention. If a visual object matches a filter, then it receives attentive processing and it is added to or updated in the pilot's working memory. Visual objects that do not match the perceptual filter receive preattentive processing, which is cheap. The visual objects processed preattentively are maintained in perceptual memory, where they remain unless they are later selected by the filter. By using a perceptual filter, the pilot can control its visual focus of attention, eliminating a large amount of wasted computation on task-irrelevant visual objects. To pass a filter, a visual object has to match all the attribute-values in the filter. Thus, by using a filter to specify values for one or more of the attributes, the pilot can limit and choose the types of agents the perceptual system will make cognitively accessible.

4. CONCLUSIONS

Perceptual groups provide a natural way of orienting attention. Many of tasks performed by the pilot use perceptual groups as a way of orienting attention—visual objects may be filtered, based on group membership—eliminating the unnecessary perceptual processing on the agents in irrelevant groups. The results of a series of experiments showed that in cases where there were 120 agents in the field of view, the average time spent per decision cycle with attention was approximately 15 milliseconds, with peaks around 40 milliseconds. Without attention, the average time was around 50 milliseconds, with peaks above 500 milliseconds. For comparison, peaks above 100 milliseconds are often fatal for the pilot, and an average rate of 50 milliseconds is a high load, so attention makes a critical difference

Several enhancements are planned for perceptual grouping: (1) merge groups that come close together; (2) represent hierarchies of groups, so as to get more abstraction; and (3) changing the clustering algorithm to group based on proximity within a visual angle rather than on absolute distances from one another. The latter approach to clustering is more natural and will automatically provide a different grain size for the cluster based on the distance from the visual objects. Our long-term goal is to implement a theory of perception that is human-like and that integrates with the Soar architecture.

5. ACKNOWLEDGMENTS

This work was sponsored by DARPA under prime contract N61339-97-K-008, via subcontract L74158 with the University of Michigan. Approved for public release, distribution unlimited.

6. REFERENCES

[1] Egeth, H. and Yantis, S. Visual Attention: Control, Representation, and Time Course. Annual Review of Psychology, 1997, 48:269-97.

[2] Hill, R., Chen, J., Gratch, J., Rosenbloom, P. Tambe, M. "Intelligent Agents for the Synthetic Battlefield: A Company of Rotary Wing Aircraft," Proceedings of Innovative Applications of Artificial Intelligence (IAAI-97), Providence, RI, July 1997.

[3] Julesz, B. and Bergen, J. R. Textons, The Fundamental Elements in Preattentive Vision and Perception of Textures. The Bell System Technical Journal, Volume 62, No. 6, July-August 1983.

[4] Moore, C., and Egeth, H. Perception without attention: Evidence of grouping under conditions of inattention. Journal of Experimental Psychology: Human Perception and Performance, Vol. 23, No. 2, 339-352, 1997.

[5] Newell, A. Unified Theories of Cognition. Cambridge, MA: Harvard University Press, 1990.