How Long Can an Agent Look Away From a Target?

Youngjun Kim, Randall W. Hill, Jr., Jonathan Gratch
Information Science Institute
University of Southern California
4676 Admiralty Way, Marina del Rey, CA 90292
310-448-8732, 8783, 8730
{yjkim,hill,gratch}@isi.edu

Keywords:
Situation awareness, Perception, Visual attention

ABSTRACT: *Situation awareness (SA) is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future [3]. Although the impact of situation awareness and assessment on humans in complex systems is clear, no one theory for SA has been developed. A critical aspect of the SA problem is that agents must construct an overall view of a dynamically changing world using limited sensor channels. For instance, a (virtual) pilot, who visually tracks the location and direction of several vehicles that he cannot see simultaneously, must shift its visual field of view to scan the environment and to sense the situation involved. How he directs his attention, for how long, and how he efficiently reacquires targets is the central question we address in this paper. We describe the perceptual coordination that helps a virtual pilot efficiently track one or more objects. In SA, it is important for a virtual pilot having a limited visual field of view to gather more information from its environment and to choose appropriate actions to take in the environment without losing the target.*

1. Introduction and Motivation

A number of researchers have focused on creating human-like virtual agents that reside in simulated worlds. Virtual agents react to various situations that occur in simulated worlds. A critical issue arising in virtual agents is how to control visual attention. Simulation worlds usually offered all information to virtual human. However, providing all information to the virtual human not only is unrealistic but can also cause perceptual overload since extracting the pertinent information from the available sensors is challenging and results in unpredictable consequences. To be psychologically realistic, the virtual agents must construct an overall view of a dynamically changing world using limited visual fields. Given a limited visual field so that an agent receives partial information about a simulated world, the agent should have a method of controlling its visual field to grasp situations happening outside of the visual field.

1.1 Visual attention and perceptual coordination

In order to reduce information overload, we reduced the amount of input to the perceptual system by limiting the virtual human's field of view. By restricting the field of view, however, the virtual human needs a way of coordinating the tracking of multiple objects when one or more of the objects are outside the field of view. This calls for a focus of attention and a way of controlling it. For instance, if a virtual human with a limited field of view (e.g., 15 degrees) is tracking two objects, one of which is not currently in view, the agent has to shift its visual attention between the objects, and it has to do it frequently enough to remain sufficiently aware of the situation to avoid disasters (like collisions). A central issue that emerges from shifting visual attention is how long the pilot can look away from the primary target without losing track of it. When it is time to shift visual attention back to a target, there has to be a reasonable prediction of where the object will be located to ease the reacquisition. We have developed a method for predicting a target object's future position, and the amount of time the prediction is valid, given the current environmental state (e.g., terrain features) and the observed motion of the object.

1.2. Why is perceptual coordination needed?

When a visual field of view is limited, a virtual pilot is not able to completely sense the current situation. How can it overcome this perceptual limitation? Two elements are needed. The first is to shift the visual field of view, thereby sensing the current situation more accurately by getting more information from other regions. This, however, introduces the problem by which the pilot may lose track of the previously observed target. The second element needed to resolve this problem is the process of mental tracking, which predicts of a target object's future position given the environmental state and the object's status (e.g.

speed, velocity, heading). When the visual field of view is held at other regions, the virtual pilot mentally tracks the primary target so that it can subsequently reacquire the target visually. In our approach, we have focused on learning how to mentally track entities whose behavior will be influenced by key elements of the environment (i.e. maintaining a representation of an object's position when outside the visual field). Based on formations of key elements of the environment, we have made the virtual pilot predict the future position of the target. Key elements of the environment are strategically important terrain features (e.g. hill, mountain, road, river, and lake). The virtual pilot can then look away from the primary target after making a prediction of its future position and search other regions for more information.

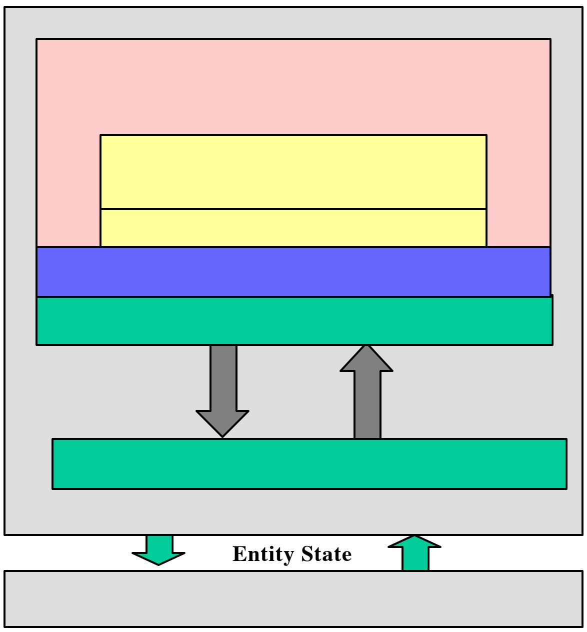


Figure 1. Virtual Pilot Architecture for visual attention

2. Modeling Perceptual Coordination in a Virtual Pilot

Our research has concentrated on developing models of a realistic virtual helicopter pilot in a distributed, interactive simulation system called ModSAF [1]. The architecture for visual attention of a virtual pilot is shown in Fig.1. To provide the interface between the simulator (ModSAF) and the virtual pilot, the perceptual and motor systems are implemented as a set of C routines. The virtual pilot can perceive the environment through a set of on-board virtual sensors that gather information about the dynamic environment. We developed a mental tracking method called *perceptual coordination* for a virtual pilot having a limited field of view to track a target, even if the virtual pilot shifts its visual attention to other regions to gather additional information. The approach we have developed for perceptual coordination consists of the following steps:

Off-line: Collect training data that is comprised of several elements – as input data, target object's information (e.g. <call-sign> <sim-time> <cell> <x> <y> <altitude-agl> <heading> <pitch> <roll> <velocity-x> <velocity-y> <velocity-z> <speed>) and key elements of the environment (e.g. hill, road, forest, lake and other environment elements), which are extracted from the target object's visual field that is divided into 27 sub-fields (3 vertical sub-fields * 9 horizontal sub-fields), and as output data, the direction of the target object - and then Produce weight data for a specific object type using a neural network (Fig 2.)

On-line: Insert the environment factors surrounding an object to neural network, then get the predicted future position of the object and the time for *looking-away* from the object.

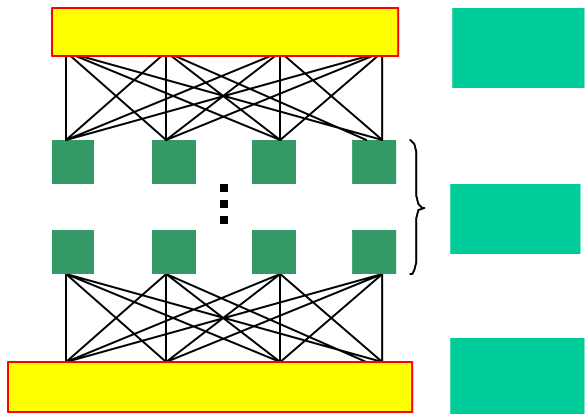
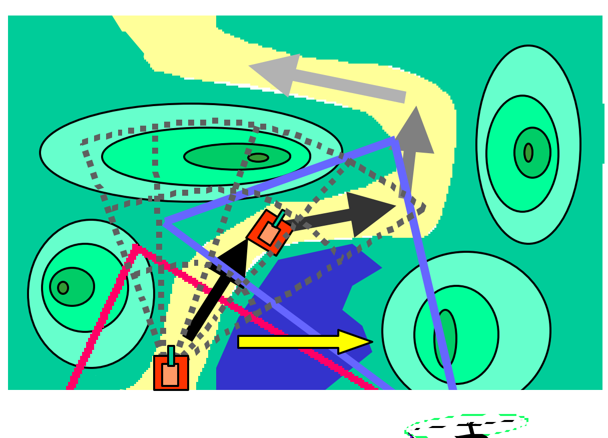


Figure 2. Virtual Pilot Neural network

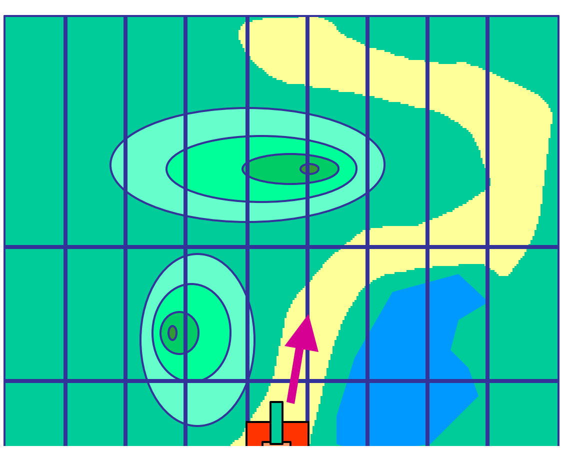
Without limiting the visual field of view of the pilot, the perceptual system is susceptible to overload by the information of external world. Given a limited field of view, the pilot has to move its visual attention in order to gather additional information, which helps the pilot more accurately reason about the current situation. To predict the amount of time the pilot look away, a neural network is applied since the expected output is how many degrees the object will turn based on the key environment elements surrounding the object and the degree of turning of the object is discrete (-35 to 35). When the virtual pilot detects an object, it checks whether it has the information about the object. If it has the information of the object, it determines how many degrees the object will turn,



then predicts the future position and how long it can look away from the object to check other regions. If the pilot does not have the information about the object, the least squares method is applied to predict the future position of the object with the object's previous movement history. The least squares method is useful for a short-time prediction (e.g. 1 or 2 secs). When the pilot predicts the future position, a trust factor is applied for decreasing the degree of uncertainty over the output of neural net with increasing time. The trust factor determines how much the virtual pilot relies on the output of the neural net, and its value is varied by the task (e.g. battle, reconnaissance). Given a number of training examples, a neural net learns how the formation of key elements of an environment surrounding a specific target object (e.g. tank) affected the movements. Given the above method, the pilot could shift its visual field of view to other regions while mentally tracking the target object.

3. Example

The virtual pilot flies a synthetic helicopter and performs tactical operations with a team of other team pilots. After detecting a tank (Fig. 3), the pilot figured out the key elements of the environment surrounding the target tank. The key elements were inserted as inputs to the neural net and the degree of turning of the target tank was produced as output. With the generated output (the degree of the turning of the tank) and the current speed of the tank, the pilot predicted the most plausible next position of the tank. In the process of predicting the future positions (e.g. each arrow), the trust factor, which is similar to uncertainty threshold [6], was applied to decrease the degree of trusting the output of the neural network. An example of predicting future position is shown in the Fig. 5. The number of network input is 196 (27 sub-visual fields * 7 key elements). Figure 4 shows an example of the sub-visual fields of a target. When the virtual pilot perceived features of the environment that were in the path of the target, it recognized that the feature affected the movement of the target and then predicted the next movement. The network outputs are the probabilities of degrees of turning (i.e. the most plausible movement (e.g. the arrow) of the target as given the formation of the environment). After getting the future position of the tank, the pilot shifted its visual field to gather more information about the environment. Whenever the pilot finished the job of checking other regions, it moved its visual attention to the changed position of the target object. The experimental results showed that the ability of predicting the future position of the target and shifting its visual attention made the virtual pilot more realistic in synthetic environments.



4. Related Work

There are a number of researchers interested in perceptual coordination in virtual agents. Chopra and Badler developed modules for gaze behavior in a virtual human with a visual attention controller. The visual attention behavior of the virtual human is not based on a realistic model of perceptual attention and cognition – the attention behaviors do not affect what the agent actually perceives, nor are the generated to service a cognitive need for information. Similarly, Rickel and Johnson developed a virtual tutor that teaches the student how to operate equipment aboard a Navy ship [4]. The virtual tutor shifts its gaze alternatively between the student and the equipment in a human-like fashion, which helps to direct the student's attention and creates a more credible tutor-student interaction. However, the difference from our work is that the tutor perceives everything regardless of the visual attention behavior. Reece and Shafer developed a robot control program that integrated perception and action selection [2]. They employed *active vision* to use optimized sensor settings, reduced field of view, and relatively simple algorithms to efficiently extract specific information from a scene. Task-based perceptual priorities and the use of visual routines drove the search for visual objects in a scene. Although they have the method of shifting visual attention with task-specific vision, there is no specific description about constraints of looking-away time (i.e. while driving, the agent always concerns about the objects in the front side.)

```
Network input is: a formation of key
elements of an environment.
1 1 0 0 0 0 0 1 1 1 0 0 0 0 1 0 1 0 0 0 0
1 0 1 0 0 0 0 1 0 1 0 0 0 0 1 0 1 0 0 0 0
```

5. Discussion and Conclusion

Visual attention is one of the most critical issues in situation awareness. Virtual agents in simulation worlds have been usually given all information about the environment. Even if all information is not given, the agents just follow task-specific movements. The agents mentioned above are more psychologically realistic - humans do not have all information about the environment in which they live. To create psychologically realistic agents, agents must construct an overall view of a dynamically changing world using limited sensor channels. The method for perceptual coordination we proposed is a way to create more realistic agents since it provides a reasonable time constraint on shifting the visual field. One of the weaknesses of our current approach is that there is not a deliberate method of describing spatial relationships between a target and key elements of an environment. In our future work, we will extend our approaches by integrating perceptual coordination with organizational and spatial structures [7]

6. References

[1] Calder, R.B., J.E. Smith, A.J.Courtemanche, J.M.F. Mar, A.Z. Ceranowicz, 1993. "ModSAF Behavior Simulation and Control," Proceedings of the Second Conference on Computer Generated Forces and Behavioral Representation, STRICOM-DMSO, July, 1993, pp. 347-356.

[2] D.Reece and S. Shafer "Control of Perceptual Attention in Robot Driving," Artificial Intelligence, 78 (1995) 461-505.

[3] Endsley M. 1988. Design and evaluation of situation awareness enhancement. In Proc of the Human Factors Society 32nd Annual Meeting, Santa Monica, CA.

[4] J. Rickel and W. Lewis Johnson. "Animated Agents for Procedural Training in Virtual Reality: Perception, Cognition, and Motor Control." To appear in Applied Artificial Intelligence.

[5] Randall W. Hill, Jr. 1999. Modeling Perceptual Attention in Virtual Humans. In Proc of 8th Conference on Computer Generated Forces and Behavioral Representation, Orlando, FL.

[6] Sonu Chopra and N.Badler. "Where to Look? Automating Visual Attending Behaviors of Virtual Human Characters," Proceedings of the Third International Conference on Autonomous Agents. May 1-5, 1999. Seattle,WA.

[7] Weixiong Zhang, Randall W. Hill, Jr. 1999. A Template-based and pattern-driven approach to situation awareness and assessment in virtual humans. To appear in the Fourth International Conference on Autonomous Agents, Agents-2000.

Author Biographies

YOUNGJUN KIM is a graduate research assistant and Ph.D. student at Information Science Institute / University of Southern California. He received B.S. from Pukyong National University,Pusan,Korea in 1997 and completed his M.S. degree in computer science at USC in 1999. His research interests are in the areas of cognitive modeling and perception.

RANDALL W. HILL, JR. is a project leader at the University of Southern California Information Sciences Institute (USC-ISI) and a research assistant professor in the computer science department at USC. He received his B.S. degree from the United States Military Academy at West

Point in 1978 and his M.S. and Ph.D. degrees in computer science from USC in 1987 and 1993, respectively. His research interests are in the areas of integrated intelligent systems, cognitive modeling, perception, and intelligent tutoring systems.

JONATHAN GRATCH is a computer scientist at University of Southern California Information Sciences Institute (USC-ISI) and a research assistant professor in the computer science department at USC. He completed his undergraduate education in computer science at the University of Texas at Austin in 1986. He received his Ph.D. in 1995 from the University of Illinois in Urbana Champaign. His research interests are in the areas of planning, learning and decision theory.