Live WWW and Its Agents: the CAS-LWWW Project

Yiming Ye[1] **John K. Tsotsos**[1] **Karen Bennet**[2]
Department of Computer Science, University of Toronto[1]
Toronto Ontario, Canada M5S 1A4
yiming@vis.toronto.edu and tsotsos@vis.toronto.edu
IBM Centre for Advanced Studies[2]
Stn. 2G, Dept. 894, 1150 Eglinton Avenue East, North York, Ontario, Canada M3C 1H7
bennet@vnet.ibm.com

Abstract

This paper[1] proposes the idea of vision agents over Internet and describes an object search agent and its communications with other agents. The goal is to reduce the traffic on the Internet for Live WWW by integrating the WWW technology, agent theory and computer vision technology together.

Vision Agents Over Internet

The WWW technology has advanced rapidly such that even live video can be integrated into the Web. But, there are some problems that make the real time transmission of the video data from the server site to the client site not appealing, one of them being the intense use of bandwidth. When the modem at the client site is slow, the real time video transmission simply will not work. One way to limit the bandwidth that is used is to lower the rate of the image transmission. But if the image rate is too low, the clients might feel inconvenient and might wait for a long time to get the wanted image. The goal of the CAS-LWWW (Centre for Advanced Studies-Live World Wide Web) project is to solve the above problems and reduce the Internet traffic by integrating computer vision technology, agent theory and WWW technology together. Instead of frequently transmitting the image data grabbed from the camera to the client and let the client to analysis the video image, we simply build agents at the server site and ask the agents to help executing the tasks required by the clients. Therefore, the image grabbed from the camera is first **pre-processed** by the agents, only those images that are of interest to the client are sent through the Internet. Thus the bandwidth used is greatly reduced. This strategy can be applied to many areas, such as security surveillance, etc. It is even more interesting when the camera need to be controlled in order to perform a certain task, such as searching for an object in the server's site. Because the client does not need to interactively control the camera and get the video data over the Internet again and again. The search agent will control the camera and perform the image processing operation at the server's site, and only the most promising images that might contain the target is transmitted to the client over the Internet.

The LWWW Object Search Agents

The LWWW object search agent architecture is an application of the agent architecture in the context of physical object search through Internet. The goal is to accommodate active vision over the Web using client server based agency. There are two agents involved, one is object search agent (OSA) which selects the camera's state parameters (field of view and viewing directions) and sends promising images over Internet, the other is camera control agent (CCA) which controls the camera's state parameters according to the signal sent by OSA and grabs images for OSA to analyze.

The OSA is a software package that assists the server to search for the target. It has the property of (A) *awareness*: it knows the search region and the target and it has a set of recognition algorithms to detect the target; (B) *reactivity and knowledge adaptation*: it has the knowledge of the influence of an action (in the form of the probability of detecting the target) and the cost of an action, it obtains the initial target distribution from the incomplete information provided by remote browser, and it updates its knowledge after each failed action using Bayes' formula; (C) *trust and proactiveness*: according to the information provided by the browser, it explicitly represents the goal of the task as maximizing the probability of detecting the target within a given time constraint and generates a practical solution; (D) *Autonomy*: it selects actions without the intervention of the browser or the server and determines the next action according to its perception of the world; (E) *Communication*: it communicates with CCA to perform the search task and it communicates with the server to know whether the browser is satisfied.

References

Ye, Yiming 1996. Sensor Planning for 3D Object Search. Ph.D. diss., Dept. of Computer Science, University of Toronto