KNOWLEDGE-BASED IMAGE ENHANCEMENT FOR COOPERATIVE TELE-ASSISTANCE

Erika Rogers, Versonya Dupont, Robin R. Murphy , and Nazir Warsi

ABSTRACT

There is an increasing need in complex environments for computerized assistance, both for the effective filtering and display of pertinent information or data, and also for the decision-making task itself. The combination of artificial intelligence techniques with image processing and graphics capabilities provides the foundation for building intelligent systems which act as mediaries between the human and the task domain. In the field of tele-assistance, this type of system enables cooperative problem-solving between a remote semi-autonomous robot and a local human supervisor. This paper describes current work on such a system, with an emphasis on the development of knowledge-based image enhancement capabilities. These allow the intelligent assistant to request particular images related to a failure state, and to automatically enhance those images in such a manner that the local supervisor may quickly and effectively make a decision.

INTRODUCTION

The purpose of this paper is to present an overview of our research on building an intelligent cooperative system for supervising a remote robot through tele-assistance. Our objective is to support the perceptual and problem-solving capabilities of humans in visual reasoning tasks by providing a mediating intelligent system or agent which can assist and enhance those capabilities.

Our original work began in the field of diagnostic radiology, and extensive human protocol studies led to the development of a cognitive model of visual interaction, which formed the basis for the first intelligent assistance program, called VIA (Visual Interaction Assistant) [8]. The blackboard-based architecture of the VIA system was chosen to facilitate the opportunistic problem-solving needed for complex and ill-structured applications, and evaluation of a preliminary prototype for radiologists, (VIA-RAD), demonstrated some promising results, both in performance and in acceptance by the users [10].

The application of this approach to the field of tele-assistance has exploited the fundamental need for more effective interaction between the local human supervisor and remote semi-autonomous robots. Advanced robotics still has a need for keeping the "human in the loop" for two main reasons: 1) to observe the evolution and state of the robot, and 2) to compensate for the information and decision-making inadequacies of the robot [1]. These robot limitations are discussed further by Pin et al:

> Both the type of tasks they can handle and the reasoning flexibility they have in performing these tasks are strictly bounded by the domain knowledge and inferencing capabilities which they have been given a priori...Although not critical under normal (i.e., as expected) operations, this fixed and static role becomes a serious drawback as soon as "off-normal" conditions appear during operation of the overall human-machine system. This occurs not only when the task or environmental conditions unexpectedly change with time, but also when a component (e.g., a sensor, an actuator) fails in the system, canceling an autonomous capability of the machine, or even when a change in human operator takes place with the new operator not knowledgeable of given procedures, roles, or capabilities which himself and/or the machine are expected to exhibit [7].

Current telesystems are challenged by problems of prohibitively high communication bandwidth and human supervisor fatigue due to repetitive tasks, poor displays, and the demands of too much data and too many simultaneous activities to monitor, especially in cases of robot failure.

Our approach treats the remote and human as computational agents possessing unique knowledge and intelligence. It relies on a third computational agent called the intelligent assistant to act as an intermediary between the human and the robot. This agent resides on the local system; it doesn't move and it doesn't perceive. Rather, it supports the perception and problem-solving capabilities of the human and the robot by selectively filtering and enhancing perceptual data obtained from the robot, as well as generating hypotheses about execution failures which cannot be solved by the remote.

The addition of such an intermediate intelligent assistant is expected to have the following advantages: 1) to improve both the speed and quality of the supervisor's problem-solving performance; 2) to reduce cognitive fatigue by managing the presentation of information; 3) to maintain low communciation bandwidths associated with semi-autonomous control by requesting only the relevant sensory data from the remote; and 4) to improve efficiency by reducing the need for supervision so that one person can control multiple robots simultaneously. Furthermore, the highly modular and adaptive nature of the systems is expected to support the incremental evolution of telesystems to full autonomy.

BACKGROUND

The approach taken in our project is to combine the autonomous perceptual and motor control abilities of the Sensor Fusion Effects (SFX) architecture for mobile robots [4] with the intelligent supervisory assistance provided by the VIA system. This work is a collaborative effort between researchers at Clark Atlanta University and Colorado School of Mines; the latter houses the mobile robot laboratory which is providing the testbed for the tele-assistance experiments.

The intelligent assistant uses a blackboard architecture to observe and manage the information posted independently by the remote and human intelligences. Blackboards have been previously used successfully for teleoperation by Edwards et al [2] in the Ground Vehicle Manager's Associate project and by Pang and Shen [6] for the high level programming and control of mobile robots to assist the operation of the emergency response team involved in a hazardous material spill. In our application of the blackboard, the remote, the supervisor, and the assistant are considered independent intelligent entities, each of which has internal routines called *knowledge sources* that read and post information to the global, asynchrononous data structure called the *blackboard*. The knowledge sources at the remote post their information about the status of the robot. TeleVIA's knowledge sources examine the status, and prepare a display of information, hypotheses and images for the local supervisor to consider. The supervisor, by definition a knowledge source, communicates with the intelligent assistant and the remote via a graphical interface managed by the assistant.

A description of the basic operation of the intelligent assistant is given in the following example. If the remote detects an anomalous situation that it cannot fix itself, it posts the nature of the alert and what progress it has made in classification and/or recovery. The intelligent assistant whose knowledge sources monitor the blackboard is alerted by this posting. The intelligent assistant responds to the alert by attempting to assess the nature of the problem, and then uses the principles of visual interaction in conjunction with task-dependent models to determine what information, sensor data, and associated levels of enhancement to display to the supervisor. The supervisor then interprets the display, posts hypotheses and may request additional information and sensor data from the remote. The intelligent assistant manages the hypotheses, reminds the supervisor of appropriate diagnostic procedures, requests sensor data from the remote, and then enhances it to highlight the attributes needed to assess the current hypotheses. The assistant must also coordinate the display of relevant contextual information such as terrain or cartographic data, imagery-related data (weather conditions, etc.), and general information (e.g., [5]).

SYSTEM OVERVIEW

A diagram of the current system design is shown in figure 1 [11]. The teleVIA blackboard is the central structure of the cooperative intelligent assistant since it is where the evolutionary results of the problem-solving effort are captured. The original logical partitioning of the blackboard into Context Panel, Perceptual Panel, Hypothesis Panel and Attention Panel was based on components of the cognitive model of visual interaction described in [8]. In the domain of tele-assistance, similar logical partitions or panels are used, but they reflect a somewhat different emphasis as described below.

In the general VIA design, the Context Panel contains information that is known about the overall problem context. For the radiology application, this

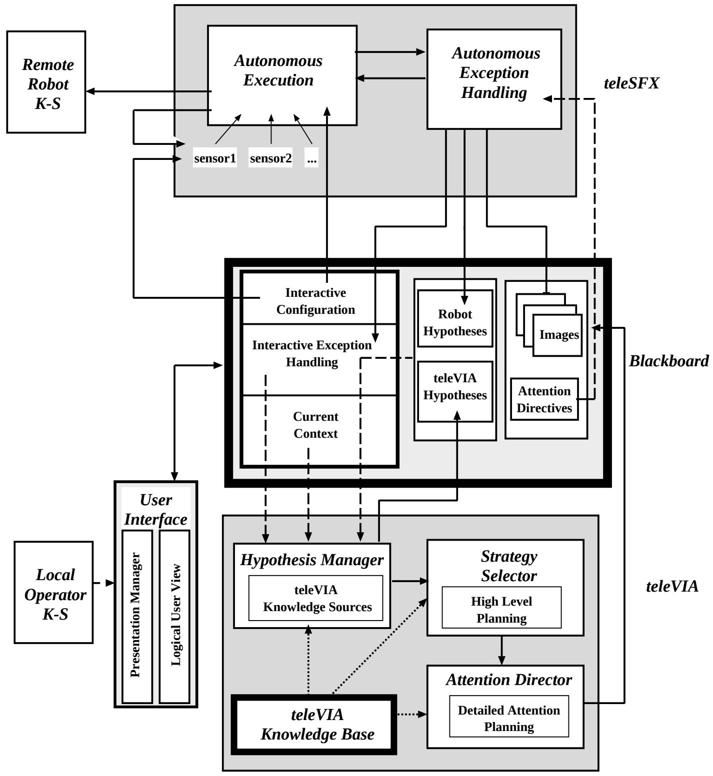

Figure 1: Cooperative Tele-Assistance System Design.

includes knowledge about anatomical landmarks that are typically expected in chest x-rays. In teleVIA, the *Current Context Panel* contains information about the current robot and its sensor configuration, the task to be performed, and the known environmental factors and conditions.

The Perceptual Panel of the VIA system was designed to capture perceptual input from the user about what is seen in the displayed image. In the VIA-RAD system, this corresponds to features which describe the anatomical landmarks and abnormal findings seen in the chest x-ray. In the tele-assistance application, this panel currently corresponds to the *Interactive Exception Handling Panel*, since it is where the perceptual status of the robot is posted when a failure is signaled. This includes the type of failure, currently active sensors, and the sensors' belief values at the time of the sensor fusion attempt.

The Hypothesis Panel contains the current hypotheses that constitute the partial (or complete) solutions that are evolving as a result of the problem-solving activity. In the radiology domain, two subpanels were needed to contain hypotheses about abnormalities in the image (findings) as well as hypotheses about disease diagnoses. In teleVIA, there are also two subpanels, reflecting hypotheses from two different problem-solving entities. The *Robot Hypotheses Panel* contains the hypotheses generated by the teleSFX system at the remote site, and reflects the diagnostic and problem-solving activities carried out autonomously by the exception handling mechanism of the robot. On the other hand, the *TeleVIA Hypotheses Panel* maintains the hypotheses generated by the knowledge sources of teleVIA. These hypotheses are based on the information posted by the remote system in combination with the more extensive knowledge retrieved from the teleVIA knowledge base. They may also be interactively changed by the human supervisor according to his/her assessment of the failure situation.

In all VIA systems, the Attention Panel is the locus of the visual focus-of-attention mechanism, and is typically divided into two parts: *Attention Directives* and *Images*. The former are issued by the teleVIA system in order to assist the local supervisor's perception of relevant data. To accomplish this, teleVIA may request particular images to be transmitted by the robot. In this way, delays due to transmission of unnecessary and/or extraneous data may be avoided. Furthermore, since the images are selected by teleVIA's knowledge sources according to the current problem, they are expected to be pertinent and useful. The directives issued to the supervisor are then aimed at guiding him/her to look at particular aspects of the data provided by the remote system. The image area is where the actual images are displayed, and may include raw data images from the robot, as well as automatically enhanced images.

Additionally, an *Interactive Configuration Panel* is provided to allow the local supervisor to select appropriate sensors, and to communicate sensing and backup plans to the robot.

The system's user interface presents two logical user views: monitor mode, which is used solely for monitoring the robot's normal behavior, and failure (or problem-solving) mode, which appears when a request for assistance is sent by the robot. In the former case, only the Current Context Panel and a restricted version of the Attention Panel are visible to the user. This allows the supervisor to request intermittent sensor images from the robot, and to track current environmental expectations. In failure mode, all of the blackboard panels become visible for full participation in the problem-solving activity.

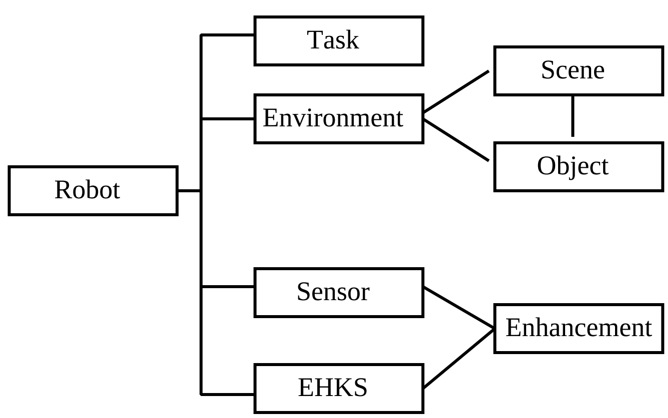

Figure 2: Knowledge Base Concept Relationships.

The basic concepts which the teleVIA system needs to know about are illustrated in figure 2, and are currently represented as frames in the system.

While most of these concepts lead to knowledge that remains static throughout the particular mission (e.g., the number and properties of the active sensors, known environment conditions, etc.), the Exception Handling Knowledge Structure (EHKS) is a dynamic entity. It comes into being when a failure is signaled by the robot, and provides the framework for the robot to transmit information related to the failure event. This knowledge structure is a modified version of the structure utilized by teleSFX in the robot's own exception handling routines. It provides not only the sensor data leading to the fusion failure, but also a history of the robot's exception handling behavior before it called the local supervisor. An example of an instantiated EHKS frame is shown in figure 3. In this example, it should be noted that not only can there be multiple hypotheses associated with a fusion failure (due to the robot's own autonomous exception handling efforts), but if the robot decides to reacquire sensor data as part of its attempt to recover, additional EHKS frames will be generated. This sequence of frames must be considered by teleVIA's problem-solving mechanism, as it provides a profile of activity since the original failure.

The raw-data slots of the EHKS frames are initially empty, since teleVIA's knowledge sources will only request selected raw image data to be transmitted. This is done by employing a perceptual strategy that is based on results from the earlier radiology study. In that work, it was shown that the radiologists demonstrated two main kinds of attention when trying to diagnose the digitized chest x-ray images: immediate visual capture occurred when an abnormality was noted as soon as the image was seen. This was characterized by a very rapid anomaly detection, even though actual identification of the abnormality may have taken much longer. Deliberate landmark search, on the other hand, reflected a systematic examination of known landmarks in the image. This occurred either in the absence of immediate visual capture, or to seek further nonobvious abnormalities [9]. In the teleVIA system, there is not a single static image presented for diagnosis. Rather, the system must first decide which images to display at all. Ideally, images would be selected which would allow the user to immediately identify the reason for failure, and therefore teleVIA's first strategy is to try to induce immediate visual capture. We are currently studying how to balance this goal with the constraints of the sensor data. For example, the ultrasonics data, which consists of a small number of real values, requires the lowest transmission bandwidth. However, a graphical representation of these numbers is needed to facilitate visual capture. On the other hand, images from the black and white video camera involve more data, but they present information that is more easily recognized without any further processing. The infrared data is also greyscale, but typically requires some processing to highlight more detailed informational features, while the color video camera produces images of the highest bandwidth. Thus one knowledge source is defined to immediately display the black and white image if available, while a competing knowledge source presents a polar plot of the ultrasonics data as shown in figure 4. Other knowledge sources examine the suspect sensor list in order to carry out further data requests and display activities. Once the data is locally available, there are a number of knowledge sources which compete to provide enhancements to that data. These are also affected by available contextual knowledge. For example, in the above ultrasonics plot, the user's decision

Exception Handling Knowledge Structure Frame	
failure-step	fusion
failure-type	high-conflict
num-bodies-of-evidence	4
body-of-evidence	
SENSOR-ID	black-&-white
EVIDENCE-FOR	0.00
EVIDENCE-AGAINST	1.00
DON'T-KNOW	0.00
RAW-DATA	
body-of-evidence	
SENSOR-ID	infra-red
EVIDENCE-FOR	0.50
EVIDENCE-AGAINST	0.00
DON'T-KNOW	0.50
RAW-DATA	
body-of-evidence	
SENSOR-ID	ultrasonics
EVIDENCE-FOR	0.24
EVIDENCE-AGAINST	0.01
DON'T-KNOW	0.75
RAW-DATA	
body-of-evidence	
SENSOR-ID	hi8-color
EVIDENCE-FOR	0.68
EVIDENCE-AGAINST	0.32
DON'T-KNOW	0.00
RAW-DATA	
suspect-sensor-list	(black-&-white hi8-color)
num-robot-hypotheses	2
robot-hypothesis	
HYPOTHESIS-NAME	sensor-malfunction
ACTION-TAKEN	(diagnostics-black-&-white diagnostics-hi8-color)
HYPOTHESIS-OUTCOME	denied
robot-hypothesis	
HYPOTHESIS-NAME	changed-environment
ACTION-TAKEN	check-modality-with-UV
HYPOTHESIS-OUTCOME	denied

Figure 3: Frame Example for Exception Handling Scenario.

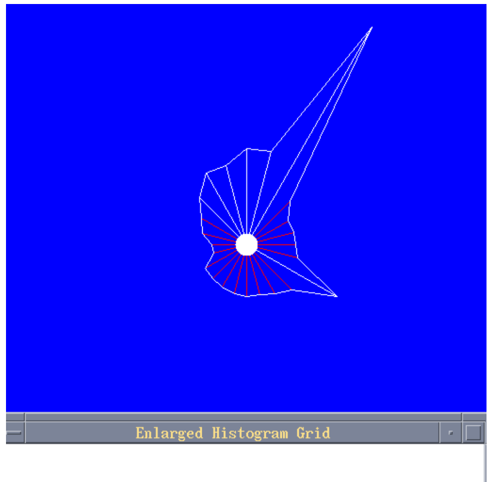

Figure 4: Polar Plot of Ultrasonics Data.

about which sensor is generating faulty data is more strongly supported if there is knowledge available about the layout of the environment. Such knowledge may be graphically presented as an overlay on the plot, thus again supporting immediate visual capture.

CONCLUSIONS

The research presented in this paper links a human supervisor, a remote robot, and an intelligent assistant in a cooperative problem-solving system which combines the human's innate ability to recognize patterns and detect anomalies with the knowledge, image and graphics processing capabilities of the computer. Work is currently underway to identify further image enhancements, and to design and implement knowledge sources which exploit those enhancements. The next stage will then be to explore problem-solving strategies which allow the system to combine its own knowledge with the perceptual feedback from the human supervisor to assist in failure classification and recovery.

Even as technology improves the autonomy of the robot, we foresee a continuing link to the human supervisor. Currently, that link is strongly motivated by diagnostic needs in failure/recovery situations, but as the robot increases its ability to handle these autonomously, the relationship is expected to move toward a cooperative exploration mode rather than just cooperative problem-solving. Systems such as Lyons and Allton's work on a planetary rover, which is designed to cooperate with a human user in the task of geological sampling and field-work already demonstrate many of these concepts [3]. The type of mediating agent discussed in this paper will still play an important role, not only by presenting the data in a meaningful fashion, but also by tagging the incoming information for "interestingness", and perhaps mining the current knowledge base for related information, again saving the human hours of tedious work, and freeing him/her for further study and discovery.

ACKNOWLEDGEMENTS

This research is currently supported by the Army Center of Excellence in Information Science at Clark Atlanta University, ARO Grant # DAAL03-92-G-0377.

References

[1] Coiffet, P. and Gravez, P., 1991. "Man-Robot Cooperation: Towards an Advanced Teleoperation Mode". In S. G. Tzafestas (ed.), *Intelligent Robotic Systems*. Marcel Dekker: New York, pp. 593-636.

[2] Edwards, G.R., Burnard, R.H., Bewley, W.L., and Bullock, B.L., 1994. "The Ground Vehicle Manager's Associate", *Tech Report AIAA-94-1249-CP*, pp. 520-526.

[3] Lyons, D.M. and Allton, J.H., 1990. "Achieving a Balance Between Autonomy and Teleoperation in Specifying Plans for a Planetary Rover". *SPIE Vol. 1387 Cooperative Intelligent Robotics in Space*, pp. 124-133.

[4] Murphy, R.R., 1993. "Robust Sensor Fusion for Teleoperations". *1993 IEEE International Conference on Robotics and Automation*, Atlanta, GA, May 2-6, 1993, vol. 2, pp. 572–577.

[5] O'Connor, R.P. and Bohling, E.H., 1991. "User interface development for semi-automated imagery exploitation". *SPIE Vol. 1472 Image Understanding and the Man-Machine Interface III*, pp. 26-37.

[6] Pang, G.K.H. and Shen, H.C., 1990. "Intelligent Control of an autonomous mobile robot in a hazardous material spill accident - a blackboard structure approach". *Robotics and Autonomous Systems 6*, pp. 351-365.

[7] Pin, F.G., Parker, L.E. and DePiero, F.W., 1992. "On the design and development of a human-robot synergistic system". *Robotics and Autonomous Systems 10*, pp. 161-184.

[8] Rogers, E., 1995. "A Cognitive Theory of Visual Interaction". In B. Chandrasekaran, J. Glasgow and N.H. Narayanan (eds.) *Diagrammatic Reasoning: Computational and Cognitive Perspectives.* AAAI/MIT Press: Menlo Park, CA.

[9] Rogers, E., 1995. "Cognitive Cooperation Through Visual Interaction". *Knowledge-Based Systems 8* Nos. 2-3, April-June 1995, pp. 117-125.

[10] Rogers, E., 1994. "VIA-RAD: A blackboard-based system for diagnostic radiology". *Artificial Intelligence in Medicine*, accepted for publication.

[11] Rogers, E. and Murphy, R.R., 1994. "Tele-Assistance for Semi-Autonomous Robots". *Proc. of AIAA Conference on Intelligent Robots in Field, Factory, Service and Space.* NASA Conference Publication 3251, pp. 500-508.