Abstract

This paper describes current work on a cooperative tele-assistance system for semi-autonomous robots. This system combines a robot architecture for limited autonomous perceptual and motor control with a knowledge-based operator assistant which provides strategic selection and enhancement of relevant data. The design of the system is presented, together with a number of exception-handling scenarios that were constructed as a result of experiments with actual sensor data collected from two mobile robots.

Introduction

Traditional telerobotics research focuses on the separate contributions of the human, computer and robot in the performance of a particular task. Historically a single human at a local system is dedicated to operating a single remote robot. However, continuous supervision may be impractical for applications where the communication bandwidth acts as a bottleneck, and/or transmission time delays make remote high-dexterity control difficult or impossible. Further, the operator may not be reliable due to fatigue, while increased environmental complexity, use of multiple sensing modalities, or the addition of more robots all may contribute to the cognitive overload of the operator.

One approach to these problems is to increase robotic autonomy through the addition of intelligent capabilities. Control schemes, such as shared control and supervisory control, reduce both the amount of communication between local and remote, and the demands on the operator by increasing the autonomy of the remote. However, there is still a need for human problem-solving capabilities, particularly to configure the remote for new tasks and to respond to unanticipated situations. Thus the teleoperation community is becoming increasingly interested in computerized assistance, both for the effective filtering and display of pertinent information or data, and also for the decision-making task itself.

The purpose of this paper is to describe current work on a cooperative tele-assistance system which combines the autonomous perceptual and motor control abilities of the Sensor Fusion Effects (SFX) architecture [4] with the intelligent operator assistance provided by the Visual Interaction Assistance (VIA) system [5]. In this approach, the remote system consists of a robot with sufficient computerized intelligence to function autonomously under a particular set of conditions, while the local system is a cooperative decision-making unit that combines both human and machine intelligence. The computerized aspect of each system enables communication to take place in a common mode, and in a common language.

The paper presents an overview of the design of the system itself, and then discusses a number of experiments and scenarios using data from two mobile robots. The scenarios demonstrate how the local tele-VIA system would assist an operator to respond to a sensing problem which could not be resolved by the exception handling mechanism of the remote robot.

Background and Related Work

Coiffet and Gravez [2] suggest that “instead of searching for an overall model including complex concepts as human behavior, it is more profitable to consider the computer role as performing assistance functions to humans”. They go on to describe three situations in which conceivably computer and human can collaborate, and compare the differences between telepresence, “a system-oriented assistance centered on machine transparency”, and teleassistance, which refers to “task-oriented assistance”. They state, further, that:

Increasing the capabilities, and therefore the complexity, of a teleoperation system soon results in a cognitive overload for the human operator. The design of a cooperative system should thus introduce strategic assistance forms that facilitate on-line symbolic
control. Generally speaking, the basis for such assistance is the selection and processing of relevant data (sensor outputs and execution reports) and the filtering of operator commands. This is reflected in a high-level structural dialogue embodying task-oriented diagnosis and proposing pertinent solutions. At the symbolic level, the computer contribution is deduced from an understanding of the work at hand, and is intended to enhance the human operator's decision-making process and to improve on-line human-machine communication.

The work presented in the remainder of this paper is in keeping with the philosophy behind these comments, and demonstrates how such a cooperative system may be designed and implemented.

**Approach**

In remote and unexplored environments where unanticipated events are likely to occur, it is important that the robot be able to handle a number of situations autonomously and in real-time. In the event of an unresolved problem, however, the remote robot must have recourse to a local human operator for assistance. However, in order for the operator to perform diagnosis effectively, the data from the robot must be appropriately selected and presented. Once the problem has been determined, the operator should then be able to re-configure the robot so that the autonomous processes may once again take over. An important consequence of this approach is that it is now conceivable for the operator to supervise more than one robot simultaneously, and furthermore, to request data from other robots working with the one in trouble.

To achieve this goal of cooperative tele-assistance, two major software systems have been joined together and modified appropriately for this application domain. The first is the Sensor Fusion Effects (SFX) architecture [4], which utilizes state-based sensor fusion to support the motor behavior of an autonomous robot. If a state failure occurs, fusion is suspended and control is passed to an exception-handling mechanism, which attempts to identify the problem and either repair or replace the sensing plan. The second system, called VIA (Visual Interaction Assistant), is designed to cooperatively assist human perception and problem-solving in a diagnostic visual reasoning task. VIA is a blackboard-style system which utilizes knowledge-based techniques to focus the user's attention on relevant parts of the image, automatically enhancing the image according to the needs of the user's problem-solving process. It further manages diagnostic hypotheses, maintaining beliefs according to current evidence, and assists the user to converge opportunistically on a solution where possible.

The advantage of linking the SFX system with the VIA paradigm is that under SFX, the robot has already attempted a certain amount of trouble-shooting itself. Thus information about what has been tried, the robot's own conclusions, and the relevant sensor images can all contribute to the decision-making process of the local operator. In order to achieve this, the teleSFX system includes an interactive exception handling component, which allows the robot to call the operator for help in the event that its own exception handling capabilities could not resolve the problem.

The central communication mechanism between the remote and local intelligent systems is the blackboard structure. The advantage of this architecture is that it allows asynchronous communication between a number of independent knowledge sources. In the general VIA system, the user is considered to be a knowledge source as well, cooperating with the knowledge-based system in the search for a solution (or partial solution). In the case of the tele-assistance system (teleVIA), the remote robotic system is also incorporated as a knowledge source, thus allowing the three entities, robot, teleVIA system, and human operator, to make contributions to the solution of the problem in a cooperative manner. This is especially important in cases where the solution set is not initially tractable, and more information must be acquired in order to quickly constrain the list of diagnostic hypotheses, so that a repair plan may be constructed. An overview of the entire system is shown in Figure 1, and further details are provided in the following subsections. In this diagram, it can be seen how the interactive configuration and interactive exception handling components of the teleSFX architecture are merged with the intelligent assistance provided by teleVIA, through the panels of the blackboard. The emphasis in this paper is on the interactive exception handling aspects of this design.

**TeleSFX**

In [3], the teleSFX control scheme was introduced, emphasizing the intelligent exception handling mechanism at the remote. Unlike configuration, exception handling must be done in real-time (for example, a robot may be moving when a sensor malfunctions). As shown in [1], autonomous exception handling is difficult because it involves domain and hardware specific information which may not always be available or correct.

TeleSFX uses a three part strategy for exception handling: detection, classification, and recovery. The first step, detection, determines that a "sensing failure" has occurred. Sensing failures are any anomalous or suspect conditions that have been previously defined by the knowledge engineer. Sensor malfunctions are one type of failure. Most sensor malfunctions manifest themselves via explicit hardware errors communicated to the controlling process (e.g., bus errors, frame grabber errors) and tend to be straightforward to classify and recover from (e.g., reset the system, request a retry). Another class of sensing failures is due to unanticipated changes in the sensing environment which degrade the performance of one or more sensors (e.g., the lights are turned off). The third and final class of failures stem from errant expectations, where the robot is interpreting the observations according to a model. If for some reason the robot has selected the wrong model at the wrong time (e.g., for mechanical reasons, the robot did not rotate fully to
the intended viewpoint), the sensor observations are unlikely to agree.

Failures in the latter two classes are difficult to detect because the sensors are operating "correctly" but their data can no longer be interpreted without accounting for the changed context. Therefore teleSFX is sensitive to inconsistencies in the evidence contributed by different sensors for a particular task. The knowledge engineer defines a set of failure conditions representing these inconsistencies for the particular implementation. Each perceptual process may have a different set of thresholds for those failure conditions, given the unique interactions between sensors.

The classification step has the remote attempt to autonomously identify a sensing failure, and adapt the sensing configuration. This involves hypothesis generation, testing and response heuristics at the remote site, and several experiments have been described in [1] which demonstrate this capability. However, the success of the classification step depends on the expert understanding of the domain and the sensors. This domain-dependence means that classification by the remote is brittle and will not always be successful. Therefore, if the remote system cannot resolve the difficulty, teleSFX must post the request for help to the blackboard, together with immediately relevant data such as current sensor data and a log of the remote's hypothesis analysis.

Figure 2 shows the details of the control system for the remote site. The local operator is involved primarily in interactive configuration, and general monitoring, until the interactive exception handling is triggered by the remote system. At that point, teleVIA takes over from teleSFX until the repair is communicated.

**Blackboard**

The Blackboard is where the evolutionary results of the problem-solving effort are captured. The original logical partitioning of the blackboard was based on components of a cognitive model of visual interaction described in [5], and was designed to facilitate transfer of information between human perception and problem-solving during a visual reasoning task. In the domain of tele-assistance, it is seen that, with one exception, the same logical partitions or panels may be used. The additional information which is contributed by the remote robotic system is accommodated in the subpanels as shown in Figure 3.

**Context Panel**

In the general VIA design, this area contains information that is known about the overall problem context. In the teleVIA mode, the Context Panel is used to monitor the robot's (or robots') current activities. It is divided conceptually into three subpanels:
Figure 2: Overview of TeleSFX.

Figure 3: Tele-VIA Blackboard.
1. **Interactive Configuration** allows the local operator to select appropriate sensors, and to communicate sensing and backup plans to the robot.

2. **Interactive Exception Handling** receives the signal for help when autonomous exception handling fails. The remote system immediately posts the type of failure, currently active sensors, and the belief table for those sensors. This tells the local operator what the perceptual status of the robot is at the time of failure, and provides initial information for teleVIA to begin formulating hypotheses, and requesting further information.

3. **Current Context** is a panel which is active during both interactive configuration and interactive exception handling. It contains information about the task underway, the known environmental factors and conditions, which sensors are active and working, and intermittent video images from the robot reinforcing the operator's knowledge of the context within which the robot is currently functioning.

In the initial design of teleVIA, the overall Context Panel is used simply as an informational tool for monitoring the robot.

**Hypothesis Panel**

This panel contains the current hypotheses that constitute the partial (or complete) solutions that are evolving as a result of the problem-solving activity. It is divided into two subpanels:

1. **Robot Hypotheses** 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.

2. **TeleVIA Hypotheses** contains the hypotheses generated by the knowledge sources of teleVIA, based on the information posted by the remote system in combination with more extensive knowledge retrieved from the teleVIA knowledge base.

**Attention Panel**

This panel is the locus of the visual focus-of-attention mechanism. It is also partitioned into two parts:

1. **Attention Directives** are issued by the teleVIA system in order to assist the local operator'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 more likely to be pertinent and useful. The directives issued to the operator are then aimed at guiding him/her to look at particular aspects of the data provided by the remote system.

2. The second area of the Attention Panel consists of one or more images, obtained from the robot by the teleVIA system. Depending on the sensory modality of the displayed images and/or data (e.g., video vs. infra-red vs. ultrasonics), teleVIA will also automatically execute appropriate image enhancements designed to facilitate the operator's perception of the feature(s) in question. In this manner, the superior perceptual capabilities of the human operator can be exploited in order to diagnose the problem more quickly.

**TeleVIA**

In Figure 4 are shown the components of the cooperative system which assists the human supervisory activities at the local site. TeleVIA contains four main control modules: Hypothesis Manager, Strategy Selector, Attention Director and User Interface, together with a knowledge base which serves as the repository of long-term information in the system. The Hypothesis Manager impacts the blackboard through the activities of hypothesis-related knowledge sources. The Strategy Selector is used to pass control from the Hypothesis Manager to the Attention Director, since the way in which attention is focused may depend on the strategy used for reducing the list of active hypotheses. The Attention Director is concerned with focusing attention by presenting and enhancing images as well as suggestions to the operator of what to look at next. The User Interface is the component through which the human operator communicates with the teleVIA system.

**Hypothesis Manager**

The purpose of the Hypothesis Manager is to percolate information through the levels of the blackboard via the activities of the knowledge sources. Each knowledge source has a set of preconditions that must be satisfied by information at a particular level of the blackboard. It then performs a transformation of the information at one or more levels. Some examples of knowledge sources which are activated by the type of sensor involved in the failure are illustrated in the following tables.

```
<table>
<thead>
<tr>
<th>K-S 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Precondition:</strong></td>
</tr>
<tr>
<td>The sensor involved is video.</td>
</tr>
<tr>
<td><strong>Action(s):</strong></td>
</tr>
<tr>
<td>Request latest image from robot.</td>
</tr>
<tr>
<td>Post image.</td>
</tr>
<tr>
<td>Retrieve and post video hypotheses.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>K-S 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Precondition:</strong></td>
</tr>
<tr>
<td>The sensor involved is infra-red.</td>
</tr>
<tr>
<td><strong>Action(s):</strong></td>
</tr>
<tr>
<td>Request latest image from robot.</td>
</tr>
<tr>
<td>Post image.</td>
</tr>
<tr>
<td>Retrieve and post infra-red hypotheses.</td>
</tr>
</tbody>
</table>
```
Figure 4: Overview of TeleVIA.

**K-S 3**

*Precondition:*
The sensor involved is infra-red and image is posted.

*Action(s):*
Retrieve model from current context.
if model = available then
    invoke model-false-color enhancement
else invoke generic-false-color enhancement.
Post image.

**Strategy Selector**

This module is invoked by the Hypothesis Manager when a knowledge source needs further information before proceeding. It examines the current blackboard configuration in order to determine an appropriate strategy for the next step in the problem-solving process. A High Level Plan is then generated to carry out the selected strategy, and is passed to the Attention Director for refinement and execution.

**Attention Director**

The Attention Director module takes the High Level Plan produced by the Strategy Selector, and constructs an Attention Plan that contains detailed instructions for focusing attention. The steps of the Attention Plan are based on the particular type of evidence that is needed to fulfill the mandate of the Strategy Selector. These steps are expanded with image enhancement procedures where appropriate, and are executed. Control is then passed to the operator for feedback. In this way, the system presents information, makes suggestions, and enhances the image(s) in such a way as to influence the direction of the operator’s problem-solving.

**User Interface**

The User Interface is divided into two parts: the *Logical User View*, which controls how much of the blackboard is visible to the user, and the *Presentation Manager*, which controls the form of the interface as it is presented to the user. The Logical User View component of the user interface allows the system to be adapted for various purposes without compromising its basic problem-solving approach. For example, when the operator is simply monitoring the robot, and performing interactive configuration, the panels involved in exception handling should be hidden from view. There may also be a certain amount of data posted to the blackboard, which is utilized by teleVIA in its hypothesis management, but which should not necessarily be visible to the operator. On the other hand, the Presentation Manager provides the actual human-machine interface of the system through a displayed representation of the Logical User View. This may take a number of forms including menus, icons, graphics, and/or direct manipulation windows, and may even extend to audio as well as visual mechanisms.

**Experiments**

The experiments described here use data for scene recognition which was collected from two sources. Scenarios 1, 2, 3 and 5 are based on sensor observations collected from the Denning DRV mobile robot, George, at the Georgia Institute of Technology. Scenario 4 is based on sensor data from Clementine, the Colorado School of Mines’ Denning MRV-3 mobile robot. It should be noted that some of the data has been used in previous experiments. The exception handling activities at the remote in the experiments described in this paper differ from previous work [3]
because they make use of the more rigorous system developed in [1].

Five different types of sensors (an InfraMetrics true infrared camera, a black and white video camera, a Hi8 color camcorder, a UV camera and ultrasonics) provided observations from George. Clementine supplied data sets from three sensors (a black and white video camera, a color camcorder, and ultrasonics). Both robots simulated security guards, where the task was to determine whether a student desk area of a cluttered room had changed since the last visit. In the following scenarios the focus is on the activities of the teleVIA system in response to the request for help from the remote system.

Scenario 1

In the first experiment, the robot collected data for the desk scene while facing a different part of the room. This resulted in a “high conflict” type of state failure during fusion. The robot then generated a hypothesis of sensor malfunction, and attempted to run diagnostics on the two conflicting sensors. These diagnostics, however, showed that both sensors were working correctly, and at this point, the robot could not proceed further, and signalled for assistance.

At this point, the robot posts a request to the interactive exception handling panel of the blackboard, indicating the type of failure it has encountered, and including the beliefs which led to this failed fusion. In the initial version of teleVIA, the images leading to this conflict are also transmitted. The first knowledge source of teleVIA is activated with the precondition that a video camera is involved in the failure. This causes the video image to be displayed first, together with a list of preliminary hypotheses of what could be the problem. Examples of hypotheses include: wrong input, sensor malfunction, sensor occlusion, sensor hardware error (missing data, self-diagnostic error), multiple sensor errors and electromagnetic interference. Since the purpose of the system is to provide assistance as quickly as possible, an assumption is made that, if applicable, images which are most easily perceived by humans (e.g., video images versus thermal images) are given priority. Thus, if by looking at the video image, the operator can immediately ascertain what the problem is and resolve it, then this most effective solution to the problem should be supported. Once the operator selects the probable diagnosis, a list of repair possibilities may be posted to the interactive configuration panel for implementation.

Scenario 2

In the second experiment, the lens of the camera was covered in opaque plastic to simulate a sensor malfunction due to external factors such as dirt on the lens, for example. In this case, the fusion process resulted in a “below minimum” type of failure, and, as a result, the exception handling mechanism generated a first hypothesis of “inadequate sensing plan”. A backup plan was then implemented, and sensor data was reacquired accordingly. The new plan added a color camera to the sensing suite, and subsequently a fusion failure of “high conflict” was encountered between the black and white camera and the color camera. As in the previous experiment, teleSFX then generated the hypothesis of sensor malfunction, performed diagnostics which denied this hypothesis, and then called for assistance.

In this case, both of the failures are posted to the blackboard, together with the beliefs generated for each attempt. Once again, the primary troubled sensor is the black and white camera, and the image generated is shown in Figure 5. In this case, however, since the second attempt introduced the conflicting image from the color camera, both the black and white, and the color video images are displayed by teleVIA for the operator to examine first. In this case as well, the operator should be able to determine the problem fairly quickly by simply comparing the black and white video image with that of the color camera.

Scenario 3

In the third experiment, the lights were turned off during data collection to simulate an unforeseen change in environment. In this case the exception handling mechanism of the robot arrived at a correct conclusion of “environmental change” by testing the visible light information. However, for this type of problem, operator assistance is still needed for recovery, and therefore a message is posted to the interactive configuration portion of the blackboard requesting intervention. The beliefs leading to the original state failure, together with the hypotheses generated and tested by the robot, are posted to the blackboard, while images and data from the relevant sensors (black and white camera, and UV sensor) are also displayed. This enables the operator to determine what type of environmental change may have occurred.

In each of these experiments, the primary sensor involved in the problem was the black and white camera. Since these experiments were originally designed to test the autonomous exception handling capabilities of the teleSFX system, the results, when extended to the teleVIA component are somewhat artificial. However, they serve the purpose to establish the type of
information which must be communicated between the remote and the local systems in even such elementary scenarios. This allows us to determine the types of knowledge sources which may be activated, the different types of hypotheses which may be needed, and how to present this information effectively using the blackboard mechanism.

Further experiments are underway which emphasize sensor data which is not as easily perceived by the human operator, and which may require enhancement before conclusions may be drawn. In these cases, teleVIA knowledge sources are activated according to the type of sensor(s) involved in the state failure. This is then combined with knowledge of the current context to select appropriate enhancements, and display the information to the local operator. The following scenarios were constructed using images acquired by the robot for a drill press scene.

**Scenario 4**

In this example, it is assumed that the ultrasonics are contributing primarily to the fusion failure. In this experiment, one out of the 24 ultrasonics transducers mounted in a ring began to report widely fluctuating readings. A sensing failure of "highly uncertain" evidence was reported, but the responsible sensor could not be isolated, thereby necessitating aid from the operator. The raw ultrasonic readings that come from a Denning mobile robot are just numbers, which represent measurements in feet. However, when this data is represented as a polar plot as in Figure 6, it is much easier to notice if one or more of the sensors is giving erroneous readings. This is further reinforced if the numerical data are examined in the light of knowledge about the current context, for example, that a room (or mine shaft) is thought to have certain dimensions. A further enhancement of the data which can aid the local operator is an occupancy grid, which presents a bird's eye view of what the robot has sensed so far. The robot builds up this grid or map as it processes ultrasonics data. With both of these types of displays, the operator is more likely to diagnose the failure of an ultrasonic transducer or board, or to detect an erroneous reading.

**Scenario 5**

When the sensor in question during exception handling is the infra-red camera, enhancements are once again needed to assist the operator's perception of the information in the image. In this case, the untouched true infra-red image is typically gray scale, and there is often not a great deal of discernable contrast in the image. It is common practise to add false color to such an image to show the heat distribution. However, certain choices of false color maps still do not enhance the image, and may obscure the details even further. In the drill-press example, dividing the grayscale into 8 equal bands of color leads to a primarily yellow image, due to the extreme heat of the drill press. A grey scale version of this image is shown in Figure 7. However, if the selection of false color map is knowledge-based, utilizing model-specific information about the drill press, for example, then a more appropriately enhanced image is produced, making it easier for the operator to see the heat profile represented as blue, green, red, yellow, and white bands. This is illustrated in the grey scale rendition in Figure 8.

**Future Work**

Current work is concentrating on constructing experiments in real-time where an operator at Clark Atlanta will interact with the remote robot (Clementine) at Colorado School of Mines. An important issue which has not been addressed in this work so far is that of learning. The robot will typically be working in hazardous and/or remote environments about which little may be known, and therefore it is difficult to anticipate the types of problems which may arise. Not only would it be desirable to increase the autonomy of the individual robot wherever possible, but the knowledge gained from solving these problems could be disseminated to other robots in the field. Further-
more, if the teleVIA system could "remember" certain interactions, these could immediately be retrieved from memory, rather than having to generate the same session over again. The technique of case-based reasoning is a natural candidate for this type of learning. Each interactive exception handling session may be captured as a case, which would be indexed on features such as particular configurations of sensors and failure types. Such a case could also include relevant images, or at least image types and enhancements used, so that teleVIA would simply use a case retrieval mechanism rather than a potentially complicated reasoning strategy. Certain aspects of the exception handling and recovery procedures might also then be communicated to the robot itself, to extend its autonomous capabilities, especially for recurrent problems.

Summary and Conclusions

This paper presents a new approach in semi-autonomous mobile robots, which reduces the level of human supervision and provides intelligent assistance for problem solving. The approach partitions problem solving responsibilities between the remote and the local machines. The remote system monitors its sensing for anomalies, called sensing failures, using teleSFX. If a failure occurs, it attempts to classify the source of the problem using a generate and test methodology. If it is successful in identifying the source, it then attempts to autonomously recover (e.g., go to back up sensors, change parameters). Otherwise, if the source cannot be classified, or if no recovery strategy is available, the local machine must provide the exception handling. Exception handling at the local is done by the operator, with the help of teleVIA. TeleVIA uses a common blackboard to cooperatively assist the operator by posting what has been done by the remote, displaying and enhancing sensor data needed in ascertaining the problem, and managing diagnostic hypotheses and beliefs. Experimental scenarios using data collected from mobile robots illustrates the operation of the system.

The advantages of this type of tele-assistance fall into three categories. First, it is practical. It reduces the need for direct human supervision and communications by having the remote monitor itself for failures. Second, it increases efficiency by freeing the operator to supervise multiple remotes, reducing cognitive overload by controlling the presentation and enhancement of sensory data from the remote, and aiding the problem-solving process through hypothesis management and expert guidance. Three, the approach supports the incremental addition of artificial intelligence as more progress is made in learning and planning.

References


