Effects of Automation on Situation Awareness in Controlling Robot Teams

Michael Lewis
School of Information Sciences
University of Pittsburgh
Pittsburgh, PA 15208, USA
ml@sis.pitt.edu

Katia Sycara
Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15208, USA
katia@cs.cmu.edu

Abstract—Declines in situation awareness (SA) often accompany automation. Some of these effects have been characterized as out-of-the-loop, complacency, and automation bias. Increasing autonomy in multi-robot control might be expected to produce similar declines in operators’ SA. In this paper we review a series of experiments in which automation is introduced in controlling robot teams. Automating path planning at a foraging task improved both target detection and localization which is closely tied to SA. Timing data, however, suggested small declines in SA for robot location and pose. Automation of image acquisition, by contrast, led to poorer localization. Findings are discussed and alternative explanations involving shifts in strategy proposed.

Keywords—human-robot interaction; robot teams; situation awareness

I. INTRODUCTION

The relation between situation awareness, performance and automation has been the subject of active investigation for more than forty years. Early researchers were concerned that isolating operators from the processes they controlled through automation might lead to deskilling [1], losing the psychomotor and cognitive skills needed to perform the tasks manually. Others have been concerned that losing touch with the moment to moment state of the process might lead to inattentiveness [2] or unfamiliarity [3] with what is going on.

Other less chronic effects are associated with the introduction of automation. The term, out-of-the-loop, has been used by [4] to describe a variety of such effects. Operators who are out-of-the-loop have been shown to have decreased situation awareness (SA) as evidenced by direct measures such a SAGAT [5] or SART [6] and indirectly through declines in performance [7]. Performance declines most often involve failures in monitoring and can arise from either undetected process anomalies [8] or failures of the automation itself [9]. The first case has been referred to as complacency [10] and occurs when an operator’s trust in automation exceeds its trustworthiness. As a consequence, the operator fails to monitor the automation as intently as needed and does not detect process anomalies or automation failures as they occur. This may result either due to sampling displays too infrequently [11] or attentional blindness in which they fail to notice the fault despite observing the evidence.

A related effect, automation bias [10], has been used to refer to a tendency to accept recommendations or actions of automation despite contrary evidence. While in complacency, the operator fails to notice, in automation bias, the operator may notice but still fail to act giving automation the benefit of the doubt. A common finding has been that users of automation have particular difficulty in dealing with novel failures [12]. Out-of-the-loop operators may fail to notice the failure [4] or noticing it, fail to properly diagnose, or diagnosing properly fail to take a correct action. These later failures may be due to either unfamiliarity with the system’s expected states or unfamiliarity with available manual actions and their effects. There are many exceptions, however, and it is not uncommon to find better performance in novel situations when automation is well designed [13].

We are conducting a series of studies to develop theories and techniques to allow individual or groups of human operators to control a larger number of robots. Controlling multiple robots substantially increases the complexity of the operator’s task because attention must be shared among robots in order to maintain situation awareness (SA) and exert control. In the simplest case, an operator controls multiple independent robots interacting with each as needed. A foraging task [14] in which each robot searches its own region would be of this category although minimal coordination might be required to avoid overlaps and prevent gaps in coverage especially if robots are in close proximity. Control performance at such tasks can be characterized by the average demand of each robot on human attention [15]. Because robots are operated independently an additional robot imposes only an additive
demand on cognitive resources. Under these conditions increasing autonomy for individual robots should allow them to be neglected for longer periods of time making it possible for a single operator to control more robots. Because robots are controlled independently, operators and robots can be added to increase system performance. For more strongly cooperative tasks individual autonomy is unlikely to suffice. The round-robin control strategy used for controlling individual robots would force an operator to plan and predict actions needed for multiple joint activities and be highly susceptible to errors in prediction, synchronization or execution.

In a recent series of experiments, we have investigated both increasing autonomy and number of operators to increase the size of robot teams that might be controlled. In an initial experiment we identified path planning as the most fruitful task to automate. In subsequent experiments we tried increasing the number of operators, alarming navigation failures, and automating the acquisition of imagery. Although absolute performance generally improved with increasing automation, our data suggest these improvements were accompanied by less dramatic decreases in SA as predicted.

A. USARSim&MrCS

The experiments reported in this paper were conducted using the USARSim robotic simulation with simulated Pioneer P2-AT robots performing Urban Search and Rescue (USAR) foraging tasks. USARSim is a high-fidelity simulation of urban search and rescue (USAR) robots and environments developed as a research tool for the study of human-robot interaction (HRI) and multi-robot coordination. USARSim supports HRI by accurately rendering user interface elements (particularly camera video), accurately representing robot automation and behavior, and accurately representing the remote environment that links the operator’s awareness with the robot’s behaviors. USARSim uses Epic Games’ UnrealEngine2 [16] to provide a high fidelity simulator at low cost and also serves as the basis for the Virtual Robots Competition of the RoboCup Rescue League. Other sensors including sonar and audio are also accurately modeled.

Validation data showing close agreement in detection of walls and associated Hough transforms for a simulated Hokuyo laser range finder are described in [17]. The current UnrealEngine2 integrates MathEngine’s Karma physics engine [18] to support high fidelity rigid body simulation. Validation studies showing close agreement in behavior between USARSim models and real robots being modeled are reported in [19,20,21,22] as well as agreement for a variety of feature extraction techniques between USARSim images and camera video reported in Carpin et al. [23].

MrCS (Multi-robot Control System), a multi-robot communications and control infrastructure with accompanying user interface(s), developed for experiments in multirobot control and RoboCup competition [24] was used in these experiments. MrCS provides facilities for starting and controlling robots in the simulation, displaying multiple camera and laser output, and supporting inter-robot communication through Machinetta, a distributed multi-agent coordination infrastructure. Figure 1 shows the elements of the baseline version of MrCS. The operator selects the robot to be controlled from the colored thumbnails at the top of the screen. To view more of the selected scene shown in the large video window the operator uses pan/tilt sliders to control the camera. The current locations and paths of the robots are shown on the Map Data Viewer (bottom right). The map is developed over the course of the run by fusing results of successive laser scans from the team of robots. Robots are tasked by assigning waypoints on one of the maps or through a teleoperation widget (upper right). The victim marking task requires the operator to first identify a victim from a video window, identify the robot and its pose on the map, and then locate corresponding features near the victim’s location in order to mark it on the laser generated map.

B. Human vs. Automated Path Planning

The foraging task can be decomposed into exploration and perceptual search subtasks corresponding to navigation of an unknown space and searching for targets by inspecting and controlling onboard cameras. An initial study [25] investigated the scaling of performance with number of robots for operators performing either the full task or only one of the subtasks to identify limiting factors. If either of the subtasks scaled at approximately the same rate as the full task while the other scaled more rapidly, the first subtask could be considered the factor limiting performance. If performance were equal across the conditions or the subtasks scaled at the same rate, automation decisions would need to be based on other factors. The logic of our approach depends upon the equivalence among these three conditions.

In the fulltask condition, operators used waypoint control to explore an office like environment. When victims were detected using the onboard cameras the robot was stopped and the operator marked the victim on the map and returned to exploration. Equating the exploration subtask was relatively straightforward. Operators were given the instruction to explore as large an area as possible with coverage judged by the extent of the laser rangefinder generated map. Because operators in the exploration condition did not need to pause to locate and mark victims the areas they explored should be slightly greater than in the fulltask condition.
Developing an equivalent perceptual search condition is more complicated. The operator’s task resembles that of the payload operator for a UAV or a passenger in a train, in that she has no control over the platform’s trajectory but can only pan and tilt the cameras to find targets. The targets the operator has an opportunity to acquire, however, depend on the trajectories taken by the robots. If an autonomous path planner is used, robots will explore continuously likely covering a wider area than when operated by a human (where pauses typically occur upon arrival at a waypoint). If human generated trajectories are taken from the fulltask condition, however, they will contain additional pauses at locations where victims were found and marked providing an undesired cue. Instead, we have chosen to use canonical trajectories from the exploration condition since they should contain pauses associated with waypoint arrival but not those associated with identifying and marking victims. As a final adjustment, operators in the perceptual search condition were allowed to pause their robots in order to identify and mark the victims they discover.

II. EXPERIMENT 1

Forty-five paid participants, 15 in each of the three conditions took part in the experiment. A large USAR environment previously used in the 2006 RoboCup Rescue Virtual Robots competition [24] was selected for use in the experiment. The environment was a maze like hall with many rooms and obstacles, such as chairs, desks, cabinets, and bricks. Victims were evenly distributed within the environment. A second simpler environment was used for training. The experiment followed a between groups repeated measures design with number (4, 8, 12) of robots defining the repeated measure. Participants in the fulltask condition performed the complete USAR task. In the subtask conditions they performed variants of the USAR task requiring only exploration or perceptual search. Participants in the fulltask condition followed instructions to use the robots to explore the environment and locate and mark on the map any victims they discovered. The exploration condition differed only in instructions. These operators were instructed to explore as much of the environment as possible without any requirement to locate or mark victims. From examination of area coverage, pausing, and other characteristics of trajectories in the fulltask and exploration conditions a representative trajectory was selected from the exploration data for each size of team. In the perceptual search condition operators’ retained control of the robots’ cameras while robots followed the representative trajectory except when individually paused by the operator.

Coverage was similar for fulltask and exploration conditions, however, perceptual search participants found more victims and reported lower workload than those in the
fulltask condition. For these tasks differences in SA can most readily be inferred from accuracy in marking victims because this task requires simultaneously considering robot location, pose, camera orientation, and victim location relative to landmarks, e.g., high SA. In this experiment accuracy in marking was better in the perceptual search condition with RMS Error (Figure 4) for perceptual search participants significantly more precise (F1,28 = 23.84, p< .0001) than fulltask participants, although accuracy in both groups declined for an increasing number of robots. This finding suggests that SA was actually higher for participants using automation.

Although area covered did not differ significantly. Accuracy was again better in the autonomous condition again suggesting that SA was higher under increased automation. Process measures, however, suggested that there might be some loss of SA associated with automated path planning.

Temporal data were analyzed to gain a more complete picture of the processes involved in victim observation, detection, and marking. A MANOVA was conducted to analyze the effects of autonomy and group organization on these monitoring and operation process measures. A significant main effect for autonomy was found for process measures (F4,52 =12.118, p<.001), but no effects were found for the group organization (F4,52 =1.781, p=.112).

The ANOVA for the time from which the victim is displayed in the camera to being successfully marked by the operator( Display to Mark), found a main effect for autonomy, F1,56=4.750, p=.034 (Figure 3). Tests for team organization and the interaction were not significant.

III. SYNCHRONOUS VS. ASYNCHRONOUS VIEWING

When considering the foraging task exploration needs to be more than simply moving the robot to different locations in the environment. For search, the process of acquiring a specific viewpoint or set of viewpoints containing a particular object is of greater concern. Because search relies on moving a viewpoint through the environment to find and better view target objects, it is an inherently ego-centric task.

<table>
<thead>
<tr>
<th>Assigned Robots</th>
<th>Shared Pool</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual</td>
<td>15 teams</td>
</tr>
<tr>
<td>Automated</td>
<td>15 teams</td>
</tr>
</tbody>
</table>

Experiment 2 compared teams of two operators controlling 24 robots with either 12 robots assigned to each or control over the 24 shared. The robots were controlled either manually or by an automated path planner demonstrated in [26] to generate paths that did not differ significantly in area covered but having higher tortuosity as measured by fractal dimension than humanly generated trajectories. Sixty teams (120 participants) were run in the experiment. While the division of responsibilities within teams affected performance the comparisons of concern to this paper involve manual and automated path planning.

The ANOVA for “Selected to Mark”, which calculated the time from when the operator selected the robot to successfully marking the victim (Figure 4), by contrast, found main effects for autonomy, F1,56=7.440, p=.009, and for team organization, F1,56=4.029, p=.050.
Multi-robot search, presents the additional problem of assuring that areas the robot has “explored” have been thoroughly searched for targets. This requirement directly conflicts with the navigation task which requires the camera to be pointed in the direction of travel in order to detect and avoid objects and steer toward its goal. These difficulties are accentuated by the need to switch attention among robots which may increase the likelihood that a view containing a target will be missed. In earlier studies [23,24] we have demonstrated that success in search at these tasks is directly related to the frequency with which the operator shifts attention between robots and hypothesized that this might be due to victims missed while servicing other robots.

To combat these problems of attentive sampling among cameras, incomplete coverage of searched areas, and difficulties in associating camera views with map locations we are investigating the potential of asynchronous control techniques previously used out of necessity in NASA applications as a solution to multi-robot search problems. Due to limited bandwidth and communication lags in interplanetary robotics, camera views are closely planned and executed. Rather than transmitting live video and moving the camera about the scene, photographs are taken from a single spot with plans to capture as much of the surrounding scene as possible. These photographs taken with either an omnidirectional overhead camera (camera faces upward to a convex mirror reflecting 360°) and dewarped [27,28] or stitched together from multiple pictures from a ptz camera [29] provide a panorama guaranteeing complete coverage of the scene from a particular point. If these points are well chosen, a collection of panoramas can cover an area to be searched with greater certainty than imagery captured with a pan-tilt-zoom (ptz) camera during navigation. For the operator searching within a saved panorama the experience is similar to controlling a ptz camera in the actual scene, a property that has been used to improve operation in a low bandwidth high latency application [30].

In this modification of the MrCS shown in figure 5 we merge map and camera views as in [31]. The operator directs navigation from the map being generated with panoramas of a chosen extent being taken at the last waypoint of a series. Due to the time required to pan the camera to acquire views to be stitched operators frequently chose to specify incomplete pans. The panoramas are stored and accessed through icons showing their locations on the map. In the Panorama interface thumbnails are blanked out with the arcs at the viewing icons indicating the viewing angles available. The operator can find victims by asynchronously panning through these stored panoramas as time becomes available. When a victim is spotted the operator uses landmarks from the image and corresponding points on the map to record the victim’s location. By changing the task from a forced paced one with camera views that must be controlled and searched on multiple robots continuously to a self paced task in which only navigation needs to be controlled in realtime we hoped to provide a control interface that would automate image acquisition and allow more thorough search with lowered mental workload. The reductions in bandwidth and communications requirements [32] are yet another potential advantage offered by this approach.

Judgments of SA for operators using asynchronous displays needs to be different from instantaneous SA such as evaluated by SAGAT. In asynchronous search the operator’s task has been transformed to depend on awareness of the products of the search to date rather than the state of the environment currently in camera view. For the asynchronous operator, awareness of the association between a marked panorama and its camera views corresponds to the relation between thumbnails and map in the synchronous case. The measure most strongly reflecting SA, therefore, would again be accuracy in marking victims.

In an initial experiment reported in [33], we compared performance for operators controlling 4 robot teams at a
simulated USAR task using either streaming or asynchronous video displays. In the scaling experiments we reported as “found” victims marked within 2 m of their actual location and reported RMS error for the distance between actual and marked locations. In the asynchronous display experiments we have reported these data differently by counting the number of victims marked within concentric circles of the actual location. Figure 6 shows these data from the initial experiment. In this experiment the average number of victims found across conditions using the 2m radius was 4.5, falling to 4.1 for a 1.5m radius, 3.4 at 1m and 2.7 when they were required to mark victims within .75m. Repeated measures ANOVAs found differences in victim detection favoring the streaming video mode at the 1.5m radius F(1,19) = 8.038, p=.01, and 2.0m radius F(1,19)=9.54, p=.006. These data suggest a pattern of greater error in marking for operators using the panoramic display, an effect indicative of reduced SA.

An alternative explanation is that the superiority of streaming video might have occurred simply because these users had the opportunity to move closer to victims thereby improving their estimates of distance in marking the map. A contrasting observation was that frequency of shifting focus between robots, a practice we have previously found related to search performance [32] was correlated with performance for streaming video participants but not for participants using asynchronous panoramas. Because operators using asynchronous video did not need to constantly switch between camera views to avoid missing victims we hypothesized that for larger team sizes where forced pace search might exceed the operator’s attentional capacity asynchronous video might offer an advantage.

In a follow-on experiment we compared panoramas with the streaming video condition from Experiment 1 with teams of 4, 8, and 12 robots to test our hypothesis that advantages might emerge for larger team sizes. Data were analyzed using a repeated measures ANOVA comparing streaming video performance with that of asynchronous panoramas. On the performance measures, victims found and area covered, the groups showed similar performance with victim identification peaking sharply at 8 robots in the streaming condition accompanied by a slightly less dramatic maximum for search coverage.

Figure 7 shows, operators in the panorama condition were again less accurate in marking victims as indicated by the 1.5 m performance which is substantially lower than the 2 m values which are very close to those of streaming video operators. This replicates the earlier finding that automating image acquisition appears to decrease SA of operators

Figure 6. Victims marked within given radius

Figure 7. Victims found for 4, 8, 12 robots

IV. DISCUSSION

While automation has frequently been associated with loss of SA and degradation of performance this does not appear to be the case with the forms of multirobot control studied. Of the performance measures expected to strongly depend on SA, marking victims which requires a variety of relational knowledge was reliably found to be more accurate at higher levels of automation for path planning but showed the expected decline for image acquisition. The process measure select-to-mark was faster under manual conditions suggesting that manual operators were more aware of the location and orientation of the robots. The faster display-to-mark times in the autonomous conditions, however, indicate operators using autonomous path planning may have better awareness of streaming video from the thumbnails. By this account automation may not lead to more or less SA but rather changes the way operators perform their tasks. In the manual conditions operators more attention to the map as they need to set waypoints and pay closer attention to the locations and orientations of the robots so they are better prepared to use this information to mark after selecting the robot. Operators in the autonomy condition, freed from the need to navigate the robots devoted more of their time to monitoring thumbnails and hence were able to achieve
shorter display-to-mark times since they were better able to catch victims as they appeared on camera.

An alternative explanation may also be possible for the greater accuracy of autonomy participants in marking victims. As operators with autonomous path planning reported lower workload and controlled larger numbers of robots, they may be presumed to have greater reserve capacity allowing them to devote greater attention to the marking task leading to greater accuracy.

ACKNOWLEDGMENT

This research has been sponsored in part by AFOSR FA955008-10356 and ONR Grant N0001409-10680.

REFERENCES


