Experiments with an Ecological Interface for Monitoring Tightly-Coordinated Robot Teams

Gal A. Kaminka and Yehuda Elmaliach*
The MAVERICK Group, Computer Science Dept.
Bar Ilan University, Israel
{galk,elmaley}@cs.biu.ac.il

***Abstract*— Many robotics applications require a human operator to monitor multiple robots that collaborate to achieve the operator's goals. Most approaches to such monitoring focus on each robot independently of its peers. When robots are tightly-coordinated, the operator is thus cognitively burdened to build a mental picture of the state of coordination. We report on extensive experiments (approximately 100 hours) with up to 25 human operators, working in two coordinated multi-robot tasks. In these, we contrasted standard displays, which assume each robot is independent, with an ecological *socially-attentive* display that makes the state of coordination explicit. The results show significant improvements in task completion time, number of failures, and the failure rate. Moreover, the display reduces the variance in operator control, thus leading to significantly more consistent operator performance.**

I. Introduction

Many robot applications require a human operator to supervise multiple coordinating robots that work together achieve the operator's goals. An important component within such tasks is to allow the operator to monitor the state of the robots. Examples of applications requiring such monitoring include search and rescue operations [4], multi-rover planetary exploration, and multi-vehicle operation [2].

Previous approaches to monitoring multiple robots use individual robot displays that are independent of each other. For instance, the operator may monitor all robots in parallel, via a split display showing each robot's individual state; or the operator may switch between such displays [1], [2].

However, independent individual displays lead to difficulties in monitoring *coordinated tasks*, requiring tight, continuous coordination between the robots; i.e., where robots are highly inter-dependent. Here, the operator must monitor the state of coordination—the relative state of robots—in addition to the state of each robot. Such monitoring is called *socially-attentive* because it focuses on inter-agent relations [3].

For example, consider the task of controlling three robots moving in formation (a task requiring tight continuous coordination between robots). Such a task can be executed by a single operator, by guiding or teleoperating the lead robot, and allowing the others to maintain the formation autonomously. To maintain the formation, the operator must monitor the formation itself—slowing down or speeding up the lead as necessary—in addition to monitoring the movement of the team towards its goal. Such monitoring can be done, in principle, by showing the camera view of each robot. However, it might be much easier to do if the operator has a bird's eye view of the formation, showing the *relative* positions of robots. Unfortunately, a bird's eye view is not always possible, e.g., for lack of a global-view camera.

To address this challenge, we develop a socially-attentive *ecological display* component—called *relation tool*—that explicitly displays the state of coordination in a team, complementing individual display. Ecological interface design emphasizes visual cues that focus on the key constraints in the user's task [8]. For coordinated tasks, these include the coordination constraints in the team [3]. The relation tool allows the operator to visualize the robots' state with respect to each other, and thus visually identify coordination failures. Since the relative state of robots may not be known directly, the relation tool fuses sensor readings from multiple robots, and reconstructs from these the state of coordination between them. In doing this, it must overcome the uncertainty and noise inherent in robot sensor data.

The graphical *socially-attentive* display complements existing displays. It allows the operator to visualize the robots' coordination—their state with respect to each other—and thus visually identify coordination failures before they become catastrophic. By showing the operator an explicit visualization of the coordination state of the team, her cognitive load would be reduced, and her performance would increase.

We empirically evaluated this approach in extensive systematic experiments with up to 25 human operators. The experiments included monitoring robots in two robotics team coordinated tasks: Movement in formation and cooperative pushing. We evaluate previous individual displays with and without the socially-attentive display. Our statistically-significant results show that the use of the ecological display (i) reduces the number of failures and task completion time in these tasks; (ii) reduces the number of failures per second; and (iii) reduces the variance in controlling robots, thus leading to more consistent performance across operators. These results indicate that the ecological socially-attentive display leads to significant qualitative improvements in the operators interaction with the robots. To our best knowledge, this is one of the largest studies done with human operators controlling multiple robots.

*This work was partially supported by BSF Grant #2002401.

II. BACKGROUND AND MOTIVATION

We focus on visual monitoring interfaces. There have been several investigations of displays (e.g., [6]) that focus on single-robot tasks, in contrast to our focus on multi-robot teams. Previously suggested displays for multi-robot teams implicitly assume robots are independent from each other, and thus focus on displaying the individual state of each robot. Our work can complement such displays.

Adams [1] investigated the use of immersive displays that allow switching control between robots, teleoperating a robot, forming navigation plans, etc. In contrast, we focus on a display that abstract away the details of robots' surroundings, focusing instead on displaying their relative state, not their absolute state with respect to some environment.

Fong et al. [2] propose a *collaborative control* system that allows robots to initiate and engage in dialog with the human operator, one robot at a time. This approach requires significant autonomy by the robots, and assumes that their monitoring need not be continuous.

Myers and Morely [5] discusses an architecture called TIGER that uses a coordinating agent that mediates between the operator and autonomous software agents. This agent centralizes the information from all agents, and can present it to the operator (or provide it to other agents). However, the way the information is presented is left unspecified.

ACTRESS [9] is an architecture including an interface for monitoring and controlling multiple robots. The operator may issue commands that affect groups or individual robots, and information is presented to the operator based on both explicit requests (from the operator to individual robots), as well as by gathering of information exchanged by the robots. However, [9] does not focus on visual presentation of the coordination, in contrast to our work.

In addition to the above conceptual differences with related work, our focus in this work has also emphasized comparing the above approaches to the new techniques, using a large set of experiments with human subjects operating physical robots in the real-world. We have utilized close to 100 operator hours in these experiments, to contrast and draw lessons from the different approaches.

III. A SOCIALLY-ATTENTIVE ECOLOGICAL DISPLAY

Ecological interface design emphasizes explicit visualization of key constraints in the user's task [8]. Socially-attentive monitoring emphasizes that in coordinated tasks, these include the relative state of robots [3]. To show these constraints, we developed the relation tool a 2D display that shows the relative state of robots by drawing a geometric shape corresponding to their state. Colored dots denote different robots. The positions of the dots denote their states, and thus the shape they make up—their relative positioning—denotes their relative states. In principle, every application requires its own method of projecting robot state onto a 2D plane, and a target shape that defines normative coordination.

The key is that the operator should be able to see, at a glance, whether the shape being maintained corresponds

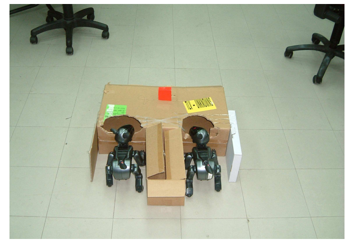


Fig. 1. **Cooperative pushing by AIBO robots**

to correct coordinated execution. When the shape deviates from ideal, the operator can easily identify coordination faults within the monitored team, with little or no need for inferring this information from the other displays. This eases the cognitive load on the operator in coordinated tasks.

We investigate the use of the relation tool concretely in two popular coordinated tasks for robots: Formation maintenance and cooperative pushing. These tasks require tight coordination between the robots, while allowing for human control. Both of these tasks are motivated by real-world applications for human-controlled multi-robot teams, e.g., convoys.

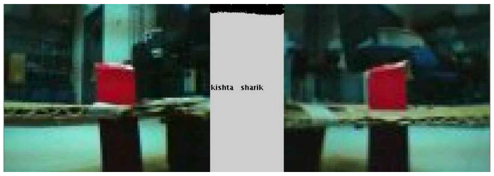
We created human-controlled versions of these tasks, and implemented them using the Tekkotsu software [7] for Sony AIBO robots. Each robot has an on-board video camera and a infra-red distance sensor pointing at the direction of the camera. They transmit their video and sensor readings to the operator's station for monitoring. The operator uses the mouse as a joystick, moving the controlled robot in the direction and speed chosen.

We begin by examining the cooperative pushing task. In this task, two AIBO robots jointly push a light-weight bar across the floor (Figure 1). One robot is teleoperated, while the other pushes the bar while maintaining a straight line with the human-controlled robot. The bar is color-marked, such that a robot can identify its position with respect to the bar. If the mark moves too much to the side, this would indicate a drift, i.e., the robot is either lagging behind or is pushing too quickly ahead. The robots do not communicate with each other.

The coordination between the robots involves a single dimension—the robots are to maintain equal velocities. One possible visualization of this relationship consists of a horizontal line that connects two dots, representing the robot. The horizontal position of the dots remains fixed, while the Y axis denotes the angle of the color mark within their view.

Figures 2 and 3 show the interfaces when executing this task. Figure 2 shows the split-camera view from the individual robots, as presented to the operator, in a successful case (Figure 2-a), and in a failure case, where the box drifts to one side (Figure 2-b). Figure 3 shows the respective relation tool displays in both cases: The successful push (Figure 3-a) and the failing push (Figure 3-b). As can be seen, it can be difficult to differentiate the split-view displays in cases of success and failure (Fig. 2). However, by showing the relative velocities of the two robots (Figure 3) the failing push is easily detected.

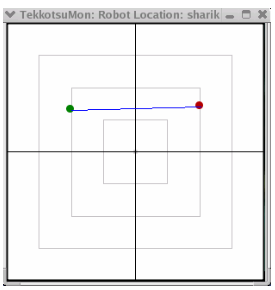
Of course, providing a visualization of the relative states



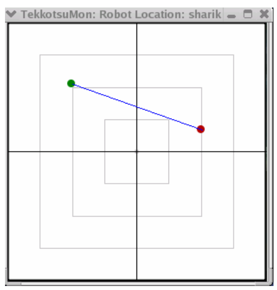
(a) Successful push.

(b) Failing push (robots push right).

Fig. 2. **Cooperative pushing: Split camera view.**



(a) Successful push.



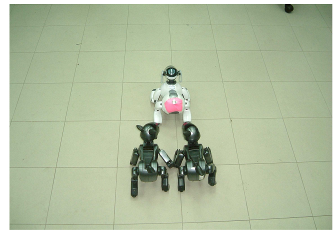
(b) Failing push (robots push right).

Fig. 3. **Cooperative pushing: Relation tool display.**

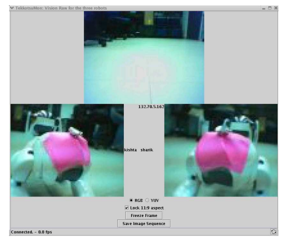
of robots is trivially done when a global world-view camera exists, or perfect global localization data is available. However, this is not often the case in real-world applications.

Thus a key challenge in developing the relation tool lies in integrating the information needed for the visualization, from the robots themselves. The approach we take analyzes the robots' own sensor readings (including camera positioning, infra-red range sensor readings, detected objects) to reconstruct the position of the robot with respect to others, from its own perspective. As a side-effect, we expose the relation tool display to the uncertainty and noise inherent in robot perception. This must be countered by noise-filtering processes within the display. In our case, a moving average filter was used on the distance and angle data to create a stable display.

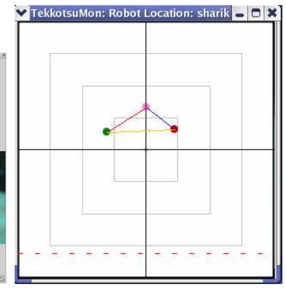
The relation tool may be used to draw the attention of the operator to specific robots that are responsible for any mis-coordination. We use the formation task to demonstrate. Here, the objective is to navigate a triangular formation (three robots), through a short obstacle course. To allow a human operator to control the formation, the lead (front) robot teleoperated by the operator, while the two follower robots maintain fixed angles and distances to this robot using their sensors. Again, the robots do not utilize any communications



(a) Ground truth



(b) Split-camera view



(c) Relation tool

Fig. 4. **Successful Formation**

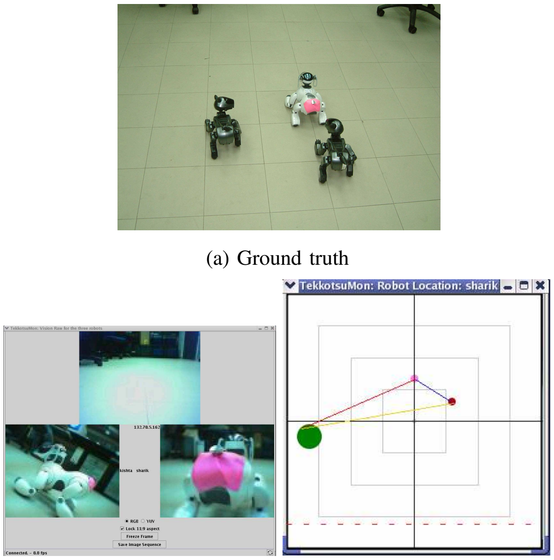
for maintaining the formation.

Figures 5 and 4 show this task in action. Figure 4 shows an example of perfect formation, while Figure 5 shows a failed formation situation. In both figures, sub-figure (a) shows the actual position of the robots on the ground; (b) shows the split-camera view from each of the individual robots; and (c) shows a screen snapshot of the relation tool.

The figures contrast the information presented to the operator with the relation tool and using existing approaches. Unlike the cooperative pushing task, the split-camera view (sub-figure (b) in Figures 5 and 4) does indeed provide indication of whether the formation is maintained. However, it is difficult to see from the split camera view to what degree the formation is maintained (i.e., the magnitude of the failure), and which robots are responsible for it (i.e., the location of the failure).

In contrast, the relation tool makes it easy, at a glance, to see not only whether the formation is maintained, but also the magnitude and location of any failures. We chose polar coordinates to describe the formation. The X axis denotes the angle to the leader, while the Y axis denotes the distance to the leader. The position of the leader is always fixed. We connected the points (that represent the robots) with lines to create a shape easily recognizable by the operator. By glancing at the shape, one can fairly quickly determine whether the formation is breaking because a robot is lagging behind (distance too great), or its angle with respect to the leader is too sharp (e.g., because of a sharp turn).

Indeed, to further assist the operator in localizing coordination problems, the display uses additional mechanisms to draw the attention of the operator where its most needed. One such fault-feedback mechanism uses the size of the dots, representing robot positions, to draw the operator's attention to failing robots. We use three sizes: regular, medium, and large. Regular size is used when the associated robot lies fulfills the constraints of the formation. Medium size is used when



(a) Ground truth

(b) Split-camera view (c) Relation tool view

Fig. 5. **Failing Formation**

the robot begins to report intermittent failures in following the lead, as these are indicative of an impending formation failure. The large size is used when the formation is essentially broken, e.g., when the robot in question completely lost track of the lead robot, and is unable to proceed.

Another fault-feedback mechanism is the dashed line drawn across the bottom of the display. This line signifies the maximum distance sensed by the robots' sensors, and thus the position in which they are likely to lose track of the leader. The operator may use this line to estimate how far it can let the robots stray away from the leader, while not getting into catastrophic failures.

We believe the the relation tool can be useful also in tasks where the coordination between the robot is not spatial. For instance, given a set of sub-tasks which are to be assigned to different robots, the relation tool could use the vertical axis to denote the load on each robot. The operator could then check whether the robots' load is balanced simply by noting the different heights of different dots (signifying different robots). It would also be trivial to use additional visual signs to show the operator the deviation of the shape from the ideal, etc. However, we leave this for future investigation.

IV. Experimental Evaluation

We evaluate the effectiveness of the relation tool in the formation maintenance and cooperative pushing tasks. Our goal is not to contrast the two tasks, but to explore the generality of the method. In the formation maintenance task, the operator leads the robots in a triangular formation towards a target destination, while avoiding obstacles. If the operator causes the lead robot to turn too sharply, or move too quickly, the formation may break. However, the operator seeks to minimize the time it takes to reach the destination. In the pushing task, the operator controls the velocity of one of the robots, while the other is pushing autonomously. The operator must be careful not to push too quickly for the other robot, nor to lag behind.

We believe that the relation tool should be used to complement, rather than replace, existing display (which focus on individual robot state). We thus conducted experiments contrasting different combinations (see below) of the socially-attentive display with individual robot display. We ran multiple experiments with novice operators, age 20–30.

19 operators were tested in the pushing task (18 males/one female, 18 students—including the female—of which 15 are in computer science). 25 operators were tested in the formation task (23 male/two female, 22 students–including the two females—of which 19 are in computer science). The students in both groups were either graduate students or undergraduates in their final year. None had previous experience controlling multiple robots of any kind.

Each operator tried all combinations available in the task she operated. However, to avoid ordering effects, the order in which each operator tried each combination was randomized. In no setting were the operators able to see the robots while operating them. In all cases, operators were given an approximate 25-minute training session in operating a single and multiple robots (including the formation and pushing tasks), until they reported they felt comfortable controlling the robots. Overall, the results below represent over 50 hours of human operation.

The first experiment examined the use of the relation tool in the cooperative pushing task. We contrasted three interfaces: a split-camera view only (representing existing approaches), a combination split-camera and relation tool, and the relation-tool alone. All 19 human operators were tested on all three interfaces, randomizing the order of their introduction to the different interfaces to prevent biasing the results due to human learning. Their performance was measured as the average absolute angle deviation from the imaginary horizontal line connecting the robots when they maintain ideal relative velocity. This angle was sampled at 20Hz during task execution.

Figure 6 shows the results of this experiment—the average absolute angle error—averaged across all operators. Clearly, both combinations that use the relation tool are significantly superior to the interface relying on camera alone. Moreover, the surprising result here is that the relation tool by itself is sufficient (in fact, even slightly better than its combination with the split view). This is due to this task being essentially a pure-coordination task: The operator does not need to worry about where the pushed object is going, as long as the relative velocity of the robots is 0 (i.e., their velocities are equal). Thus even a socially-attentive display by itself is sufficient. On the other hand, the non-social split-camera view (by itself), is difficult to use for coordination. A one-tailed t-test (assuming unequal variance) shows that the difference between using the tool by itself, and using the split-camera view, is statistically significant (we use a 0.05 significance level). The probability of the null hypothesis is $p < 0.014$ when looking at the difference in the number of failures.

In the formation maintenance task, we compare three interfaces. The first presented the operator with the split-view

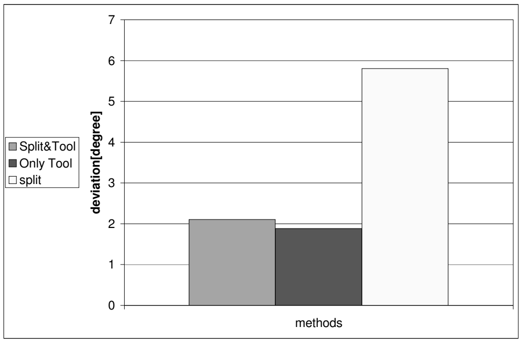
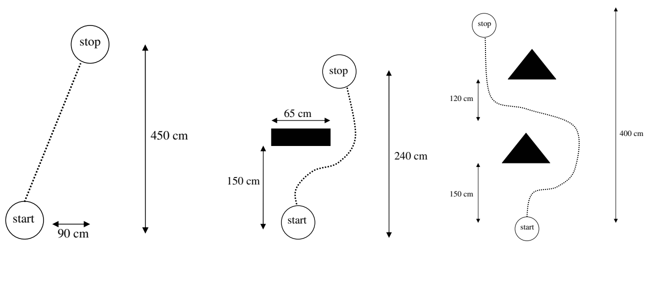


Fig. 6. **Cooperative Pushing: Total failures**



(a) (b) (c)

Fig. 7. **Formation obstacle courses.**

video streams from all robots (e.g., Figure 4-c). The second combined the this split-view with the socially-attentive display previously described. The final interface consisted of a single camera (the lead robot's) and the socially-attentive display. Each of the interfaces was tried with three different obstacle courses, varying in difficulty (a total of 9 different configurations). The *simple* course consisted of an open space with no obstacles at all (Figure 7-a). The *medium* course consisted of a single obstacle that had to be by-passed (7-b). In the *difficult* course, the operator was to lead the robots between the two obstacles (7-c). To verify the relative difficulty of the path, we sampled 7 of the experiments for the number of times a robot hit an obstacle: The simple course had no such hits (as there are no obstacles). The medium course had only a single hit in all experiments. The difficult course had 2–3 hits per method.

Again, all 25 operators tried all nine different settings, in randomized order (to prevent learning effects). For each of the trials, we recorded the number of non-catastrophic formation failures, and time to complete the task. Non-catastrophic formation failures were measured as the number of times a follower robot has temporarily lost track of the lead. These are indicative of the quality of the operator's control. Too many of them result in permanent tracking failures, which lead to total breakdown of the formation. When such failures occurred, the operator would have to teleoperate the straying robot until the formation was re-established.

Figures 8,9,10 show the results of these experiments in terms of the average number of non-catastrophic failures per operator, versus the average task completion time. The

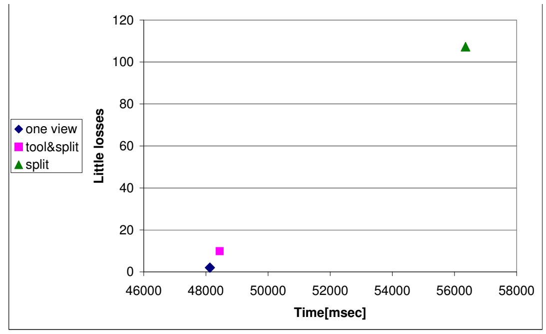


Fig. 8. **Formation failures in *simple* course**

horizontal axis shows the time (in milliseconds). The range of the horizontal axis in these figures is fixed at 12 seconds, though the offset is different, as the more difficult courses took longer. The vertical axis shows the average number of non-catastrophic failures that took place during each trial.

The results show that in all course difficulty settings, the use of the relation tool is preferable to using only individual displays. This lends support to the hypothesis that socially-attentive ecological displays can significantly improve monitoring of robots in coordinated tasks.

In particular, both course completion time and the number of failures during execution were generally reduced using the socially-attentive display. In the simple- and medium-difficulty courses, the best monitoring approach was single camera and the socially-attentive display. It was significantly better than the split camera interface, at a 0.05 significance level. In the easy course, a one-tailed t-test (assuming unequal variances—see below) shows a significant difference these method, both in the number of failures (the probability of the null hypothesis being $p < 0.011$), and in the time ($p < 0.015$). Similarly, in the medium course, there are significant differences between these two methods, both in the number of failures ($p < 0.04$) and in task completion time ($p < 0.02$).

However, in the difficult course the best monitoring approach used both the split-view and the relation tool, in spite of the additional information displayed to the operators. The difference between this approach and the split view interface was not significant in time ($p = 0.48$), but was significantly different in the number of failures ($p < 0.014$). The difference n the number of failures between the split view interface and the interface using single camera and relation tool was only moderate ($p \approx 0.15$). We believe that this is due to the operator using the split-camera view to look at obstacles that have been bypassed by the lead (see [6] for an ecological interface approach to this problem). Such obstacles were not much of a problem in the other, easier, courses. We leave further investigation of this to future work.

While the results show significant improvements in task completion time and number of failures, a question may be raised as to whether a socially-attentive ecological display qualitatively changes the way the operator interacts with the team. For instance, the experiment results above could also be indicative of the team going slower or faster, but maintaining

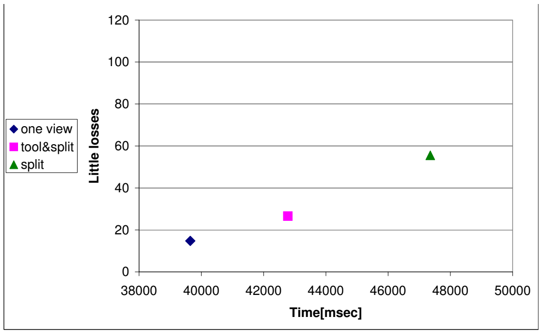


Fig. 9. **Formation failures in *medium* course**

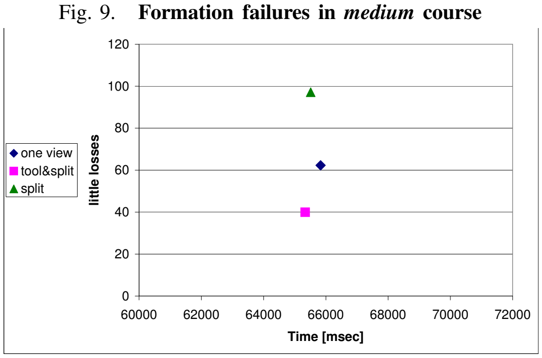


Fig. 10. **Formation failures in *difficult* course**

the same number of failures per second—thus indicating that the drop in failures is due to the team moving faster, rather than to a qualitative change in operator control.

Additional results show that rather, the use of the relation tool leads to qualitative differences in the the way the operator controls the robot team. Figure 11 shows the average number of failures per second, in the different courses. Clearly, the easy course is indeed easier than the medium-difficulty course, which is easier than the difficult course. What we see in the results is that the use of the socially-attentive display leads to a significant reduction not just in the time and total number of failures (as evident from the previous figures), but also reduces the *failure rate*.

Additional evidence for this qualitative improvement in operator control is found when we examine the standard deviation values for the number of failures and task completion time. Table I displays the standard deviation of the number of failures, for the different courses. As can be seen, the standard deviation values for the methods using the relation tool are generally much smaller than for the split camera display. This

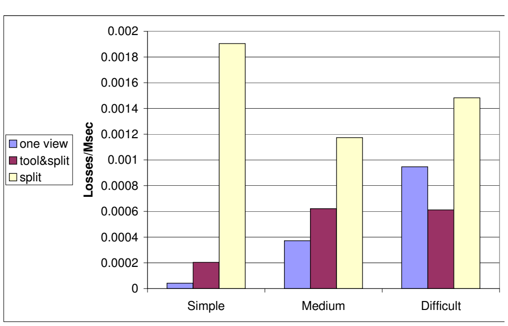


Fig. 11. **Formation failures per millisecond.**

Course	Split View	Split View and Relation Tool	Single View and Relation Tool
Easy	319.45	32.65	**6.59**
Medium	141.85	51.30	**50.56**
Difficult	144.97	**66.93**	138.65

TABLE I

Standard deviation in number of failures.

indicates more consistent values, i.e., less variance between operators in terms of ability to control the robots. In the difficult path, the single camera view with the relation tool has a large standard deviation (though smaller than the one for the split camera view by itself), but the relation tool with the split camera view has smaller standard deviation.

V. Summary and Future Work

This paper takes a step towards allowing a single human operator to effectively monitor a team of robots that are tightly coordinated. The socially-attentive *relation tool* display is an ecological interface display addressing this challenge. It has three principal advantages over previous work. First, it significantly reduces the amount of inference needed by the operator to infer the state of coordination between robots. Second, its dimensions can be used to directly provide the operator with information about failures. Third, it can easily complement other types of displays useful for the task.

Extensive experiments with 25 human operators, on real robots, show that the relation tool significantly reduces the total number of failures, and task completion time in two tight-coordination tasks. Furthermore, we have shown that the use of the relation tool leads to qualitative change in the capabilities of the operator: Not only do failures and completion time decrease, but the failure rate (failures per second) improves significantly as well. In addition, methods utilizing the relation tool lead to more consistent operator performance.

Acknowledgments. We thank Avi Rosenfeld for useful comments, and Ruti Glick for help in organizing the experiments. Special thanks to K. Ushi.

References

[1] J. A. Adams. *Human Management of a Hierarchical System for the Control of Multiple Mobile Robots*. PhD thesis, University of Pennsylvania, 1995.

[2] T. Fong, C. Thorpe, and C. Baur. Multi-robot remote driving with collaborative control. *IEEE Transactions on Industrial Electronics*, 50(4):699–704, August 2003.

[3] G. A. Kaminka and M. Tambe. Robust multi-agent teams via socially-attentive monitoring. *JAIR*, 12:105–147, 2000.

[4] R. Murphy. Human-robot interaction in rescue robotics. *IEEE Systems, Man and Cybernetics Part C: Applications and Reviews*, 34(2), May 2004. Special issue on Human-Robot Interaction.

[5] K. L. Myers and D. N. Morely. Human directability of agents. In *Proceedings of the First International Conference on Knowledge Capture, K-CAP 2001*, Canada, 2001.

[6] B. Ricks, C. W. Nielsen, and M. A. Goodrich. Ecological displays for robot interaction: A new perspective. In *IROS-04*, 2004.

[7] The Tekkotsu Homepage. www.tekkotsu.org, 2002.

[8] K. A. Vicente. Ecological interface design: progress and challenges. *Human Factors*, 44(1):62–78, 2002.

[9] K. Yokota, T. Suzuki, H. Asama, A. Masumoto, and I. Endo. A human interface system fot the multi-agent robotic system. In *ICRA-94*, pages 1039–1044, 1994.