Asynchronous Displays for multi-UV Search Tasks

Michael Lewis[1] and Huadong Wang[2]
University of Pittsburgh, Pittsburgh, PA, 15260

Andreas Kolling[3], Katia Sycara[4], Nathan Brooks[5], and Paul Scerri[6]
Carnegie Mellon University, Pittsburgh, PA, 15213

Synchronous video has long been the preferred mode for controlling remote robots with other modes such as asynchronous control only used when unavoidable as in the case of interplanetary robotics. We identify two basic problems for controlling multiple robots using synchronous displays: operator overload and information fusion. Synchronous displays from multiple robots can easily overwhelm an operator who must search video for targets. If targets are plentiful, the operator will likely miss targets that enter and leave unattended views while dealing with others that were noticed. The related fusion problem arises because robots' multiple fields of view may overlap forcing the operator to reconcile different views from different perspectives and form an awareness of the environment by "piecing them together". We have conducted a series of experiments investigating the suitability of asynchronous displays for multi-UV search. Our first experiments involved static panoramas in which operators selected locations at which robots halted and panned their camera to capture a record of what could be seen from that location. A subsequent experiment investigated the hypothesis that the relative performance of the panoramic display would improve as the number of robots was increased causing greater overload and fusion problems. In a subsequent Image Queue system we used automated path planning and also automated the selection of imagery for presentation by choosing a greedy selection of non-overlapping views. A fourth set of experiments used the SUAVE display, an asynchronous variant of the picture-in-picture technique for video from multiple UAVs. The panoramic displays which addressed only the overload problem led to performance similar to synchronous video while the Image Queue and SUAVE displays which addressed fusion as well led to improved performance on a number of measures. In this paper we will review our experiences in designing and testing asynchronous displays and discuss challenges to their use including tracking dynamic targets.

I. Introduction

CONTROLLING multi-robot systems (MrCS) has become a prominent area of research in human-robot interaction (HRI). These efforts have included theoretical and applied development of the Neglect Tolerance model and Fan-out model to characterize the control of independently operating robots[1,2], predefined rules to coordinate cooperating robots as in Playbook[3] and Machinetta[4], and techniques for influencing teams obeying biologically inspired control laws[5,6,7]. While our efforts to increase span of control over unmanned vehicle (UV) teams appear to be making progress, the asymmetry between what we can command and what we can comprehend has been growing. Automation can reduce excessive demands for human input, but exploiting the information being collected and returned poses a greater problem. A human is frequently included in the loop of a MrCS expressly to monitor and interpret video being gathered by UVs. This can be a difficult task for even a single camera[8] and begins exceeding operator capability well before reaching ten cameras[9,10]. With increasing autonomy of robot teams and plans for biologically inspired swarms of much greater size the problem of absorbing and benefiting from their product has become more pressing than improving ways to command them.

We identify two basic problems for exploiting video from multiple robots: operator overload and information fusion. Synchronous displays from multiple robots can easily overwhelm an operator who must search multiple

---

[1] Professor, School of Information Sciences, 135 N. Bellefield Ave.
[2] Ph.D. Student, School of Information Sciences, 135 N. Bellefield Ave.
[3] Postdoctoral Fellow, Robotics Institute, 5000 Forbes Ave.
[4] Research Professor, Robotics Institute, 5000 Forbes Ave.
[5] Ph.D. Student, Robotics Institute, 5000 Forbes Ave.
[6] Associate Research Professor, Robotics Institute, 5000 Forbes Ave.

1

video streams for targets. If targets are plentiful, the operator will likely miss targets that enter and leave unattended views while occupied with targets in attended views.

The related fusion problem arises because robots' multiple fields of view may overlap forcing the operator to reconcile views from different perspectives to form an awareness of the environment by "piecing them together". Simply performing the mental rotations needed to match a single camera view to map coordinates has been found to be difficult and error prone for UAV video[11]. Taken together the problems of multi-camera monitoring and fusion of viewpoints can pose insurmountable difficulties to operators who must develop and maintain situation awareness (SA) through remote camera views.

Our original motivation for investigating asynchronous displays for MrCS was to address bandwidth restrictions introduced into RoboCup Rescue competition to more realistically simulate obstacles to multi-robot control. Although the panorama mode was designed to drastically reduce bandwidth and allow operation despite intermittent communications, our system was so effective we decided to test it under conditions favorable to conventional streaming video interfaces. We hypothesized that a solution to the difficulty of monitoring multiple camera views might be to make the task asynchronous by removing the requirement for simultaneity and supporting sequential viewing. In foraging tasks such as urban search and rescue (USAR) targets (victims) are presumed to remain static while foragers (robots) explore the environment to find them. Because performing the search task requires only that a target be detected in some image, the many redundant frames of video covering a region containing a target from a variety of distances and angles are unnecessary for performing the task. The operator needs only a single frame covering the area containing the target. The problem for this approach then becomes one of choosing some set of images that best cover the regions explored by the robots. Because targets are presumed static these images can be taken at any time but need to be selected so as to include as much of the explored region as possible. This approach separates the problem of controlling the robots from that of searching the environment by focusing operator attention on views of regions of space rather than camera views of robots. Once a set of high coverage images has been selected, the operator can peruse them sequentially to find targets. This transforms a difficult mentally-loading forced-pace task of monitoring multiple changing views into a self-paced task in which the operator can spend as much time as needed to search each image. The information fusion problem arises out of the need to understand the relationship among the regions portrayed in the imagery in order to maintain SA. This is similar to the fusion problem of integrating video from the cameras of multiple moving robots. The display design problem is to develop mechanisms for storing and retrieving imagery that supports information fusion.

In this paper we review a series of four experiments investigating the use of asynchronous displays for a multi-UGV USAR task and a fifth experiment using similar techniques for aerial search. The first two experiments involve static panoramas in which operators selected locations at which robots halted and panned their camera to capture a record of what could be seen from that location. In these experiments operators controlled the robots by specifying waypoints with the final waypoint determining the images that were selected to "cover" the space. The locations at which panoramas had been taken were displayed and accessed through a map being constructed as the robots searched the space. The second experiment investigated the hypothesis that the relative performance of the panoramic display would improve as the number of robots was increased causing greater overload and fusion problems. The Image Queue system extended search of asynchronous imagery to automated path planning also automating the selection of imagery for presentation by choosing a greedy selection of non-overlapping views. A second experiment incorporated automatic target recognition (ATR) in the selection of imagery. A final experiment compared viewing strategies using the SUAVE display, an asynchronous variant of the picture-in-picture technique for fusing video from multiple UAVs.

A. Issues in Asynchronous Control

Alberts et al.[12] introduced the distinction between network centric and platform centric orientations to help explain the advantages they felt might accrue to highly networked military forces. In conventional forces a platform and its sensors operate with relative independence. A pilot, for example, would be able to see very close aircraft through his canopy, nearby aircraft through onboard radar, and hear with less precision about more distant aircraft over his radio. A commander directing the mission would only see what was available through her own sensors plus reports coming in from other aircraft. In a network centric version the aircraft could share their targets so each pilot would see what everyone else was sensing as well as data from his own instruments. This sharing of data cuts the bond between platform and information allowing the commander to focus on the mission rather than "who reported what". The pilots still must fly their planes, however, this local platform-specific task is now divorced from the commander's problem of maintaining global situation awareness. According to 12 the focus on task rather than platform should improve SA for both the commander and team. An alternative perspective[13] holds that involvement

in navigation and path planning are crucial to developing and maintaining SA and that separating these functions from target search could lead to reduced SA.

Both aspects of SA are present in the USAR task. In the panorama experiments operators needed to plan paths and coordinate robots as well as search panoramas for victims. In the Image Queue experiments, by contrast, path planning was automated so operators needed to develop SA by associating viewpoints on a map with imagery they were viewing. The issue is further clouded by the effects of reduced workload from automation on operators' ability to direct resources to information fusion and SA. In reviewing results we will focus on differences in victim detection rates and errors that may help understand how divorcing platform from search has affected SA.

II. Experimental Environments

A. USARSim and MrCS

Four of the five experiments were conducted using the USARSim robotic simulation with simulated Pioneer P3-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. Validation data showing close agreement in detection of walls and associated Hough transforms for a simulated Hokuyo laser range finder are described[14]. Validation studies showing close agreement in behavior between USARSim models and real robots being modeled are reported[15,16,17,18,19] as well as agreement for a variety of feature extraction techniques between USARSim images and camera video reported by Carpin et al.[20].

MrCS (Multi-robot Control System), a multi-robot communications and control infrastructure with accompanying user interface, developed for experiments in multirobot control and RoboCup competition was used in these experiments. MrCS provides facilities for starting and controlling robots in the simulation, displaying multiple camera, and laser output, as well as maps, and supporting inter-robot communication through Machinetta, a distributed multi-agent coordination infrastructure.

The operator selects the robot to be controlled from the colored thumbnails at the top right of the screen. These thumbnails also show the live video streams from all robots. To view more of the selected scene shown in the large video window the operator uses pan/tilt sliders to control the camera of the selected robot. The current locations and paths of the robots are shown on the Map Viewer (bottom left). Under manual control, robots are tasked by assigning waypoints on a heading-up map on the Map Viewer or through the teleoperation widget (lower right).

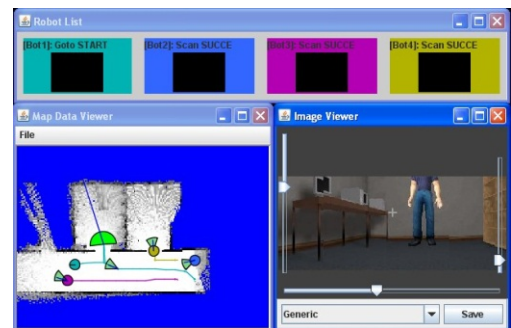**Figure 1. Base MrCS Display**

Figure 2. Panorama Display

The asynchronous displays reviewed in this paper are extensions of this basic interface. In the panorama studies (Figure 2) thumbnail video is eliminated and operators control solely through the map by specifying waypoints and the viewing angles to be captured at the terminal waypoint. Panoramas are retrieved from icons located at terminal waypoints and displayed in a window at the bottom right. This view can be panned and tilted (as on the live robot) to view the full extent of the panorama which is indicated by the size of the "pie slice" showing the range of viewing angles requested by the operator.

The Image Queue replaces operator specified paths and selection of imagery with automated path planning and selection of imagery to be viewed. The Image Queue (Figure 3) augments the basic interface with a filmstrip viewer designed to present the operator with a filtered view of what has passed before the team's cameras. A filtered view is

beneficial because the video taken contains a high proportion of redundant images from sequential frames and overlapping coverage by multiple robots. The filter attempts to reduce redundancy by only showing highly relevant images from the video stream. Relevance is scored by computing a utility for every image that determines its priority in the queue displayed in the filmstrip viewer. To achieve this we store every frame from all video streams in a database together with associated robot poses and laser scans taken at the time of capture. Visual coverage is computed by referencing the image in the map as seen in Figure 4. Images with larger areas receive higher utility scores. Areas that have already been seen by other images in the filmstrip viewer do not count towards utility. In colloquial terms this kind of utility picks images that cover large areas with minimal overlap. Figure 4 illustrates this concept of utility with a simple example. While the Image Queue filmstrip replaces live video feeds in the thumbnails the ability to select a robot from the map to teleoperate and view streaming video remains to allow operators to intervene to free entangled robots.
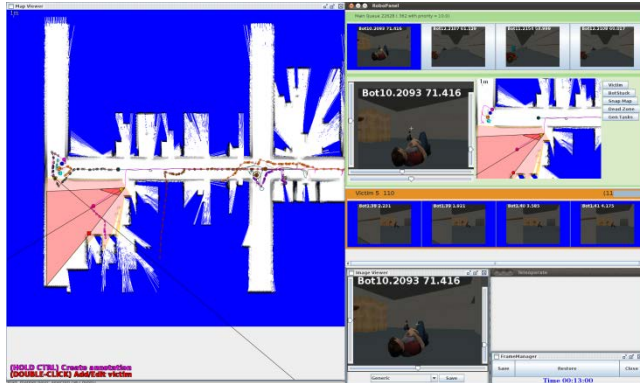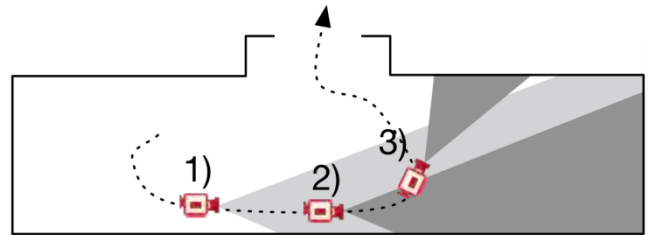

Figure 4. An illustration of the utility of individual frames from a video stream. The frame taken at 1) has the largest visual coverage and highest utility while the frame at 2) has no utility since it is entirely overlapped by 1). Frame 3) has some utility since it provides coverage in an area not covered by 1).
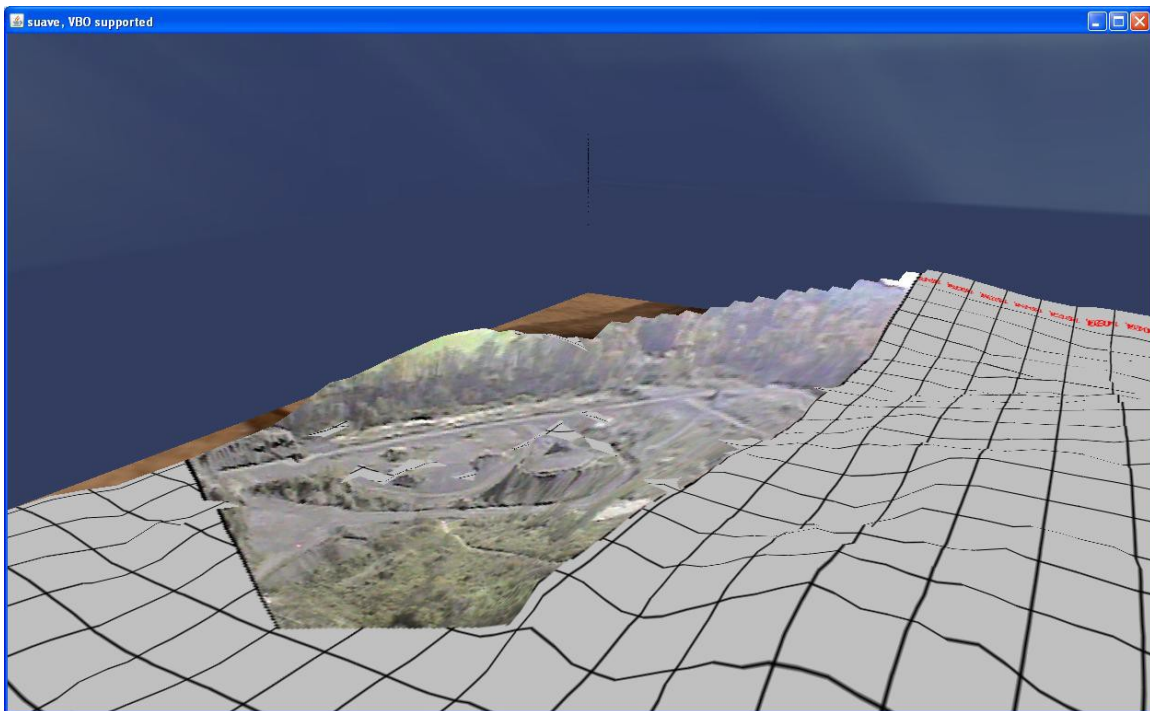
Figure 3. Image Queue Display with "filmstrip" on right

Figure 5. SUAVE Display showing imagery textured on height map

B. SUAVE

American Institute of Aeronautics and Astronautics

Simple UAV Environment (SUAVE) described at a previous InfoTech[21] is an experimental system developed to investigate the use of model-inspection techniques to exploit real-time video feeds. SUAVE uses a previously acquired height map to present imagery. As UAVs stream video, individual geo-referenced frames are selected and projected onto the height map replacing the old texture. To search the video returned by the UAVs the operator inspects the terrain model for targets. Our method resembles picture-in-picture (PiP) presentation[22,23,24,25], in which the terrain model is scaled and transformed so that video may be viewed in context but is truly asynchronous relying on inspection of sampled imagery rather than viewing video to find and mark targets. The reported experiment uses a height map and video obtained from the Virtual Battle Space II (VBS2) military simulation.

III. Experiments

Table 1 shows the studies being reviewed. All experiments used control conditions in which operators were presented with live video streams for comparison with asynchronous data presentations. The Panorama-2 experiment additionally tests effects related to the number of robots. The Image Queue-2 experiment includes ATR results in the utility function selecting imagery. Cuing for ATR recognized targets is tested within the ATR condition. SUAVE and the steaming video control were counterbalanced. The third SUAVE-22 condition tests the effects of doubling the number of UAVs to increase the resolution of the model (SUAVE).

Table 1. Reviewed Experiments

Study	N	Design	robots	Comparisons
Panorama-1[26]	21	repeated measures	4	panorama vs. streaming video
Panorama-2[27]	29	2 groups repeated measures	4,8,12	panorama vs. streaming video (groups) number of robots (repeated measure)
Image Queue-1[28]	32	2 groups	12	Image Queue vs. streaming video
Image Queue-2[29]	30	repeated measures 2 groups	12	Image Queue vs. streaming video ATR cued vs. no cue (within Image Queue)
SUAVE[30]	12	repeated measures	11,22	SUAVE vs. streaming video/ SUAVE-22 (last condition)

IV. Results

A. Targets and Coverage

Performance on primary task measures of area covered and targets found were strikingly similar across synchronous and asynchronous displays. In Panorama-1 under conditions allowing uninterrupted, noise free, streaming video from 4 robots the synchronous interface led to only somewhat better (5 vs. 4 victims) search performance. A repeated measures ANOVA found this difference significant when detection was defined by a mark within a 2 m radius of the target $F(1,19)=9.54$, $p=.006$.. This similarity in performance was replicated in Panorama-2. For victims marked within 2 m, the average number of victims found in the panorama condition was 5.36 using 4 robots, 5.50 for 8 robots, but dropping back to 4.71 when using 12 robots. Participants in the Streaming condition were slightly more successful at this range, $F_{1,29} = 3.563$, $p < .028$, finding 4.8, 7.07 and 4.73 victims respectively. As Figure 6 shows this difference is due to a spike in performance for streaming video at 8 robots. Number of robots had a significant effect on every dependent measure collected except waypoints per mission.

Similar results were found for found for the Image Queue. For Image Queue-1 participants in both conditions were successful in searching through the environment. On average participants in the streaming video condition found 9.10 victims while those in the Image Queue condition found 8.51 (Figure 7) without a significant difference between conditions ($F_{1,28} = .733$, $p = .387$). The area explored (Figure 7) for both conditions did also not differ significantly ($F_{1,28} = 2.147$, $p = .154$).

In Image-Queue-2 Participants were again successful in searching the environment with no significant differences between conditions ($F_{1,28} = .181$, $p = .674$) or groups ($F_{1,28} = .103$, $p = .751$). On average, participants in the streaming video condition found 9.03 victims, while those in the Image Queue conditions found 8.73. The area explored by the 12 robots also showed no significant differences among displays ($F_{1,28} = 0.479$, $p = .495$).
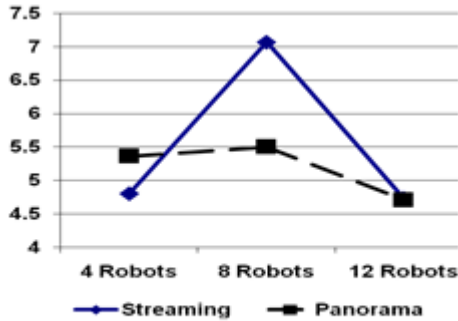
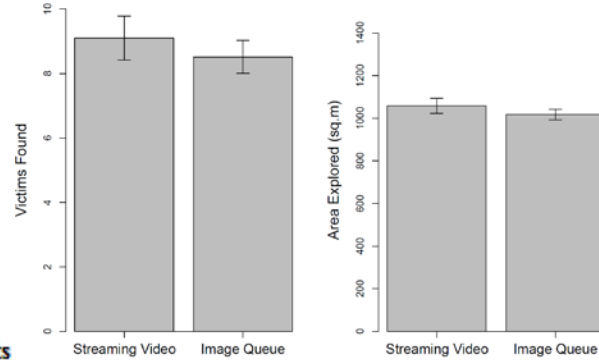**Figure 6. Victims Found as a function of N robots**

Figure 7. Image Queue-1 Victims and Coverage

Differences were found in SUAVE when targets marked within 50 meters were scored as correct, the resulting ANOVA showed a significant advantage for the SUAVE condition ($F_{1,11} = 14.438$, $p = 0.003$). When accuracy is relaxed to a 100 meter radius, however, this advantage is erased with participants in the video stream condition successfully marking 5.50 targets while those in the SUAVE condition mark 5.67 (Figure 8) ($F_{1,11} = 0.035$, $p = 0.854$). A general finding across these experiments was that asynchronous displays did not lead to marking more targets as might be expected given the reduction in operator load in changing from a forced-pace to self-paced task. Conversely, despite drastic reductions in information presented and loss of connection between robots and camera views operators were able to perform just as well searching asynchronously.
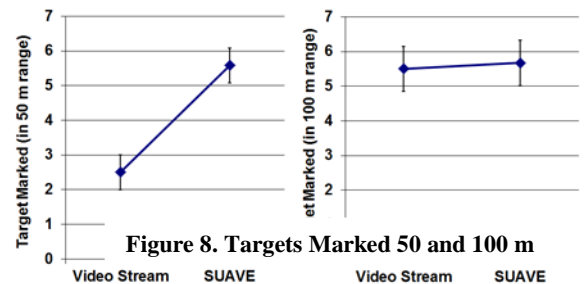

Figure 8. Targets Marked 50 and 100 m

B. Accuracy

Every mark a participant made for a victim was compared to ground truth to determine whether there was in fact a victim at the location. A mark made further than the criterion distance from any victim or multiple marks for one victim were counted as false positives. Victims that were missed, but present in the video feed, and not marked were counted as false negatives. Because the number of victims "found" is confounded with the accuracy with which they have been marked we considered data for radii of different sizes in the Panorama and SUAVE experiments. A radius of 2 m was adopted in the Image Queue experiments because data showed low rates of false positives (marking a victim beyond the 2 m radius) of between 1 (Image Queue) and 2.4 (streaming video) while successful markings were 8.5 and 9.1 suggesting that the 2 m criterion does well at associating most marks with victims.

No differences in accuracy were found in either of the Panorama studies where results remained the same across the range of radii used to define marked targets. Significant advantages favoring the asynchronous displays, however, were found in both Image Queue experiments and for SUAVE. For the Image Queue false negatives may include targets that have been observed by robots but not presented as well as those actually missed by the operator. In Image Queue-1 the data showed a significant advantage for both types of errors. There were significantly fewer false positives ($F_{1,28} = 13.032$, $p = .001$) as well as fewer false negatives ($F_{1,28} = 5.526$, $p = .026$) with an average in the Image Queue condition of 7.48 while participants in the streaming video condition missed 9.34 victims on average (Figure 9).

For Image Queue-2 the number of false positives showed no significant difference between the Image Queue conditions and streaming video ($F_{1,28} = .053$, $p = .819$). A one-way ANOVA, however, found a significant advantage for the *no-cue* group over the *cued* group ($F_{1,28} = 4.974$, $p = .034$) within the Image Queue conditions. The Image Queue did, however, show a significant improvement over the streaming video condition ($F_{1,28} = 7.292$, $p = .012$) for false negatives, with the average number of missed victims dropping to 7.17 from the 8.67 missed in the streaming video condition.

As shown in Figure 8 SUAVE also substantially improved operators' accuracy allowing them to mark within 50 m rather than the 100 m radius needed when marking directly from video.
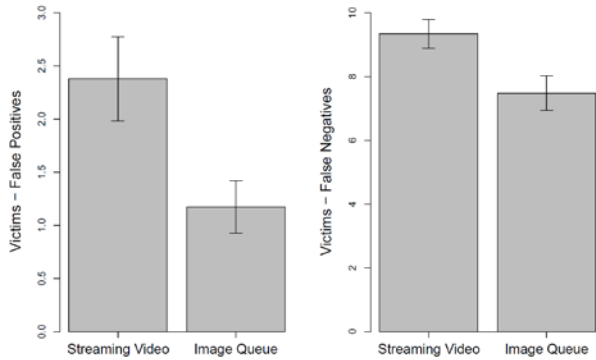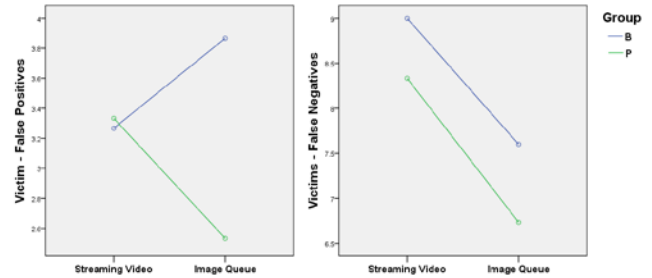

Figure 9. Image Queue-1 Errors in Marking

Figure 10. Image Queue-2 Errors in Marking (note: because separate groups were used for the cueing condition two groups are shown for the Streaming Video control according to this division of participants)

C. Workload

An initial premise of our investigation of asynchronous displays was that changing forced-pace search tasks into self-paced asynchronous ones would reduce operator workload and lead to better performance. Another empirical aspect of workload is additional actions that may be required by one system and not the other. Because finding targets using the streaming video display requires associating the robot in the thumbnail with the one on the map an operator may need to teleoperate the in-focus robot to help locate it increasing his workload. No differences were found in NASA-TLX workload ratings for either of the Panorama experiments. In Image Queue-1 a repeated measures ANOVA showed a significant difference for teleoperation instances between the streaming video and Image Queue conditions (Figure 11) . Participants in the streaming video condition teleoperated on average 21.24 times while participants in the Image Queue conditon teleoperated only an average of 4.97 times ($F_{1,28}$ = 150.719, p < .001). The full scale NASA-TLX workload measure also revealed a significant advantage in workload ($F_{1,28}$ = 7.347, p = .001). Examining individual dimensions of workload we found significant differences for temporal demand ($F_{1,28}$ = 6.503, p = .016) and effort ($F_{1,28}$ = 4.576, p = .040) .

For Image Queue-2 a repeated measures ANOVA showed a significant difference ($F_{1,28}$ = 176.845, p < .001) for the count of teleoperation times between the streaming video and Image Queue condition with participants in the streaming video condition teleoperating an average of 16.07 times while they chose to teleoperate only 0.87 times in the Image Queue condition. While the full-scale NASA-TLX workload measure revealed no advantage for either the Image Queue or streaming video conditions, the no-cue version of the Image Queue was judged significantly less taxing than the cued version ($F_{1,28}$ = 5.364, p = .028). Workload followed a similar pattern in the SUAVE experiment with the full scale NASA-TLX workload measure finding no advantage for either condition. ($F_{1,11}$ =0.139, p = 0.716).
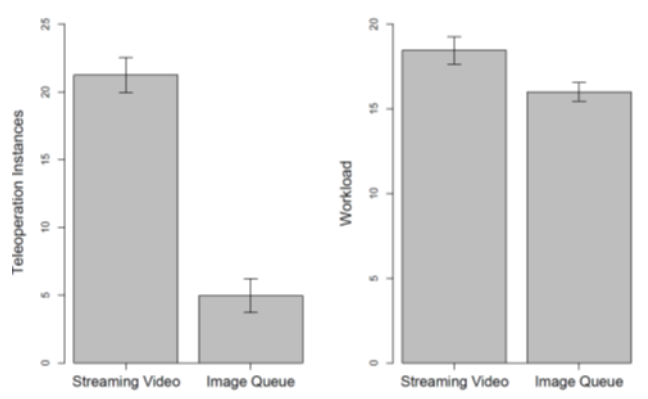

Figure 11. Teleoperation and NASA-TLX

V. Discussion

Our review of five recent experiments investigating the use of asynchronous displays to combine video from multiple robots shows promise in the approach but also a need for further research. While asynchronous displays did not lead to finding more targets they did in most cases lead to greater accuracy in marking targets and somewhat lower workload. In principle, integrative displays such as Image Queue or SUAVE should be fully scalable improving performance while imposing no additional load on the operator as the number of robots increases. In the case of Image Queue additional robots should increase the probability of obtaining higher utility imagery making the operator's search more productive in reviewing the same number of images. For SUAVE adding more UAVs could either increase the spatial resolution of the display by using more cameras at higher resolutions or increase its temporal resolution by refreshing areas more frequently. This should improve the "quality" of imagery being reviewed without requiring any additional effort on the part of the operator. Our studies, however, have not yet found this expected scaling advantage for asynchronous displays. Although operator workload in Panorama-2 would be expected to increase with N robots due to the need to set waypoints and specify viewing cones, this disadvantage should have been offset by scalable gains in visual search of stored panoramas. Instead, operators showed the same decline in performance with 12 robots as those monitoring streaming video and failed to benefit going from 4 to 8 as users of the conventional display did. A similar failure to show an expected improvement was found in the SUAVE experiment where going from 12 to 22 UAVs (which should have improved both coverage and resolution) had no effect on performance. While we remain confident in the rationale for scalability of asynchronous displays we have not yet demonstrated it in our experiments.

Suitability for multi-operator control is another potential advantage for asynchronous displays such as the Image Queue. Operators attempting to control or monitor robot teams in real time would be faced not only with the daunting task of controlling and coordinating their own robots but with coordinating with others trying to perform the same difficult tasks. Asynchronous control such as the Image Queue provides convenient ways to divide tasks functionally among operators, such as allocating exploration and target identification to different operators. Shifting focus from platforms and camera video to the network and regions being explored allows searchers to concentrate on their primary search task rather than on driving or monitoring robots. Just as our Image Queue operators were called upon to teleoperate robots out of trouble from time to time, we envision future systems which are controlled at both network and platform levels. To realize this kind of control architecture, we propose a call center approach in which some operators address independent control needs for monitoring and exploration of UVs, while other operators address independent location-based images in a queue for victim marking and other perceptual tasks. While this approach seems theoretically plausible it remains untested.

A final challenge to the use of asynchronous displays is making them suitable for dynamic targets and environments. Our current asynchronous displays trade-off temporal resolution for spatial resolution. You can either see what is going on at a few dispersed points in the environment by viewing streaming video or view a wider swath of the environment but not know the temporal relations among the imagery you are viewing. We are currently investigating several potential solutions to these problems including coordinating path planning to incorporate techniques from the pursuit-evasion, monitoring boundaries to build probabilistic models of target locations, and providing operators control allowing them to trade-off temporal and spatial resolution.

References
[1]Olsen, D. and Wood, S. 2004. Fan-out: Measuring Human Control of Multiple Robots. In *Proceedings of Human Factors in Computing Systems* (CHI'04), Vienna, Austria: ACM Press, 231-238.

[2]Crandall, J., Goodrich, M., Olsen, D. and Nielsen, C. 2005. Validating human-robot interaction schemes in multitasking environments. *IEEE Transactions on Systems, Man, and Cybernetics, Part A,* 35(4), 438–449.

[3]Miller, C. and Parasuraman, R. 2007. Designing for flexible interaction between humans and automation: Delegation interfaces for supervisory control. *Human Factors* 49(1), 57-75.

[4]Scerri, P., Liao, E., Lai, L., Sycara, K., Xu, Y. and Lewis, M. 2005. Coordinating very large groups of wide area search munitions. *Theory and Algorithms for Cooperative Systems*, World Scientific, 451-480.

[5]Kira, Z. and Potter, M. 2010. Exerting Human Control Over Decentralized Robot Swarms. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*. IROS'10, 566-571.

[6]McLurkin, J., Smith, J., Frankel, J., Sotkowitz, D., Blau, D. and Schmidt, B. 2006. Speaking Swarmish: Human-Robot Interface Design for Large Swarms of Autonomous Mobile Robots. In *Proceedings of AAAI Spring Symposium*.

[7]Ding, X., Powers, M., Egerstedt, M., Young, S. and Balch, T. 2009. Executive Decision Support: Single-Agent Control of Multiple UAVs. *IEEE Robotics & Automation Magazine*, 16 (2), 73-81.

[8]Cooke, N., Pringle, H., Pedersen, H. and Connor, O. (Ed.) 2006. *Human Factors of Remotely Operated Vehicles*. Amsterdam, NL: Elsevier.

[9]Humphrey, C., Henk, C., Sewell, G., Williams, B. and Adams, J. 2007. "Assessing the Scalability of a Multiple Robot Interface," In *Proceedings of the 2nd ACM/IEEE International Conference on Human-Robotic Interaction (HRI'07)*, 239-246.

[10]Lewis, M., Wang, H., Chien, S., Velagapudi, P., Scerri, P. and Sycara, K. 2010. Choosing autonomy modes for multirobot search, *Human Factors*, 52(2), 225-233.

[11]Gugerty, L. and Brooks, J. 2001. Seeing Where You Are Heading: Integrating Environmental and Egocentric Reference Frames in Cardinal Direction Judgements. Journal of Experimental Psychology: Applied, 7 (3), 251-66.

[12]Alberts, D., Garstka, J., & Stein, F. (1999) *Network Centric Warfare: Developing and leveraging information superiority*, the Command and Control Research Program.

[13]Peruch, P., J. Vercher and G. Guthier (1995). "Acquisition of Spatial Knowledge through Visual Exploration of Simulated Environments." *Ecological Psychology* **7**(1): 1-20.

[14]Carpin, S., Wang, J., Lewis, M., Birk, A., and Jacoff, A. 2006. High fidelity tools for rescue robotics: Results and perspectives, *Robocup 2005 :Robot Soccer World Cup IX,* Springer LNAI vol 4020, Springer, 301-311.

[15] Carpin, S., Lewis, M., Wang, J., Balakirsky, S. and Scrapper, C. 2006b. Bridging the gap between simulation and reality in urban search and rescue. *Robocup 2006: Robot Soccer World Cup X, Springer, Lecture Notes in Artificial Intelligence*, 1-12.

[16]Lewis, M., Hughes, S., Wang, J., Koes, M. and Carpin, S. 2005. Validating USARsim for use in HRI research, In *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society*, Orlando, FL, 2005, 457-461.

[17]Pepper, C., Balakirsky, S. and Scrapper, C. 2007. Robot Simulation Physics Validation, In *Proceedings of PerMIS'07*.

[18]Taylor, B., Balakirsky, S., Messina, E. and Quinn, R. 2007. Design and Validation of a Whegs Robot in USARSim, In *Proceedings of PerMIS'07*.

[18]Zaratti, M., Fratarcangeli, M and Iocchi, L. 2006. A 3D Simulator of Multiple Legged Robots based on USARSim. *Robocup 2006: Robot Soccer World Cup X, Springer, LNAI*, 13-24.

[20]Carpin, S., Stoyanov, T., Nevatia, Y., Lewis, M. and Wang, J. 2006. Quantitative assessments of USARSim accuracy. In *Proceedings of PerMIS*.

[21]Owens, S., Sycara, K. & Scerri, P. 2009. Using immersive 3D terrain models for fusion of UAV surveillance imagery, Proceedings of AIAA InfoTech 2009.

[22]Draper, M., Calhoun, G., Nelson, J & Ruff, J. 2006"Evaluation of synthetic vision overlay concepts for UAV sensor operations: landmark cues and picture-in-picture". Proceedings of the Human Factors and Ergonomic Society 50th Annual Meeting, San Francisco, CA. 2006.

[23]Hunn, B. 2005. "The Human Challenges of Command and Control with Multiple Unmanned Aerial Vehicles".Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting. 2005.

[24]Calhoun, G., Draper, M, Abernathy, M. Patzek, M. & Delgado, F. 2005. "Synthetic vision system for improving unmanned aerial vehicle operator situation awareness". Enhanced and Synthetic Vision 2005, 5802(1), 2005.

[25]Drury, J. Richer, J., Racklie, N. & Goodrich, M. 2006.. "Comparing situation awareness for two unmanned aerial vehicle human interface approaches". In Proceedings of the IEEE International Workshop on Safety, Security and Rescue Robotics, 2006

[26]Velagapudi, P., Wang, J., Wang, H., Scerri, P., Lewis, M., and Sycara K. 2008. Synchronous vs. Asynchronous Video in Multi-Robot Search, Proceedings of first *International Conference on Advances in Computer-Human Interaction* (ACHI'08), 224-229.

[27]Velagapudi, P., Wang, H., Scerri, P., Lewis, M. & Katia Sycara 2009. Scaling Effects for Streaming Video vs. Static Panorama in Multirobot Search, *2009 International Conference on Intelligent Robots and Systems* (IROS'09), IEEE, 5874-5879.

[28]Wang, H., Kolling, A., Abedin, S., Lee, P., Chien, S., Lewis, M., Brooks, N., Owens, S., Scerri, P. & Sycara, K. 2011 Scalable target detection for large robot teams, Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction, 363-370.

[29]Brooks, N., Wang, H., Chien, S., Lewis, M., Scerri, P.,& Sycara, K. (2011) Asynchronous Control with ATR for Large Robot Teams, Proceedings of the 55th Annual Meeting Human Factors and Ergonomics Society (HFES'11).

[30]Abedin, S., Wang, H., Lee, P., Lewis, M., Brooks, N., Owens, S., Scerri, P. and Sycara, K. (2011) SUAVE: Integrating UAV Video Using a 3D Model 2011 IEEE International Conference on Systems, Man, and Cybernetics, (SMC'11)