Towards Effective User-Guided Robot Search (Extended Abstract)

Shahar Kosti
Computer Science Dept
Bar-Ilan University
shahar.kosti@gmail.com

Gal A. Kaminka
Computer Science Dept
Bar-Ilan University
galk@cs.biu.ac.il

David Sarne
Computer Science Dept
Bar-Ilan University
sarned@cs.biu.ac.il

Categories and Subject Descriptors

I.2.9 [**Artificial Intelligence**]: Robotics—*operator interfaces*

General Terms

Experimentation, Human Factors, Performance

Keywords

Exploration, Autonomous agents, Asynchronous interface, Human-robot interaction, Experimentation, Robotics

1. INTRODUCTION

In recent years there has been a growing interest amongst Human-Robot Interaction (HRI) researchers in the usage of robotic systems for exploration or search missions [6, 3, 4].

In the traditional approach to robot supervision by humans, knows as *camera guided teleoperation* [3], the operator navigates the robot while interpreting camera imagery. Both tasks require constant attention, and contribute to operator's workload. Automating the robot navigation task, as suggested in [5], is one possible step to reduce workload. Another step is to eliminate the need to view live video, thus turning the interface to use *asynchronous* video.

Recent work [3] on multi-robot interfaces investigated the usage of asynchronous operation schemes in the context of Urban Search And Rescue (USAR) missions [2]. With the asynchronous approach, robots explore the environment autonomously, while the operator is presented with recorded imagery (*asynchronous*) from the robots' camera, rather than live imagery (*synchronous*).

Asynchronous interfaces have several advantages. First, scaling up the number of robots is easier, since the operator no longer has to observe live video feeds from multiple robots. This can help to reduce operator's workload when the number of robots is increased [4]. Moreover, communication constraints can limit the ability to stream live video from the robots and to control them in real-time. This could severely affect the operator's ability to extract useful information and maintain spatial orientation [1]. With the asynchronous approach however, transmission of recorded images can be much more flexible. Images may be transmitted with a delay (as time no longer affects the user exprience), transmitted in bulks (e.g. where wireless reception is better), or even not transmitted at all (the operator would view the recorded images after the robot explored the area autonomously).

Two major challenges of asynchronous interfaces for robot-based search, are how to select the most relevant imagery to display, and how to display it to the operator. A recent approach suggests storing the images in a database and ordering them according to a utility value, which is computed by the image area that was not already seen [4]. Users can then view images according to that order, and navigate through images that were recorded near the selected image. This is a system-guided approach, as the image sequence to view is mainly determined by the system.

In contrast, we propose an asynchronous *user-guided* interface, in which the operator can view the recorded imagery of a given indoor environment, by selecting *points of interest* (POI) on a map. The interface provides the operator with highly-relevant images of the POI from several view points, after applying a dynamic filtering and ranking process over the recorded images. To enable selection of images based on points that appear in them, we develop efficient methods for storage and retrieval of images, possibly indexed by areas covered.

2. RELATED WORK

Multi-robot interfaces have been studied in various contexts, in particular USAR missions. Among these, we note the work of Velagapudi et al. [3], that compared two interfaces: *synchronous* and *asynchronous*. The former was similar to previous multi-robot interfaces (live video), while the latter had no live video display. Instead, an operator directed the search task by assigning robots with waypoints. After reaching the final waypoint, the robot would take a panoramic image of that location, which appeared as a new symbol on a map. The user would then click the symbol to view a panoramic view of the recorded location. Results showed better performance for the *asynchronous* interface when the number of robots increased.

A more recent work [4], which we used as a benchmark, compared two interfaces for multi-robot search: *streaming video* and *image queue*. The robots were autonomous but manual control was allowed. In the *image queue* interface, images from all robots are stored in a database and sorted by utility, which is calculated by the size of visible area that wasn't already seen. The images with the highest utility

Appears in: *Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France.*

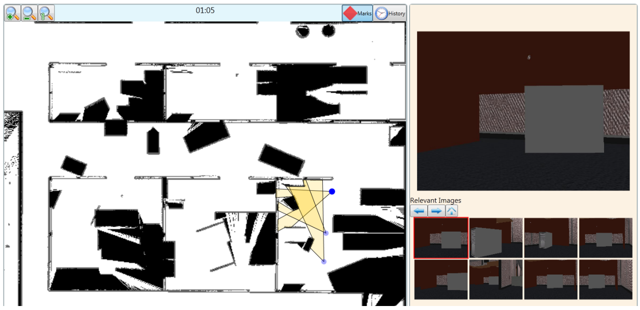

Figure 1: The interface after a POI (point of interest) was selected.

are presented on a "filmstrip". The user can select an image and view a "sub-queue" of nearby images (to get a better perspective). Results show similar performance in terms of the total number of victims found. However, with the "image queue" interface there were less errors (false positives and negatives) and the workload was lower.

3. POI-BASED USER INTERFACE

Our asynchronous interface enables users to view recorded imagery and mark locations on a map. In the context of a USAR mission, the marked locations can specify the approximate positions of disaster victims. The input to the interface is an image database collected by one or more autonomous robots, that provide an on-going stream of camera images and range scanner readings.

Navigating between recorded images is done in the following manner. The user selects a POI (point of interest) on the map by clicking on it. The system finds all camera images that cover the selected point, and ranks them. Figure 1 illustrates the interface after a POI was selected. The highest-ranked image is displayed in the upper-right corner. The bold dot on the map represents the robot location while recording current image. Other highly-ranked images are available for selection on the map, and displayed as smaller dots. All other images that cover the POI are displayed as thumbnails, below the current image.

Recorded images from all robots are saved in a single database. Each database entry contains the image and auxiliary information regarding the robot at the time of recording, such as location, orientation and range scanner readings. During retrieval, as the user clicks on a point p, the system must find all polygons that cover (contain) the point of interest, p. A specialized data structure allows to store the images efficiently and to quickly respond to such queries.

Normally, the point selected by the user would be covered by many images, possibly too many. In order to decrease the number of images the operator has to view, we apply a ranking process. The ranking process consists of grouping images by their view angle of the POI, and ranking the images in each group using a utility function. This allows to provide the user with relevant images of the POI from different perspectives. The utility function ranks images by considering parameters such as the area size covered by the image and the distance from the POI.

4. EVALUATION

To evaluate our interface we compared it with the state-of-the-art asynchronous interface for USAR missions by Wang et al. [4], originally called *image queue*. We have implemented this interface as described by the authors. The main difference between the interfaces is how the user navigates between images. Unlike our interface, the map is not used for this purpose, but only to identify and mark victims. Navigating between images is done with "Next" and "Previous" buttons, and the images are ordered by a utility value. The utility value is computed by the size of the unseen image area.

An experiment was conducted in a simulated environment with pre-recorded data, generated by USARSim. Participants were given the entire map and image database of an environment, after a robot explored it. Two simulated indoor environments were created, based on an office environment from the 2006 RoboCup Rescue competition finals. Human-like characters ("victims") in various postures were placed in each environment. The environments differed in area size, the number of recorded images and the average image space covered by a victim.

32 adults were recruited to this study. The experiment followed a between-subjects design, with each participant using only one of the interfaces. After a training session, participants were given 10 minutes in each environment to locate and mark as much victims as possible. We measured the number of *found* victims (correct marks), as well as several kinds of marking errors. We found our interface to perform better than the existing interface, in terms of the number of found targets for relatively large maps, suggesting that it might scale better.

5. REFERENCES

[1] R. P. Darken, K. Kempster, M. K. Kempster, and B. Peterson. Effects of streaming video quality of service on spatial comprehension in a reconnaisance task. In *Proceedings of the meeting of I/ITSEC*, Orlando, FL, 2001.

[2] I. Nourbakhsh, K. Sycara, M. Koes, M. Yong, M. Lewis, and S. Burion. Human-robot teaming for search and rescue. *Pervasive Computing, IEEE*, 4(1):72–79, 2005.

[3] P. Velagapudi, J. Wang, H. Wang, P. Scerri, M. Lewis, and K. Sycara. Synchronous vs. asynchronous video in multi-robot search. In *Proceedings of the First International Conference on Advances in Computer-Human Interaction*, pages 224–229. IEEE Computer Society, 2008.

[4] H. Wang, A. Kolling, N. Brooks, S. Owens, S. Abedin, P. Scerri, P.-j. Lee, S.-Y. Chien, M. Lewis, and K. Sycara. Scalable target detection for large robot teams. In *Proceedings of the 6th international conference on Human-robot interaction*, HRI '11, pages 363–370, New York, NY, USA, 2011. ACM.

[5] H. Wang, M. Lewis, P. Velagapudi, P. Scerri, and K. Sycara. How search and its subtasks scale in n robots. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction - HRI '09*, page 141, La Jolla, California, USA, 2009.

[6] H. A. Yanco and J. L. Drury. Rescuing interfaces: A multi-year study of human-robot interaction at the AAAI robot rescue competition. *Autonomous Robots*, 22(4):333–352, Dec. 2006.