Establishing Human Situation Awareness Using a Multi-Modal Operator Control Unit In An Urban Search & Rescue Human-Robot Team

Benoit Larochelle, Geert-Jan M. Kruijff
German Research Center for Artificial Intelligence
DFKI GmbH, Saarbrücken (Germany)

Nanja Smets, Tina Mioch, Peter Groenewegen
TNO
Soesterberg (The Netherlands)

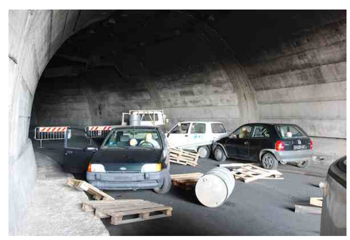
***Abstract*— Early on in a disaster it is crucial for humans to make an assessment of the situation, to help determine further action. Robots can potentially assist humans here, particularly when the hotzone is too dangerous for humans. Crucial in this human-robot team effort is that the system of robot and means of interacting with it afford the human to build up a sufficient awareness of the situation, to make an assessment. The paper investigates this issue from the viewpoint of the operator control unit (OCU). The paper presents the principles, design, implementation, and high-fidelity field trial results of an OCU for a human-robot team in an Urban Search & Rescue mission.**

I. INTRODUCTION

Robots have the potential to aid humans in the domain of Urban Search & Rescue (USAR), particularly in *reconnaissance and mapping* tasks [1]. In such tasks, robots support humans to make a situational assessment of the disaster site, early on in the operation. Important to making an assessment is that a human operator is able to gain sufficient awareness of the situation, through the deployment of one or more robots in the disaster hotzone.

In this paper, we discuss the design and implementation of an operator control unit (OCU) for remotely interacting with a single ground rover (UGV), and the results of field trials in high-fidelity setups. The human operator is located remotely, outside of the hotzone. His sense of the environment is thus entirely mediated by the OCU. The OCU combines spoken dialogue with a Graphical User Interface (GUI). The GUI displays a map visualization (from a 2D laser), video channels (from an omni-directional camera and from a pan-tilt monocular camera), and it provides an overview of the ongoing action and interaction (from planning and situated dialogue). Different views of the situation can be displayed at once. The OCU enables the operator to navigate the robot through tele-operation, verbal commands (small movement) and instructions (destination description), or waypoint selection in a map. While exploring an environment, the robot can use spoken dialogue to keep up a running commentary of what it sees and does.

We performed field trials with the OCU and a UGV in a high-fidelity simulation of a tunnel accident, analyzing the degree of operator situation awareness and the difficulty in establishing it (cognitive load). The tunnel accident use case presents a disaster in which a lorry lost its load while traveling through a tunnel, causing an accident involving multiple cars. The load included barrels containing possibly hazardous material, pipes, and pallets. Smoke at the accident scene prevents humans from going in, before having ascertained the source of the smoke. We recreated comparable instantiations of this use case at two emergency services training grounds, namely the *Scuola di Formazione Operativa* (SFO) in Montelibretti (Italy) of the Vigili del Fuoco, the Italian national rescue organization; and the training grounds of the Fire Department of the City of Dortmund (FDDO). See Figure 1. At SFO we performed initial field exercises (September 2010), the evaluations took place at FDDO (January 2011).



(a) SFO (Italy)

(b) FDDO (Germany)

Fig. 1. Tunnel accident use case

Contributions: Several studies have reported on the use of OCUs in USAR mission, e.g. [2], [3], [4], the issue of situation awareness in OCUs for human-robot teaming [5], and the typical problems that occur when deploying robots in complex environments such as urban disasters [6], [7]. In this paper, we build on these insights and investigate the difficulty for a human operator to build up situation awareness while using a robot *under time pressure*. We particularly look at the operator's cognitive load and his ability to make a situational assessment, while performing under stress and having to deal with occasional disruptions in system functionality.

II. BACKGROUND

Experience in the field with a variety of OCUs for operating tele-operated and semi-autonomous robots in human-robot teams has led to several sets of guidelines as to how "user interfaces" should be designed, e.g. [8], [9], [5]. These guidelines focus on making interaction *efficient*. Efficiency is seen as a min/max balance: we try to minimize negative impact on a human's cognitive work load, and maximize the task performance of the human-robot team.

The guidelines are formulated against a large body of observations on design issues for OCUs. Chen et al [10], [9] provide a detailed overview of these issues, with a particular focus on tele-operation. Human performance issues in tele-operation typically arise in remote perception, and remote manipulation (which includes navigation). Limiting factors include the operator's ability to maintain situational awareness by building up a mental model of a remote environment, and his motor skills to manually control the robot [11], [12]. Operating under realistic circumstances like USAR typically adds to the problems, as one has to content with system failures. [6] provide a detailed overview of such failures. The authors conclude that reliability of UGV performance in the field tends to be low, with common causes including unstable control systems, low scalability, limited wireless communications range, and insufficient bandwidth for video feedback. These issues cut across remote perception and manipulation. Chen et al [9] discuss a wide variety of possible solutions to these problems, particularly for OCUs for tele-operation. The design principles behind these solutions are captured by the guidelines proposed by Goodrich & Olsen [8] and Riley et al [5]. Both sets of guidelines deal with information management (including attentional processes) and aspects of decision-management. Also, both assume the possibility for adaptive automation, and its effect on shared control.

Goodrich & Olsen [8] proposed seven principles to make interactions between a robot and a human more efficient. The human is assumed to be at a remote location, i.e. these principles first of all apply to OCU design. The guidelines particularly focus on the presentation and use of information about the environment: The OCU should mediate between the *environment* and the human. The robot ideally acts as an extension of the human, providing a sense of remote presence. Riley et al [5] also adopt this view that a remotely operated robot facilitates the projection of human presence and intent into the environment of the robot. They focus particularly on how an OCU can then support situation awareness. Riley et al argue that situation awareness is the crucial factor in the successfully performing a mission. They adopt a model of human-robot interaction that encapsulates Endsley's definition of situation awareness (SA; Level 1: perception, Level 2: comprehension, Level 3: prediction) [13]. The model makes explicit the interaction between SA and the basis that it provides for decision-making and acting, and various factors that make up the ecology of the interaction: task/mission, environment, system, individual & team, and external world & user interface design. Riley et al then present guidelines for the design of user interfaces to appropriately deal with each of these factors, from the viewpoint of how they impact SA and performance.

In the next section, we present user requirements, our design, and a first implementation of the design, which are informed by these guidelines.

III. Design and Implementation

We adopt a user-centric methodology for designing and testing the OCU, working closely with end users in emergency services (e.g. FDDO and Vigili del Fuoco). Early meetings with VVFF and FDDO resulted in the elicitation of design requirements for the OCU, the most significant presented in Table I. This process was done before a software technology was chosen, and before a GUI was first imagined/drawn. This methodology helped to avoid biases towards pre-conceived designs, thus encouraging designs that matched more closely the user requirements.

Software Requirements
Give clear fluid camera images for navigation
Provide high-quality images for inspection
Allow multiple operators
Allow multiple robots
Require little training or specialized knowledge
OCU Hardware Requirements
Resist Western European weather (heat, cold, rain, snow)
Resist USAR environments (rocks, dust, humidity, high heat)
Usable with fire-fighting gloves
Viewable in the sun or in the dark
Robot Hardware Requirements
Provide pan-tilt-zoom camera
Provide a substitute to cameras during low-visibility (laser, infrared)
Navigate on USAR terrain (rubble, oil, small obstacles, stairs)
Have sufficient battery power for typical interventions
Able to communicate with ad hoc wireless networks

TABLE I
Main User Requirements

Although in the tunnel accident use case humans operate outside of the hotzone, they can still be in a hostile environment. Thus robust hardware is required for using the OCU in the field. We selected the Panasonic Toughbook CF series, for its higher level of certification, extra protection against rain, dust, shocks, for its extra battery life, and for its versatile hardware configurations. The Toughbook CF-19 and CF-31 offer a 1024 X 768 single-touch monitor, 10.4" and 13.1" in size. The OCU design included these physical constraints as requirements, allowing now for a use with a small screen and without a keyboard or mouse.

The basic design principle of the OCU is that the users are not necessarily accustomed to using computers and that they will be using the system under less than ideal conditions (i.e. varying cognitive load, high stress, no comfortable "office setting"). Natural and intuitive ways of interacting with the system must therefore be possible. Ideally, all modalities (text, speech, clicking, etc.) should have the same salience, so that users are not biased to use one over the other. Therefore, the term multi-modal was the main focus all along the design phase, rather than adding extra modalities to a previously existing "menus-and-buttons" GUI. A sketch of the main screen is illustrated in Figure 2.

The two main modes of input are by voice and by touch (finger, mouse). For example, making the robot move forward can be done in several ways. The user can click a motor control widget to manually move the robot. The user can also say "Move forward" and the robot will move an appropriate distance. Finally, the user can click on the desired destination on a map or video feed and the robot will go to that location. These different modes implicitly change

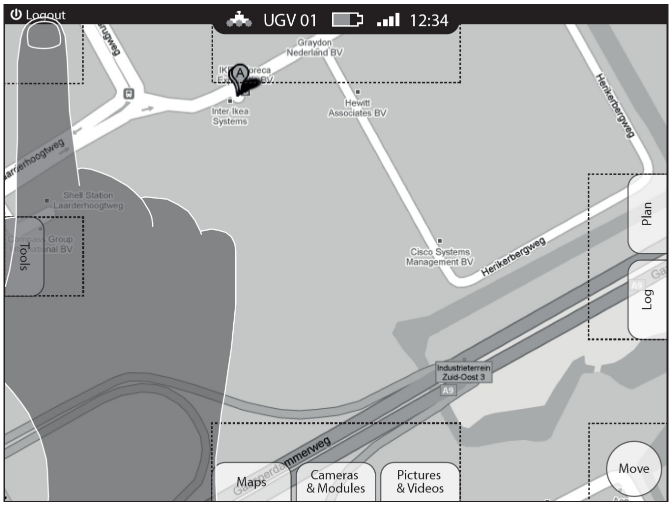


Fig. 2. Overall GUI design, with hand to show scale

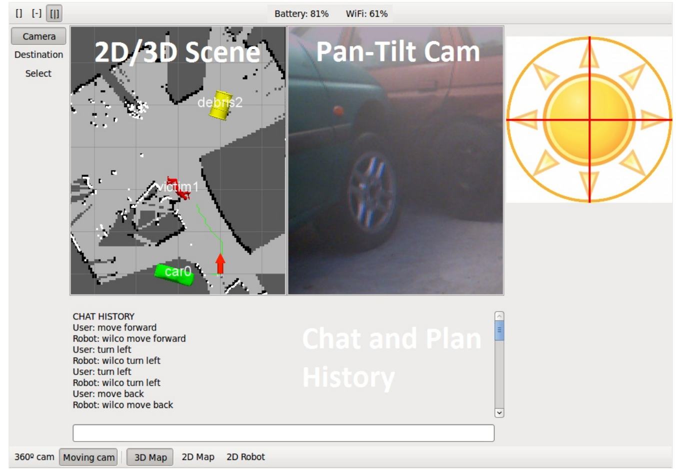


Fig. 3. Current State of the OCU Implementation

the robot's autonomy level, from semi-autonomous when clicking on the map to fully tele-operated when using the control widget. Since no explicit switching between modes is required, cognitive load is minimized, respecting the first and third principle of [8].

The three main parts of the GUI, with respect to information exchange, are the video feeds, which present the area surrounding the robot, the map, which is a 3D representation of the environment, and the robot's plan (its tasks and goals). Figure 3 shows the current state of implementation of the OCU. The video cameras typically are the main mode for gathering new information. They are meant for the user to navigate the robot and to detect new information, while allowing also the robot to detect specific items, such as cars and victims. The map is used to keep a memory of the environment and to represent it under any angle, in 2D or 3D. This representation reduces the cognitive load on the user by reducing what has to be remembered, and by representing it from the point of view that is the most useful for the user. This feature follows the sixth principle of Goodrich & Olsen [8]. The plan also reduces the cognitive load because the user does not need to micro-manage every action of the robot, and does not have to wonder why the robot is performing a certain action. A higher-level SA is thus built, as suggested by the fifth guideline of Riley et al [5]. It is crucial that the plan be always available, easily understood, synchronized with the robot's actions, and user-modifiable; otherwise, a problem of human-out-of-the-loop could develop and increase the cognitive load and reduce operator performance [14], [8].

NIFTi aims for a single integrated cognitive robot architecture that is above the various underlying technologies and software component distribution. Thus, most system components use the same data structures, which are implemented in both the Robotics Operating System (ROS) [15] message language and in the Specification Language for the Internet Communication Engine (SLICE). Then, C++, Java, and Python equivalents are generated for use by the system components. Beyond a central data structure, the system is highly distributed (see Figure 4). ROS is used to exchange data between components, allowing the components to be started and stopped at any time and from any computer on the network. Most components are completely independent, in that they run even if other parts of the system is down. For example, if the mapping algorithm stops working, the current map will continue being displayed in the OCU. In the same manner, if the operator shuts down the OCU, the robot will keep on sending maps. The data is sent in a peer-to-peer manner, but a ROS dispatcher (`roscore`) must be running at all times. This is a serious problem because a loss of signal or a crash would halt the system: multi-core solutions are being investigated.

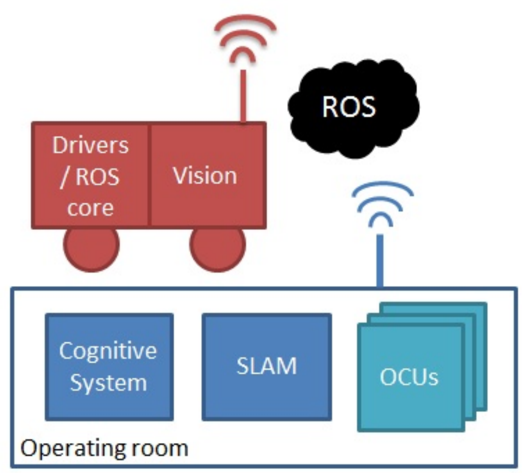


Fig. 4. Overall System Architecture

The OCU is thus an independent component, which connects to different aspects of the architecture through ROS. Multiple instances of the OCU can run simultaneously for multiple operators. The implementation of the

OCU is based on the ROS `rviz` tool, to which different visualization and operation functionalities were added. In addition, two computers run in the operating room and two on the robot. The first one in the room uses the CAST [16] technology for the cognitive sub-systems (e.g. situated dialogue, memory, reasoning, and planning). The other one runs mapping algorithms, such as Simultaneous Localization and Mapping (SLAM). On the robot, one computer runs the hardware drivers and control components, as well as the ROS dispatcher. The other one runs visual detection algorithms, such as victim detection.

IV. Field Trials and Results

At the training area of FDDO, we organized a week-long field trial with end users. Firemen controlled the robot from a remote location outside of the hotzone to explore a tunnel accident. The purpose of this field trial was to gain a better insight in the ways in which humans and robots might collaborate when exploring disaster areas. The focus was on remote interaction. We analyzed how the operator's cognitive task load and situation awareness vary, as he explores the accident under time pressure. We provide here results based on field trials with three operators, for an overall runtime of 38 minutes. Although the data does not provide sufficient ground for statistically significant observations, it is commensurate to many other real-life field trials like [2].

The physical environment was created inside a garage that can normally hold 3 fire trucks (approx. $10m \times 15m$). One third was used for training users with the robots, and the remaining space was closed off. 15 key elements were carefully inserted in the scene: 3 moderately damaged cars, 1 motorcycle, 5 dummies, 3 back-pack-sized yellow containers, and 3 danger signs (flammable/explosive) (see Figure 1b). Each operator was given 15 minutes to identify and locate in the environment as many elements as possible. The operators spent 30-60 minutes before each run training with the robot in exercises such as slalom and object detection.

During the field trials, often cited technical problems [6] also occurred here (uncontrolled). These included signal loss, insufficient power, insufficient bandwidth for video-based feedback, and robot blockage in the environment. Due to these problems, two of the three runs were cut short, thus providing us with data for runs lasting 10, 13, and 15 minutes. The results must be interpreted carefully, as scenario duration was a critical performance factor. Table II presents the raw data about user performance. The three columns refer to the three runs, and show how many elements were detected. The table also indicates how well the operators thought that they performed.

We also analyzed the operators' tasks during the runs. Burke et al [2] showed that typically only 44% of the operator's activities are directly related to navigation, and that there is great variability in how much time a robot is actually moving. Table III shows similar figures were observed for our operators.

Figures 5-7 present a deeper insight into what happened during one of the scenarios. On the horizontal axis is plotted the time in minutes. Below this axis are shown the main three inputs to the OCU and the moments during which they were functional. Two vertical axes help to show three pieces of information. On the right, the axis is gradated from 0 to 120 seconds. The 'navigation time' bars show for how many seconds the user was navigating the robot in the previous 2-minute block. The left axis ranges from 0 to 8 and serves a dual purpose. During the scenario, the operators had to indicate their workload on a hand-held device every 2 minutes. In the graphs, the answers given by each operator are shown, along with his response time in seconds.

Element	**10min**	**13min**	**15min**
Cars	3 / 3	3 / 3	3 / 3
Motorcycle	0 / 1	1 / 1	1 / 1
Victims	1 / 5	1 / 5	2 / 5
Containers	1 / 3	2 / 3	3 / 3
Danger signs	0 / 3	1 / 3	2 / 3
Total	5 / 15	8 / 15	11 / 15
Thinks he did well (Lickert 1-7)	3	3	6

TABLE II
Elements Found in the Scenario

Observations	**10min**	**13min**	**15min**
% time moving	62%	47%	52%
Average time between moves	42.9s	33.3s	30.0s
Average duration of moves	26.5s	20.2s	15.7s

TABLE III
Time when the robot moved vs. remained stable

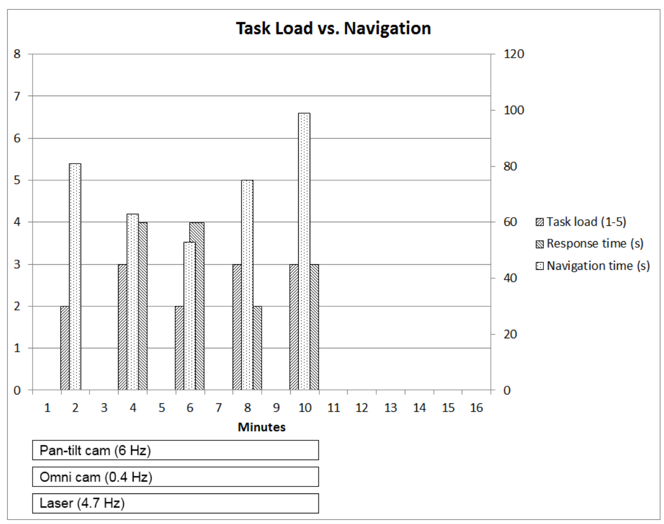


Fig. 5. Task load during the 10-minute run

During the runs, the operators were given a whiteboard and asked to draw a map of the scene and indicate the elements found. They were not asked to draw to scale or even to draw the size of the objects, but only to show the qualitative spatial relationships between the elements. Figure 8 shows the drawings, super-imposed with the actual positions of the elements. 'M' indicates the motorcycle, 'C' a dangerous container, and 'V' a victim.

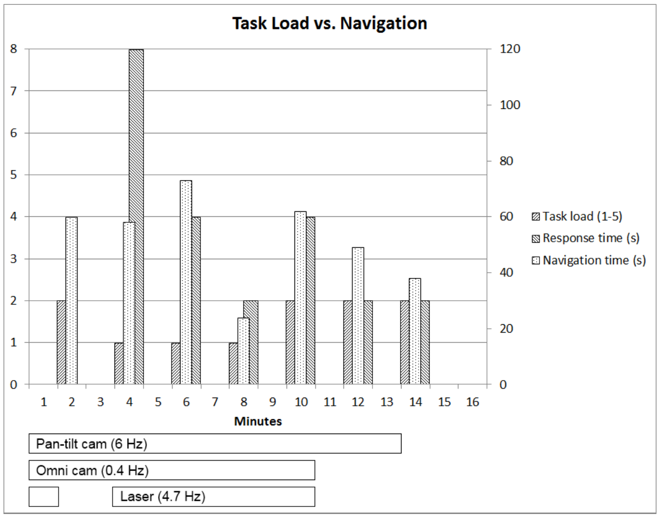


Fig. 6. Task load during the 13-minute run

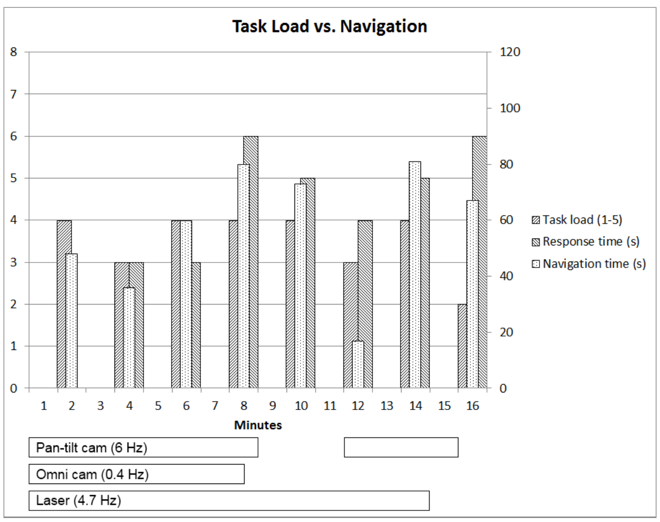


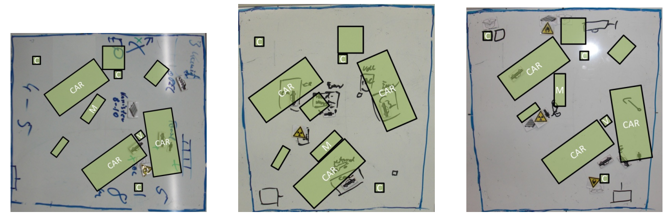
Fig. 7. Task load during the 15-minute run

The robot also built up a map as it was navigating. With a 2D laser, it generated a metrical map that is shown in Figure 9. A line displays the path taken by the user, and an arrow indicates the final position. The 13-minute run was interrupted by a laser failure, so the metrical map shown represents only the last 6 minutes.

In addition to these observed results, questionnaires were given to the users to assess their feelings and observations concerning the NIFTi system. Table IV shows the questions and answers about the operator's situation awareness. The answers represent a Lickert scale from '1 Totally disagree' to '7 Totally agree'.

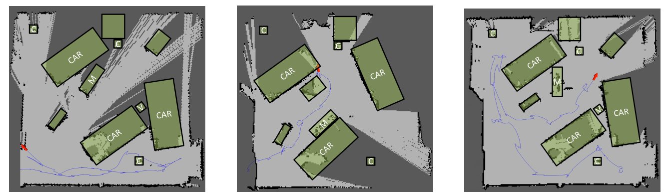
V. Reflections

Cognitive Load: To analyse the cognitive load, we must look at what the operators were doing during the scenarios. As shown in Table III, they spent around half of the time navigating in stretches of 15 to 30 seconds, pausing for 30 to 45 seconds at a time. Further analysis [17] shows that



(a) 10-minute run (b) 13-minute run (c) 15-minute run

Fig. 8. Sketch maps with overlays



(a) 10-minute run (b) 13-minute run (c) 15-minute run

Fig. 9. Metrical maps with overlays

out of these 30-45 second long "observations time blocks", at least two thirds of the time is spent looking directly at the environment elements, such as danger signs or car windows that could display trapped victims. The rest of the time is spent searching for viable navigation pathways. In absolute numbers, Table III shows that users spent 6-8 minutes navigating, and had only 4-7 minutes left for observation. During this observation time, up to two minutes was spent on navigation-related issues, which brings the total time devoted to navigation tasks to 8-10 minutes and the total time for real observation to 2-5 minutes.

We expect that by giving more autonomy to our robot, the operators will be able to spend significantly more time looking at the environment. During the field trials, the path planner seemed to have difficulties handling the complex environment, forcing the operators to manually tele-operate. Crandall and Goodrich [18] demonstrate that more navigational autonomy makes the system more robust to the user attending other tasks. In addition, we expect that robot autonomy will decrease the operator's cognitive load. Although the data are sparse, Figures 5–7 seems to indicate a correlation between the amount of time spent navigating and the task load. These results are in line with [19], [20], which conclude that manually tele-operating a robot leads to a higher workload than managing a fully autonomous robot. Birk & Carpin [21] also call for more autonomy to overcome manual navigation problems.

Reconnaissance Results: Table II shows how many key elements the operators found. It is clear that the robot's 50 cm high cameras do not offer a great vantage point, especially for looking into car windows, as only 1 or 2 victims out of 5 were found. These results indicate that some help is necessary to increase performance. We are currently developing different kinds of user support features, such as object detection and a UAV to be used as roving sensor. These objects, detected by the UGV or UAV, will be highlighted in the camera views, and permanently shown in

Question	10min	13min	15min
In general, I had a good idea of the environment	4	5	4
In the beginning, it was difficult to build a picture of the environment	4	4	4
Sometimes, I lost track of what was going on in the environment	1	2	2
I knew the whole time where the robot was	4	5	5
I know the whole time in which direction the robot moved	4	5	5
When I saw an object, I knew its position relative to me	4	5	4
When I saw an object, I knew its position in space (e.g. on the map)	5	5	5

TABLE IV

QUESTIONNAIRE ON SITUATION AWARENESS (LICKERT 1-7)

the 2D/3D map. The latter will allow the operator to release the cognitive resources that would otherwise be required to remember the existence and position of a detected element [8].

Situation awareness: Although it is difficult to evaluate sketch maps, Figure 8 and Figure 9 seem to indicate that the users had a relatively good mental map of the environment. Other than the first user who misplaced a car on his map, other objects are mostly drawn within one meter of their actual locations. We did not expect this level of precision, but we believe that these results are due to the combination of simultaneous camera feeds and the 3D scene. The results agree with the perceived situation awareness from the operators, as shown in Table IV.

CONCLUSIONS

Combining OCU design principles from Chen et al, Riley et al, and Goodrich & Olsen with requirements and feedback elicited from end-users at VVFF and FDDO, we designed a multi-modal OCU to be used in USAR scenarios. The design plans for 2D and 3D views of the scene, as well as video feeds from two cameras. We organised field trials and let fire fighters operate the robot in high fidelity USAR scenarios. The results show that victims are difficult to find only with the UGV, but other key elements are more visible. The robot had to be manually tele-operated, which forced the operators to spend most their time on navigation-related tasks. The operators found the system useful and easy to use, and claimed to have a relatively good situation awareness. However, their SA could be improved by a better display, a richer feature set for the robot-generated metrical map, and more autonomy.

ACKNOWLEDGEMENTS

This research has been financed by the European Integrated Project NIFTi under contract number FP7 ICT 247870, Cognitive Systems and Robotics. The authors would also like to thank the end-user organizations, Vigili del Fuoco and FDDO, for their continuing support.

REFERENCES

[1] R. Murphy, S. Tadokoro, D. Nardi, A. Jacoff, P. Fiorini, H. Choset, and A. Erkmen, "Search and rescue robotics," in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Springer Verlag, 2008, pp. Part F, 1151–1173.

[2] J. Burke, R. Murphy, M. Coovert, and D. Riddle, "Moonlight in Miami: An ethnographic study of human-robot interaction in USAR," *Human Computer Interaction*, vol. 19, no. (1–2), pp. 85–116, 2004.

[3] R. Murphy, "Human-robot interaction in rescue robotics," *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, vol. 34, no. 2, pp. 138–153, 2004.

[4] R. Murphy and J. Burke, "The safe human-robot ratio," in *Human-Robot Interactions in Future Military Operations*, ser. Human Factors in Defence, M. Barnes and F. Jentsch, Eds. Ashgate, 2010, pp. 31–49.

[5] J. Riley, L. Strater, S. Chappell, E. Connors, and M. Endsley, "Situation awareness in human-robot interaction: Challenges and user interface requirements," in *Human-Robot Interactions in Future Military Operations*, ser. Human Factors in Defence, M. Barnes and F. Jentsch, Eds. Ashgate, 2010, pp. 171–192.

[6] J. Carlson and R. Murphy, "How UGVs physically fail in the field," *IEEE Transactions on Robotics*, vol. 21, no. 3, pp. 423–437, 2005.

[7] J. Chen, P. Durlach, J. Sloan, and L. Bowens, "Human robot interaction in the context of simulated route guidance missions," *Military Psychology*, vol. 20, no. 3, pp. 135–149, 2008.

[8] M. A. Goodrich and J. D. R. Olsen, "Seven principles of efficient interaction," in *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, 2003, pp. 3943–3948.

[9] J. Chen, E. Haas, and M. Barnes, "Human performance issues and user interface design for teleoperated robots," *IEEE Transactions on Systems, Man and Cybernetics: Part C: Applications and Reviews*, vol. 37, no. 6, pp. 1231–1245, 2007.

[10] J. Chen, E. Haas, K. Pillalamarri, and C. Jacobson, "Human-robot interface: Issues in operator performance, interface design, and technologies," Army Research Laboratory, Tech. Rep. ARL-TR-3834, July 2006.

[11] T. Fong, C. Thorpe, and C. Baur, "Multi-robot remote driving with collaborative control," *IEEE Transactions on Industrial Electronics*, vol. 54, no. 4, pp. 699–704, 2003.

[12] D. Woods, J. Tittle, M. Feil, and A. Roesler, "Envisioning human-robot coordination for future operations," *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, vol. 34, no. 2, pp. 210–218, 2004.

[13] M. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors*, vol. 37, no. 1, pp. 32–64, 1995.

[14] R. Parasuraman, M. Barnes, and K. Cosenzo, "Adaptive automation for human-robot teaming in future command and control systems," *International Journal of Command and Control*, vol. 1, no. 2, pp. 43–68, 2007.

[15] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Ng.

[16] N. Hawes and J. Wyatt, "Engineering intelligent information-processing systems with cast," *Advanced Engineering Informatics*, vol. 24, no. 1, pp. 27 – 39, 2010.

[17] S. Keshavdas, H. Zender, G. Kruijff, M. Liu, F. Colas, K. Zimmermann, T. Petricek, and T. Svoboda, "Use of functional mapping in supporting operators in human-robot teams for assessing urban disasters," February 2011, unpublished report.

[18] J. W. Crandall and M. A. Goodrich, "Measuring the intelligence of a robot and its interface," in *Proceedings of PERMIS 2003*, 2003.

[19] S. R. Dixon, C. D. Wickens, and D. Chang, "Comparing quantitative model predictions to experimental data in multiple-uav flight control," in *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting*, 2003, pp. 104–108.

[20] S. P. Schipani, "An evaluation of operator workload during partially-autonomous vehicle operations," in *Proceedings of PerMIS*, 2003.

[21] A. Birk and S. Carpin, "Rescue robotics - a crucial milestone on the road to autonomous systems," *Advanced Robotics*, vol. 20, no. 5, pp. 595–605, 2006.